

ÍNDICE:

GLOSARIO

INTRODUCCIÓN, METODOLOGÍA Y OBJETIVOS XIX**PRIMERA PARTE: LOS RETOS DEL PERIODISTA DIGITAL****CAPÍTULO 1- ¿QUÉ ES INTERNET?**

1.1 BREVE HISTORIA DE INTERNET.....	49
1.2 FUNCIONAMIENTO DE LA RED.....	52
1.3 SERVICIOS DISPONIBLES EN LA RED:	
CORREO ELECTRÓNICO, FTP, TELNET, NEWS, IRC.....	56
1.4 GOPHER.....	66
1.5 HISTORIA DEL HIPERTEXTO.....	67
1.6 LA WORLD WIDE WEB.....	69
1.7 . DIFERENCIAS ENTRE LA RED E INTERNET.....	73
1.8 INTERNET2.....	75
1.9 INTERNET EN ESPAÑA.....	79

CAPÍTULO 2- EL PERIODISMO A TRAVÉS DE INTERNET

2.1 INTRODUCCIÓN.....	83
2.2 DE LA GALAXIA DE GUTEMBERG A LA GALAXIA DE MCLUHAN.....	86
2.3 ELEMENTOS DEFINITORIOS DEL PERIODISMO <i>ONLINE</i> :.....	96
2.4 LA PERSONALIZACIÓN DE LA INFORMACIÓN.....	114
2.5 LA REDACCIÓN <i>ONLINE</i>	121
2.6 INCONVENIENTES DEL PERIODISMO ELECTRÓNICO EN LA ACTUALIDAD.....	132
2.7 PERIÓDICOS EN INTERNET EN ESPAÑA. ¿GRATUITOS O DE PAGO?	134

CAPÍTULO 3- EL PAPEL DEL PERIODISTA DIGITAL EN LA SOCIEDAD ACTUAL

3.1 INTRODUCCIÓN.....	138
3.2 EL PERIODISTA DIGITAL FRENTE A LAS FUENTES DE INFORMACIÓN.....	142
3.3 EL PERIODISTA DIGITAL FRENTE AL RECEPTOR DE LA INFORMACIÓN.....	143
3.4 EL PERIODISTA DIGITAL FRENTE AL NUEVO MEDIO.....	144
3.5 EL PERIODISTA DIGITAL FRENTE A LOS CONTENIDOS.....	148
3.6 INTERNET COMO HERRAMIENTA DE TRABAJO PARA EL PERIODISTA ACTUAL.....	155

SEGUNDA PARTE: INTERNET COMO FUENTE DE INFORMACIÓN PARA EL PERIODISTA

CAPÍTULO 4- CÓMO BUSCAR EN LA WORLD WIDE WEB

4.1 INTRODUCCIÓN.....	159
4.2 LA BÚSQUEDA DE INFORMACIÓN EN INTERNET.....	163
4.3 ¿QUÉ CONTIENE LA WORLD WIDE WEB?.....	168
4.4 LA CREDIBILIDAD Y LA VALORACIÓN DE LA INFORMACIÓN HALLADA EN LA WEB.....	170
4.5. LOS COMIENZOS DE LA NAVEGACIÓN POR LA RED.....	171

CAPÍTULO 5-RECURSOS DE BÚSQUEDA

5.1. INTRODUCCIÓN.....	173
5.2. LAS PRIMERAS HERRAMIENTAS DE BÚSQUEDA.....	173
5.3.NAVEGAR Y BUSCAR.....	176
	180
	185

5.4. BUSCADORES GENERALES DE INTERNET.....	
5.5. MOTORES DE BÚSQUEDA.....	
5.5.1 FUNCIONAMIENTO DE LOS MOTORES DE BÚSQUEDA.....	185
5.5.2. LOS BUSCADORES GENERALES Y EL TAMAÑO DE LA RED	208
5.5.3. METABUSCADORES.....	210
5.5.4. SOLAPAMIENTO DE LOS MOTORES DE BÚSQUEDA.....	214
5.6 DIRECTORIOS TEMÁTICOS GENERALES.....	215
5.7 ANILLOS WEB.....	218
5.8 PORTALES HORIZONTALES, VERTICALES Y CANALES TEMÁTICOS.....	221
5.9 OTROS RECURSOS DE BÚSQUEDA.....	223
5.9.1 GUÍAS TEMÁTICAS ESPECIALIZADAS.....	223
5.9.2 BUSCADORES ESPECIALIZADOS.....	224
5.9.3 BUSCADORES AUTOMÁTICOS DE SERVICIOS DE INTERNET.....	231
5.9.4 BUSCADORES POR SOPORTES DE INFORMACIÓN.....	234
5.9.5 FUENTES PERIODÍSTICAS EN INTERNET.....	236
5.9.6 OTROS RECURSOS DE REFERENCIA.....	242
5.9.7 MEJORES RECURSOS DE BÚSQUEDA EN CADA CASO.....	247
5.10 EL FUTURO DE LOS BUSCADORES.....	259

**TERCERA PARTE: LA RED PROFUNDA Y SU UTILIDAD PARA EL
PERIODISTA**

CAPÍTULO 6- LA RED PROFUNDA

6.1.INTRODUCCIÓN.....	
6.2 EL TAMAÑO DE LA RED.....	
6.3 RED PROFUNDA VERSUS RED SUPERFICIAL.....	276
6.4 LA RED PROFUNDA O INTERNET INVISIBLE.....	280
6.5 LOS MOTORES DE BÚSQUEDA Y LA RED PROFUNDA.....	286
6.6 SEIS NIVELES DE PROFUNDIDAD DE LA INFORMACIÓN DIGITAL.....	295
6.7 LOS ‘AGUJEROS NEGROS’ DE INTERNET.....	300
6.8 LA BÚSQUEDA EN LA RED EN BASE A LA NECESIDAD DEL USUARIO.....	303
6.9 ANÁLISIS DE LOS MAYORES SITIOS WEB DISPONIBLES EN LA RED PROFUNDA.....	310
6.10 EL FUTURO DE LA RED PROFUNDA.....	325

CUARTA PARTE: ANEXOS, CONCLUSIONES, GLOSARIO Y BIBLIOGRAFÍA:

ANEXOS:

1. LISTADO DE RECURSOS DE BÚSQUEDA DE LA RED PROFUNDA.....	333
2. ENTREVISTAS REALIZADAS POR EL AUTOR	
2.1 INTRODUCCIÓN A LAS ENTREVISTAS.....	451
2.2 MARIO TASCÓN: DIRECTOR DE CONTENIDOS DE PRISACOM Y EXDIRECTOR DE LA EDICIÓN DIGITAL DEL DIARIO <i>EL MUNDO</i>	454
2.3 MARILÓ RUIZ DE ELVIRA: DIRECTORA DE CONTENIDOS DE INICIA (PORTAL DEL GRUPO PRISA) Y EXDIRECTORA DE LA EDICIÓN DIGITAL DEL DIARIO <i>EL PAIS</i>	461

2.4 RICARDO FORNÁS: CREADOR DEL BUSCADOR DE BUSCADORES “BUSCOPIO”	
2.5 OTRAS ENTREVISTAS:	
2.5.1 RAMON SALAVERRIA: : DIRECTOR DEL LABORATORIO DE COMUNICACIÓN MULTIMEDIA (MMLAB) DE LA UNIVERSIDAD DE NAVARRA.....	471
2.5.2 LEONARD KLEINROCK • 'PADRE' DE LA RED.....	477
3. CRONOLOGÍA DE INTERNET DE HOBBS.....	480
CONCLUSIONES	522
BIBLIOGRAFÍA	526

GLOSARIO DE TÉRMINOS RELACIONADOS CON INTERNET

ActiveX.- Concepto de arquitectura de sistemas desarrollado por Microsoft como sucesor de OLE y COM y presentar una alternativa al lenguaje Java para el desarrollo de programas para la red Internet. Presenta mayores flexibilidad y rendimiento que el java pero no es independiente de la plataforma pues sólo funciona en los sistemas operativos de Microsoft.

ADN.- Traducido literalmente como Red Digital Avanzada, se refiere a las líneas dedicadas de 56 Kbps.

Anfitrión.- [En inglés, Host.] Una computadora en una red. Se aplica en vez del término en desuso "nodo" que se utiliza en el lenguaje de definición de documentos.

Applet.- Programa en lenguaje java que se utiliza en las páginas de Internet para conseguir efectos especiales que el lenguaje html no puede realizar.

Arañas.- [Spiders] Programa automatizado que busca en el Internet.

Archie.- Herramienta de software para localizar archivos almacenados en sitios FTP anónimos.

ARPANet.- Siglas de la expresión inglesa Advanced Research Projects Agency Network (Red de la Agencia de Proyectos Avanzados de Investigación), red precursora a la Internet. Se desarrolló a finales de la década de los años 60 por parte del Departamento de Defensa de los Estados Unidos en la experimentación de una amplia red funcionara tras un ataque o guerra nuclear.

ASCII.- Siglas de la expresión inglesa American Standard Code for Information Interchange (Codificación Americana Normalizada para el intercambio de Información). Es la norma mundial para la codificación usada en las computadoras a fin de representar los caracteres requeridos para la comunicación entre máquinas. Hay 128

códigos normalizados ASCII, cada uno de los cuales se puede representar con un número binario de 7 dígitos.

Autoridad de Certificación.- En inglés "Certificate Authority". Un emisor de Certificados de Seguridad para las conexiones SSL

ASP.- [Active Server Pages]. Sistema de programación desarrollado y propiedad de Microsoft para realizar las tareas antes realizadas por los CGI.

Atado.- En inglés Attachment, también llamado Anexo, es un documento adicional incluido en un correo electrónico.

Banner.- Anuncio promocional colocado intencionalmente en una página de Internet.

Banner dinámico.- Banner con imágenes en movimiento para atraer la atención del usuario de la página.

B2B.- Del inglés "Business to business", de negocio a negocio, o bien, de empresa a empresa.

B2C.- Del inglés "Business to consumer", de negocio al consumidor.

Baud.-Unidad que representa la velocidad de transferencia de la información. Es equivalente a bytes por segundo.

bit.- Unidad elemental de la información. El nombre proviene del inglés, "binary digit" o dígito binario. Originalmente explicada por Sócrates en los Diálogos de Platón, habiéndole llamado "diada" que sería su denominación óptima.

byte.-Conjunto de 8 bits. Puesto que 8 bits es la mínima cantidad requerida para representar los símbolos alfanuméricos Originalmente explicada por Sócrates en los Diálogos de Platón, habiéndole llamado "diada" que sería su denominación óptima.

bps.-Iniciales de bits por segundo.

Buscador.- [Search site] Sitio de Internet que contiene una amplia base de datos sobre las páginas que se encuentran en la red. El mas popular es Yahoo, que opera por afinidad semántica, aunque sean más efectivos otros buscadores como Altavista o Excite que operan por afinidad textual. Para lograr su objetivo los buscadores utilizan arañas.

Cache.- Espacio de almacenamiento temporal que el navegador emplea para almacenar los archivos (textos e imágenes) que recibe del Internet. Cuando se vuelve a visitar una página, el navegador rápidamente obtiene los archivos desde el cache en lugar de obtenerlos de la localidad remota donde originalmente los encontró. Se habla del cache de disco cuando los datos se guardan en el disco duro y de cache de memoria cuando se almacenan en la RAM de la computadora.

Certificado de Seguridad.- Archivo de texto usado por el protocolo SSL para establecer una conexión segura. La información en los certificados de seguridad incluye a quien pertenecen, al emisor, un número único de identificación, fechas de validez y una "huella" encriptada que se puede usar para verificar el contenido del certificado.

Para que exista una conexión SSL ambas partes deben poseer un Certificado de Seguridad válido.

Cliente- [Client] Máquina que conectada a una red solicita acciones a otra que actúa como servidor.

CGI.- Interface de la Compuerta Común (Common Gateway Interface). Es un conjunto de reglas que describe como se comunica un servidor de la red con otros programas en la misma máquina y como otros programas (programas cgi) se comunican con el servidor. Cualquier programa se puede considerar como CGI si maneja la entrada y salida de información de acuerdo con la norma CGI.

Un programa CGI es aquel que se ejecuta en el servidor y no en el navegador del cliente. De esta manera el programa tiene acceso tanto a los datos del servidor como a todos los archivos que hay en él.

cgi-bin.- Nombre común del directorio del servidor en el que se almacenan los programas CGI.

Código de retorno- [Return Code] El estado de respuesta a una solicitud especificando el resultado de una solicitud.

Los más comunes son los códigos de error:

- 400 = Failed: Bad Request (Mal requerido)
- 401 = Failed: Unauthorized (No autorizado)
- 402 = Failed: Payment Required (Se requiere pagar)
- 403 = Failed: Forbidden (Prohibido)
- 404 = Failed: Not Found (No encontrado)
- 500 = Failed: Internal Error (Error interno)
- 501 = Failed: Not Implemented (No implementado)
- 502 = Failed: Overloaded Temporarily (Sobrecargado temporalmente)
- 503 = Failed: Gateway Timeout (Tiempo terminado para el "gateway")

Compatibilidad.- (De un navegador) Este término se refiere a que la página por su creación (programación) contiene instrucciones que pueden interpretarse por cualquier navegador.

Compresión.- [Compression] El proceso de reducir el tamaño en bytes de un archivo para reducir el tiempo de su transferencia entre máquinas.

Ciberespacio.- Del término en inglés cyberspace, originado por el escritor William Gibson en su novela "Neuromancer the word Cyberspace"¹ y que se usa para describir la totalidad de los recursos informáticos disponibles a través de las redes de cómputo.

CSS.- [en inglés Cascading Style Sheets, literalmente hojas de estilo en cascada.] Método que permite definir por separado las reglas para definir las características de los elementos HTML, DHTML y XML. Sus versiones se distinguen por el número de edición: CSS1, CSS2, etc.

Daemon.- Programa que corre independientemente del navegador. Los Daemones pueden realizar varias tareas administrativas como las de construir índices, resúmenes y

¹ Gibson, William: *Neuromante*. Ediciones Minotauro. Barcelona 1989.

retroenlaces. En Unix se utiliza el término por el de servidor debido a que los servidores operan independientemente.

DHTML.-[Dynamic html o html dinámico]. Combinación de html, hojas de estilo y Javascript que permiten modificaciones automáticas en los elementos de las páginas.

DOM.-[Document Object Model o Modelo de objetos en documentos]. Es una interface independiente de la plataforma y del lenguaje que permite que los programas y scripts tengan acceso dinamicamente y actualicen el contenido, la estructura y estilo de los documentos.

Dominio.- [Domain] El nombre asociado a una dirección IP de una computadora en el Internet.

Espinazo.- (En inglés Backbone). Una línea de alta velocidad o serie de conexiones que forman una trayectoria principal en una red.

FAQ.- De la frase en inglés Frequently Asked Questions (o Preguntas más frecuentemente contestadas). Son documentos que listan las preguntas más comunes sobre un tema así como las respuestas a las mismas. Las FAQ generalmente se escriben por gente cansada de contestar continuamente las mismas preguntas.

Finger.- Herramienta de programación utilizada para comprobar la presencia de una persona atendiendo una cuenta de internet.

Extranet.- La red usada por una empresa para conectarse con sus clientes y socios de negocios.

Firewall.- En español, barrera de fuego. Programa o equipo que separa a una red local (LAN) en dos o más partes con propósitos de seguridad.

FTP.- [File Transfer Protocol] Protocolo de transferencia de archivos. Es el método normal de enviar archivos entre computadoras en el Internet.

Galletas.- [Cookies] Son archivos que contienen información respecto a los visitantes de un sitio (por ejemplo, el nombre del usuario y sus preferencias). Esta información la proporciona el usuario en su primera visita al servidor. El servidor registra esta información en un archivo de texto y la guarda en el disco duro del usuario. Al regresar al sitio, el servidor busca la galleta y la utiliza.

Gateway.- O compuerta, es un programa o equipo que se encarga de traducir la información contenida en dos protocolos diferentes.

Hit.- Acción de solicitar una acción a un servidor, tal como requerir la visualización de una página o la transferencia de un archivo.

Hipertexto.- [Hypertext] es el término que se dio al enlace que permite el salto rápido entre dos textos por afinidad conceptual.

html.- Abreviación del término en inglés HyperText Marking Language (Lenguaje de marcado de hipertextos), es el lenguaje de programación que permite la inclusión de textos, imágenes fijas y móviles, video, archivos, etc. y su enlace mediante hipertexto por el usuario de la computadora, independientemente de la estructura de la máquina o del sistema.

HTTP.- Abreviación de la designación inglesa para Protocolo de transferencia de hipertexto. Se trata del protocolo más utilizado para transferir datos entre un servidor y otra máquina.

Internet.- [De inter, internacional y net, en inglés, red].- Todas las computadoras del mundo conectadas entre si. como si se tratara de una red. En su primera etapa la conexión de las computadora es a través de la red telefónica existente. En su última etapa la conexión será por medio de fibra óptica, si es que no aparecen tecnologías que le permitan hacerlo vía inalámbrica.

IGU.- [En inglés, GUI. Interface gráfica del usuario.] Presentación en pantalla del programa en cuestión que permite que el usuario interactúe con éste. Lo constituyen las imágenes, los íconos y los menús.

Internic.- [Internet Network Information Center. Centro de Información de la Red Internet]. Organización privada responsable del registrar los nombres de los dominios de la red.

Intranet.- [De intra, interno y net, en inglés, red].- Red interna de una empresa, que parcialmente puede exponer información al exterior via Internet. Es el concepto moderno con el que se manejan los sistemas internos de una empresa, tales como inventarios, requisiciones, liberaciones; ordenes de entrada y salida de almacen; ordenes de trabajo, de venta y de compra; facturación, requisiciones; documentación MRP I y II, SPC; documentación técnica y de producto, etc. permitiendo que los empleados accedan al sistema a través de un sistema de accesos controlados.

IP o dirección IP.- [IP Address] Dirección en el protocolo del Internet que identifica a una máquina conectada.

ISP.- Siglas de Internet Service Provider (Proveedor del servicio de Internet). Empresa que proporciona el servicio de acceso a la red Internet.

Java.- Lenguaje de programación de computadoras, cuyo creador, la compañía Norteamericana Sun, ideó como un lenguaje que puede usarse en todas las computadoras, independientemente de sus diferencias o plataformas. El lenguaje java permite que el mismo programa que se ejecuta en una Apple se pueda ejecutar también en una máquina compatible con las PC (Personal Computer), originalmente desarrolladas por IBM. En las páginas se pueden incluir programas escritos en el lenguaje java como applets

Javascript.- Lenguaje para realizar programas que logran efectos especiales en las páginas desarrollado por la compañía Netscape. Hoy en día tiene soporte suficiente para ser interpretado por una gran cantidad de navegadores de Internet.

LOG (Archivo).- [Log File] Archivo creado por un servidor que contiene toda la información relativa al acceso a un sitio.

Macromedia.- Empresa que desarrolla programas multimedia de gran espectacularidad por los efectos combinados de sonidos, movimiento y transformación de imágenes para alcanzar efectos especiales en las páginas del Internet. Requiere que el usuario halla instalado una extensión especial a su navegador para observar estos efectos. Estos programas entran en decadencia con el desarrollo del formato SVG.

Marcos.- [En inglés frames]. Procedimiento por el cual se divide una página en varias secciones, cada una de las actúa como una página por separado.

META (Etiquetas META).- Elementos de programación HTML que permiten clasificar el contenido de las páginas en los sitios. Se definen así el lenguaje a utilizar, la descripción del contenido de la página, el conjunto de caracteres a utilizar, entre otros datos útiles.

MIME.- Siglas de la expresión inglesa Multipurpose Internet Mail Extensions (Extensiones de correo Internet multipropósito). Norma para anexar archivos no textuales a los mensajes de correo normales de la Internet. Tales archivos pueden ser gráficos, hojas de cálculo, documentos de procesadores de texto con formato, audio o vídeo.

Se dice que un programa de correo es compatible con MIME si puede enviar y recibir archivos usando esta norma. Al enviar estos archivos se codifican como si se

tratara de texto, aunque este no sea legible. La norma MIME se usa también para que los navegadores identifiquen los archivos enviados al buscarlos en una lista de tipos MIME que los relaciona con los programas específicos para manejar cada tipo de archivo.

Multimedia.- El término, creado fundamentalmente para propósitos comerciales, se refiere a la utilización de los diferentes medios de comunicación con el usuario que tiene una computadora, permitiéndolo combinar imágenes en movimiento con sonidos.

Navegador.- [Browser, literalmente "paginador"] Es el programa de computadora que permite interpretar y presentar la información en lenguaje html.

NIC.- Siglas de la designación en inglés Networked Information Center (Centro de Información de Red). Cualquier oficina que maneje la información sobre una red. La más famosa es la InterNIC, encargada de la administración de los nombres de dominio registrados.

Nodo: Segmento de información, que entra en relación con otro u otros nodos (entre los que se establecen los enlaces hipertextuales que facilitan la navegación). Es importante recordar que un nodo pertenece únicamente a un mismo documento. Otra definición de nodo es la de "origen o destino de un enlace".

P3P.- [Platform for Privacy Preferences Project o Proyecto de plataforma para preferencias privadas]. Norma del WC3 para proporcionar a los usuarios una manera automatizada por la que los usuarios toman mayor control sobre el uso de su información personal en los sitios que visitan.

Página.- [Page] Documento de computadora que se presenta mediante un navegador.

Página Principal.- [Home Page, que literalmente significa "página del hogar".] Se trata de la página más importante de un sitio. Proporciona un resumen y los enlaces al resto del sitio. Frecuentemente contiene una tabla del contenido del sitio o el enlaces a la tabla.

Plataforma.- [Platform] El sistema operativo de la máquina, tal como Windows 95, Windows NT, UNIX, LYNUS, etc.)

Plug-in.- Programa de computación que se agrega al Navegador para manejar en éste cierto tipo de archivos.

PICS.- [Platform for Internet Content Selection o Plataforma para la selección de contenido en la red Internet]. Esta especificación permite etiquetas (metadatos) que asocian el contenido de la red Internet. Originalmente se diseñó para controlar el acceso infantil a la red, pero también facilita la utilización de firmas y otros aspectos de privacidad.

POP.- [Post Office Protocol o Protocolo de Oficina Postal]. Un protocolo por el que un servidor de correo permite recoger los mensajes electrónicos y descargarlos en su computadora. Un servidor POP es la computadora en la que se encuentran los mensajes de correo electrónico.

Portal.- Página utilizada para comenzar una sesión de Internet. Los portales se caracterizan por incluir información útil tal como noticias, el clima, servicio de correo electrónico y en general cualquier información relevante para el usuario.

Protocolo.- [Protocol] El conjunto de reglas que permite intercambiar datos entre dos máquinas.

Puente.- En inglés Bridge, es un dispositivo que se usa para enlazar dos redes dando como resultado una sola red.

RDF.- [Resource Description Framework o Marco de trabajo para la descripción de recursos]. Esquema que integra diversos metadatos, incluyendo mapas de sitios, calificación de contenido, definiciones de los canales con flujo (streaming), las colecciones de datos para los buscadores y otros conceptos, empleando la sintaxis del XML.

Retroenlace.- [Back link.] Un enlace en una dirección implica la existencia de un enlace correspondiente en la otra dirección.

Ruteador.- [En inglés, Router]. Dispositivo que enruta los paquetes de información electrónica tomando decisiones de tráfico, en base a las condiciones de la red.

Script.- Pequeño programa para realizar efectos especiales en las páginas.

Servidor.- [Server] Máquina conectada a otras que ejecuta una acción a solicitud de las otras (clientes).

Servlet Aplicación Java que opera como un módulo en un servidor.

SGML.- [Standard Generalized Markup Language o Lenguaje de marcado normativo generalizado] Es la recomendación general (ISO 8879) de la ISO (International Organization for Standardization, u Organización Internacional de Normas) para la

creación de métodos de representación de textos en forma electrónica independientes de la máquina o plataforma (1986).

Sitio .- [Site] Este término se aplica a la ubicación donde se encuentra la información personal (sitio personal) o de la compañía (sitio empresarial). A todo sitio está asociada cuando menos una dirección de Internet (url) y una IP.

SMIL.-[Synchronized Multimedia Integration Language o Lenguaje de integración sincronizado de multimedia]. Es un lenguaje basado en XML que permite mezclar presentaciones en varios medios y sincronizarlas. Aunque puede parecer complicado, es relativamente sencillo para alguien familiarizado con html y javascript.

SMTP.- [Simple Mail Transfer Protocol o Protocolo Sencillo de transferencia de correo.] El protocolo con el que se transmite un mensaje de correo electrónico de una máquina a otra.

SSL.- Protocolo para permitir comunicaciones encriptadas y autenticadas a través de la red Internet. La aplicación del protocolo se inicia generalmente con la llamada a una página con el protocolo "https". El protocolo proporciona privacidad, autenticación e integridad en el mensaje. En una conexión segura o SSL cada una de las partes envía información a la otra del Certificado de seguridad propio, que se utiliza para codificar la información. Para decodificar esta información se requiere por lo tanto dos claves, una del emisor y otra del receptor, garantizando la seguridad de los mensajes.

Sufijo del dominio.- [Suffix of Domain Name.] Sufijo de tres dígitos para identificar el tipo de organización.

- .com = Comercial
- .edu = Educacional
- .int = Internacional
- .gov = Gubernamental
- .mil = Militar
- .mx Usado solo identifica a una organización gubernamental mexicana. Acompañado por otro sufijo explicita que se encuentra en México.
- .net = Red
- .org = Organization

SVG.- [Scalable Vector Graphics o Gráficos vectoriales escalables] es un formato de imágenes basado en el XML desarrollado por el W3C que permite transferir las imágenes más rápidamente, hacerlas escalables y buscables, además de otras características.

TCP/IP.- Tomado de la expresión en inglés Transmission Control Protocol/Internet Protocol (Protocolo de control de transmisiones y protocolo de la Internet). Es el conjunto de Protocolos que definen la comunicación Internet.

Tiempo de carga.- [Load Time] El tiempo requerido, generalmente expresado en segundos, para que el contenido total de una página de Internet se transmita por entero al usuario. Los tiempos de carga deben ser reducidos a fin de no impacientar al usuario. La regla de que suceda algo antes de 5 segundos después de la acción del usuario proviene de la especificación del antiguo sistema operativo MS-DOS.

T1.- Norma norteamericana para líneas de transmisión de señales telefónicas que operan a 1.544 Millones de bps.

URI.-[Uniform Resource Identifier o Identificador Uniforme del recurso] Es el conjunto genérico de todos los nombres y direcciones en forma de denotaciones cortas que se refieren a un recurso.

URL.-[Uniform Resource Locator o Localizador uniforme del recurso] Es el mecanismo para identificar una ubicación exacta en el internet.

Vbscript.-Subconjunto de la programación en el lenguaje Visual Basic de la compañía Microsoft para realizar efectos especiales en las páginas.

WC3.- [World Wide Web Consortium o Consorcio para la amplia red mundial] Es la organización que normaliza a través de recomendaciones las reglas de escritura e interpretación de las páginas. Se ubica en www.w3.org

VRML.- [Virtual Reality Modeling Language o Lenguaje para el modelaje de la Realidad Virtual.] Término definido por Dage Ragget en las primeras conferencias de la

WWW en Ginebra, en mayo de 1994. El VRML está propuesto como formato de marcado lógico para la realidad virtual independiente de la plataforma.

XML.-[Extensible Markup Language o Lenguaje de marcado extendible]. Norma recomendada por el WC3 para definir nuevos tipos de documentos, permitiéndolo al usuario definir sus propias etiquetas de marcado para extender las capacidades del html. Este lenguaje facilita las transacciones en las bases de datos.

XSL.- [Extensible Stylesheet Language o Lenguaje de hojas de estilo extendible]. Recomendación general para el lenguaje de las hojas de estilo que permite extender las capacidades de este procedimiento.

INTRODUCCIÓN, METODOLOGÍA Y OBJETIVOS

El periodismo es en la actualidad una de las profesiones que más cambios ha sufrido tras la irrupción de Internet. Los profesionales de la información han visto multiplicadas sus posibilidades de realizar su trabajo, llegando a poder incluir en una misma noticia texto, gráficos, vídeo y audio. Asimismo se ha facilitado enormemente la accesibilidad a las fuentes de información a nivel mundial. A través del correo electrónico el periodista puede ponerse en contacto directo e inmediato con fuentes de primera mano, sin necesidad de pasar por intermediarios.

Otro de los grandes beneficiados son los lectores ya que gracias a esta tecnología pueden participar e interactuar con el medio de comunicación haciendo saber a los periodistas sus opiniones respecto a la noticia en cuestión, aportando ideas o incluso actuando ellos mismos como fuente de una noticia.

La comunicación ha jugado un papel esencial en el desarrollo de la sociedad, desde el primer momento en que apareció el hombre en la Tierra. La Era de la información, en la que nos encontramos ahora, es el momento álgido de una paulatina evolución que comenzó con las pinturas en las paredes de los prehistóricos y que ha ido desarrollándose ha medida que ha evolucionado el conocimiento humano. En la actualidad, nuestra sociedad está experimentando, un profundo cambio en todos los ámbitos gracias a las nuevas tecnologías.

Sin olvidarnos de la invención del teléfono por Bell en 1876, de la radio por Marconi en 1898, y del tubo de vacío por De Forest en 1906, fue durante la Segunda Guerra Mundial y el periodo subsiguiente, cuando tuvieron lugar los principales avances tecnológicos en la electrónica: el primer ordenador programable y el transistor, fuente de la microelectrónica, el verdadero núcleo de la Revolución de la tecnología e la información en el siglo XX. Sin embargo, según afirma Manuel Castells en su libro *La era de la información*², podemos decir que esta revolución, como tal revolución, nació en la década de 1970, sobre todo si se incluye en ella el surgimiento y difusión paralelos

² Castells, Manuel (1996): *La Era de la Información (Economía, sociedad y cultura)*. Vol.1. La sociedad Red. Madrid. Alianza Editorial.

de los primeros pasos de Internet y de la ingeniería genética, en torno a las mismas fechas y lugares.

Ya en el siglo XXI la predicción de Mc Luhan de ver transformado el mundo en de una “aldea global” ha cobrado forma por el poder de la comunicación.

En 1948, el físico ruso nacionalizado estadounidense George Gamow, planteó que el Universo se creó en una explosión gigantesca y que diversos elementos que hoy se observan se produjeron durante los primeros minutos, después del “Big Bang”. Según la antigua teoría griega de la creación, el “caos” inicial, era un oscuro y silencioso abismo de donde procede la existencia de todas las cosas. De Caos nació la negra Noche y Erebo, la región oscura e insondable donde habita la muerte. Estos dos hijos de la primitiva oscuridad se unieron a su vez para producir el Amor, que originó la Luz y el Día. En la mitología posterior, Caos pasó a ser la materia informe de la que fue creado el cosmos u orden armonioso³.

Volviendo al presente, y utilizando de metáfora la teoría explicada, otro “Big Bang” ha creado algo inimaginable para la sociedad hace tan sólo unos años: un nuevo mundo, o más bien, un “microcosmos” intangible que tiene como base el intercambio de información. La red Internet nos ha dado acceso a una realidad paralela cuyas posibilidades aún se están explorando. Al igual que, en el principio de los tiempos, en el que existía una libertad absoluta propiciada por el caos y todo lo que existe ahora se fue formando regido por el propio orden de la naturaleza, Internet nació libre de barreras, caótica e incontrolable, lo cual fue un hecho esencial para su exponencial crecimiento y para su rápida introducción y aceptación por parte de la sociedad. Por poner un ejemplo, fue vía Internet como el subcomandante Marcos, jefe de los zapatistas de Chiapas (México), se comunicó con el mundo y con los medios desde las profundidades de la selva Lacandona durante su retirada en febrero de 1995.

Junto a las enormes ventajas que ha aportado por esta naturaleza, se han unido inconvenientes por el surgimiento de nuevos delitos y actividades perjudiciales por el individuo. De aquí viene el debate el debate de la regulación necesaria en Internet. De hecho, ya está empezando a haber leyes que rigen ciertas “actividades virtuales”. Poco a poco el “caos” de Internet está siendo organizado y dando paso al “cosmos”.

³ Enciclopedia Interactiva Encarta 98.

Tal y como se explica en el primer capítulo de esta tesis, Internet fue concebida, en sus inicios, como un sistema para mantener la conexión entre los diferentes centros estratégicos para la defensa de los Estados Unidos, en el caso de que se produjera un ataque por parte de otras potencias extranjeras. Sin embargo, finalmente fueron las universidades y otros centros de investigación quienes desarrollaron este proyecto, facilitaron su uso creando la llamada *World Wide Web* y lo dieron a conocer al público en general. Un ordenador conectado, mediante una línea telefónica, a un servidor bastaba para tener acceso libre a una cantidad de información inimaginable hasta la fecha, además de para poner en contacto e intercambiar documentos con cualquier persona a nivel mundial que también se hallara conectada. Del mismo modo, la Red fue creada para que cualquiera pudiera publicar en ella ya que se encuentra libre de barreras culturales y políticas.

En este primer capítulo, hemos intentado también aclarar ciertos conceptos relacionados con Internet, necesarios para la contextualización y comprensión del resto de la tesis. Entre ellos, se explican las diferencias entre Internet y World Wide Web, ambos conceptos completamente diferentes, que se suelen prestar a usarse como sinónimos por parte de muchos usuarios de la Red. Internet es la mayor red de ordenadores del mundo que permite a todo tipo de ordenadores conectarse y comunicarse entre ellos a nivel mundial. Sus orígenes datan de 1969 y es fruto de un proyecto elaborado por la Agencia de Defensa estadounidense DARPA (U.S. Defense Advanced Research Agency)

Por otro lado, la World Wide Web es uno de los muchos protocolos que permiten el acceso a Internet y que facilita a los usuarios internautas el acceso a los archivos que se encuentran en Internet. La Red (o Web) fue creada en 1990 por Tim Berners-Lee, un programador de ordenadores que trabajaba para el CERN (European Organizatin for Nuclear Research). Con anterioridad a la aparición de la Red el acceso a los archivos que se encontraban en Internet era una tarea bastante ardua que requería conocimientos especializados. Gracias a la World Wide Web podemos acceder fácilmente a archivos de texto, imágenes, audio y vídeo a través del sencillo mecanismo de los enlaces de hipertexto.

Evolución constante

La evolución de Internet no cesa. En la actualidad se están desarrollando nuevos proyectos que mejoran la capacidad y la rapidez de la Red de redes. Entre ellos está la llamada Internet 2 que comenzó a elaborarse en octubre de 1996. El 4 de febrero de 1997, el presidente de EEUU, Bill Clinton, y su Administración deciden apoyar también la creación de esta nueva generación de Internet, con el objetivo de que las universidades y los laboratorios nacionales pudieran comunicarse a velocidades mil veces más rápidas que hoy día.

Este proyecto tiene previsto incluir nuevas aplicaciones derivadas de las necesidades académicas en cuanto a búsqueda de información, enseñanza y aprendizaje; es decir, se está diseñando pensando en los objetivos y retos de las futuras generaciones universitarias.

Internet2 creará, en un principio, una Red mucho más eficaz entre las universidades y el campo de la docencia y permitirá una investigación más avanzada. Asimismo, cuando el proyecto se halle lo suficientemente desarrollado, se espera que se extienda a todos los usuarios de Internet. Según Irving Wladawsky-Berger, Director General de la División de Internet de IBM, la nueva Red no se pondrá en funcionamiento hasta aproximadamente el año 2005.

El periodismo digital

Una vez explicados los conceptos básicos de Internet, soporte del nuevo estilo de periodismo que nos ocupa, hemos analizado detalladamente el llamado “periodismo digital”, analizado sus elementos, su situación actual y sus perspectivas de futuro.

Es un hecho que la Red ha comenzado a formar parte de nuestra realidad diaria y se ha extendido rápidamente entre la población mundial. La *World Wide Web* se ha convertido en una inmensa nube de información cada vez mayor que necesita ser organizada, seleccionada y servida al usuario que busca en ella datos veraces, razonados y útiles.

Internet es en la actualidad un nuevo medio de comunicación con grandes oportunidades no sólo para los profesionales de la información, sino para la sociedad en

general. El hecho de que no sea “propiedad de nadie”, ha potenciado la utilización de la *World Wide Web* por usuarios “normales”, particulares y empresas deslindadas del mundo científico y universitario de los comienzos de Internet.

A diferencia de los medios tradicionales, como la radio y la televisión, en la Red son los usuarios los que deciden la mayor parte de los contenidos disponibles. Su facilidad de uso y su accesibilidad permite a cualquier persona publicar absolutamente cualquier tipo de información que desee, con la posibilidad añadida de ser visible en cualquier lugar del mundo que tenga un ordenador conectado a la Red de redes. Otra de las diferencias, es que Internet admite diversos modos de comunicación: entre un emisor y un receptor individuales, entre un emisor individual y múltiples receptores y entre múltiples emisores y receptores. En cualquier momento, un receptor puede convertirse en suministrador de contenidos. Por consiguiente, la Red Internet es radicalmente distinta de la transmisión tradicional y también de los servicios tradicionales de comunicación.

Internet ya ha conseguido poner en contacto a más de 150 millones de personas. Se trata de un medio con un público de masas, y es el medio que más rápidamente ha alcanzado un público de masas en toda la historia de la humanidad. La radio tardó 38 años en alcanzar 50 millones de radioyentes, la televisión tardó 13 años en tener 50 millones de telespectadores e Internet en tan sólo cinco años ha conseguido la cifra de 150 millones de usuarios. Con respecto al número de internautas y según el estudio *eGlobal Report* realizado por *eMarketer*⁴, el número de usuarios de la Red llegará a los 350 millones en el año 2003. Esto equivale a un crecimiento del 267% con relación a los 95,6 millones de internautas contabilizados a finales de 1998. A nivel europeo, las cifras de Morgan Stanley Dean Witter indican que la Unión Europea puede contar con más de 100 millones de navegantes gracias a la reducción de los costes de acceso y al incremento de las velocidades de transmisión.

Sin embargo no son las cifras de usuarios lo que le está convirtiendo en el cuarto gran medio de comunicación, sino sus posibilidades. Internet permite incluir hipertexto, fotografías, imágenes animadas, sonido, vídeo, etc. Permite comunicarse desde cualquier lugar del mundo. Permite recibir y entregar todo tipo de información. Pero si algo hace a este medio especial es su interactividad. El usuario de la gran Red pasa de

⁴ E-marketer: <http://www.e-marketer.com>

ser un sujeto pasivo a tener un papel activo eligiendo lo que quiere y lo que no. Esto da a la comunicación un nuevo carácter que se establece a través de un nuevo tipo de relación entre emisor y receptor.

La nueva empresa periodística

Lo que sí es cierto es que Internet es en la actualidad un sinónimo de cambio. Las empresas que deseen adaptarse a esta nueva tecnología tendrán que estructurar sus organismos eliminando puestos de trabajo obsoletos por otros de nueva creación. Ejemplos de ello son las figuras del webmaster, el director de web marketing o los integrantes del departamento de atención al cliente *online*. Estas empresas han incorporado sus catálogos de productos a sitios *web* donde dan toda la información necesaria tanto a clientes como a distribuidores, proveedores, etc., a tiempo real. Además todo el conjunto de procesos de comunicación realizados mediante fax, cartas o llamadas de teléfono entre otros, comienzan a ser realizados ya a través de las llamadas Intranets corporativas, lo que supondrá la creación de empresas capacitadas para interactuar con otras y compartir información y recursos. Esto supondrá para las empresas del futuro su conversión en estructuras flexibles y muy competitivas.

Adaptación al cambio

Como veremos a lo largo de esta parte de la tesis, el periodismo electrónico ha ido encontrando en la propia Internet su fuente de noticias. Aunque inicialmente se haya limitado a reproducir los contenidos habituales de sus versiones tradicionales, poco a poco los contenidos y los formatos están siendo modificados, y están empezando a aprovechar las grandes ventajas que ofrece este soporte.

Marshall McLuhan afirmó en su *Galaxia Gutenberg* que todo nuevo medio trata, en una primera instancia, de integrar a los medios precedentes y referirse a ellos. Esto resulta muy claro en el caso de Internet, puesto que tanto la prensa como la radio y la televisión van tomando su sitio en la Red y abriendo sus respectivos “sitios web”. Sin embargo, Internet posee una estructura interna muy distinta a la de los medios de comunicación precedentes, derivada en parte de su origen, y en parte de su evolución.

Cabe decir que el periodismo electrónico no comienza con la aparición de Internet. Podemos encontrar claros precedentes en el videotexto y teletexto. Otros antecedentes más lejanos, nos llevan a las supuestas alternativas al periódico tradicional, desarrolladas con el objetivo de cumplir uno de los mayores deseos de los medios de comunicación: la inmediatez. Destacan entre ellos: el periódico por fax, los servicios en línea (bases de datos, correo electrónico mediante el que los lectores pueden comunicarse, servicios de compra-venta a distancia), o el audiotexto. En cuanto a este último, cabe decir que ya a principios del siglo XX encontramos un precedente. Se trata del “periódico-teléfono”, puesto en marcha en la ciudad de Nueva York en 1913. Al parecer contaba con 2.500 suscriptores que pagaban al mes el equivalente a 7,50 pesetas de entonces. El servicio consistía en que los clientes recibían las noticias de viva voz, mediante un artefacto similar a un teléfono que sólo servía para hablar con la redacción de ese medio de comunicación en concreto⁵.

Con la llegada de Internet, se incorpora al ideal perseguido por sus precedentes un concepto básico a la hora de hablar de la Red hoy en día: la interactividad. Y no sólo éste, sino además la posibilidad de integrar elementos textuales y audiovisuales en un mismo documento. Esto es lo que el profesor Jean Cloutier ha denominado “l’audioscriptovisuel”⁶.

Díaz Noci, en su artículo “El nuevo periódico electrónico: redefinición del mensaje periodístico tradicional como producto interactivo y multimedia⁷”, señala cinco causas principales que según él han dado lugar al periódico interactivo y multimedia. Entre estas se encuentran: Incapacidad de los periódicos de aumentar e incluso mantener sustancialmente el número de lectores; el aumento de la información, la diversificación de la actividad empresarial comunicativa, la escasez del papel y

⁵ Díaz Noci, Javier: *El nuevo periódico electrónico: redefinición del mensaje periodístico tradicional como producto interactivo y multimedia*. Mayo 2000.

⁶ Cloutier, Jean (1994): *L’audioscriptovisuel et le multimedia*, Communication et langages, nº 99, París: Edicions Retz, pp.99.

⁷ Díaz Noci, Javier: *El nuevo periódico electrónico: redefinición del mensaje periodístico tradicional como producto interactivo y multimedia*. Mayo 2000.
<http://www.andalucia24horas.com/caracola/doc002.htm>

aumento de la sensibilidad ecologista; y los avances técnicos especialmente en el terreno de los ordenadores.

Una nueva forma de informar

En la actualidad, Internet se ha convertido en un medio de comunicación de masas, en el que intentan estar presentes los antiguos medios tradicionales. Cada vez es mayor el número de periódicos que decide sacar una edición on-line. Estos medios han creado su "imagen" en Internet, aprovechando así su capacidad y facilidad de difusión. Sin embargo, hay que tener en cuenta la esencia de Internet como medio de comunicación en sí mismo. No se trata simplemente de un soporte en el que adaptar los medios tradicionales. Se trata de una nueva forma de informar y de distribuir dicha información.

La nueva forma de informar que implica Internet, está determinada por las posibilidades que ofrece el propio medio; y para explotar estas posibilidades los nuevos medios de comunicación, y por tanto las informaciones que se ofrecen a través de Internet, deben adoptar una serie de características específicas. Éstas son las siguientes:

- a) **La interactividad:** Según Nicoletta Vittadini, en la recopilación de textos realizada por Gianfranco Bettetini y Fausto Colombo, titulada *Las nuevas tecnologías de la comunicación* (Paidós, 1995), una de las características más innovadoras de los nuevos *media* lo constituye la instauración de una modalidad comunicativa no permitida antes por los *media*: la comunicación interactiva.

- b) **El hipertexto, hipermedia:** La aparición del *World Wide Web* ha sido decisiva para la expansión y la difusión mundial de Internet, y ello por múltiples motivos, entre los que se encuentra lo que podemos llamar la arquitectura del WWW, es decir el entramado de conexiones que lo sostiene y que está basado en el hipertexto. Dado que el periodismo electrónico ha de adaptarse en el momento actual a dicha arquitectura, se deben tener presentes algunas nociones básicas al respecto, así como las consecuencias que tienen en el caso del periodismo electrónico.

- c) **Elementos multimedia:** Con la palabra multimedia podemos designar dos conceptos diferentes. En primer lugar el referente a empresas o grupos de empresas de comunicación que engloban a distintos medios informativos, escritos y audiovisuales. El segundo concepto hace alusión a la unión en un mismo producto de materiales escritos, audio, imágenes fijas o en movimiento y vídeo. Ésta última significación, se encuentra íntimamente relacionada a los grandes y rápidos avances de las nuevas tecnologías. En la actualidad, prácticamente cualquier ordenador doméstico es capaz de procesar información gráfica y sonora, a pesar de que ésta ocupa más memoria que el simple texto.
- d) **Lenguaje html:** El html es un lenguaje de marcas basado en hipertexto (HyperText Markup Language) que se usa para formatear documentos que puedan ser interpretados por cualquier navegador compatible con el protocolo HTTP. Su origen está en la necesidad de disponer de un código que permitiera general ficheros válidos para el Web. Este lenguaje de programación fue pensado como una semilla que pudiera desarrollarse en función de la evolución de la Web. De esta forma, a la vez que el WWW ha ido demandando nuevos servicios, el html ha ido evolucionando para incluir nuevas prestaciones. Todo esto manteniendo la compatibilidad con las versiones anteriores del lenguaje y con todo tipo de plataformas informáticas.
- e) **Información personalizada:** Se trata de que el usuario obtenga únicamente aquella información que le interesa, de la forma que él crea más conveniente, de acuerdo a sus propios gustos. Para alcanzar este objetivo, se suele incluir en la publicación una página-registro en la que el lector tiene la oportunidad de determinar los contenidos y el diseño del periódico, según sus intereses. Sin embargo, hay que tener en cuenta que aunque sea el usuario quien supuestamente selecciona las noticias que le

llegan, siempre necesitará la ayuda de los profesionales de la información para orientar al lector ante la gran cantidad de datos, muchos de ellos poco fiables, que se encuentran disponibles en la Red.

- f) **Información documentada:** Otro aspecto importante a tener en cuenta es la ilimitada cantidad de documentación que puede acompañar a una noticia "on-line", por lo que su contextualización está asegurada. Este acto de documentar se lleva a cabo a través del establecimiento de links (enlaces) que llevan al archivo del periódico, a otras bases de datos especializadas en el tema en cuestión, o a otras páginas simples especializadas en el tema. El usuario tendrá a su disposición todas las fuentes fiables (ya que han de haber sido comprobadas por los periodistas del medio en cuestión) sobre ese tema disponibles en Internet.

- g) **Actualización. Capacidad para procesar datos en tiempo real:** El periódico electrónico permitirá a su vez publicar información con una periodicidad mucho mayor que en los medios tradicionales escritos. Del mismo modo, se podrán subsanar los errores en cuanto sean detectados. Ya no se hablará de un único cierre ya que las noticias serán continuamente renovadas. Hay que tener en cuenta que estamos ante un nuevo medio de comunicación que permite la integración de elementos multimedia, junto con el texto, como audio, video, gráficos interactivos ... etc. El único problema actual para la utilización de estos elementos integrados en una noticia, o de aquellos que el periodista cree que puede ayudar al lector a un mejor entendimiento de la información, es el ancho de banda, problema que a veces se soluciona a través de formatos como "Schockwave".

- h) **Nuevas formas de diseño:** El diseño de las publicaciones "online" es parte esencial de la misma. Se trata de incitar al usuario a interactuar con la publicación. Para ello hay que ir más allá de la belleza puramente

estética, y preocuparse de la funcionalidad de la página, dando preferencia a los elementos gráficos. La navegación dentro de la publicación debe ser sencilla y fácil de seguir.

El papel del periodista digital

El auge de los medios tradicionales en Internet ha desatado todo un cúmulo de reflexiones sobre cómo presentar de manera más efectiva e idónea la información periodística. Se trata de averiguar la forma de hacer “más atractiva” *la interfaz* gráfica, cómo mantener la atención del lector-internauta y otras especificaciones referidas a aspectos meramente técnicos. En el capítulo dedicado al “periodista digital” nos centramos en la persona encargada de recopilar y proveer la información, es decir, de las nuevas funciones del periodista.

En un medio como Internet en el que todo aquel que posea un ordenador conectado a una línea telefónica tiene acceso, cualquiera se puede convertir en informador potencial sin que nadie se lo impida. Por lo tanto, está claro que las funciones del periodista *online* van a ser diferentes a las que tendría en los medios tradicionales.

Ha habido muchos autores que se han aventurado en definir los perfiles profesionales del periodista digital. En 1995 Vicent Partal veía al nuevo periodista como un guardia urbano que dirige el tráfico de la información, dando preferencia a la circulación de determinados carriles, siendo un punto central en la ordenación del tráfico en las autopistas de la información.

Dos años más tarde, Luis Ángel Fernández Hermana y Alfons Cornella propusieron dos visiones diferentes. Fernández Hermana comparó al nuevo periodista con un cartógrafo, afirmando que, en el medio digital, las carreteras y autopistas están por construir, y que posiblemente nunca llegarían a estar establecidas. Para este autor, la función del periodista digital es la siguiente⁸: “Consiste en darle sentido a un mapa informativo desordenado que pide un diseño comprensible para la intervención del internauta. Los accidentes geográficos –el entramado en el caso del guardia- deben ocupar el lugar que les corresponde a partir de la acción de este nuevo operador de la

⁸ Fernández Hermana, L.A. *De exploradores y cartógrafos*. Editorial de En.red.ando (1997). <http://enredando.com/cas/cgi-bin/editorial/plantilla.pl?ident=252>

Red. Para ello, como sucede en el caso del explorador y del cartógrafo, debe servirse de las herramientas apropiadas en cada caso y en función de los objetivos, independientemente de las diferencias propias entre publicaciones por razón de sus finalidades. Son éstas las que determinarán las proporciones entre ayudas a la navegación, contenidos audiovisuales, relaciones entre sistemas informativos, creación de comunidades de interés, calidad de su información o el uso que de ella hagan los usuarios (...). El cartógrafo de la Red requerirá aunar, por lo menos, conocimientos y experiencias que proceden de diferentes campos del conocimiento, como la ingeniería, las telecomunicaciones, el diseño, los flujos de la información y la comunicación, las necesidades coyunturales de los usuarios (demanda y oferta) y la interacción con contenidos de diverso origen cuya síntesis generen productos nuevos”.

Asimismo, Alfons Cornella⁹ nos definió al periodista digital como “infonomista”, un profesional que estudiaba las estructuras comunicacionales y diseñaba economías y flujos óptimos de información.

En países donde el periodismo electrónico es una fuerte y estable fuente de trabajo, la imagen del reportero-periodista ha dado paso al denominado "proveedor de contenido", una especie de *Blade Runner* de la información que toma su pauta, busca los datos, hace las entrevistas, toma las fotografías y asiste a los eventos, todo a través de Internet¹⁰.

En el periodismo tradicional el periodista tenía que ir en busca de la información. En pocas ocasiones las noticias venían a él. Sin embargo, en este nuevo medio de comunicación, bajo el soporte de Internet, el profesional de la información se encuentra justo con el problema opuesto. En cuanto se conecta a la Red tiene a su disposición un mar de información que debe seleccionar y verificar. Debido a esta gran cantidad de información de la que se dispone, el valor del periodista digital se encuentra en su propia valoración de cada fuente y en su ética personal.

La búsqueda de información en Internet

⁹ Cornella, A. *Eureka: es la infonomía, y somos infonomistas*. Artículo de Extra!Net (1997) <http://www.extra-net.net/articulos/en971108.htm>

¹⁰ Núñez Aldasoro, Antonio J: *Los retos del periodista digital*. <http://www.saladeprensa.org>

La segunda parte de la investigación, se centra en la búsqueda de información a través de la Red, dado el hecho de que Internet es, en la actualidad, una fuente esencial para el trabajo de los periodistas debido al volumen de información que contiene y genera diariamente, además de por su utilidad potencial para facilitar el contacto con fuentes personales (como pudieran ser expertos o testigos de una información) a nivel mundial.

Uno de los retos más importantes que plantea este nuevo medio al periodista, es la necesidad de conocer lo mejor posible todos los recursos de búsqueda de información en Internet en aras de explotar mejor todas las posibilidades que ofrece la Red. Estos sistemas de búsqueda han evolucionado bastante a lo largo de la historia de Internet; una evolución necesaria dado el creciente aumento de información. En 1989 había un pequeñísimo número de páginas web, lo que hacía simple recuperar los datos que se alojaban en esas páginas. Con el paso del tiempo fue necesario crear nuevos recursos que ayudaran a recuperar y a indexar los datos que contenían el creciente número de servidores.

Actualmente el número de buscadores se ha multiplicado (hoy hay más de 400 motores de búsqueda entre buscadores generalistas y especializados). Los amantes de la música, por ejemplo, pueden buscar los artistas o canciones que deseen en cualquier buscador de música digital. Tan sólo tienen que escribir el título de un tema y obtendrán todos los cantantes y bandas que lo interpretaron. Sin embargo, a pesar de toda esta gran cantidad de información disponible, en numerosas ocasiones el periodista que no conoce bien los recursos de búsqueda se siente perdido ante la dificultad de encontrar un dato preciso.

Una aguja en un pajar

Ramón Salaverría, director del Laboratorio de Comunicación Multimedia (Universidad de Navarra), suele emplear la siguiente metáfora a la hora de definir los buscadores en Internet, que ilustra acertadamente la tarea de los mismos¹¹:

“Imaginemos un pajar del tamaño de un gran estadio de fútbol, repleto de pequeñas briznas de heno hasta diez metros de altura, y lancemos desde el aire una aguja de coser. A continuación, pidamos a alguien que encuentre esa aguja. Ese desventurado individuo estará perdido si pretende hallar la aguja revolviendo desordenadamente la paja.

¹¹ Salaverría, Ramón: *Internet para periodistas*. Escuela de Medios de *La Voz de Galicia*. A Coruña, 20-22 de septiembre de 2000.

Invertirá horas y horas en vano, hasta la desesperación. Por el contrario, si emplea el utillaje adecuado —un gran imán, por ejemplo— encontrar la aguja puede ser cuestión de segundos”.

La llamada World Wide Web contiene aproximadamente mil millones de documentos y crece a un ritmo diario de 1,5 millones de páginas¹². Los motores de búsqueda más avanzados no llegan a alcanzar más que una ínfima parte del contenido de esta Web. Específicamente cabe reseñar, que los buscadores que más páginas tienen indexadas son Fast (<http://www.alltheweb.com>) con aproximadamente 300 millones de documentos y Northern Light (<http://www.northernlight.com>) con 218 millones. Además hay que tener en cuenta que la World Wide Web es solo una parte de Internet. Existe también la llamada Web profunda o Internet Invisible, a la que no se puede acceder mediante los buscadores, que es aproximadamente 500 veces mayor que la visible y sobre todo, de una mayor calidad, y que analizamos en profundidad en la tercera parte de esta tesis.

A la hora de enfrentarse a este mar de información disponible en Internet, el periodista debe de tener un buen conocimiento de los distintos recursos de búsqueda que le ayudarán a alcanzar su propósito y de cómo realizar dicha búsqueda.

Según el periodista Miguel Ángel Díaz Ferreira de la Facultad de Ciencias de la Información de la Universidad Complutense de Madrid, existen cinco formas diferentes de localizar recursos por Internet¹³:

- a) Por medio de **publicaciones impresas** que recogen direcciones de la Red. Puede tratarse de trabajos de investigación, artículos de publicaciones especializadas o incluso de periódicos o revistas de información general y, sobre todo, de libros escritos *ad hoc* para las búsquedas. Pueden ser manuales sobre la Red o las “Páginas Amarillas” de Internet. Sus dos principales desventajas son las siguientes:
 - Su temprana obsolescencia.

¹² *The Deep Web: Surfacing Hidden Value*. Informe de BrightPlanet.com:
http://www.completeplanet.com/tutorials/deepweb/introduction_howsearch07.asp.

¹³ Citado en: Pérez Luque, María José: *El periodismo y las nuevas tecnologías*. Newbook Ediciones Navarra, 1998. pp: 81-82

- La necesidad de introducir los datos necesarios a mano (por tratarse de publicaciones impresas), sin aprovechar la utilidad del hipertexto para moverse por sus páginas.
- b) A través de los **directorios de recursos y clasificaciones temáticas**, que recogen y organizan los recursos de la Red sobre la base de un ordenamiento previo. Su problema es que no siempre están lo suficientemente actualizados y que no recogen todos los recursos de la Red, sino tan sólo los que los autores de la información hayan dado de alta en el servicio. Su ventaja es su propia organización que ayuda a localizar las direcciones con mucha facilidad y de forma muy esquemática.
- c) Utilizando los **motores de búsqueda e indización automáticos**, que pueden localizar cualquier recurso de la Red. Éstos tienen la desventaja de que, si no acotamos mucho nuestra búsqueda, nos pueden proporcionar demasiados enlaces, la mayoría de ellos irrelevantes para nuestra investigación. Su gran ventaja consiste en que la información no necesita ser dada de alta para que el motor la localice, por lo que están más actualizados y suelen ser más completos que los directorios y clasificaciones temáticas.
- d) **Consultando los enlaces que ponen a disposición del internauta** otras páginas relacionadas con el tema de nuestra búsqueda. Una vez localizada cualquier página relativa a la materia que se pretende encontrar, bastará con conocer esa dirección, porque es casi seguro que esa página contará con interesantes enlaces que llevarán a los puntos más importantes donde localizar toda la información necesaria. El problema es que no todas las páginas cuentan con esos enlaces.
- e) **Preguntando a los propios usuarios de la Red**. Ellos son, sin duda, los que están mejor informados, los que conocen las últimas páginas aparecidas no accesibles mediante buscadores. Su información es siempre la más actualizada y suele ser pertinente. El éxito de la búsqueda es precisamente encontrar a algún

internauta dispuesto a ayudarnos y que además esté interesado o conozca el asunto sobre el que buscamos información. Hay diversos métodos para localizar a entablar conversación entre los internautas: listas de distribución, grupos de noticias, IRC, Chats, etc.

Las fórmulas de búsqueda en Internet son múltiples y, muchas de ellas, específicas para cada tipo de búsqueda. Por este motivo es necesario conocer las características de las herramientas para este propósito disponibles, sus ventajas e inconvenientes, y también sus limitaciones.

El contenido de la Red

En la actualidad la World Wide Web contiene cerca de 5 millones de sitios web, que que aumentan cada día. La mayoría de ellos, aproximadamente un 86%, se encuentran escritos en inglés¹⁴.

Pero esto no es todo, la World Wide Web es únicamente una ínfima parte de los contenidos que se hallan y que se vuelcan realmente en Internet a través de bases de datos, listas de distribución (news groups), correo electrónico, etc, que pueden ser hasta 500 veces mayor que el contenido de la Web.

Los documentos “web” que se van creando son accesibles a cualquier usuario de la Red gracias a las llamadas “arañas”, una tecnología de los buscadores cuyo cometido es leer cada página que encuentran en su camino, y mandar las palabras que contiene, junto con la información de su ubicación a enormes bases de datos. A continuación, estas arañas siguen los enlaces que se encuentran en esta página hasta llegar a un nuevo sitio, y así sucesivamente. Sin embargo, según afirma *Search Engine Watch*¹⁵, ningún buscador indiza más del 50% de la Red. Según un estudio de la revista *Nature*¹⁶ de febrero de 1999, en esta fecha entre todos los buscadores no se cubría más que el 42% de la Red. A pesar de que la tecnología avanza y de que cada vez se amplía más esta cifra, continúa habiendo una cierta tendencia a indizar sobre todo sitios de EEUU, principalmente los más visitados y los sitios comerciales más que los relacionados con

¹⁴ *The Deep Web: Surfacing Hidden Value*. Informe de BrightPlanet.com:
http://www.completeplanet.com/tutorials/deepweb/introduction_howsearch07.asp.

¹⁵ Search Engine Watch: <http://www.searchenginewatch.com/reports/sizes.html>

¹⁶ Lawrence, Steve y Giles, C. Lee, del NEC Research Institute. Citado en *Nature* del 08/07/1999 (resumen accesible en <http://www.wwwmetrics.com>).

la educación. También son poco privilegiados en este aspecto los documentos escritos en lenguas minoritarias que podemos encontrar en la Red en una baja proporción.

Según la empresa *Wordtracker*¹⁷, los usuarios realizan diariamente un total de 250 millones de consultas a los buscadores. Gracias a estas herramientas no es tan difícil encontrar algo en la Web, tan sólo hay que saber cómo utilizarlos adecuadamente, usar los llamados operadores booleanos, combinando palabras para definir la búsqueda e intentar que la palabra clave a utilizar no pueda hallarse en otro campo.

A la hora de analizar el contenido disponible en la Red debemos tener también en cuenta lo efímeros que son algunos enlaces. La media de cualquier enlace, hace algunos años, era de 44 días. En febrero del 2000, según un estudio de Inktomi¹⁸, frente a 4,2 millones de sitios accesibles que se examinaron, 0.7 millones eran inalcanzables debido a la desaparición de su servidor o a su traslado. Con el propósito de preservar la mayor parte de la Web posible, incluyendo a los grupos de discusión, en 1995 Brewster Kahle comenzó a desarrollar el proyecto “The Internet Archive¹⁹”. Hasta el momento lleva almacenados 15, 5 terabytes. La mayoría de las webs que contiene son obtenidas mediante arañas, aunque también se aceptan páginas cedidas por sus propietarios.

Por lo tanto a la pregunta de qué contiene la World Wide Web, la respuesta es una infinidad de información que el usuario tiene que aprender a encontrar. Que no se encuentre lo que se busca, no significa necesariamente que no esté ahí. Juan José Millán, en su artículo “El libro de medio billón de páginas²⁰”, identifica a los buscadores como intermediarios, es decir, como el filtro por el que los contenidos de las *webs* llegan al público. “Imaginemos que la única vía de acceso a todas las publicaciones mundiales fueran los ficheros de una decena de bibliotecas. En la Web estamos así, con la diferencia de que en los buscadores no ha habido bibliotecarios que apliquen sus saberes clasificadores: sus programas actúan ciegamente, analizando el código de la página, y clasificando los sitios según criterios formales. Este último aspecto es básico: cuando puede haber cientos o miles de sitios que responden a una

¹⁷ Wordtracker: <http://www.wordtracker.com/>

¹⁸ Inktomi: <http://www.inktomi.com/webmap/>

¹⁹ The Internet Archive: <http://www.archive.org/>

²⁰ Millán, Juan José: *El libro de medio billón de páginas*: <http://jamillan.com/ecoling.htm>

determinada búsqueda, figurar en los primeros puestos de la lista de respuestas de un buscador puede ser la diferencia entre existir o no. No extrañará, por tanto, que haya quien lo intente por todos los medios”, afirma.

La valoración de una información hallada en Internet

Como hemos señalado con anterioridad, la Red es un gran mar de información completamente abierto y accesible para que cualquier persona que posea un ordenador con conexión a Internet pueda hacer su aportación. Por ahora no existe ningún filtro capaz de valorar la veracidad o ética de las miles de páginas que se vuelcan diariamente en la Web.

En la actualidad todo son facilidades a la hora de crear una página en Internet. Los portales llegan a ofrecer de 25 a más megabytes de forma gratuita por darse de alta en sus servidores. Un usuario, una clave y un espacio completamente libre de ataduras y censuras para publicar. Por ahora todo parece bien. Sin embargo, esta misma frase que simula una situación “perfecta”, conduce a un cierto problema. La Red se encuentra repleta de información útil para sus usuarios, pero del mismo modo también se pueden encontrar, con la misma facilidad, documentos falsos, con una falta absoluta de ética o incluso que rayan el delito.

No existe ningún filtro que, en un principio impida publicar en la Red. Lo que sí es posible es utilizar un *software* (ej: censorware o cibernanny) para restringir el acceso a ciertos sitios que podrían resultar perjudiciales, sobre todo para los niños.

En el caso que nos ocupa, es decir, el de los periodistas que utilizan Internet como medio para buscar sus fuentes, éstos deben comprobar cualquier información hallada en la Red ya que cualquiera puede haberla publicado bajo nombre falso. Del mismo modo, a la hora de publicar en la Web, el periodista debe regirse por la misma ética de veracidad y exactitud que tendría en un medio tradicional.

La navegación por la Red

Una vez analizados los aspectos más generales de la búsqueda de información en Internet, pasaremos a adentrarnos en la navegación por la Red y en las herramientas y técnicas de búsqueda. En esta parte de la tesis, hemos intentado ofrecer una visión

global de los recursos que debe conocer el periodista a la hora de buscar eficazmente información por Internet.

La World Wide Web tuvo un éxito inmediato gracias a su naturaleza descentralizada. Aunque los primeros servidores estuvieran alojados en el CERN, ni siquiera Berners-Lee ejercía ningún tipo de control sobre quien decidía establecer un nuevo servidor en cualquier sitio de Internet. Lo único que se requería para esta tarea era enlazar con otros servidores e informar a otros usuarios de la Red sobre los nuevos servidores de manera que éstos pudieran crear a su vez enlaces de hipertexto de vuelta. Este hecho propició el crecimiento exponencial de sitios en Internet que fue dificultando cada vez más el hecho de hallar un dato preciso.

En la actualidad, a la hora de realizar una búsqueda en Internet, el usuario tiene a su disposición una serie de recursos que le facilitan encontrar lo que desea. Es importante conocerlos para no perderse en el mar de información, cada vez mayor, disponible para el usuario. Principalmente son los siguientes:

- **Directorios**
- **Motores de búsqueda**
- **Anillos Web**
- **Portales (Horizontales y verticales)**
- **Bases de datos**

Directorios

Los directorios temáticos generales se diferencian de los motores de búsqueda en que los primeros son elaborados y organizados por personas. Ésta es la razón por la que ofrecen su base de datos por temas clasificados jerárquicamente. No precisan de motor de búsqueda ya que descendiendo por categorías y subcategorías se puede llegar a la información deseada. Permite a los responsables de un sitio web registrarse mediante un formulario.

Entre las ventajas de los directorios se encuentra, además de la calidad de las

webs indexadas, la posibilidad de hacerse una idea de los sitios más relevantes sobre un determinado tema. Dos de sus principales inconvenientes es su lentitud y su reducido catálogo, si los comparamos con un motor de búsqueda.

Motores de búsqueda

Los motores de búsqueda son bases de datos que contienen el texto completo de páginas web indexado. Cuando un usuario utiliza un buscador, no está buscando en toda la World Wide Web sino en la base de datos específica del motor. Esto permite que se obtengan rápidos resultados, hecho que no sería posible si cada vez que se realiza una consulta se tuvieran que rastrear en tiempo real los millones y millones de documentos que se encuentran disponibles en la Red.

A la hora de crear estas bases de datos, los motores de búsqueda, con una periodicidad normalmente mensual (aunque depende del motor de búsqueda en cuestión) utilizan las llamadas “arañas” o robots inteligentes que “saltan” de una web a otra a través de los enlaces de hipertexto registrando la información allí disponible. Esta base de datos también se alimenta de las *webs* que ceden sus propietarios, los cuales indican al motor de búsqueda, mediante un formulario, la existencia de su servicio. Otro mecanismo de búsqueda que poseen los motores es el que activan los usuarios de Internet. Su función es recuperar información de la base de datos que previamente se ha elaborado mediante el rastreo de la “araña”.

Anillos web

Los *webRing* o anillos web son grupos de webs relacionadas sobre un mismo tema. Cada Web inserta en sus páginas un código de programación gracias al cual aparecen agrupadas distintas webs por temas. Existen anillos sobre, Informática, Marketing, Medicina, Vídeo Juegos, Alpinismo, y un largo etcétera.

El WebRing fue creado por Sage Weil, a partir del proyecto EUROPA²¹ (Expanding Unidirectional Ring of Pages), que era un conjunto de páginas enlazadas a modo de anillo. Este proyecto había sido ideado por Denis Howe, en el Imperial College de

²¹ Accesible en: <http://www.wombat.doc.ic.ac.uk/europa.html>

Londres.

Portales

Los portales son sitios web que actúan como distribuidores a otros sitios vinculados a un tema o campo. Se distinguen dos tipos: los portales horizontales y los verticales. Los primeros ofrecen información y servicios a un universo amplio. Sin embargo, los portales verticales o vortales, ofrecen información específica y seleccionada sobre un tema.

Por su parte, los canales temáticos se centran en materias específicas que suelen interesar a un amplio número de usuarios. Se diferencian de los portales verticales principalmente en que estos últimos tienen una dirección común a varios de ellos, que se suelen integrar en un portal generalista, mientras que los canales temáticos son independientes empresarialmente hablando.

Bases de datos

Tal y como analizamos en profundidad en la tercera parte de esta tesis, en Internet existen mucha información a la que los buscadores generalistas, que normalmente utilizan los usuarios, no son capaces de acceder. La amplitud y la calidad de la información disponible en esta parte de Internet hace que su conocimiento sea indispensable para el periodista.

A parte de estos recursos principales existen otros muchos como preguntar a otros usuarios de la Red a través del e-mail o de los foros de discusión.

La Red profunda

La tercera parte de la tesis está dedicada al estudio y análisis de la Red profunda o aquella parte de Internet no accesible mediante buscadores convencionales.

Cuando se quiere encontrar un dato entre los más de 8,5 millones de sitios web disponibles en Internet, lo normal es acudir a buscadores del estilo de Yahoo o Google,

pero estas populares herramientas pasan por alto una importante parte de la Red que contiene entre 400 y 550 veces más información. En 1994, la doctora Jill Ellsworth, especializada en el estudio de la Red, acuñó el término 'Internet invisible' para referirse a la información que no podían encontrar los buscadores más comunes por razones técnicas o simplemente por conveniencia. En 2000, un estudio de la consultora estadounidense BrightPlanet, elaborado por Michael Bergman, confirmaba y explicaba la existencia de una Red profunda, que tendría aproximadamente 7.500 terabytes (equivalente a 7.500 billones de bytes) de información frente a los 19 de la superficial o parte de la Red accesible mediante los buscadores convencionales.

Estas cifras suponían superar entre 400 y 550 veces el volumen de información en el que un usuario que solamente utilizase, por poner un ejemplo, Google o Altavista era capaz de buscar. Catálogos de bibliotecas, bases de datos, revistas electrónicas y archivos de documentos que no pueden ser indexados por su formato componen este mar de información desconocido por la mayoría de los usuarios que se limitan a usar las herramientas de búsqueda más comunes.

Esta supuesta invisibilidad de ciertos sitios de Internet es fruto de la tecnología actual que utilizan los motores de búsqueda. Esto no significa que sea imposible acceder a estos documentos si se utiliza una tecnología más avanzada.

Conocer el funcionamiento de los motores de búsqueda es vital para comprender el porqué de la existencia de la Red profunda. Google, Altavista o cualquier otro motor son esencialmente enormes bases de datos que contienen el texto completo de páginas web. Cuando un usuario utiliza un buscador, no está buscando en toda la World Wide Web, sino en la base de datos específica del motor. Esto permite que se obtengan rápidos resultados, hecho que no sería posible si cada vez que se realiza una consulta se tuvieran que rastrear en tiempo real los millones y millones de documentos que se encuentran disponibles en la web.

A la hora de crear estas bases de datos, los motores de búsqueda, con una periodicidad normalmente mensual, utilizan las llamadas arañas o robots inteligentes que saltan de una web a otra a través de los enlaces de hipertexto registrando la información allí disponible.

Si una web no tiene enlaces, la única forma que tiene de ser registrada en un buscador es que su autor la incluya manualmente rellenando un formulario. En caso

contrario, esta web resultara invisible para todos aquellos usuarios de Internet que no conozcan la URL o dirección web concreta. Sin embargo, el que una página carezca de enlaces no es la única causa que puede llevar a una web a permanecer en las profundidades. Los buscadores suelen indexar páginas estáticas que incluyen textos simples y que están programadas en HTML, el lenguaje de programación más común con el que están escritos los sitios web. Ésta es la información que, sin ninguna duda, un usuario encontraría en la Red superficial utilizando los buscadores más comunes, el resto pasa a formar parte de la Red profunda.

Actualmente, la posibilidad de buscar e indexar cierto tipo de archivos no textuales, como imágenes, audio, vídeo, archivos PDF o Postscript, programas ejecutables o archivos comprimidos, se ha convertido en un reto para los motores de búsqueda más generales. Técnicamente la mayoría de estos formatos pueden ser indexados, pero los motores eligen no hacerlo por razones de negocio. En primer lugar, hay mucha menos demanda de este tipo de formatos que de textos HTML y, además, estos formatos son más difíciles de archivar y organizar, por lo que requieren mayores recursos del servidor. Estos archivos no suponen una parte muy grande de la Red profunda, sin embargo, se va extendiendo rápidamente su uso a medida que pasa el tiempo, principalmente para documentos de alta calidad, concernientes a los Gobiernos o alguna otra autoridad. A pesar de estos inconvenientes, la demanda se está imponiendo y algunos motores de búsqueda como Google o Altavista ya están incorporando entre sus opciones la posibilidad de buscar imágenes o archivos PDF.

En realidad, el mayor reto al que se enfrentan los motores de búsqueda recae en el acceso a la información que se encuentra en las bases de datos, que conforman la mayor parte de la Red profunda.

El problema radica en que cada una de las bases de datos es única en diseño, estructura, herramientas y método de búsqueda. Sólo presentan dinamismo en respuesta a una pregunta directa, cosa que hasta el momento sólo es capaz de hacer el propio usuario y en ningún caso los programas araña que usan los buscadores en sus tareas de catalogación. Si el usuario pretende realizar una búsqueda en ellas, debe usar las propias herramientas que la base de datos ofrece y que, al estar creadas específicamente para este propósito, proporcionarán los mejores resultados.

El inconveniente está en que es necesario encontrar la base de datos en primer lugar, tarea en la que a veces los motores de búsqueda ayudan y otras no. Hay diferentes tipos de bases de datos a los que se puede acceder a través de la Red. Sin embargo, se debe tener en cuenta que el hecho de que los documentos web estén almacenados en una base de datos no los hace automáticamente parte de la Red profunda. A veces se usan simplemente porque permiten un mantenimiento más fácil del sitio.

La Red profunda abre un sinnúmero de nuevas posibilidades. Su supuesta invisibilidad depende del conocimiento de su existencia y de no conformarnos con los recursos convencionales que nos brinda Internet. El hecho de que Google o Altavista no encuentren el dato que buscamos no significa que no esté en la Red, como veremos en esta parte de la tesis. Tan sólo hay que sumergirse algo más para comprobar si ese dato está en la telaraña mundial o es que simplemente es invisible a los ojos de ciertos buscadores.

Como anexo hemos incluido una relación de los principales recursos disponibles en la Red profunda, organizados temáticamente, que consideramos de una importancia clave ya que aporta una visión práctica de la teoría elaborada en la tercera parte de la tesis.

Entrevistas realizadas

De gran importancia para la elaboración de esta tesis han sido las entrevistas realizadas por el autor a algunos de los máximos expertos en esta materia, los cuales han aportado la visión práctica necesaria para completar la investigación. Los entrevistados son los siguientes:

- **Mariló Ruiz de Elvira:** Fue directora de la edición digital del diario *El País* (<http://www.elpais.es>) desde sus comienzos. En 1999 pasa a ser directora de producto de Proel (Promotora de Ediciones Electrónicas del Grupo PRISA). A partir de esta empresa se crea Inicia, (<http://www.inicia.es>) portal de Internet del Grupo PRISA en el que Ruiz de Elvira ocupa hasta la actualidad el cargo de directora de contenidos.
- **Mario Tascón:** Fue director de la edición digital del diario *El Mundo* (<http://www.elmundo.es>) desde sus comienzos. En el año 2000 pasa a ser

director de contenidos de Prisacom, empresa creada para agrupar todos los contenidos digitales del Grupo PRISA.

- **Ricardo Fornás:** Creador del buscador de buscadores “Buscopio” (<http://www.buscopio.com>).
- **Ramón Salaverría:** Es director del Laboratorio de Comunicación Multimedia (MMLab) de la Universidad de Navarra (<http://www.unav.es/fcom/mmlab/>). Es además autor de la *Guía de Internet para periodistas* (<http://www.unav.es/fcom/guia/>) y editor del Boletín MMLab (<http://www.unav.es/fcom/mmlab/boletin/>), publicación difundida por correo electrónico que recopila noticias nacionales e internacionales sobre comunicación.

La enorme importancia del usuario en el proceso de la información, especialmente en su actuación como fuente, es una de sus afirmaciones comunes más interesantes a nuestro juicio. La credibilidad de los datos que puede recoger el periodista de la Red es otro de los temas polémicos que han contestado los entrevistados resaltando una de las principales reglas del periodista tradicional: siempre hay que comprobar las fuentes de referencia, se trate del medio del que se trate.

Asimismo, los entrevistados se han referido al coste de los medios de comunicación en Internet. Desde el primer momento el acceso a la información de los medios de comunicación que comenzaban a surgir en Internet fue gratuito. La reducción de costes que suponía la ausencia de papel y de distribución y los pocos recursos humanos que exigían las primeras ediciones digitales, ya que prácticamente eran un volcado de los medios impresos, supusieron la viabilidad temporal de esta tendencia mundial. Sin embargo, la evolución del periodismo digital, ligado al propio desarrollo de Internet, ha puesto de manifiesto la creciente necesidad de cobrar por la información de calidad o contenidos específicos que haya elaborado el medio digital en cuestión, con el objetivo de mantener su viabilidad. Mientras que las informaciones generales que poseen todos los medios serían gratuitas, aquellas más elaboradas supondrían un coste adicional al usuario. Al parecer esta tendencia, corroborada por los entrevistados, se está

manifestando cada vez más en Internet como única vía para salir de la crisis en la que actualmente se hayas sumidas las empresas dedicadas a la comunicación digital.

Como podremos ver a lo largo de esta investigación, ciertamente Internet ha abierto múltiples opciones para el periodista, sin embargo en la actualidad la tecnología impide o dificulta la realización práctica de algunas de ellas. El ancho de banda actual es uno de los principales problemas que provoca que en numerosas ocasiones el usuario se sienta defraudado ante la larga espera de carga de una determinada página web. A medida que avance la técnica, y sobre todo que ésta pueda ser adquirida por la mayor parte de los usuarios, sin que un elevado precio se lo impida, surgirá realmente y de forma práctica esta nueva y revolucionaria forma de hacer periodismo.

Debido a los términos técnicos necesarios para la explicación de diversas cuestiones relacionadas con el mundo de Internet hemos incluido entre los anexos un glosario como apoyo de nuestro trabajo, donde se podrá recurrir para encontrar una definición precisa.

En definitiva, a lo largo de la investigación hemos analizado los retos que supone Internet para el periodista que trabaja en este nuevo medio, poniendo especial atención en la especificación y descripción de las distintas fórmulas y herramientas que tiene a su disposición el profesional de la información a la hora de realizar una búsqueda en la Red. Tanto la primera, como la segunda parte de la tesis sirven para comprender la relativa a la Red Profunda (tercera parte) en la que nos hemos centrado en un análisis más exhaustivo dada la importancia de su conocimiento para el profesional de la información y su elevado grado de desconocimiento actual en nuestro país.

DESARROLLO DE LA INVESTIGACIÓN

En esta tesis, nos proponemos demostrar cómo Internet ha incidido de forma significativa en la profesión del periodista suponiendo para la profesión grandes posibilidades y nuevos retos. Nuestro objetivo ha sido, además, analizar las posibilidades que ofrece Internet como fuente de información para el periodista que trabaja a través de este nuevo medio y que, de la misma forma, pretende difundir su conocimiento, ofreciendo también en sus informaciones las ventajas que la Red de redes dispone. Hemos puesto especial atención en la explicación y análisis de aquella parte de Internet no accesible mediante buscadores generalistas, conocida como Red profunda, dado el gran volumen y calidad de información que contiene.

Partimos de las siguientes premisas

- La Red de redes ofrece al campo de la comunicación grandes oportunidades derivadas de la propia estructura de Internet, entre las que se encuentra la universalidad (cualquier persona que posea un ordenador conectado a la Red puede acceder a la información allí disponible), la posibilidad de que el público-usuario entre en contacto directo con la fuente de información o el acceso de los llamados internautas a sugerentes bases de datos cuya publicación en otro medio sería impensable.
- Precisamente la gran cantidad de información disponible en la Red y su necesidad de ser profesionalmente verificada y eficazmente seleccionada, abre un nuevo campo de especialización para el periodista.
- En numerosas ocasiones la tarea de buscar datos concretos en la llamada telaraña mundial y también mal llamada autopista de la información (ya que no posee precisamente un tráfico de datos muy ordenado), puede resultar bastante ardua para un usuario que no conoce los distintos sistemas de búsqueda ideados con este propósito. Una de las tareas primordiales del periodista debe de ser conocer

a la perfección estas herramientas y, de esa manera, servir de mediador entre la masa caótica de información de la telaraña mundial y el público-usuario.

- La Red profunda, parte de Internet no accesible mediante buscadores generalistas, abre un campo inmenso de gran utilidad para la investigación para el periodista, ya que posee 500 veces más información y, sobre todo, una mayor calidad. Conocer de su existencia es esencial ya que permite una búsqueda más especializada y fiable.

Fuentes consultadas

El objetivo de la presente tesis condiciona su metodología. Ya que tratamos de explicar como Internet es una fuente clave de información para el periodista digital y precisamente para probar la tesis que aquí se afirma, la mayor parte de los documentos que hemos utilizado para la realización de este trabajo, han sido extraídos de la propia Red.

Desde hace aproximadamente 6 años, coincidiendo con la explosión de Internet como medio de comunicación social (1995), el autor comienza a interesarse por el nuevo medio que comenzaba a surgir. Enseguida advierte sus posibilidades y sobre todo se siente impresionado por la gran cantidad de información allí disponible, ya entonces, a cualquier persona que dispusiera de un ordenador conectado a Internet. A partir de este momento comienza a recopilar material referente a este tema, principalmente procedente de revistas, agencias de noticias, periódicos y medios de actualidad en general. Las fuentes bibliográficas eran todavía escasas dada la novedad y rápido avance de esta nueva tecnología.

Asimismo, hemos tenido acceso a la base de datos de la Agencia EFE, las fichas que recogen todas las ediciones del diario *El País* desde sus Inicios (1976), los CD Rom del diario *El Mundo* que, del mismo modo, contienen la mayor parte de las noticias aparecidas en él, y sobre todo, la mayor fuente de datos mundial: la Red Internet. El acceso a todos estos medios, permitió al autor, desde los comienzos de su Investigación, seguir la trayectoria de esta nueva tecnología, observar su rápida evolución y sobre todo, su progresivo impacto social.

De gran ayuda a la hora de resolver los problemas y dudas que se iban presentando han sido las listas de distribución en las que participamos (Red IRIS, Enredando...), ya que, de esta forma, hemos tenido la oportunidad de entrar en contacto, a través del correo electrónico, con diversos investigadores e interesados en la misma materia, a nivel mundial, de forma rápida y directa. Ésta es una de las grandes ventajas de la Red de redes, la desaparición de las fronteras geográficas. Gracias a Internet es posible entrar en contacto con expertos de todo el mundo y obtener fácil y cómodamente una visión global de tu investigación.

Asimismo tanto la experiencia que hemos acumulado de dos años en el medio digital como las diversas entrevistas realizadas a varios de los máximos expertos en esta materia, que se incluyen en el apartado de apéndices, han sido de gran utilidad a la hora de dar una visión práctica a este proyecto.

Quizá lo más difícil de la investigación haya sido, no los problemas comunes que suelen presentarse en estos casos como la falta de accesibilidad a una información en concreto o los reducidos datos disponibles sobre un tema, sino precisamente lo contrario. La gran masa de información que podemos encontrar en Internet produce precisamente el efecto contrario: la desinformación, es decir, grandes contrastes y opiniones diferentes sobre un mismo tema, documentos aparentemente interesantes cuya fuente desconocida es poco fiable, y cada vez mayor multitud de datos derivado de la facilidad de publicación en la Red. Para solventar este problema hemos tenido que llevar a cabo un elaborado proceso de selección y comprobación de dichas fuentes.

Asimismo, otro de los problemas que se han presentado, ha sido la rápida evolución del medio base de la investigación. Muy a menudo surgen nuevos avances que modifican las posibilidades de Internet y por tanto del periodismo *online*. Debido a esto, la actualización y revisión de los distintos capítulos de este proyecto de investigación ha tenido que ser continua. Sin embargo, aunque por un lado pueda suponer un inconveniente, por otro este hecho ha aportado a la investigación un gran dinamismo y nosotros gran curiosidad por el futuro inmediato de esta nueva tecnología y de sus implicaciones sociales.

AGRADECIMIENTOS

A lo largo de la realización de esta tesis han sido muchos los que me han apoyado y ayudado en las dudas y dificultades que me iban surgiendo. Me gustaría, en primer lugar, mostrar mi agradecimiento a mi director de tesis Leopoldo Seijas sin cuya guía y consejo habría estado perdida. Asimismo, quisiera agradecer especialmente su apoyo y ánimo incondicional a mis padres, a mi novio, y en general a todo mi familia que han tenido que soportarme hablando casi continuamente de un trabajo que para ellos (y para mí en ocasiones también) se mostraba eterno. Quisiera hacer mención en estos agradecimientos a las personas a las que he robado su tiempo en diversas entrevistas, entre ellos: Mariló Ruiz de Elvira, Mario Tascón, Ramón Salaverría y Ricardo Fornás. En un periodo de varios años son muchos los que, en un momento u otro, te prestan su apoyo. Al no poder citarlos a todos me conformaré con aludir a mis compañeros de trabajo de Inicia (empresa del grupo PRISA ahora desaparecida) en especial a Xabier Meilán y a Paco Olivares, que han sido claves en los momentos más difíciles.

PRIMERA PARTE: LOS RETOS DEL PERIODISTA DIGITAL

CAPITULO 1: ¿QUÉ ES INTERNET?

1.1 BREVE HISTORIA DE INTERNET

Internet ha revolucionado en poco tiempo el mundo de las comunicaciones pasando de ser un sistema de comunicaciones de elite usado por el Departamento de Defensa del Gobierno de los Estados Unidos y los más selectos grupos de investigación de las principales ciudades del mundo, a formar parte de nuestras vidas tanto desde el punto de vista social, como laboral y económico. La historia de este nuevo medio de comunicación es reciente y compleja debido a las propias condiciones de sus inicios.

Tras la Segunda Guerra Mundial, Estados Unidos comenzó a crear una serie de organismos cuyo objetivo era recuperar nuevamente la ventaja tecnológica después del lanzamiento soviético del Sputnik, y estar prevenida frente a nuevos avances tecnológicos de enemigos potenciales. Fue así como nació en 1958 el “Advanced Research Project Agency”, o ARPA. Su organización interna estaba dividida en departamentos que trabajaban en diferentes campos, pero no disponían de medios propios de investigación. La actividad principal, que se conoce, era la financiación de cualquier iniciativa, tanto pública como privada, que se considerase útil para futuras aplicaciones en materia de defensa.

Los fondos que aportaba el ARPA dotaban de unas extraordinarias infraestructuras a las organizaciones que los recogían, que eran normalmente universidades y centros estatales de investigación. Por lo tanto, aunque el origen fue militar, las iniciativas y los creadores materiales de Internet procedían de los entornos académicos del país.

En esta época, las comunicaciones militares en EEUU estaban basadas en enlaces que intercomunicaban los diferentes centros de mando de las autoridades con todas las ciudades y bases del país. Fue en mayo de 1961 cuando fueron detenidos unos miembros del Ejército Republicano Americano acusados de volar tres torres de comunicaciones en el Estado de Utah. Tras el atentado parte del sistema de comunicaciones norteamericano quedó interrumpido y puso en evidencia la vulnerabilidad del mismo. Dado este hecho, el Departamento de Defensa de los

Estados Unidos encargó a la RAND Corporation²² (una de las primeras organizaciones dedicadas a realizar estudios en materia de defensa), el análisis de este problema y sus posibles soluciones. Al año siguiente, Paul Baran, responsable del análisis, entregó un informe basado en once documentos que detallaban una solución teórica denominada “Distributed Adaptive Message Block Switching”. Estos documentos fueron la primera piedra de la construcción de las redes de comunicaciones entre ordenadores y por lo tanto, de Internet.

Paralelamente al encargo realizado a la RAND Corporation a raíz del atentado, el departamento de Defensa abrió en el seno de ARPA una línea de investigación sobre sistemas de mando y control (Command and Control Research, CCR), en la que participa J.C.R. Licklider, un psicólogo formado en el Instituto Tecnológico de Massachussets. Éste dedicó dos años a este empeño y convirtió el CCR en una oficina dentro del ARPA encargada de investigar técnicas para procesar la información, denominada IPTO (Information Processing Techniques Office) y del cual fue su primer director entre 1962 y 1964. En agosto de 1962 Licklider da a conocer, en una serie de notas, la primera descripción de las interacciones sociales que se podrían permitir a través del establecimiento de una red. En estos documentos se preveía un conjunto global interconectado de los ordenadores a través de los cuales cada uno podría tener acceso rápido a datos y a programas de cualquier sitio.

En 1966 el IPTO disponía de más de una docena de centros de investigación subvencionados. La mayoría de ellos estaban dotados de instalaciones muy costosas y de grandes ordenadores. La cuestión económica se convierte en un problema, por lo que el entonces director del IPTO, Robert Taylor, toma las ideas de su predecesor Licklider, y se plantea la posibilidad de crear algún sistema que permitiese compartir todos los recursos disponibles con aquellas personas que participaban en los diferentes programas del mismo, independientemente del lugar donde se encontrase su centro de trabajo.

Desde ese momento se acuña el término “networking”, que intentaba describir un grupo de ordenadores intercomunicados entre sí, formando una red de tal manera que se pudiesen compartir los recursos. Inicialmente se trataba sólo de software; la idea de

²² RAND Corporation: <http://www.rand.org>

compartir el hardware todavía estaba lejana. Una de las pocas personas que habían realizado esta clase de experimentos, era Larry Roberts. Éste había conseguido ejecutar programas remotamente en un ordenador, utilizando una línea telefónica convencional a una velocidad de 1200 bps (bits por segundo, equivalente a unos 130 caracteres por segundo). Taylor fue la primera persona contratada específicamente para trabajar en la Red.

En abril de 1967, los representantes de este nuevo proyecto asistieron al congreso anual de responsables de investigación que ARPA organiza en la Universidad de Michigan. La propuesta que presentaron consistía en la creación de un sistema de comunicaciones que interconectase los distintos superordenadores financiados por ARPA con el objetivo de poder extender su uso y rentabilizar así al máximo la inversión realizada. El primer inconveniente con el que se encuentran es que cada uno de los ordenadores de los que disponían, eran de un fabricante diferente, tenían diferentes sistemas operativos y no tenían prácticamente nada en común. Para solventarlo crean el denominado “Communication Group” que se encargó, en un primer lugar, de construir la red de datos físicamente, es decir, definir los canales de comunicaciones que se iban a utilizar para intercomunicar los diferentes ordenadores, qué velocidades de transmisión se usarían y su coste. Otro de sus cometidos era conectar cada ordenador con la red, decidir qué sistema operativo usarían y qué “protocolo” sería el elegido para su intercomunicación.

Tres meses después, el Grupo de Comunicación, consiguió solucionar el primer problema, adoptando como soporte de comunicación las líneas telefónicas que estarían permanentemente conectadas entre los ordenadores, sobre las que se transmitirían “paquetes” de información. El segundo se solventa construyendo un dispositivo que tuviese dos partes diferentes. Una de ellas sería la conectada a las líneas de comunicaciones y al resto de la red. Estaría compuesta por equipos electrónicos y un protocolo de comunicaciones idéntico en todos los sitios. Se entiende por protocolo al conjunto de normas de diseño electrónico y de programación que permitieran abrir un canal de comunicación entre dos ordenadores, hacer fluir a través de él la información y, una vez finalizada la transmisión, cerrarlo. La segunda parte sería específica para cada uno de los sitios y se construiría en función de los ordenadores allí instalados. Desde entonces se conocen como “hosts”.

A partir de estas resoluciones el ARPA proporciona a la futura Red la infraestructura mínima compuesta de líneas de comunicaciones y de parte común de los dispositivos que comienzan a llamarse IMP (Interface Message Processor), predecesor del que hoy denominan “routers”.

El 3 de junio de 1968, el director de IPTO, Robert Taylor, presenta al director de ARPA un plan denominado “Resource Sharing Computer Network” (Red de Ordenadores para Compartición de Recursos), que contenía el resultado de los últimos trabajos realizados para construir la red. El 21 de este mismo mes, ARPA asigna 500.000 dólares a la construcción de su red de datos denominada “ARPA NETwork” (ARPANET).

Anterior al protocolo TCP/IP existió otro denominado NCP (Network Control Program), que fue desplazado por el actual debido a sus problemas internos de funcionamiento. Con el objetivo de buscar una solución Bob Kahn y Victor Cerf, propusieron la creación de un nuevo protocolo encargado de sustituir al que había entonces dando cabida a todas las redes existentes en ese momento.

Cerf , en aquella época ya había sido nombrado presidente del INWG (International Network Working Group) y se encontraba en la Universidad de Stanford, lugar donde comenzó a trabajar y experimentar en un tema que empezaba a ser conocido con el término “Internetting”. La investigación que Kahn y Cerf estaban realizando, concluyó en los dos protocolos conocidos actualmente (TCP/IP; Transmission Control Protocol e Internet Protocol respectivamente), que se encargan de transmitir paquetes de información entre dos ordenadores.

De esta manera, llegamos a la World Wide Web actual, y al auge del usuario privado, el cual puede utilizar el canal telefónico convencional para acceder a Internet. Pero en este caso, aparte del gasto que supone la llamada, se tendrá que pagar también el uso de un determinado ancho de banda, para que el proveedor le deje circular por su canal que le comunica con el resto de la red a través de los “backbone”. Un “backbone”, es una línea de comunicaciones principal donde se conectan infinidad de ordenadores compartiendo posibilidades de comunicación, así como los equipos electrónicos puestos a su servicio.

El ancho de banda citado anteriormente, es muy importante a la hora de calibrar la velocidad y calidad de transmisión de los datos. Actualmente la Red tiene un ancho de banda pequeño, comparable a las carreteras locales, mientras que un futuro próximo se espera que se llegue a “circular por carreteras nacionales”, es decir, con un ancho de banda bastante mayor. Sin embargo, por mucho que evolucione, siempre se demandará más, debido a que seguirán saliendo muchas aplicaciones de imagen y sonido complejos que requieran una mayor capacidad de la Red.

1.2 FUNCIONAMIENTO DE LA RED

La principal novedad que planteaba Arpanet era la integración de un protocolo de comunicaciones basado en la conmutación de paquetes y no el método antiguo de conmutación de circuitos, que generaba redes muy inestables y poco robustas. Precisamente éste fue el trabajo realizado por Leonard Kleinrock cuando en julio de 1961 demostró la eficacia del método de conmutación de paquetes. En la primitiva Arpanet, los diferentes nodos de la red eran muy similares pero la evolución de Internet vaticinaba la incorporación de diferentes tipos de redes: Redes de paquetes por satélite, por radio y otros, para conseguir este objetivo era necesario disponer de un protocolo de comunicaciones muy superior al original que usaba Arpanet.

El primer protocolo que permitió la conexión de dos nodos fue NCP (Network Control Protocol) desarrollado por S. Crocker en diciembre de 1970 en el Network Working Group. La implementación terminó a principios de 1972. Mientras el tipo de nodos fue similar, el sistema funcionó a la perfección.

Sin embargo, el sistema estaba calculado para funcionar en una red limitada y con no demasiados nodos, como era Arpanet. El protocolo NCP tenía serias limitaciones inherentes al sistema que hacían que fuera necesario cambiar toda la estructura de la red para superarlas. NCP sobrevivió casi hasta mediados de los años 80. El cambio a un nuevo protocolo de comunicación entre los *host*, TCP/IP, se convirtió en una complicada transición.

Robert Khan, ingeniero de BBN (Bolt Beranek and Newman Ltd) durante los comienzos de la Red en 1968, decidió desarrollar una nueva versión del protocolo de

comunicación que permitiera, por un lado, corregir los errores de comunicación y por otro, permitir interactuar con otras redes de forma abierta en un nivel de trabajo en redes que se conocía como *internetworking*.

El cambio de un protocolo de comunicación a otro terminó siendo algo más que un simple cambio técnico. El advenimiento del TCP/IP supuso la ruptura con la primera estructura de la Red. Arpanet tuvo que evolucionar hacia una estructura mucho más abierta que permitiera la inclusión de nuevas redes de forma sencilla.

Cuatro reglas básicas fueron críticas en el pensamiento de Kahn en esta primera época:

- a) Cada red distinta tenía que permanecer igual y no se requerían cambios internos en cualquier red para conectar a Internet.
- b) Las comunicaciones serían sobre la base del mejor esfuerzo. Si un paquete no llegaba al destino final, rápidamente sería retransmitido desde el origen.
- c) Se usarían cajas negras para conectar las redes; posteriormente se les llamaría pasarelas (gateways) y enrutadores (routers). Las pasarelas no retendrían información acerca de los flujos individuales de paquetes que pasan por ellas, de este modo se las mantenía simples y se evitaba la complicada adaptación y recuperación desde distintos modos de error.
- d) No habría control global al nivel de operaciones.

Después de estos cambios Arpanet se convirtió en Internet. Los promotores de esta evolución fueron dos ingenieros que habían trabajado en el proyecto desde los primeros días: Robert Khan y Victor Cerf.

A la hora de entender el significado de Internet hay que tener en cuenta que la conexión que se produce es entre multitud de redes compuestas por ordenadores de todo tipo. Una red que conecta computadoras con diferentes sistemas operativos (Unix, MacOS, Windows, MSDos, VMS, etc) y que maneja millones de datos de información por segundo, necesita un protocolo de comunicación que mantenga en todo momento al sistema firmemente cohesionado. Éste es el llamado protocolo TCP/IP, que se deriva de los dos protocolos más importantes de Internet: *Transmisión*

Control Protocol (TCP) e *Internet Protocol* (IP), sin los cuales una red de nodos descentralizados sería impensable.

La información en Internet no viaja como una cadena continua de caracteres de nodo a nodo. Los datos se envían en pequeños trozos de información que llamamos paquetes. Por ejemplo: si queremos mandar un documento a otra persona, que vive en cualquier lugar del mundo, mediante el correo electrónico, TCP dividirá el mensaje en paquetes, a cada paquete se le asigna un número secuencial y la dirección del destinatario, junto a determinada información de control de errores. En entonces cuando entra en funcionamiento el trabajo de IP, que debe encargarse de llevar los paquetes hasta el nodo de destino, usando el camino más seguro. Si en algún momento el tráfico resultara demasiado denso en alguna parte de la Red, IP está capacitado para alterar la ruta y buscar otro camino alternativo.

Cuando a un nodo llega un paquete que no es para él, se lo pasa al siguiente nodo de la cadena hasta que llega al *host* de destino. Cuando esto ocurre TCP retoma toda la información, reordena los paquetes y comprueba que no hay ningún error.

El TCP/IP tiene una serie de características que lo diferencian de otros protocolos de comunicación. En primer lugar, las líneas de comunicación pueden ser compartidas por muchos usuarios, todos los paquetes de comunicación están perfectamente identificados. En segundo lugar, si por alguna causa una ruta queda fuera de servicio, las computadoras que regulan el flujo de datos (*routers*) pueden encontrar un servicio alternativo. Y por último, si algo falla en la transmisión, sólo hay que reenviar una pequeña parte de la información, ya que ésta se encuentra dividida en paquetes.

La gran ventaja de Internet consiste en poder compartir aplicaciones entre máquinas diferentes conectadas por un protocolo de comunicaciones único. Para que Internet pueda funcionar como medio de comunicación válido, es necesario poder intercambiar datos entre todo tipo de plataformas informáticas. Para ello hay que evitar las incompatibilidades inherentes a los distintos sistemas operativos. La arquitectura cliente/servidor trata de evitar este problema a la vez que posibilita el crecimiento del sistema y la creación de nuevas aplicaciones.

Básicamente un servidor es un ordenador con conexión a Internet, especializado en realizar un trabajo determinado. Un cliente es un *software* que es capaz de conectarse a un servidor y obtener los datos que éste aloja. Cada cliente está especializado en un determinado servicio ofrecido por un servidor. Existen clientes de correo electrónico, de FTP, de IRC, etc. A su vez estos clientes pueden haber sido escritos para cualquier sistema operativo conocido; de esta forma cualquier sistema operativo que se use en futuro podrá permitir que se programe un cliente para usar cualquier recurso ofrecido por un servidor conectado a la Red.

Durante su estancia en BBN, Robert Kahn investigó en los principios que harían funcionar el nuevo protocolo. Sus primeras ideas dieron lugar a un memorándum que lleva por título "*Communications Principles for Operating Systems*". En primavera de 1973 se puso en contacto con Vinton Cerf, de la Universidad de Standford, para que le ayudara a llevar a cabo el proyecto. Cerf ya había trabajado en el desarrollo del NCP y era un experto en la construcción de interfaces para todo tipo de sistemas operativos.

En septiembre de 1974 se dio a conocer una primera versión del protocolo en una sesión especial del INWG (International Network Working Group) celebrada en Sussex. Diez años más tarde, el 1 de enero de 1983, se produjo el cambio definitivo de NCP a TCP/IP, y, por tanto, el comienzo de Internet como la conocemos hoy en día.

Los protocolos son una serie de reglas que describen técnicamente cómo deben realizarse determinadas tareas. Por ejemplo, hay un protocolo que explica exactamente el formato que debe de tener un mensaje. Todos los programas de correo de Internet seguirán este protocolo para enviar un mensaje. TCP/IP es el nombre genérico de más de 100 protocolos que se usan para conectar ordenadores y redes. Cada nodo en la Red está identificado por un número único compuesto por cuatro secuencias de 8 bits. Este número se llama dirección IP y sirve para administrar el tráfico de datos en la Red.

En la actualidad, Internet es un conjunto de redes locales conectadas entre sí a través de un ordenador especial por cada red, conocido como "gateway". Las interconexiones entre "gateways" se efectúan a través de diversas vías de

comunicación, entre las que figuran líneas telefónicas, fibras ópticas y enlaces por radio. Como hemos explicado anteriormente, la información que debe enviarse a una máquina remota se etiqueta con la dirección computerizada del ordenador en cuestión. Internet no tiene un control central, es decir, ningún ordenador individual que dirija el flujo de información. Esto diferencia a Internet y a los sistemas de redes semejantes de otros tipos de servicios informáticos de red como CompuServe, America Online o Microsoft Network.

Las redes situadas fuera de Estados Unidos utilizan sufijos que indican el país, por ejemplo (.es) para España o (.ar) para Argentina. Dentro de Estados Unidos, el sufijo anterior especifica el tipo de organización a que pertenece la red informática en cuestión, que por ejemplo puede ser una institución educativa (.edu), un centro militar (.mil), una oficina del Gobierno (.gov) o una organización sin ánimo de lucro (.org).

1.3 SERVICIOS DISPONIBLES EN INTERNET

El correo electrónico

Una vez conseguido el primer objetivo de compartir recursos informáticos en EEUU, quedaba aún algo por lograr: la posibilidad de que los científicos de los centros de investigación que compartían sus recursos pudieran enviarse fácilmente mensajes entre sí, sin necesidad de recurrir al teléfono. A pesar de no existir correo electrónico como tal, ya en los años 60 un sistema para intercambiar mensajes entre usuarios de un mismo ordenador que se conocía con el nombre de *Intra-computer mail*.

En 1971 Ray Tomlinson, ingeniero Bolt Beranek and Newman (BBN) comenzó a investigar la posibilidad de intercambiar mensajes entre los usuarios de la red. Entre otras, a él se debe la inclusión del símbolo @ para representar la inclusión en una determinada máquina. Su propuesta radicaba en el desarrollo de ficheros de texto donde los usuarios escribiesen sus mensajes y posteriormente fueran enviados utilizando protocolos de transferencia de ficheros. Este programa lo diseñó para funcionar bajo el

sistema operativo TENEX y lo llamó SNDMSG. Éste era una versión desarrollada a partir del sistema de correo *Intra-computer* que proporcionaba un rudimentario *interface* de usuario en forma de editor de textos y de una aplicación para la transferencia de ficheros denominada CPYNET. Mezclando las dos aplicaciones, un usuario generaba un mensaje de la misma manra que si fuese a ser enviado a otro usuario de su misma máquina, pero en lugar de depositar dicho mensaje en el casillero local (mailbox) determinado, el CPYNET se encargaba de realizar una copia en una máquina remota. Lo podemos ver en el siguiente ejemplo:

```
Arpavax::!>
Arpavax::!>mail mary
Arpavax::!>Subjet: Cumpleaños Mary
Arpavax::!>Write the message and press a dot in a single
line!!
Arpavax::!>Hola Mary
Arpavax::!>Quería felicitarte por tu cumpleaños.
Arpavax::!>Espero verte pronto
Arpavax::!>
Arpavax::!>Un abrazo
Arpavax::!>.
Arpavax::!>
Arpavax::!>Message 030575@1276345 Send OK
Arpavax::!>
Arpavax::!>
```

Otro ejemplo de cómo ya en los principios de ARPANET los investigadores hacían uso intensivo del correo electrónico podemos encontrarlo en este extracto de uno de los textos fundamentales en la historia de Internet: “Pequeña historia de Internet” del escritor Bruce Sterling (1993)

“En su segundo año de operatividad, sin embargo, algo extraño se hizo patente. Los usuarios de ARPANET habían convertido la red en una oficina de correos electrónica de alta velocidad subvencionada federalmente. La mayor parte del tráfico de ARPANET no era el proceso de datos a largas distancias. En vez de eso, lo que se movía por allí eran noticias y mensajes personales. Los investigadores estaban usando ARPANET para colaborar en proyectos, intercambiar notas sobre sus trabajos y, eventualmente, chismorrear. La gente tenían sus propias cuentas personales en los

ordenadores de ARPANET y sus direcciones personales de correo electrónico. No es que sólo utilizaran ARPANET para la comunicación de persona a persona, pero había mucho entusiasmo por esta posibilidad, mucho más que por la computación a larga distancia²³.”

En un principio por correo electrónico sólo podían mandarse datos consistentes en caracteres ASCII. Al no poder ser enviados datos binarios, era imposible mandar por correo electrónico ningún tipo de programas. Mediante un software llamado UUENCODE se podían convertir los datos binarios en ASCII. Mediante el software UUDECODE se podía invertir el proceso. Este mismo trabajo lo hace hoy el protocolo MIME (*Multipurpose Internet Mail Extensions*) y BinHex. Gracias a esta facilidad es posible enviar por correo electrónico cualquier tipo de datos: imágenes, sonido, etc. Habitualmente MIME es el formato que se usa en entornos PC y BinHex en los Macintosh.

El primer programa de correo electrónico que permitía guardar, contestar y reenviar los mensajes fue escrito en 1975 por John Vittal, su nombre fue MSG.

Direcciones de correo electrónico

Las primeras direcciones de correo electrónico se parecían a las actuales pero no aportaban información acerca del dominio al que pertenecía el usuario debido a que aún no se había inventado el sistema de dominios. Este sistema resultaba problemático ya que resultaba muy ambiguo incluso para los pocos usuarios que había entonces interconectados. Los nombres de las máquinas a veces coincidían al igual que el nombre de los usuarios.

Poco después comenzaron a desarrollarse otras versiones que proporcionaron herramientas de correo similares a las que existen en la actualidad. Algunos de estos programas fueron: READMAIL, NRD, WRD, BANANAR o HERMES. Sin embargo, la utilización de distintos programas, cada uno con un lenguaje distinto era un claro inconveniente. Por este motivo, en julio de 1971 se publicó RFC 196 titulado “Mail Box Protocol”, cuyo objetivo era acordar un formato de correo único para garantizar la comunicación entre todas las máquinas.

²³ Sterling Bruce: *Pequeña historia de Internet*. Revista de fantasía y ciencia ficción. CT (USA) Febrero

El sistema de correo que utilizamos en la actualidad está basado en el RFC 822 de agosto de 1982. A partir de éste se han ido añadiendo modificaciones menores para estandarizar el uso de los nuevo contenidos de mensajes.

Hoy en día una dirección de correo electrónico consta de tres partes: usuario, sistema al que está conectado y dominio de esa máquina. La separación entre el nombre del usuario y su sistema se hace por medio del símbolo @. El sistema y las diferentes máquinas se separan del dominio por medio de puntos.

javier-alvarez@inicia.es

En la siguiente tabla se exponen algunos de los dominios que más se utilizan y su significado.

DOMINIOS DE PRIMER NIVEL Tipo organización	DOMINIOS DE PRIMER NIVEL Tipo geográfico
com organización comercial edu institución educativa gov gobierno int organización internacional mil organización militar net gestión de redes org organización no lucrativa	at Austria au Australia ca Canadá de Alemania dk Dinamarca es España fr Francia...

A continuación exponemos diversos conceptos relacionados con el correo electrónico:

- **Agentes de usuario:**

Son los programas que se encargan de tratar con el agente de transporte de la computadora encargada del correo de nuestro nodo: Mail, Ean, Pegasus, Eudora. Actualmente se tiende a usar una suite que integra diversos servicios todo-en-uno sobre

la base de un navegador, así los agentes de usuario más usados en Internet son Outlook Express que viene integrado en la suite Explorer de Microsoft y Messenger de la suite Communicator de Netscape.

- **SMTP**

Simple Mail Transfer Protocol. Es el protocolo mediante el cual los nodos se intercambian el correo entre sí. El agente de transporte es el programa que se encarga de la gestión de correo, en segundo plano, dentro de los nodos. *Sendmail* es uno de los programas más conocidos para realizar esta función.

- **POP**

Post Office Protocol. Es un protocolo más moderno que SMTP y permite recibir en nuestro propio ordenador el correo. La mayoría de los agentes de usuarios que se usan en ordenadores personales trabajan con este protocolo. En los archivos de configuración suele encontrarse este parámetro como dirección de la máquina que se encargará de entregarnos el correo.

- **Netiquette**

El uso del correo electrónico conlleva una determinada forma de comportamiento en red. En muchos casos se tiende a escribir de la misma manera en que se habla, pero la palabra hablada transcrita al papel pierde los matices de la conversación. Para dotar al correo electrónico de mayor facilidad para la comunicación se han establecido una serie de convenciones:

- No enviar nunca información confidencial por e-mail.
- No usar mayúsculas CONTINUAMENTE, da mala impresión.
- Para contestar a un mensaje se suele acotar la parte del mensaje que se responde con este símbolo >, pero no se debe reenviar al destinatario su mensaje original completo. Tampoco se debe contestar a un mensaje sin hacer ningún tipo de referencia al mensaje original.
- Se debe enviar un acuse de recibo en el caso de mensajes de cierta importancia. Si se ha recibido un mensaje que requiere una contestación por nuestra parte y vamos a demorarnos en su respuesta, se debe enviar un acuse de recibo y una

fecha aproximada de contestación. El correo electrónico es un medio de comunicación muy veloz y lo habitual es la contestación a vuelta de correo.

- Pueden usarse *smileys* sin abusar, para matizar algunas expresiones. Los *smileys* o caritas como se conocen en castellano sirven para dotar a la palabra escrita de la frescura de la conversación. El origen de esta costumbre estuvo en un mensaje enviado el 12 de abril de 1979 por Kevin McKenzie a MsgGroup. En este mensaje se sugiere añadir “sentimientos” a los textos agregando algunos símbolos como por ejemplo -) para sugerir ironía. A continuación pueden verse algunos de los más usados.

:-) Alegre
 :-(Enfadado
 :-/ Mosqueado
 :-| Indiferente
 ;-) Guiño
 :,-(Lloroso

- **Listas de Correo:**

(*Mailing list*) Una serie de suscriptores reciben todos los mensajes escritos por el resto de los participantes de la lista. Por lo general se trata temas muy concretos moderados por uno de los suscriptores. El cliente más conocido para gestionar este tipo de listas es Majordomo, el cual permite al moderador automatizar la gestión de altas y bajas en la lista.

Otros conceptos relacionados con Internet que merece la pena mencionar son los siguientes:

- **FTP**

FTP es uno de los protocolos más potentes de Internet, nos permite enviar y recibir archivos desde cualquier máquina que esté conectada a la Red, El *File Transfer Protocol*, junto con el correo electrónico contribuyó a la popularización de Internet.

Para poder cargar (*upload*) o descargar (*download*) ficheros es necesario tener en nuestro ordenador un cliente FTP. Los primeros clientes FTP que corrían sobre máquinas Unix o VMS eran difíciles de manejar obligando al usuario a teclear complicadas combinaciones de comandos, hoy en día los clientes que funcionan sobre ordenadores personales tienen un funcionamiento intuitivo que no hace necesario recordar los antiguos *get*, *open* o *cdup*. Algunos de los más populares son WS-FTP o CuteFTP, aunque toda esta labor ya suele hacerse habitualmente de forma transparente para el usuario desde cualquier navegador Web.

Algunos servidores FTP admiten conexiones anónimas usando como *login* bien *anonymous* o bien *guest*. Esto significa que podemos conectarnos a ellos para obtener ficheros de libre distribución, controladores, documentos de texto, etc. A continuación se muestran algunas de las más conocidas direcciones de FTP anónimo:

garbo.uwasa.fi
ftp.cica.indiana.edu
ftp.coast.net

- **TELNET**

Telnet es el recurso que permite la conexión a una computadora remota en Internet, permitiendo usarla desde nuestro puesto como si realmente estuviéramos al teclado de la otra máquina. La primera experiencia relacionada con la conexión de dos ordenadores en lugares remotos fue, de hecho, una conexión vía Telnet. Aunque hoy es un servicio muy poco usado, durante años fue el motivo de la existencia de Internet: poder compartir la potencia de proceso de ordenadores remotos para cálculos científicos realmente complicados. Junto al correo electrónico y al FTP, fueron, durante años, los tres recursos que componían el núcleo de la Red.

Posteriormente Telnet fue la forma habitual de conectarse con BBS's de Internet y de acceder a un gran número de bases de datos públicas y privadas.

El acceso a servidores Telnet se realiza mediante la entrada de un nombre de usuario y de una clave. Al igual que los servidores FTP, muchos servidores Telnet permiten el acceso público a través de cuentas de invitado o anónimo. La configuración de estas cuentas suelen ser complicadas debido al tipo de terminal que debe emular para

poder acceder al servicio.

Para simplificar la comunicación con los servidores Telnet existe un programa llamado HYTELNET (*Hypertext Telnet*), un programa basado en menús de muy fácil manejo. Este programa canadiense proporciona información ordenada, no sólo de los nombres de los anfitriones de acceso público y los *login* y *password* necesarios para acceder, sino que se indica el tipo de información y la manera de acceder a ella. Se obtiene gratis en numerosos servidores FTP en la Red, aunque el uso masivo del Web ha hecho casi desaparecer los servidores Telnet en Internet.

- **NEWS**

Existen más de 60.000 áreas de discusión en las que se tratan todo tipo de temas, relacionados o no con la informática y los diversos campos de investigación de las distintas facultades. Los *newsgroups* pertenecen a un sector de la Red llamado USENET (Unix User Network). Usenet no es una red sino una gran BBS (Bulletin Board System) a la que se puede acceder vía Internet. Por otra parte puede darse el caso de participar en Usenet sin estar conectado a Internet, por ejemplo a través de un sistema Unix conectado por UUCP a otra computadora que recibe los artículos de los *newsgroups*.

Usenet se creó en 1979 para intercambiar artículos y noticias entre las universidades de Carolina del Norte y Duke mediante el protocolo UUCP, gracias al esfuerzo de los estudiantes Tom Truscott, Jim Ellis y Steve Bellovin. El protocolo UUCP (Unix to Unix Copy Protocol) fue desarrollado por el AT&T Bell Lab en 1976 y desde el año siguiente se incluía en el sistema operativo Unix..

Los grupos de noticias están moderados, es decir, es necesario que haya una persona que se encargue de organizar y moderar la participación de los usuarios. Esta persona puede coincidir con el administrador del sistema aunque no es obligatorio.

Para poder acceder a las *news* es necesario conocer la dirección de la máquina destinada en nuestro nodo a almacenar los artículos y tener instalado un cliente lector de noticias. Hasta hace poco se usaba Free Agent en entornos de ordenadores personales y nn, rn, tin o trn en los sistemas Unix. Actualmente las suites más usadas para conectarse a la Red (Explorer y Netscape) incluyen lector de noticias. Para un uso intensivo o especializado sigue siendo recomendable usar un cliente específico.

Los grupos de interés se identifican de forma jerárquica en función de las primeras letras de su nombre. Estos son algunos de los más habituales:

ALT. Actividades alternativas: Cajón de sastre al que van aquellos grupos que no tienen cabida en otras categorías.

COMP. Grupo informático, computadoras: Lenguajes de programación, sistemas operativos, seguridad informática, etc.

REC. Actividades recreativas: Música, deportes, cine, etc.

SCI. Grupo científico: Todo tipo de temas basados en la ciencia.

SOC. Social: Cada país tiene un foro asociado en el que se reúnen principalmente extranjeros interesados en esa cultura y nativos desplazados a otros países.

ES. Así empiezan todos los grupos españoles gestionados por Red Iris. Estos grupos están escritos en español.

- **IRC**

(*Internet Chat Relay*) - Es un sistema de charla en tiempo real, basado en canales temáticos a los que los usuarios pueden conectarse y participar en vivo desde cualquier lugar del mundo. Se trata de un servicio independiente que se basa en servidores dedicados y en un software específico. Sin embargo la confluencia hacia la Web hace que cada vez sea más habitual encontrar sistemas de *chat* basados en Web que no necesitan más que un explorador Web.

El programa que popularizó el chat fue IRC-Chat, su facilidad de manejo y sus posibilidades de configuración lo hicieron tremendamente popular. Otro de los programas de chat más populares es ICQ (*I seek you*, literalmente “te busco” en inglés), está soportado por una amplia comunidad de usuarios.

Los sistemas de charla en Internet nos permiten comunicación por texto o voz con los usuarios de la Red, de hecho algunos programas permiten incluso la interconexión con la red telefónica para poder comunicarse con cualquier persona del mundo que tenga un teléfono.

En los sistemas clásicos de *chat* basados en texto, se pueden encontrar programas de

mundos virtuales que permiten una experiencia. Este servicio cada vez ofrece más posibilidades que nos permiten estar informados de cuándo se conecta algún conocido y poder hablar o escribirnos con estas personas. Los programas que admiten este tipo de prestaciones se conocen como mensajería más completa e interactiva instantánea. Entre los más conocidos podemos destacar el Messenger de Yahoo, el de Netscape, empleado por los usuarios de America On Line, y Odigo. Los programas de mensajería instantánea permiten ya la comunicación vocal mediante un pequeño micrófono conectado a la entrada de audio del ordenador.

Iphone es un programa específico para realizar comunicaciones por voz, durante un tiempo fue una gran novedad, pero sus prestaciones son ofrecida por la mayoría de programa de mensajería como hemos visto más arriba. Se trataba de una versión moderna de *chat* con una comunicación similar a la telefónica que puede revolucionar las comunicaciones globales. Sobre esta base se pueden montar videoconferencias que se desarrollarán cuando el ancho de banda sea superior y permita velocidades de conexión mayor. Entre tanto han comenzado a aparecer empresas que aprovechan la potencia de Internet para comercializar telefonía básica a bajo precio entre cualquier usuario de telefonía del mundo.

1.4 GOPHER

A comienzos de los noventa, la cantidad de documentos que había en Internet era ya lo bastante amplia como para empezar a pensar en sacar a la luz lo que estaba oculto. En esta época palabras como catalogar, explorar o indexar comienzan a unirse a Internet. Aparecen utilidades como Whois, Verónica o Archie, que servían para encontrar recursos en la Red. Ordenar el caos en que se estaba empezando a convertir la Red fue algo prioritario.

Una de las ideas que estaba en la mente de los creadores de Internet era la posibilidad de poder saltar de un documento a otro, sin que importara el lugar en el que se encontrara ese documento. Sobre esta base surgió uno de los pocos recursos de Internet que han desaparecido con el paso del tiempo: Gopher.

Paul Linder y Mark P. McCahill de la Universidad de Minnesota crearon en 1991 un sistema que permitía enlazar documentos, disponibles en Internet, mediante un sencillo menú ordenado numéricamente. La gran novedad del sistema era que posibilitaba saltar de un servidor Gopher a otro con mucha facilidad.

El funcionamiento mediante menús de Gopher permitía de forma muy interactiva seleccionar diversas posibilidades entre las ofrecidas. Algunos menús permitían obtener directamente un gráfico (que se podía visualizar si se tenía instalada en la máquina la aplicación apropiada), un documento de texto u otro menú con nuevas opciones, siempre teniendo en cuenta que cada una de las diferentes peticiones podía ser dirigida a cualquier parte del mundo con el objetivo de traer el fichero demandado.

Entre los más famosos servidores Gopher que llegaron a estar en activo, habría que destacar los de las universidades americanas, el de la NASA o el de la Biblioteca del Congreso de los Estados Unidos. En España tuvo mucha repercusión el servidor Gopher de la Universidad de Oviedo que permitía el acceso desde España al resto de los servidores Gopher del mundo.

Las posibilidades de búsquedas en el llamado gopherespacio estaban aseguradas por Verónica, un índice de recursos Gopher al que se podía interrogar y devolvía una lista de directorios que contenían el elemento buscado.

El rápido crecimiento de Gopher se vio interrumpido y en muy poco tiempo terminó siendo abandonado. Se ha dado por buena la interpretación de que la desaparición de Gopher fue la evolución natural ante la aparición del Web, sin embargo el propio

creador de la *World Wide Web*, Tim Berners-Lee, afirma en su libro *Tejiendo la Red* que Gopher desapareció por la intención de la Universidad de Minesotta de comercializar su *software* de servidor Gopher.

“Fue por esa época, la primavera de 1993, cuando la Universidad de Minnesota decidió pedir una tasa de licencia a cierta clase de usuarios que querían usar el Gopher. Como el software de Gopher se estaba expandiendo tanto, la universidad iba a cobrar una tasa anual. El navegador, y el acto de navegar, serían gratuitos, y el software del servidor seguiría siendo gratuito para instituciones educativas y no lucrativas. Pero cualquier otro usuario, sobre todo las empresas, tendría que pagar para usar el software del servidor Gopher. [...] La industria abandonó Gopher como una patata caliente. Los promotores sabían que no podían hacer nada que se pudiera relacionar con el protocolo Gopher sin preguntar antes a sus abogados para negociar los derechos. Incluso si una compañía escribía su propio cliente o servidor Gopher, la universidad podía demandarla más tarde por infringir algún derecho de propiedad intelectual.”²⁴.

1.5 HISTORIA DEL HIPERTEXTO

El hipertexto es una forma de saltarse las rígidas normas que regulan el uso típico de la escritura. Se trata de una forma de escritura no lineal que permite al lector determinar en qué sentido va a dirigirse la lectura.

El término de hipertexto fue acuñado por Ted Nelson en 1965, quien lo definió en su publicación *Literary Machines* como “escritura no secuencial”. Añadía lo siguiente: “la escritura tradicional es secuencial por dos razones. Primero, se deriva del discurso hablado, que es secuencial, y segundo, porque los libros están escritos para leerse de forma secuencial. Sin embargo, las estructuras de las ideas no son lineales. Están interrelacionadas en múltiples direcciones. Y cuando escribimos siempre tratamos

²⁴ Berners-Lee, Tim: *Tejiendo la Red*. Ed: Siglo XXI. 2000. pp: 68

de relacionar cosas de forma no secuencial”²⁵. Según citan Caridad y Moscoso, el hipertexto es por tanto “un modelo basado en la idea de que el pensamiento humano funciona mediante asociaciones, y permite que, gracias a los sistemas de almacenamiento y recuperación electrónica de la información, se supere la tradicional organización jerárquica y lineal de la información”. Nelson en su obra *Literary Machines* incluía en un capítulo, un artículo anterior de Vannevar Bush titulado *As We May Think*. Vannevar Bush es una figura clave en el desarrollo del hipertexto. Nacido en 1890, sus avances en el campo del funcionamiento de la mente humana y su relación con las matemáticas, le llevaron al Departamento de Energía Eléctrica del MIT, donde permaneció 25 años y del que llegó a ser Decano. Durante la Segunda Guerra Mundial fue el director de la agencia de investigación y desarrollo científico. En esos años se encargó de la realización de diseños de antenas de radar y tablas de cálculo para el fuego de artillería. Su trabajo también estuvo relacionado con la óptica y la fotocomposición, los microfilms y las máquinas de calcular.

A finales de los años 30, Bush diseñó un prototipo para una máquina capaz de establecer relaciones entre documentos microfilmados llamada MEMEX. Ésta permitía que un individuo organizara libros, discos o toda clase de informes de forma automática y que estos pudieran ser consultados con flexibilidad y rapidez. MEMEX permitía además En 1945 Bush escribió sobre este proyecto en el artículo anteriormente mencionado *As We May Think*. Ted Nelson y Douglas Engelbart, cada uno por su parte, siguieron y aplicaron en lo posible las ideas de Vannevar Bush.

Engelbart fue el primero en llevar a la práctica un proyecto de hipertexto. En compañía de otros colegas creó durante los 60 en el Stanford Research Institute, un Centro para la investigación del engrandecimiento de la mente humana, un entorno de desarrollo en el que crearon *On-line Sistem* (NLS). Se trataba de un espacio de colaboración para el incremento de las habilidades humanas a través del uso de los ordenadores.

A parte de estos tres investigadores, muchos otros han investigado sobre el uso del hipertexto en años posteriores. Hemos de mencionar, por la importancia de sus

²⁵ Caridad, M.; Moscoso, P. (1991): *Los sistemas de hipertexto e hipermedios: una nueva aplicación en informática documental*, Madrid, Salamanca, Fundación Germán Sánchez Ruipérez; Madrid, Pirámide, 1991.

investigaciones, a George Landow y Paul Delany que editaron *Hypermedia and Literary Studies*, una colección en la que se unen diferentes autores en función de la relevancia de sus estudios. Jacob Nielsen es el autor, entre otras obras, de *Hypertext and Hypermedia*, un estudio fundamental sobre este tema.

1.6 LA WORLD WIDE WEB

Toda la evolución de Internet que hemos explicado hasta ahora, puede considerarse la prehistoria de la Red. La más reciente historia de Internet comenzó a principios de los noventa ante la necesidad de crear una herramienta de trabajo cuyo objetivo fuese el de realizar los vínculos comunicativos sencillos entre todos los componentes de un grupo de investigadores del Laboratorio de Física de Partículas del Consejo Europeo de Investigación Nuclear (CERN) en Ginebra, Suiza. Otro de los motivos, fue que los responsables de este centro constataron que mucha de la documentación generada por proyectos ya concluidos, quedaba archivada u olvidada haciéndose casi imposible su recuperación a la hora de ser reutilizada en nuevas iniciativas.

La solución a estos problemas fue propuesta en marzo de 1989 por Tim Berners-Lee, un físico graduado por la Universidad de Oxford (Inglaterra). Este estudio denominado “Information Management: A Proposal” proponía la utilización de la tecnología “HyperText” (Hipertexto) ideada en la década de los años cincuenta por Ted Nelson. El hipertexto introducía dos nuevos conceptos, el *link* (enlace) que llevaba a otra parte del documento o incluso a documentos diferentes, y el concepto de multimedia, es decir, la capacidad de un documento para alojar, además de información, gráficos, voz y vídeo.

Uno de los motivos que animaba la investigación de Berners-Lee eran las conexiones, el modo en que se unen las cosas. La aplicación que terminaría convirtiéndose en el Web no fue la primera que Tim escribió para relacionar ideas y conceptos. Al menos hubo dos anteriores: Enquire y Tangle. De hecho Enquire fue usado y en parte desarrollado durante la primera estancia de Berners-Lee en el CERN. Sobre el segundo de estos programas, un pequeño juego, el mismo autor escribe:

“La idea era la siguiente: lo que importa está en las conexiones. No está en las letras, está en el modo en que se juntan para formar palabras. No está en las palabras, está en el modo en que se juntan para formar frases. No está en las frases, está en el modo en que se juntan para formar un documento. Pensé en el modo de introducir en el ordenador una enciclopedia de ese modo, y luego hacer al Tangle una pregunta. La pregunta sería descompuesta en nodos, que sucesivamente se referirían a los mismos nodos que aparecieran en la enciclopedia. El enredo resultante contendría todas las respuestas relevantes.”²⁶

Tim Berners-Lee llegó por primera vez al CERN en 1980. Su misión era apoyar uno de los departamentos de informática que daba soporte a los complicados experimentos sobre física de partículas que allí se realizaban. Cuando comenzó su trabajo en el laboratorio lo primero que descubrió es que el CERN era un sitio grande en el que había muchos proyectos diferentes que en cierta forma estaban interrelacionados. Los investigadores que trabajaban en estos proyectos venían de todas partes del mundo y muchos de ellos traían sus propios ordenadores. De esta forma se reunían allí una gran cantidad de sistemas operativos y de *software* diferente e incompatible entre ellos, precisamente en un entorno en el que era clave poder compartir la información.

Por motivos laborales a Berners-Lee le interesaba poder llevar un cierto control de qué personas hacían qué experimentos con qué máquinas. Pero este tipo de relación podía complementarse con otras informaciones como por ejemplo informes técnicos, manuales o resúmenes de reuniones.

Ya existían sistemas que permitían compartir datos entre diferentes usuarios, el problema es que estos sistemas obligaban a organizar el trabajo de determinada forma, de tal manera que todos los usuarios tenían que cambiar su forma habitual de trabajar. El sistema que iba a desarrollar Berners-Lee buscaba estar lo más cerca posible de no tener ningún tipo de reglas y de no tener que obligar a nadie a cambiar sus hábitos. El modelo que mejor sustentaba esta idea era el hipertexto.

²⁶ Berners-Lee, Tim: *Tejiendo la Red*. Ed: Siglo XXI. 2000. pp: 12

La Red que se estaba preparando debía tener una dimensión universal, sin autoridad central y tenía que permitir que cualquier usuario pudiese comenzar a introducir documentos sin pedir permiso a nadie.

Berners-Lee estuvo durante más de una década en el CERN en dos periodos de tiempo. Durante su primera estancia en el CERN, en 1980, apenas pudo poner en prácticas sus ideas sobre las conexiones y sólo logró hacer algunas pruebas de funcionamiento de Enquire. En 1984 Berners-Lee vuelve al CERN con una beca. En esta ocasión permanece allí casi diez años consecutivos, lo que le permite poder retomar su proyecto y poner los medios para llevarlo a cabo.

Finalmente Berners-Lee presenta una propuesta para poder realizar el desarrollo del Web, siempre sobre la base de que era una herramienta para su uso exclusivo en el Laboratorio y de que Internet no estaba incluido. Dicho estudio se presenta en marzo de 1989 y no tuvo ningún tipo de respuesta por parte de la Dirección del CERN.

En estos mismos años se procedió en el CERN a la compra de una de las primeras unidades de una máquina NeXT. Estos ordenadores eran parte de un ambicioso proyecto dirigido por Steve Jobs, unos de los fundadores de Apple. En el CERN siempre habían estado muy interesados en conocer todas las novedades que se producían en la industria relacionadas con nuevos ordenadores. Con la excusa de poder evaluar las posibilidades de la nueva adquisición, Berners-Lee pudo comenzar a escribir las primeras líneas de código del Web en el NeXT.

En un principio, la idea de Berners-Lee era encontrar algún editor comercial de hipertexto que pudiera adaptarse al proyecto del Web y que a la vez funcionara como navegador. Era del mismo modo necesario escribir los programas del Protocolo de Transferencia de Hipertexto (HTTP), del Identificador de Recursos Universal (URI) y del servidor de HTTP que corría en el NeXT y que de momento era el encargado de servir los documentos que iban escribiendo. En diciembre de 1990 comenzó a funcionar un editor/navegador de WWW. Este editor formateaba las páginas en HTML, una simplificación del SGML, el lenguaje habitual en el se elaboraban los documentos de hipertexto. La adaptación del HTML también fue obra de Tim Berners-Lee.

El primer proyecto que se realizó con las nuevas herramientas fue la adaptación de la base de datos de teléfonos del CERN. Se trataba de una vieja base de datos que se mantenía de forma manual y que había que adaptar a cada uno de los sistemas

operativos que estaban funcionando en el Laboratorio. Convertir esta base de datos en un proyecto Web significaba que cualquier usuario que tuviera un navegador podría consultarla sin necesidad de solicitar una adaptación para su sistema.

El siguiente paso fue dar a conocer el proyecto a escala internacional mediante charlas y conferencias en Universidades y empresas. También había que dotar al Web de contenidos para que hubiera algo que ver. Mientras tanto la dirección del CERN seguía sin dar autorización a un proyecto que ya empezaba a crecer casi sin control.

Para evitar el problema de licencias que había suscitado Gopher, la Red fue diseñada para poder ser usada libremente por todo el mundo. Ni siquiera se había planteado la posibilidad de usar una Licencia Pública General (GPL) que habría imposibilitado su uso en algunas grandes corporaciones como IBM. El libre uso del software no impedía que empresas comerciales hicieran uso del Web. La ausencia de esta combinación de libre uso y de interés comercial fue la que frenó el avance de recursos como Gopher y que estaban limitando el uso de Internet a Universidades y grandes empresas. Podemos afirmar que fue la aparición y la rápida extensión del WWW lo que permitió que Internet llegara a un número amplio de usuarios, que se conectaban a la Red desde cualquier sitio.

A principios de 1993 apareció un navegador comercial desarrollado por un empleado de HP. Este navegador llamado Arena ya permitía el uso de tablas, imágenes y otras mejoras que aumentaban las prestaciones del recién creado HTML. Por otra parte la Universidad de Kansas había escrito un navegador de Web en modo pantalla, Lynx, que mejoraba el prototipo original creado por Berners-Lee en el CERN unos años antes. Una de las organizaciones que manifestó un mayor interés por el WWW fue el NCSA (Centro Nacional de Aplicaciones Superinformáticas de la Universidad de Illinois). Dos personas de este centro, Marc Andreessen, un estudiante, y Eric Bina, miembro del personal, habían desarrollado un navegador para Xwindows de Unix al que pusieron por nombre Mosaic. Este navegador era muy fácil de cargar y de instalar.

El éxito del nuevo producto no se hizo esperar y rápidamente se convirtió en el estándar para acceder al Web. La prensa especializada empezó a fijarse en Mosaic y el NCSA, al contrario de la actitud que todavía mantenía el CERN, dio todo su apoyo al equipo que había desarrollado el navegador.

El caso del Mosaic junto a otros anteriores y la rápida extensión de WWW habían hecho que Tim Berners-Lee se plateara ahora más que nunca la creación de un consorcio independiente de ninguna organización con ánimo de lucro para controlar el crecimiento del Web. La labor de este consorcio sería la de coordinar su crecimiento y ampliación impidiendo que ninguna otra organización pretendiera convertir al Web en algo particular. En un principio a Berners-Lee le parecía que el consorcio podría estar ubicado en el CERN, de forma que el WWW no saliera de Europa. Sin embargo mientras que el CERN seguía sin mostrar su apoyo al Web, en EEUU se mostraban muy interesados ante la posibilidad de poder generar documentos de hipertexto a través de la Red.

Michael Dertouzos del Laboratorio de Ciencias Informáticas del MIT no tardó en proponer la idea de crear el consorcio en el MIT. Berners-Lee estaba de acuerdo con esta posibilidad aunque seguía insistiendo en que al menos debería existir una sede europea del consorcio en el CERN. Se organizó una gran conferencia mundial sobre el WWW para el mes de mayo de 1994, que se celebraría en las instalaciones del CERN en Ginebra. De esta conferencia saldría ya el Web convertido en el recurso de Internet que todos conocemos hoy. El nombre de su artífice, Tim Berners-Lee, entraría desde ese momento en la historia de Internet.

Durante la primavera de 1994 parte del equipo creador de Mosaic con Marc Andreessen a la cabeza se marchan del NCSA para formar su propia compañía Mosaic Communication Corp. Por problemas con los derechos del nombre que pertenecía a la Universidad de Illinois, en abril de 1994 la empresa cambio su nombre a Netscape y sacó al mercado la primera versión de su navegador llamada Mozilla. Meses más tarde Microsoft sacaba a la luz Windows95 que llevaba incluido su propio navegador Web: Explorer.

1.7. DIFERENCIAS ENTRE LA WORLD WIDE WEB E INTERNET

Una vez explicada la evolución de Internet y de la llamada World Wide Web merece la pena resaltar las diferencias entre uno y otro término con el objetivo de establecer una clara diferencia entre ambos (ya que en numerosas ocasiones, en la actualidad, se están utilizando indistintamente).

Internet es la mayor red de ordenadores del mundo que permite a todo tipo de ordenadores conectarse y comunicarse entre ellos a nivel mundial. Sus orígenes, como hemos señalado con anterioridad, datan de 1969 y es fruto de un proyecto elaborado por la Agencia de Defensa estadounidense DARPA (U.S. Defense Advanced Research Agency)

Por otro lado, la World Wide Web es uno de los muchos protocolos que permiten el acceso a Internet y que facilita a los usuarios internautas el acceso a los archivos que se encuentran en Internet. La Red (o Web) fue creada en 1990 por Tim Berners-Lee, un programador de ordenadores que trabajaba para el CERN (European Organizatin for Nuclear Research). Con anterioridad a la aparición de la Red el acceso a los archivos que se encontraban en Internet era una tarea bastante ardua que requería conocimientos especializados. Gracias a la World Wide Web podemos acceder fácilmente a archivos de texto, imágenes, audio y vídeo a través del sencillo mecanismo de los enlaces de hipertexto. Otros de los protocolos que conviven con la Wold Wide Web, son:

- E-Mail
- Foros y Tablones de anuncios
- Listas de distribución
- Grupos de Noticias
- Sistemas de intercambio de archivos (como Napster o Gnutella).
- Bases de Datos accesibles mediante la Red

1.8. INTERNET2

Cuando comenzó el proyecto de Internet2, en octubre de 1996, algunos colaboradores propusieron utilizar números romanos (es decir Internet II) para designar la nueva “telaraña mundial” de mayor potencial que se superpondría a la actual Internet. Se

trataba de una razón histórica basada en que la primera persona que utilizó este término fue un general romano de la época de Cesar Augusto, llamado Internettus²⁷. Más tarde se descubrió que esta razón era falsa, y que Internettus no era más que un funcionario menor a cargo de redes de mosquitos.

Según la investigación del término, un científico distinguido de Arabia, llamado Omar de Ephesus (quien hizo grandes descubrimientos en el campo del álgebra), ya había mencionado “Internet2” en sus discursos sobre “la geometría Euclidian en el siglo VII”, especialmente la referente a las dos jorobas de los camellos y su habilidad para transportar dos paquetes a la vez. Este último motivo dio pie a adoptar la forma arábica y designar al proyecto con el nombre de Internet2.

Descripción del proyecto

En una reunión celebrada en Chicago en octubre de 1996, representante de treinta y cuatro universidades accedieron a incorporarse al proyecto, comprometiéndose a realizar las investigaciones necesarias. En agosto de 1997 había más de cien centros adscritos cooperando en cuestiones de investigación y financieras. El 4 de febrero de 1997, el presidente de EEUU, Bill Clinton, y su Administración deciden apoyar también la creación de esta nueva generación de Internet, con el objetivo de que las universidades y los laboratorios nacionales pudieran comunicarse a velocidades mil veces más rápidas que hoy día.

Este proyecto tiene previsto incluir nuevas aplicaciones derivadas de las necesidades académicas en cuanto a búsqueda de información, enseñanza y aprendizaje; es decir, se está diseñando pensando en los objetivos y retos de las futuras generaciones universitarias.

Internet2 creará, en un principio, una Red mucho más eficaz entre las universidades y el campo de la docencia y permitirá una investigación más avanzada. Asimismo, cuando el proyecto se halle lo suficientemente desarrollado, se espera que se extienda a todos los usuarios de Internet. Según Irving Wladawsky-Berger, Director General de la División de Internet de IBM, la nueva Red no se pondrá en funcionamiento hasta aproximadamente el año 2005.

²⁷ Internet2: http://www.internet2.edu/html/i2_history.html

Los principales propulsores del proyecto son, además de las universidades, agencias federales y gran cantidad de empresas de ordenadores y telecomunicaciones, como: Ameritech, Cisco Systems, Digital Equipment Corporation, IBM, MCI, Spring y Sun Microsystems. La agencia aeronáutica norteamericana, NASA, es el mayor inversor y colaborador, debido principalmente a que esta organización necesita redes de comunicación con gran ancho de banda para transmitir los grandes volúmenes de datos que manejan entre todos los centros que tiene repartidos por todo el mundo, que a su vez reciben datos de los satélites que investigan el espacio exterior.

A pesar de que en la actualidad es un proyecto meramente norteamericano, se espera que en un futuro sea incorporado también a Europa y al resto del mundo.

La nueva arquitectura de Internet2

El problema con el que se han encontrado los creadores de la nueva versión de Internet, es que los protocolos de comunicación empleados estaban pensados inicialmente para redes mucho más pequeñas que la actual Red .

El protocolo con el que se tiene previsto que funcione Internet2, seguirá siendo IP, aunque en una versión mejorada y adaptada a las nuevas redes: el Ipv6, cuando en la actualidad es el Ipv4.

Además se producirán nuevos elementos. Uno de ellos será el “gigapop”, que significa “gigabit capacity point of presence” (punto de presencia con capacidad de gigabit). Este parece ser el elemento clave para el desarrollo de Internet2. Los servidores de cada universidad se unirán a diferentes “gigapops”, los cuales estarán interconectados a su vez consiguiendo una red de altas prestaciones. En un principio estarán gestionados por las universidades, pero se espera que llegue un momento en que interfieran las empresas privadas, y sean éstas las que impulsen la nueva tecnología entre usuarios comunes.

Internet2 tendrá las siglas “QoS” (Quality of Service guarantee- garantía de Calidad de Servicio). Consiste en asegurarse conexiones con una velocidad y ancho de banda estables en un largo periodo de tiempo. Actualmente todas las transmisiones viajan a través de la Red con la misma prioridad, es decir, es lo mismo que un usuario

normal esté mandando un e-mail que estén transmitiendo datos científicos. El QoS permite que se puedan establecer esta clase de prioridades. Gracias a esto la Red podrá utilizarse para aplicaciones en tiempo real, como pueden ser la telemedicina o la telepresencia, las videoconferencias o la educación a distancia (mediante audio y vídeo a tiempo real e interactividad).

Sin embargo, con esta nueva aplicación, “la Red solidaria” que es ahora Internet comenzará a hacer distinciones. ¿Quién decidirá lo que es prioritario transmitir?. Bill Gates afirmó en una ocasión, en referencia a los delitos y la legislación en la Red, que es imposible controlarla. Y justamente era aquí donde residía su atractivo. Nos encontramos ante un “producto salvaje” que está dando mucho de sí gracias a la libertad que conlleva. Pocas veces han existido y existen prioridades por necesidad, normalmente se compran (un millonario o una persona influyente puede comprar su prioridad). Un ejemplo de lo que podría llegar a ocurrir lo encontramos en las PYMES. En la actualidad las pequeñas y medianas empresas encuentran en Internet el medio ideal para entrar a competir, en una supuesta igualdad de condiciones, con las grandes empresas. Pues bien, si el sistema de prioridades llega a funcionar, nos encontraríamos ante el comienzo de la ruptura de este tópico que hasta ahora venía anexo al concepto de Internet.

En definitiva, la nueva Red se crea con el objetivo de superar el ancho de banda existente, que permite poca velocidad de transmisión. Internet2 permitirá nuevos servicios, pero con el tiempo la sociedad y la industria demandarán aplicaciones mejores y de mayor calidad que exigirán un ancho de banda aún mayor, por lo que puede que se cree Internet 3. Y así sucesivamente se crearán: Internet 4, Internet 5, Internet 6...Será pues la demanda de la sociedad y de las empresas las que hagan evolucionar esta tecnología. Ocurre algo parecido al fenómeno Windows, Windows 3.11, Windows 95, Windows 98, Windows 2000..., es decir, se van creando entornos en base a las nuevas exigencias.

Expansión de Internet2

A pesar de que en la actualidad es un proyecto meramente americano, se espera que en un futuro sea incorporado a Europa y al resto del mundo. Sin embargo, son muchos los

que ven la concentración norteamericana actual como la preparación para un futuro monopolio.

La Unión Europea está financiando, a su vez, un proyecto de redes de alta velocidad (igualando en rapidez a la Internet2 norteamericana) conocido con el nombre de “TEN-34” (Trans-European Network Interconnect al 34 Mbps). Esta Red conectará los nodos principales de los países de la comunidad de forma que se puedan alcanzar velocidades de hasta 155 Mbps. El proyecto tiene dos elementos principales: una red “backbone” de 34 Mbps, que cubre toda Europa, y otro derivado de su colaboración con el proyecto JAMES, que consiste en la unión de los operadores de telecomunicaciones europeos que han establecido una red experimental para probar la nueva tecnología ATM. Ésta permite asignar unos anchos de banda de manera más dinámica, por lo que se puede utilizar para aplicaciones en tiempo real, videoconferencias...etc.

Web Oficial de Internet2 (www.internet2.edu)

INTERNET² Search

[About Internet2](#) | [News](#) | [Members](#) | [Activities](#) | [Contact](#)

[Applications](#) | [Middleware](#) | [Networks](#) | [Engineering](#) | [Partnerships](#)

Showcase

Internet2 Government Partnerships

AAAS

Highlights

- [Internet2 Campus Focused Workshop on Advanced Networks, Netcast](#)
- [Fall 2001 Internet2 Fall Member Meeting Call for Proposals, Form](#)
- [Internet2 End-to-End Performance Initiative, Design Document \(PDF\)](#)
- [Information on Internet2 Backbone Connectivity for NSF HPNC Program Participants](#)
- [Job Opportunities](#)

1.9 INTERNET EN ESPAÑA

En 1985, El Ministerio de Educación y Ciencia de España, observando la evolución de las nuevas tecnologías en el exterior, empezó a ver la necesidad de interconectar los centros de investigación con el ordenador central, instalado en la calle Vitrubio de Madrid. Asimismo, FUNDESCO comenzó a potenciar esta idea entre las universidades, tanto para la comunicación entre estudiante y profesores, como para que los ordenadores situados en diferentes emplazamientos pudiesen trabajar en los mismos proyectos de investigación, compartiendo los mismos datos, y comunicándose entre ellos a tiempo real.

En 1987, se crea el Plan Nacional de Investigación Científica y Desarrollo Tecnológico (PNI), que creó el marco organizativo y financiero, para el desarrollo del “proyecto IRIS”, según se había previsto en un informe anterior realizado en 1985.

La RedIris se elabora siguiendo el modelo de otras redes de investigación que se estaban creando, o que ya estaban operativas en Europa, como puede ser la inglesa “JANET”. FUNDESCO fue la encargada de materializar el proyecto, dotándole de una entidad jurídica propia; asimismo, estuvo gestionando la Red desde 1989 hasta 1994.

Durante 1993, se introdujo en RedIris el protocolo que había revolucionado Internet en otros lugares del mundo: el Word Wide Web. Fue en esta misma época, en la que FUNDESCO cedió la gestión de la RED al CSIC (Consejo Superior de Investigaciones Científicas).

Hasta entonces había treinta y cinco universidades españolas conectadas. A partir de 1994, Internet comenzó a salir del ámbito universitario, y empezaron a crearse empresas proveedoras de acceso, que ofrecían fundamentalmente correo electrónico y otros servicios básicos. “Goya Servicios Telemáticos”²⁸, como representante española de la internacional EUNET, fue la primera empresa española que explotó Internet de forma comercial.

Durante 1995, aparecieron en España dos figuras empresariales dedicadas al comercio virtual: Los Proveedores de Accesos, que proporcionaban acceso a la Red a los usuarios; y los Proveedores de Servicios. Estos últimos se dedicaban, entre otras

²⁸ Goya Servicios Telemáticos: <http://www.goya.es/>

cosas, a gestionar a las empresas un dominio de segundo nivel, es decir “.es” en cualquier dirección de Internet que no hubiera sido previamente asignada. El dominio permitía a las empresas recibir correo electrónico y poner un servidor Web con páginas web, con el fin de extender la imagen corporativa de la empresa al nuevo medio.

En 1996 se dispara el uso de Internet en España, en todos los ámbitos. Aparecieron los buscadores de páginas españolas, y numerosas publicaciones dedicadas a este tema. Con la llegada de la competencia entre los proveedores de acceso, los precios comenzaron a bajar, hasta colocarse en un nivel accesible para prácticamente cualquier persona.

Sin embargo, el entusiasmo general que creó, trajo también desventajas. Debido a la fuerte comercialización de la Red, ésta se saturó, por lo que en la actualidad (debido al poco ancho de banda del que se dispone), muchas veces es difícil la conexión con algunos servidores, la creación de gráficos o imágenes complejas, o simplemente navegar de forma rápida. A pesar de la saturación, la evolución del número de dominios en España sigue subiendo, y de manera paulatina aunque con retraso, se va igualando a la del resto del mundo que conforma actualmente “la aldea global virtual”.

Direcciones web de interés relacionadas con la historia de Internet:

- Bolt Beranek and Newman Ltd: <http://www.bbn.com>
- The San Francisco Bay Area Regional Research Network:
<http://www.barnet.net>
- Committee on Institutional Cooperation Network: <http://www.cic.net>
- INDnet: <http://www.ind.net/INDNET/>
- Michigan Educational Research Instructional Network: <http://www.mith.net/>
- Midnet: <http://www.mis.net>
- Minnesota Regional Network: <http://www.MR.Net/>
- The North East Academic Research Network: <http://www.near.net/>
- NETIllinois: <http://www.illinois.net/>

- New York State Education and Research Network: <http://www.nysernet.org/>
- The Ohio Academic Research Network: <http://www.oar.net>
- Sesquinet: <http://www.sesqui.net/>
- Westnet: <http://www.westnet.net/>
- Fundación para el Desarrollo de las Comunicaciones: <http://www.fundesco.ess>
- Hobbes' Internet Timeline v2.1:
<http://www.handshake.de/infobase/netze/internet/hit.htm>
- BBN Timeline: http://www.bbn.com/customer_connection/timeline.htm
- Motorcade Timeline: <http://www.motorcade.com/webauth/timeline.htm>
- The Roads and Crossroads of Internet's History:
<http://www.internetvalley.com/intvalconcl.html>
- A Hypertextual Tool for Understanding How the Global Village Grew:
http://infoserver.etl.vt.edu/coe/COE_Students/Laughon/london.html
- Internet Domain Survey: <http://www.nw.com/zone/WWW/top.html>
- Reseaux IP Europeens: <http://www.ripe.net/>
- Internet Society: <http://info.isoc.org>
- Matrix Information and Directory Services, Inc: <http://www.mids.org>
- The Netcraft Web Server Survey: <http://www.netcraft.com/Survey/>
- A Short History of the Net: <http://www.ocean.ic.net/ftp/doc/snethistnew.html>
- RFC 1000: <http://www.internic.net/rfc/rfc1000.txt>
- RedIris: <http://www.rediris.es>
- How the Internet Works:
<http://hosting.homecom.com/iw-articles/howitworks.tml>
- InterNIC: <http://www.internic.net/>
- Internet Engineering Task Force (IETF): <http://www.ietf.org>
- Internet International Ad Hoc Committee: <http://www.iahc.org>
- Computer Emergency Response Team: <http://www.cert.org>
- The Internet Society: <http://www.isoc.org>
- The Internet Assigned Numbers Authority: <http://www.iana.org/iana>
- The World Wide Web Consortium: <http://www.w3c.org>
- European Laboratory for Particle Physics: <http://www.cern.ch>

- The National Science Foundation: <http://www.nsf.gov>
- The Advanced Research Projects Agency: <http://www.arpa.mil>
- How the Internet Came to Be:
<http://virtualschool.edu/mon/Internet/CerfHowInternetCame2B.html>
- Internet 2: <http://www.internet2.edu>

CAPITULO 2: EL PERIODISMO A TRAVES DE INTERNET

"El fenómeno Internet no es más que la punta del iceberg, es el paradigma de la sociedad digital. Internet es un nuevo medio de comunicación. Primero fue la prensa, luego la radio, después la televisión. Ahora asistimos al nacimiento de un nuevo soporte para la información, que será la materia prima más valiosa del siglo XXI. Son las redes de telecomunicaciones, que forman un nuevo espacio cultural y social. También es un nuevo territorio para la comunicación y el periodismo".²⁹

2.1 INTRODUCCIÓN

Internet es en la actualidad un nuevo medio de comunicación con grandes oportunidades no sólo para los profesionales de la información, sino para la sociedad en general. El hecho de que no sea "propiedad de nadie", ha potenciado la utilización de la *World Wide Web* por usuarios "normales", particulares y empresas deslindadas del mundo científico y universitario de los comienzos de Internet.

A diferencia de los medios tradicionales, como la radio y la televisión, en la Red son los usuarios los que deciden la mayor parte de los contenidos disponibles. Su facilidad de uso y su accesibilidad permite a cualquier persona publicar absolutamente cualquier tipo de información que desee, con la posibilidad añadida de ser visible en cualquier lugar del mundo que tenga un ordenador conectado a la Red de redes.

A diferencia de los medios tradicionales, Internet admite diversos modos de comunicación: entre un emisor y un receptor individuales, entre un emisor individual y múltiples receptores y entre múltiples emisores y receptores. En cualquier momento, un

²⁹ Luis Foix es director adjunto de *La Vanguardia* y miembro del Consejo Editorial de la revista *Web*.

receptor puede convertirse en suministrador de contenidos. Por consiguiente, la Red Internet es radicalmente distinta de la transmisión tradicional y también de los servicios tradicionales de comunicación.

Internet ya ha conseguido poner en contacto a más de 150 millones de personas. Se trata, pues, de un medio con un público de masas, y es el medio que más rápidamente ha alcanzado un público de masas en toda la historia de la humanidad. La radio tardó 38 años en alcanzar 50 millones de radioyentes, la televisión tardó 13 años en tener 50 millones de telespectadores e Internet en tan sólo cinco años ha conseguido la cifra de 150 millones de usuarios.

Sin embargo no son las cifras de usuarios lo que le está convirtiendo en el cuarto gran medio de comunicación, sino sus posibilidades. Internet permite incluir hipertexto, fotografías, imágenes animadas, sonido, vídeo, etc. Permite comunicarse desde cualquier lugar del mundo. Permite recibir y entregar todo tipo de información. Pero si algo hace a este medio especial es su interactividad. El usuario de la gran Red pasa de ser un sujeto pasivo a tener un papel activo eligiendo lo que quiere y lo que no. Esto da a la comunicación un nuevo carácter que se establece a través de un nuevo tipo de relación entre emisor y receptor.

Lo que sí es cierto es que Internet es en la actualidad un sinónimo de cambio. Las empresas que deseen adaptarse a esta nueva tecnología tendrán que estructurar sus organismos eliminando puestos de trabajo obsoletos por otros de nueva creación. Ejemplos de ello son las figuras del webmaster, el director de web marketing o los integrantes del departamento de atención al cliente *online*. Estas empresas han incorporado sus catálogos de productos a sitios *web* donde dan toda la información necesaria tanto a clientes como a distribuidores, proveedores, etc., a tiempo real. Además todo el conjunto de procesos de comunicación realizados mediante fax, cartas o llamadas de teléfono entre otros, comienzan a ser realizados ya a través de las llamadas Intranets corporativas, lo que supondrá la creación de empresas capacitadas para interactuar con otras y compartir información y recursos. Esto supondrá para las empresas del futuro su conversión en estructuras flexibles y muy competitivas.

Con respecto al número de internautas y según el estudio *eGlobal Report* realizado por *eMarketer*³⁰, el número de usuarios de la Red llegará a los 350 millones en el año 2003. Esto equivale a un crecimiento del 267% con relación a los 95,6 millones de internautas contabilizados a finales de 1998. A nivel europeo, las cifras de Morgan Stanley Dean Witter indican que la Unión Europea puede contar con más de 100 millones de navegantes gracias a la reducción de los costes de acceso y al incremento de las velocidades de transmisión.

El perfil de la audiencia en Internet es bastante homogéneo y bien definido. Aún así, la cantidad real de gente conectada a Internet es difícil de medir exactamente aunque se estima que el número de ordenadores conectados a la Red a nivel mundial era de 179 millones en junio de 1999. De estos, 42,69 millones se encontraban en Europa.

En España, el número de usuarios en la segunda ola de 1999 del Estudio General de Medios era de 3,1 millones de ordenadores con acceso a Internet. Esto significa un aumento de alrededor de un 40% con respecto a la misma ola del año anterior. El perfil mayoritario de los usuarios españoles es básicamente de hombres, adultos (el grueso se encuentra entre los 20 y los 44 años) de clase social media-alta.

Como vemos, la Red ha comenzado a formar parte de nuestra realidad diaria y se ha extendido rápidamente entre la población mundial. La *World Wide Web* se ha convertido en una inmensa nube de información cada vez mayor que necesita, de alguna forma, ser organizada, seleccionada y servida al usuario que busca en ella datos veraces, razonados y útiles. Es éste, pues, el nuevo reto de los profesionales de la información que deberán adaptarse a este nuevo medio aprovechando sus ventajas y superando sus inconvenientes.

Como veremos a lo largo de este capítulo, el periodismo electrónico irá encontrando en la propia Internet su fuente de noticias. Aunque inicialmente se limite a reproducir en la Red sus contenidos habituales, diseñados para otro medio de comunicación y para otro espacio social, en la medida en que aumente el número de personas que circulan ella, los contenidos y los formatos se irán modificando, adaptándose al nuevo espacio social.

³⁰ E-marketer: <http://www.e-marketer.com>

2.2. DE LA GALAXIA DE GUTENBERG A LA GALAXIA DE MCLUHAN

Marshall McLuhan afirmó en su *Galaxia Gutenberg* que todo nuevo medio trata, en una primera instancia, de integrar a los medios precedentes y referirse a ellos. Esto resulta muy claro en el caso de Internet, puesto que tanto la prensa como la radio y la televisión van tomando su sitio en la Red y abriendo sus respectivas “páginas web”. Sin embargo, Internet posee una estructura interna muy distinta a la de los medios de comunicación precedentes, derivada en parte de su origen, y en parte de su evolución.

Cabe decir que el periodismo electrónico no comienza con la aparición de Internet. Podemos encontrar claros precedentes en el videotexto y teletexto. Otros antecedentes más lejanos, nos llevan a las supuestas alternativas al periódico tradicional, desarrolladas con el objetivo de cumplir uno de los mayores deseos de los medios de comunicación: la inmediatez. Destacan entre ellos: el periódico por fax, los servicios en línea (bases de datos, correo electrónico mediante el que los lectores pueden comunicarse, servicios de compra-venta a distancia), o el audiotexto. En cuanto a este último, cabe decir que ya a principios del siglo XX encontramos un precedente. Se trata del “periódico-teléfono”, puesto en marcha en la ciudad de Nueva York en 1913. Al parecer contaba con 2.500 suscriptores que pagaban al mes el equivalente a 7,50 pesetas de entonces. El servicio consistía en que los clientes recibían las noticias de viva voz, mediante un artefacto similar a un teléfono que sólo servía para hablar con la redacción de ese medio de comunicación en concreto³¹.

Con la llegada de Internet, se incorpora al ideal perseguido por sus precedentes un concepto básico a la hora de hablar de la Red hoy en día: la interactividad. Y no sólo éste, sino además la posibilidad de integrar elementos textuales y audiovisuales en un mismo documento. Esto es lo que el profesor Jean Cloutier ha denominado “l’audioscriptovisuel”³².

³¹ Díaz Noci, Javier: *El nuevo periódico electrónico: redefinición del mensaje periodístico tradicional como producto interactivo y multimedia*. Mayo 2000.

³² Cloutier, Jean (1994): *L’audioscriptovisuel et le multimedia*, Communication et langages, n° 99, París: Edicions Retz, pp.99.

Como hemos podido ver en capítulos precedentes, Internet procede de Arpanet, que fue creada en los EEUU para poder mantener interconectados los centros de decisión más importantes en el plano político, militar y científico en caso de un holocausto nuclear. Tras la separación de *Milnet*, la Red dedicada a uso estrictamente militar, fue constituyéndose la actual Internet, manteniendo su estructura descentralizada inicial, e integrando sucesivamente a las diversas redes que habían sido creadas en base a otro esquema, tanto por países concretos como por grandes empresas. Es importante destacar que el carácter abierto y descentralizado de la Red, así como el alto grado de autonomía que cada red integrada en la Red global cuenta para su propio funcionamiento, han sido claves en el éxito de Internet, cuyos usuarios crecen aceleradamente mes tras mes, expandiéndose por todos los países.³³

Internet es, en definitiva una Red horizontal, multidireccional, descentralizada e interactiva, lo cual resulta poco frecuente en el caso de los medios de comunicación anteriores. En el caso del libro, el paradigma de *la Galaxia Gutenberg*, que aportó un medio de comunicación multidireccional, la posición del autor es central y asimétrica con respecto a la del lector, y la interactividad apenas existe. Y otro tanto cabría decir de la radio y la televisión, a diferencia del teléfono, que sí es interactivo, pero sólo es bidireccional, al menos en su formato actual. Esta estructura de la Red telemática mundial ha sido determinante en la formación de una cultura específica en el ciberespacio.

A lo largo de la Historia, las nuevas tecnologías han propiciado cambios en los medios de comunicación existentes y la formación de otros nuevos, que no siempre han sido bien acogidos por las empresas de comunicación. A este respecto cabe citar a Humberto Eco³⁴ y a su dicotomía entre apocalípticos e integrados. La posición apocalíptica sería, en este caso, la de aquellas personas o tendencias que únicamente ven el lado negativo de la Red, es decir, el crecimiento de la difusión de pornografía infantil, la dificultad de persecución de determinados delitos, la saturación informativa, etc. En cuanto a los integrados, éstos serían los que manifiestan el lado positivo de Internet, valorando sus ventajas y posibilidades.

³³ Jesús Ranz Abad: *Breve historia de Internet*. Ed: Anaya. Madrid 1997.

³⁴ Eco, Humberto: *Apocalípticos e integrados*. Ed. Lumen

Algunas personas, como el lingüista italiano Raffaele Simone, profesor de Lingüística en la Universidad de Roma Tres, y autor de *La Tercera Fase, formas de saber que estamos perdiendo* (editado por Taurus), ven a este nuevo medio que es Internet como "el principal enemigo del libro y de la lectura, a pesar de su apariencia de estar hecho para leer y escribir"³⁵. Asimismo, Simone señala que ha habido cuatro cambios fundamentales en estos últimos años que han traído "la disolución de un paradigma de cultura, de información y de educación". Para Simone, ha cambiado la jerarquía de los sentidos (ahora la visión natural prevalece sobre la alfabética), ha aumentado el valor de la imagen (y con ella la supremacía de lo menos estructurado sobre lo más estructurado), ha cambiado la naturaleza de la escritura y la tipología de los textos, (que son ilimitadamente modificables) y, por último, ha originado una nueva forma de elaborar la información que el lingüista bautizó como "no proposicional". Esta nueva forma de crear información, según Simone, ha perdido los rasgos tradicionales de ser analítica, estructurada, contextualizada y referencial, para convertirse en "una masa indiferenciada donde todo está en todo" que desprecia el análisis y experiencia³⁶.

Muchas empresas periodísticas han temido a lo largo de la Historia la suplantación de los medios tradicionales por los más modernos. Sin embargo, en la mayoría de los casos no se produce sino una coexistencia, normalmente beneficiosa para todos. Por este motivo se puede decir que Internet no anulará a los periódicos al menos en un futuro próximo.

Es en la década de los 90 cuando nacen los primeros periódicos electrónicos multimedia e interactivos. Las primeras empresas que comienzan a publicar sus diarios en Internet son las estadounidenses, aunque en Europa ya comenzaban a desarrollarse proyectos como el INES-IFRA³⁷. Javier Díaz Noci en su artículo "Tendencias del periodismo electrónico hace una aproximación a la investigación sobre medios de

³⁵ Gómez, Juan J. Gómez: Entrevista publicada en el diario *El País*, con fecha del 14 de febrero de 2001

³⁶ Gómez, Juan J. Gómez: Entrevista publicada en el diario *El País*, con fecha del 14 de febrero de 2001

³⁷ INES-IFRA: <http://www.darmstadt.gmd.de/INES>

comunicación en Internet³⁸”, recogiendo estas primeras experiencias estadounidenses, que serían la base del periodismo electrónico actual:

“La compañía Knight-Ridder, que tuvo a Roger Fidler en la década anterior investigando en las capacidades de un sistema de videotex llamado *Viewtron* (que fue un fracaso comercial), lanzó una versión de un pequeño periódico de Silicon Valley (California), el *San Jose Mercury News*, al que denominó *San José Mercury Center*, todavía hoy uno de los más influyentes periódicos electrónicos. Este medio ha marcado pautas, como por ejemplo la convivencia de zonas de libre acceso (similares a la información que suministra la edición en papel) con otras de pago, para segmentos más especializados, el facilitar servicios a la población, la búsqueda de anuncios clasificados, etc, que marcan la diferencia respecto al tradicional periódico impreso. Otras iniciativas estadounidenses dignas de mención son NandoNet, toda un área de servicios informativos de todo tipo, puesta en marcha por la empresa que edita el *Wall Street Journal* en versión electrónica, un informativo de pago que explora a fondo las posibilidades de la transmisión inmediata de información sensible como es la económica, el servicio (renovado y actualizado constantemente, y cada vez con más elementos multimedia) de la cadena de televisión CNN, *CNN Interactive*, o el galardonado por la revista *Editor & Publisher* como el mejor entre 1530 periódicos electrónicos, el *New York Times*”.

En cuanto al caso español fue pionera a la hora de incluir su periódico en la Red la revista cultural *El Temps* de Valencia (1994). Pronto se unió el Boletín Oficial del Estado y varios diarios de Barcelona: *Avui*, *La Vanguardia* y *El Periódico de Catalunya*. Este último lidera una investigación en el seno de la Unión Europea sobre el desarrollo del periódico electrónico accesible mediante una tableta portátil. De los Nacionales, el primero en editarse en Internet fue *ABC*; le siguieron *El Mundo* (1995) y *El País Digital* (mayo 1996). Asimismo, la prensa regional comenzó a incorporarse en breve siendo sus primeras versiones, copias de la edición impresa. Los bajos costes de publicación en la Red y su cada vez mayor facilidad de acceso y uso hicieron que

³⁸ Díaz Noci: Javier: *Tendencias del periodismo electrónico. Una aproximación a la investigación sobre medios de comunicación en Internet*. Marzo 2000. <http://www.ehu.es/zer/6artdiaz.html>

pequeñas empresas de comunicación se aventuraran a publicar todo tipo de revistas, boletines y otros servicios en Internet.

Díaz Noci, en su artículo “El nuevo periódico electrónico: redefinición del mensaje periodístico tradicional como producto interactivo y multimedia³⁹”, señala cinco causas principales que según él han dado lugar al periódico interactivo y multimedia. Éstas son las siguientes:

- a) Incapacidad de los periódicos de aumentar e incluso mantener sustancialmente el número de lectores, debido en gran parte al agotamiento de las fórmulas tradicionales, como ponen continuamente de manifiesto los órganos de expresión de la profesión periodística en Estados Unidos⁴⁰. Se trata, fundamentalmente, de poder ofrecer al lector una información que él, de forma interactiva, pueda seleccionar a su gusto, adaptándola a sus necesidades y referencias personales.
- b) Aumento de la información, que el formato limitado de un periódico tradicional, a pesar del cada vez mayor número de páginas y suplementos que ofrece, no puede ya albergar, y que ha dado lugar a servicios que pueden considerarse antecedentes del periódico electrónico: acceso a bases de datos, correo electrónico, acceso a archivos del medio en cuestión a través de ordenadores, etc. Asimismo, el acceso a la información, muchas veces remoto, es cada vez más fácil, y a ello han contribuido redes como Internet, Prodigy o American Online, que cuentan con millones de usuarios (especialmente Internet).
- c) Diversificación de la actividad empresarial comunicativa; las empresas se convierten en multimedia, conciertan alianzas con otras empresas, extienden su actividad a otros campos e investigan nuevas posibilidades que les aseguren presencia y pujanza en el difícil mundo de los negocios.
- d) Escasez del papel y aumento de la sensibilidad ecologista: el problema se planteó ya en los primeros años 80, y todavía hoy, aunque la

³⁹ Díaz Noci, Javier: *El nuevo periódico electrónico: redefinición del mensaje periodístico tradicional como producto interactivo y multimedia*. Mayo 2000.

<http://www.andalucia24horas.com/caracola/doc002.htm>

⁴⁰ Fitzgerald, Mark: *Envisioning the electronic newspaper*. Editor & Publisher. 19 de febrero 1994

desaparición del soporte hecho a base de fibra de madera no es un hecho inmediato, son muchos los que comprueban con terror como menguan los bosques del planeta y cómo, por otro lado, el coste del papel supone uno de los mayores gastos para un periódico⁴¹.

- e) Avances técnicos especialmente en el terreno de los ordenadores, en los últimos años: mayor velocidad de los coprocesadores, mayor capacidad de almacenamiento, mejor definición de las pantallas, abaratamiento de los costes, que han hecho que la informática personal se popularice.

Maria José Pérez Luque⁴² establece que desde la aparición de Internet se han producido dos importantes evoluciones en este nuevo medio:

- Un proceso de homogeneización entre los medios *online*.
- Cambios en los contenidos ofrecidos.

Con respecto al proceso de homogeneización señala que la irrupción de los primeros periódicos electrónicos a principios de la década de los 90, coincidiendo con el momento en que Internet era un medio principalmente textual, con pocos elementos gráficos, fue debido a su semejanza con el medio escrito y la crítica situación del mercado del papel.

Más adelante comenzaron a descubrirse las verdaderas posibilidades de Internet, primordialmente la interactividad que permitía y la opción de integrar audio, vídeo y otros elementos gráficos que aportaban un valor añadido inestimable para la difusión de información. Rompía además las fronteras de tiempo y espacio que limitaban a los medios tradicionales.

⁴¹ En 1991, según la Canadian Pulp & Paper Association, se consumieron 32.783 toneladas de papel prensa, mientras que la cifra fue de algo más de 13.000 en 1960. En 1990, el coste del papel prensa suponía más de la cuarta parte de los gastos de los periódicos, según el Newsprint Information Comité (datos recogidos de Marchak, William: "Del papiro al papel prensa", en AEDE nº 18, 1992, pp. 104-110).

⁴² Pérez Luque, María José: *El periodismo y las nuevas tecnologías*. Ed: Newbook Ediciones. Pamplona, 1998.

Según Pérez Luque, los efectos que produce este fenómeno de homogeneización son los siguientes:

- a) Se difuminan las líneas que tradicionalmente delimitaban a los medios y a sus audiencias. Por ello, el término más apropiado para denominar a estos medios de comunicación es el de “servicios informativos *online*”. Hasta ahora éste era el término para denominar a los periódicos *online*. Sin embargo, debe ser igualmente aplicado a todas las empresas informativas en la Red, sin importar la procedencia de la empresa madre. Todas se dedican a ofrecer información *online*, sin importar qué formato de información (texto, imagen, audio, etc) predomina en sus páginas web.

Como consecuencia, los usuarios ya no se conectan a un servicio informativo *online* porque se trate de un periódico o de una televisión en la Red; sino porque ese *website*, de entre toda la oferta existente en Internet, es el que mejores contenidos le ofrece.

- b) Se incrementa el número de competidores que luchan por los mismos lectores y anunciantes en una determinada región.

Con respecto a la evolución de los contenidos ofrecidos, John Pavlik, creador del Center for New Media de la Universidad de Columbia, distingue las siguientes tres tendencias⁴³:

- a) En un primer momento los periodistas *online* se limitaban a reproducir los contenidos de la versión impresa. Esta primera tendencia aún pervive en algunos medios de comunicación en la Red. V. Giuliano, director de The Electronic Publishing Group, denomina a esta fase “shovelware”.

⁴³ Pavlink, J: *The future of online journalism*. Columbia Journalism Review. Julio/Agosto 1997. <http://www.cjr.com>

- b) La segunda etapa, surgida a finales del año 1996, se caracterizó por la presencia de *webs* informativos de mayor calidad. Los periodistas crean contenidos originales e incluso lo mejoran con *hyperlinks*, secciones de máquinas de búsqueda e incluso con cierto grado de personalización.
- c) La fase actual presenta un tipo de contenido que únicamente se encuentra en determinados *websites*: son los artículos noticiosos originales diseñados específicamente para el *web* como nuevo medio de comunicación. Este estadio destaca por la voluntad de replantearse la naturaleza de la comunidad *online*, y sobre todo, experimentar con las nuevas formas de narrar.

A parte de estas tres, Pérez Luque añade una cuarta fase en la que los servicios informativos *online* ofrecen un tipo de contenido que se aleja de lo puramente noticioso: informaciones sobre la ciudad de la que es originario el servicio informativo *online* (restaurantes, cines, espectáculos, carreteras, etc). Y en todos estos casos, manteniendo las posibilidades de interactividad que ofrece este nuevo medio.

Asimismo Pérez Luque⁴⁴, establece dos clases de factores que podrían haber llevado a la industria periodística a explorar las posibilidades de Internet. Éstos son factores externos e internos, englobados bajo un mismo techo: los cambios tecnológicos.

Según Pérez Luque, los factores externos vienen marcados por el propio mercado. Actualmente la industria periodística tiene puestas sus esperanzas en los medios *online* como respuesta estratégica a las transformaciones que se han ido produciendo en el mercado. Estos cambios han situado a los periódicos ante un panorama poco favorable:

- Incremento de los costes asociados con la producción y la distribución de los periódicos. La aparición de este nuevo medio se produjo cuando el precio del papel estaba alcanzando cotas más altas.

⁴⁴ Pérez Luque, María José: *El periodismo y las nuevas tecnologías*. Ed: Newbook Ediciones. Pamplona, 1998.

- Disminución del número de lectores, especialmente jóvenes que se habían alejado del hábito de leer prensa.
- Pérdida de la primacía como distribuidores de *breaking news* frente a los medios audiovisuales.
- Huida de la publicidad hacia otros medios, especialmente la televisión.
- La competencia (otros medios) había comenzado a hacerse presente en este nuevo medio.

La búsqueda de rentabilidad del producto de la empresa informativa es principalmente la idea que motiva los siguientes factores internos que explican la presencia de los periódicos en Internet:

- Incrementar gradualmente los ingresos mediante la nueva venta de un producto ya existente, con o sin servicios de valor añadido.
- Promocionar su producto base mediante la creación de réplicas *online*.
- Ganar experiencia en la creación de publicaciones *online*.

Como muestra cabe destacar un ejemplo, que puede parecer anecdótico, pero resulta muy significativo. Como es sabido, uno de los sitios del World Wide Web más visitados es el de la revista *Play Boy*. Los directivos de esta revista se plantearon en un momento dado cobrar a cada uno de los internautas que ojeara sus ficheros. En los días previos a la adopción de la medida, se produjo una auténtica avalancha de consultas, posiblemente con el fin de archivar el mayor número de imágenes posibles. Esto dio lugar a que la decisión prevista fuera reconsiderada. Los directivos de *Play Boy* decidieron adaptarse a la gratuidad tradicional de los contenidos de la red, añadiendo a cambio publicidad a sus imágenes. Parece ser que actualmente los ingresos por publicidad de la edición electrónica superan ya a los ingresos que se obtienen por toda la edición europea de la revista. Se trata pues de un buen ejemplo de adaptación a la estructura de la red y a las costumbres que se han derivado de ella.

Algunos de los principales medios de comunicación en Internet

The image displays a collage of four prominent news websites from the early 2000s:

- SER (Superespectáculos y Noticias):** Located at the top left, it features a blue and yellow color scheme. The main headline reads "Decenas de miles de vascos se manifiestan contra ETA" (Thousands of Basques demonstrate against ETA). A search bar and navigation menu are visible at the top.
- elmundo.es:** Located at the top right, it has a dark blue header with the site's logo. The main headline is "Los partidos trasladan la campaña a Jaca y PP y PSOE piden acabar con ETA en las urnas" (The parties move the campaign to Jaca and PP and PSOE demand to end ETA at the polls). It includes a search bar and a navigation menu.
- CNN.com:** Located at the bottom left, it features a red and white color scheme. The main headline is "Study finds that 50% of voters..." (Study finds that 50% of voters...). It includes a search bar and a navigation menu.
- The New York Times:** Located at the bottom right, it has a white background with a black header. The main headline is "U.S. Reconnaissance Flights Resumed Off China's Coast" (U.S. Reconnaissance Flights Resumed Off China's Coast). It includes a search bar and a navigation menu.

2.3 ELEMENTOS DEFINITORIOS DEL PERIODISMO *ONLINE*

En la actualidad, Internet se ha convertido en un medio de comunicación de masas, en el que intentan estar presentes los antiguos medios tradicionales. Cada vez es mayor el número de periódicos que decide sacar una edición on-line. Incluso la radio y la televisión parecen haber descubierto las grandes cualidades y posibilidades que la Red de redes les ofrece. Estos medios han creado su "imagen" en Internet, aprovechando así su capacidad y facilidad de difusión. Sin embargo, hay que tener en cuenta la esencia de Internet como medio de comunicación en sí mismo. No se trata simplemente de un soporte en el que adaptar los medios tradicionales. Se trata de una nueva forma de informar y de distribuir dicha información.

La nueva forma de informar que implica Internet, está determinada por las posibilidades que ofrece el propio medio; y para explotar estas posibilidades los nuevos medios de comunicación, y por tanto las informaciones que se ofrecen a través de Internet, deben adoptar una serie de características específicas. Éstas son las siguientes:

- i) La interactividad
- j) El hipertexto, hipermedia
- k) Elementos multimedia
- l) Lenguaje html
- m) Información personalizada
- n) Información documentada
- o) Actualización. Capacidad para procesar datos en tiempo real
- p) Nuevas formas de diseño

a) Interactividad

Según Nicoletta Vittadini, en la recopilación de textos realizada por Gianfranco Bettetini y Fausto Colombo, titulada *Las nuevas tecnologías de la comunicación* (Paidós, 1995), una de las características más innovadoras de los nuevos *media* lo constituye la instauración de una modalidad comunicativa no permitida antes por los *media*: la comunicación interactiva.

"En efecto –asegura Vittadini- la configuración tecnológica de los nuevos *media* determina la superación de uno de los caracteres de la comunicación unidireccional y difusiva propia de los *media* tradicionales: la ausencia de un *feedback* inmediato, simultáneo a la emisión y enviado al receptor por el mismo canal de transmisión. La interactividad, por consiguiente, se define como una evolución respecto de la comunicación permitida por los *media* unidireccionales y se caracteriza por la sustitución de la canalización pasiva por una utilización activa del medio."

En otras palabras, no sólo los periodistas del futuro digital se enfrentan al reto de comunicar a un público hipersegmentado, sino también a un público que tiene el derecho de réplica.

El ansiado *feedback* comunicacional ya es posible gracias a los medios interactivos basados en soportes multimedia. Tanto el periodismo informativo como el de opinión se prestan a que el usuario-receptor del mensaje pueda comentarlo, criticarlo, incluso modificarlo o aportar datos nuevos. El uso del correo electrónico y la lectura digital de diarios *on line* permiten una auténtica e inmediata interactividad entre el periodista y el lector, lo cual antes resultaba casi imposible.

Este nuevo elemento también representa un desafío para el periodista digital, pues ahora casi todas las informaciones que comunique no sólo serán analizadas por los lectores y afectados, sino también respondidas y comentadas.

A una mayor intervención de la interactividad en los medios de comunicación electrónicos, habrá una modificación en la forma de hacerlos. Esto establece un doble acuerdo al ofrecer el medio la posibilidad de que sus receptores interactúen con él y determinen en cierta medida la manera de diseñarlos y presentarlos.

Los comunicadores en general (editores y periodistas) deberán estar conscientes de que, en el nuevo esquema de la información digital e interactiva, la emisión de los mensajes ya no proviene del "la información estrictamente veraz y comprobada", pues ya no existirán barreras para hacer llegar una comunicación personal hasta el director o redactor.

En este nuevo medio de comunicación es posible encontrar informaciones de todas las partes del mundo, y de tu propio barrio con la misma accesibilidad. Las posibilidades son infinitas, y todas están al servicio del lector, quien, sin ningún tipo de intermediarios tiene la opción de manipular la información según su propia línea ideológica sin tener que seguir una corriente de opinión predefinida. Esto tiene múltiples consecuencias en el caso del periodismo electrónico que se pueden ilustrar con dos ejemplos:

Supongamos que un usuario de Internet quiere obtener información sobre el último desastre natural ocurrido en Nicaragua. Tras introducir los índices de búsqueda correspondiente, la información a la que puede acceder no se limita a las de las agencias de noticias: también incluye los rumores y los comentarios que pudieran circular por la Red. En una palabra, el lector del periódico electrónico no sólo podrá acceder a lo que es de actualidad, tal y como esta actualidad es realizada por los medios digitales, sino que podrá disponer de todo cuanto se comenta en Internet sobre el tema, así como antecedentes. E incluso este lector podrá publicar su propia opinión.

Un segundo ejemplo con el que podríamos ilustrar la interactividad, esta vez en un ámbito que afecta más directamente al usuario, sería el siguiente: Supongamos que una persona que vive en Madrid está organizando sus vacaciones y quiere ir a Valencia. Al ser usuario de Internet, este turista no se limitará a consultar los prospectos de las agencias de viaje y las guías turísticas oficiales. Es muy probable que indague previamente a través de la Red sobre los hoteles, los restaurantes, las actividades culturales, e incluso que visite los museos y monumentos virtuales correspondientes. Probablemente reservará sus plazas y comprará sus entradas a través de la Red, lo cual va ligado al problema de la seguridad para las transacciones económicas. En definitiva, esta persona podrá acumular datos sobre Valencia a través de múltiples circuitos, tanto públicos como privados. Los medios digitales que proporcionen unos buenos servicios al respecto podrán constituirse como un referente fundamental para este tipo de “consultas virtuales”.

Una de las novedades principales planteadas por el periodismo electrónico en Internet se deriva de la interactividad. Secciones tales como las Cartas de los lectores,

los anuncios por palabras o las tribunas de opinión pueden verse fuertemente reforzadas. Así como los clubs y las tertulias fueron en su día ámbitos fundamentales para la formación de la opinión pública, Internet está mostrando que la opinión vuelve a formarse en estos espacios, y no sólo en las grandes tribunas mediáticas. La manifestación de más de 2.500.000 personas en Internet para protestar por la firma por parte del presidente Clinton de unas normas que suponían la implantación de formas de censura a la libertad de circulación por la Red, da idea de que este tipo de actividad en la Red va reproduciendo muchas de sus formas clásicas, aunque ésta vez sea a distancia.

Los periodistas deben intentar buscar nuevas formas de narrar que se ajusten a las necesidades y capacidades del medio, manteniendo como punto esencial de partida esta interactividad con el lector.

M^a José Pérez Luque, en su libro *El periodismo y las nuevas tecnologías*⁴⁵, establece los siguientes tipos de interacción en tiempo real que genera un medio de comunicación online:

- Fuente – redacción
- Redacción - lectores
- Lectores - anunciantes
- Lectores - lectores

Siendo pues la interactividad una de las principales esencias de la comunicación on-line, merece la pena destacar las principales características de estos aspectos:

- Las ilimitadas posibilidades de acceso a todo tipo de información a nivel mundial ofrece al periodista la opción de interactuar con unas fuentes inestimables. Sin embargo, las características del propio medio acrecientan aún más la eterna discusión sobre la fiabilidad de dichas fuentes. Internet, hasta ahora, proporciona la ventaja o desventaja del anonimato, para aquel que así lo desee. A pesar de que cada vez aparecen nuevas tecnologías, como puede ser la huella digital, que permiten la identificación del interlocutor, aún están fuera del alcance y conocimiento del usuario de a pie.

⁴⁵ Pérez Luque, M^a José: *El periodismo y las nuevas tecnologías*. Ed: Newbook Ediciones. 1998. pp: 110-111

- La interactividad permite asimismo, una relación activa entre los periodistas y los lectores, los cuales se pueden dirigir directamente al redactor de la información, en tiempo real, si así lo desean, y compartir sus opiniones sobre el caso. Del mismo modo, merece la pena reseñar la potencialidad que tiene el usuario para controlar los elementos de los que constará la publicación que él reciba y cómo van a ser distribuidos dentro de la página.
- Los publicistas han visto en este nuevo medio un gran potencial para sus anuncios. Así, los profesionales de la publicidad pueden interactuar con el lector de una forma mucho más directa e influyente. No se trata tan sólo de la posibilidad de anunciar un producto, sino también de ofrecer un servicio, ya que el usuario puede adquirirlo desde la página *web* del anunciante en cuestión. La publicidad interactiva *online* permite el contacto directo con el anunciante, acabando por tanto con los grandes gastos de las campañas publicitarias tradicionales.
- Del mismo modo, los lectores de una publicación en la web, pueden compartir sus opiniones entre ellos, en tiempo real, en los chats o foros de discusión del medio, llegando a crear la llamada "cibercomunidad". Este hecho es de enorme interés para el medio ya que permite conocer mejor a sus usuarios y adaptar aún más su oferta. También ofrece a los publicistas una información mucho más contrastada sobre los intereses de la audiencia y de su perfil.

Otra dimensión a este respecto que cita Pérez Luque con respecto a la interactividad es la interacción con la información. Ella la define de lo siguiente forma:

“La capacidad que tiene el usuario final de controlar qué elementos van a ser difundidos y cuándo. Como consecuencia de esta facultad que tiene el usuario de definir y redefinir la estructura de las noticias, el contexto queda alterado⁴⁶”.

b) El hipertexto

⁴⁶ Pérez Luque, M^a José: *El periodismo y las nuevas tecnologías*. Ed: Newbook Ediciones. 1998. pp: 111

La aparición del *World Wide Web* ha sido decisiva para la expansión y la difusión mundial de Internet, y ello por múltiples motivos, entre los que se encuentra lo que podemos llamar la arquitectura del WWW, es decir el entramado de conexiones que lo sostiene y que está basado en el hipertexto. Dado que el periodismo electrónico ha de adaptarse en el momento actual a dicha arquitectura, se deben tener presentes algunas nociones básicas al respecto, así como las consecuencias que tienen en el caso del periodismo electrónico.

Cuando Theodor H. Nelson acuñó en 1974 el término 'hipertexto', lo definió de la manera siguiente:

"Con 'hipertexto' me refiero a una escritura no secuencial, a un texto que bifurca, que permite que el lector elija y que se lea mejor en una pantalla interactiva (...). La escritura tradicional es secuencial por dos razones. Primero, se deriva del discurso hablado, que es secuencial, y segundo, porque los libros están escritos para leerse de forma secuencial. Sin embargo, las estructuras de las ideas no son lineales. Están interrelacionadas en múltiples direcciones, y cuando escribimos siempre tratamos de relacionar cosas de forma no secuencial”.

El concepto de pensamiento secuencial deriva del pensamiento filosófico de Thomas Hobbes, quien en los primeros capítulos de su obra *Leviatán* afirma lo siguiente: “Por consecuencia o encadenamiento de pensamiento entiendo esa sucesión con que un pensamiento sigue a otro, y que, para distinguirla del discurso de palabras, recibe el nombre de discurso mental”⁴⁷. Sin embargo, parece bastante claro que el modo en que fluyen estos pensamientos no siempre es la misma. Es la propia persona la que, en virtud de su situación personal, gustos, intereses, etc, determina el camino a seguir.

Aunque tradicionalmente se considerara la construcción de un texto de forma secuencial, ya entonces existían en algunos escritos ciertas fórmulas que rompían dicha secuencia. Por ejemplo: los libros científicos incluían notas a pie de página. Cada vez

⁴⁷ Hobbes, Thomas: *Leviatán*. Ed. Sarpe. Madrid 1983.

que aparecía una de estas llamadas había que saltar en la lectura del texto principal al secundario de la nota. En la actualidad una página web está llena de enlaces de hipertexto que conectan unas partes del texto principal con otras páginas relacionadas. Éstos a su vez enlazan con otros... y así de forma continuada.

Merece la pena remontarse al significado etimológico de la palabra “Texto”, con el fin de comprender mayor medida el término que ahora nos ocupa. Ésta proviene de la palabra latina “tejido”. Se trata de una palabra polisémica. Según afirma Rogelio Gallo M⁴⁸, médico ginecólogo y doctorado en psicología, en su artículo *El libro virtual*, lingüistas, gramáticos, científicos, semánticos y semióticos se refieren al término *texto* como un conjunto de palabras que componen un escrito. La definición depende de la corriente o escuela científica que la utilice. En la actualidad este significado está virando y ampliándose debido a la terminología de la información.

Gallo afirma que ahora texto abarca no sólo la literatura, sino el arte, la gráfica, el texto hablado, comunicable. Según la teoría de la comunicación humana, texto es en la actualidad “un punto donde se intersectan varios códigos o sistemas que configuran una compleja red de relaciones intertextuales”.

Esta evolución en el significado amplio de la palabra texto ha originado un abanico de matices, una familia importante de términos, tales como intertextual, contextual, subtexto, etc. Incluso las tecnologías de la información y en particular, este nuevo medio que es Internet, se ha apropiado del término añadiendo a esta lista el “hipertexto” que ahora nos ocupa.

Antes incluso del texto tan clarificador de Nelson (1970) que hemos citado anteriormente, Roland Barthes describe un ideal de textualidad que coincide exactamente con lo que se conoce como hipertexto electrónico, un texto compuesto de bloques de palabras (o de imágenes) electrónicamente unidos en múltiples trayectos, cadenas o recorridos en una textualidad abierta, eternamente inacabada y descrita con términos como *nexo (enlace), nodo, red, trama y trayecto*. Barthes afirma lo siguiente⁴⁹:

⁴⁸ Gallo, Rogelio M. *El libro virtual*.: <http://lawyersinmexico.com/portal/article.php3?sid=10>.

⁴⁹ Barthes, Roland. *S/Z*. Éditions du Seuil. París. 1970. Pp: 15

“Con *hipertexto* me referiré a un medio informático que relaciona información verbal como no verbal. Los enlaces unen lexias tanto "externas" a una obra, por ejemplo un comentario de ésta por otro autor, o textos paralelos o comparativos, como internas y así crean un texto que el lector experimenta como no lineal o, mejor dicho, como multilínea o multiseccional. Si bien los hábitos de lectura convencionales siguen válidos dentro de cada lexia, una vez que se dejan atrás los oscuros límites de cualquier unidad de texto, entran en vigor nuevas reglas y experiencias (...). En este texto ideal, abundan las redes (réseaux) que actúan entre sí sin que ninguna pueda imponerse a las demás; este texto es una galaxia de significantes y no una estructura de significados; no tiene principio, pero sí diversas vías de acceso, sin que ninguna de ellas pueda calificarse de principal; los códigos que moviliza se extienden hasta donde alcance la vista; son indeterminables...; los sistemas de significados pueden imponerse a este texto absolutamente plural, pero su número nunca está limitado, ya que está basado en la infinitud del lenguaje”.

Coincidiendo con la idea de Barthes, Michel Foucault concibe el texto en forma de redes y enlaces. En *Archeology of Knowledge*, afirma que "las fronteras de un libro nunca están claramente definidas", ya que se encuentra "atrapado en un sistema de referencias a otros libros, otros textos, otras frases: es un nodo dentro de una red... una red de referencias"⁵⁰.

Con la aparición de los ordenadores multimedia, esos nexos no sólo tienen lugar entre palabras, sino también entre sonidos, imágenes y datos. Ello permite un tipo de recuperación de información transversal. En lugar de seguir el orden de las páginas de un libro, de las hojas de un periódico, de los compases de una melodía o de las secuencias de una película, el hipertexto proporciona múltiples hilos para recuperar la información en función de los intereses concretos del usuario.

José B. Terceiro, catedrático español y autor del libro *Sociedad Digital*⁵¹, explica la connotación del término hipertexto, trasladado al uso de las Nuevas

⁵⁰ Foucault, Michel: *The Archeology of Knowledge*. 1976. (trad. A.M. Sheridan Smith, Nueva York, Harper Colophon).

⁵¹ Terceiro, José B.: *Sociedad digital (del homo sapiens al homo digitalis)*. Alianza Editorial. Madrid 1996

Tecnologías de la Comunicación y la Información: "Tradicionalmente la información se ha producido y transmitido de manera secuencial debido a los requerimientos impuestos por las sucesivas tecnologías utilizadas. El texto impreso ha venido organizando la información de forma lineal o jerárquica, diferente de las formas de relación basada en las conexiones que se establecen entre las distintas partes de aquella. En el hipertexto, palabras, frases o documentos se asocian en toda su extensión con información del mismo o de otros documentos a través de los adecuados enlaces, salvando así las limitaciones impuestas por la naturaleza lineal del texto impreso. Esa capacidad de remisión a fuentes colaterales de conocimiento que un texto ofrece a sus lectores, la materializa el hipertexto en los hiperenlaces (*hyperlinks*) incluidos en el texto y distinguidos con un subrayado o un tipo distinto de letra. Cuando con el ratón del computador activamos un hiperenlace para saltar del texto actual al texto que el hiperenlace invoca, lo que estamos haciendo realmente es saltar en el hiperespacio de un documento a otro".

La hipertextualidad digital a través de Internet y otros medios electrónicos obliga a los comunicadores sociales a dos cosas. Por una parte, a desarrollar una capacidad de vinculación entre conocimientos o hechos comunes, pero en algunas ocasiones tan distantes que anteriormente se creían totalmente desvinculados entre sí. Es imperativo en la información periodística digital lograr una habilidad de "hiperenlazar" realidades distantes (pero comunes) para poder establecerlas también entre documentos (textos, imágenes, sonidos, videos, gráficos, etc.). Esto permite presentar informaciones "contextualizadas" a los usuarios.

En segundo lugar, la hipertextualidad impone al comunicador una capacidad de síntesis, dada la brevedad del usuario ante la pantalla de la computadora personal y la avalancha de información que inunda la Red. Sobre esto, Terceiro apunta lo siguiente: "Tendrán éxito los grandes comunicadores que, capaces de condensar en pocas palabras la complejidad actual, pasarán a ser la referencia en la avalancha informativa. El columnismo en la prensa actuará como aportador de criterio e inteligencia para ayudarnos a dar sentido al caudal de información que nos llega, con el riesgo indudable del tremendo poder que ciertos medios tendrán en la dirección de los grandes asuntos.

Como ha señalado un importante editor de prensa: ya no habrá más lugares comunes, frases hechas, criptología de camarilla, periodismos endogmáticos y provinciano, cuyo *revival* de los últimos tiempos no es más que su canto de cisne. Las redes globales serán las dueñas de todas las primicias y los periodistas tendrán que aplicarse a la tarea de informar bien, proporcionando noticias contextualizadas, diseccionadas y valoradas. Los lectores dispondrán así de una información más libre, plural y rigurosa".

Parece clara, pues, la afirmación de Caridad y Moscoso al señalar que el hipertexto es la mejor solución para un periódico electrónico, ya que "puede asemejarse más a la estructura del pensamiento humano, creando una red de nudos y enlaces que permite al usuario desplazarse a través del sistema de forma multidimensional". En el sistema de hipertexto, determinadas partes del discurso textual (palabras, imágenes, etc) "se asocian con información del mismo o de los documentos"⁵².

Según el artículo de Díaz Noci titulado "Conceptos en torno al periodismo electrónico"⁵³, los enlaces se crean con diferentes objetivos y pueden ser de distintos tipos.

- Se puede tratar de un enlace textual, palabras o frases que llevan a otras palabras, frases o textos.
- Otro tipo de enlace son los iconos, que al ser pulsados por el ratón, o incluso por el dedo en el caso de pantallas táctiles, remiten a otras informaciones (textuales, icónicas o sonoras). Mientras que el texto-enlace aparece destacado mediante el color, tipo de letra o subrayado, el icono ofrece la ventaja de ser mucho más fácilmente identificado como enlace. Además, en numerosos casos se les da el aspecto de botones o mandos. Por ejemplo, se puede activar una película o un archivo de audio, mediante un sistema de botones similares a los que tiene un vídeo o un magnetófono, aunque estos botones sólo tengan presencia virtual.

⁵²Caridad, M.; Moscoso, P: *Los sistemas de hipertexto e hipermedios: una nueva aplicación en informática documental*, Madrid, Salamanca, Fundación Germán Sánchez Ruipérez; Madrid, Pirámide, 1991. PP 35-38.

⁵³ Díaz Noci, Javier: *Conceptos en torno al periodismo electrónico interactivo multimedia*. XV Cursos de verano, Universidad del País Vasco. <http://www.andalucia24horas.com/caracola/doc009.htm>.

- En otros casos basta con pulsar sobre una imagen fija para que se active un gráfico animado.

Según señala el profesor Raymond Colle, de la Universidad Pontificia Católica de Chile, en su “Curso de periodismo digital”⁵⁴, existen tres tipos de estructuras que podríamos encontrar en cuanto al hipertexto se refiere:

- **Estructura básica:**

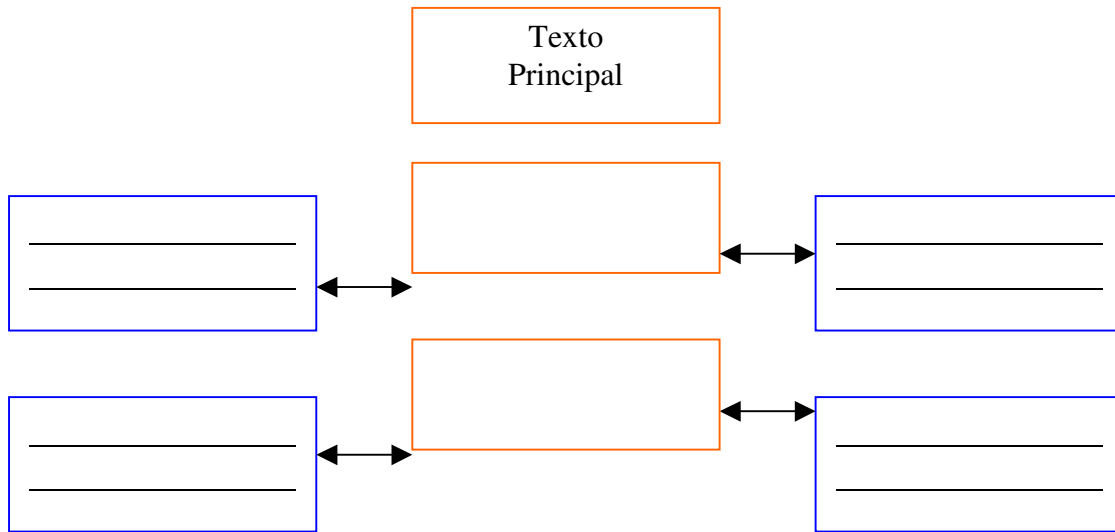
De acuerdo con las investigaciones de G. Landow⁵⁵, las primeras aplicaciones del hipertexto corresponden a:

“transliteraciones al hipertexto de poesía, de ficción y de otras materias originalmente concebidas para la tecnología del libro. La forma más sencilla y limitada de esta transliteración preserva el texto lineal, con su orden e inalterabilidad, y luego añade, a modo de apéndices, críticas variantes textuales u otros textos, cronológicamente anteriores o posteriores. En estos casos, el texto original, que conserva su forma antigua, se convierte en un eje fijo del cual irradian los textos conectados, y ello modifica la experiencia del lector de este original texto en un nuevo contexto”.

En la siguiente ilustración podemos ver como el texto principal se divide en diversas piezas que mantienen su estructura secuencial. Pero a cada uno de estos se vinculan notas y comentarios que pueden ser leídos paralelamente.

⁵⁴ Colle, Raymond: *Curso de periodismo digital*. Servicio de Computación, Informática y Comunicaciones, Pontificia Universidad Católica de Chile:
http://www.puc.cl/curso_dist/dmw/teoria/hitex2.html.

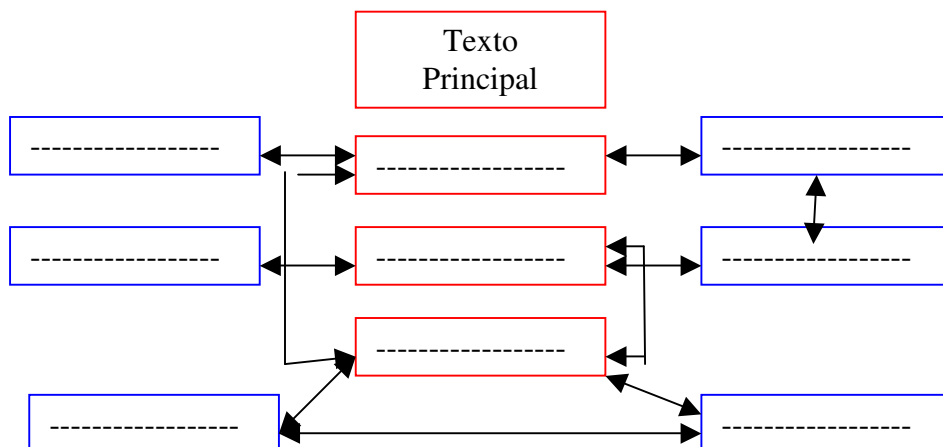
⁵⁵ Landow, George P. *Hipertexto: la convergencia de la teoría crítica contemporánea y la tecnología*. Barcelona: Paidós Ibérica, 1995. 288 p. (Paidós hipermedia, 2).



En este ejemplo podemos ver cinco páginas (una que integra el texto principal y seis anexas) unidas entre sí por vínculos.

- **Estructura mediana:**

En ciertos casos, el texto original ya estaba dividido en distintas piezas y el autor invita a cambiar el orden de lectura. En este caso, el traspaso al hipertexto cumple mejor que el libro los designios del autor o de una comunidad que requiere de una lectura multisequencial para interpretar mejor el texto. Un ejemplo de este tipo de hipertexto es el del “CD WORD: The interactive Bible Library” del Seminario Teológico de Dallas, que incluye cuatro versiones inglesas de la Biblia, más algunos textos griegos, léxicos, etc. En este caso la estructura es más compleja.

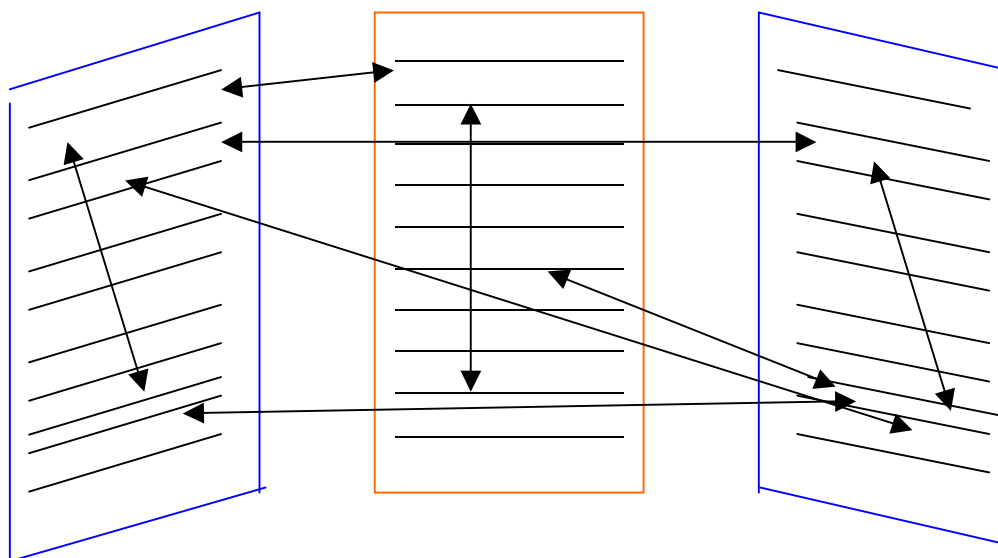


En este gráfico continúa habiendo siete páginas, pero se han añadido vínculos o *links* dentro de la página del texto principal y entre páginas anexas. Como podemos observar, esta estructura mediana permite leer los capítulos o apartados en el orden que se desee y, si el usuario-lector está interesado, consultar los textos complementarios.

- **Estructura compleja:**

Un ejemplo de este tipo de estructura, es el sistema usado con el “In Memoriam” de Tennyson, en un hipertexto elaborado por la Universidad de Brown, ya que “se vale de los nexos electrónicos para elaborar un mapa de alusiones y referencias del texto, tanto internas como externas”⁵⁶.

⁵⁶ Landow, George P. *Hipertexto: la convergencia de la teoría crítica contemporánea y la tecnología*. Barcelona: Paidós Ibérica, 1995. 288 p. (Paidós hipermedia, 2). pp 53.



En cuanto a su aplicación al periodismo electrónico, estos hipervínculos sustituirán a las citas parciales y servirán como apoyo de documentación, sin necesidad que este texto secundario se halle en el cuerpo principal de la noticia. Asimismo, tal y como afirma Andrés Restrepo en su artículo: “La unidad fundamental de la nueva mecánica informativa”⁵⁷, la calidad del redactor no será evaluada por su peripecia en conseguir “exclusivas”. La pauta de su capacidad estará sujeta a la habilidad para remitir a su lector a los vínculos electrónicos que pueden ampliar la información que ha tratado en su artículo.

Según Restrepo, estos tramados informativos que Barthes llama *lexias*, “reemplazarán la unidad fundamental del periodismo tradicional que todos conocemos como *lead*. La *lead* adicción es un vicio informativo atribuido al ritmo de vida frenético de las sociedades modernas. Sin embargo las nuevas tendencias parecen indicar que los espectadores quieren estar mejor informados que nunca para conseguir, o al menos intentarlo, entender lo que está pasando en el mundo”.

⁵⁷ Restrepo, Andrés: *La unidad fundamental de la nueva mecánica informativa*: <http://members.nbci.com/perelec/art-ar01.html>.

c)Elementos multimedia

Con la palabra multimedia podemos designar dos conceptos diferentes. En primer lugar el referente a empresas o grupos de empresas de comunicación que engloban a distintos medios informativos, escritos y audiovisuales. El segundo concepto hace alusión a la unión en un mismo producto de materiales escritos, audio, imágenes fijas o en movimiento y vídeo. Ésta última significación, se encuentra íntimamente relacionada a los grandes y rápidos avances de las nuevas tecnologías. En la actualidad, prácticamente cualquier ordenador doméstico es capaz de procesar información gráfica y sonora, a pesar de que ésta ocupa más memoria que el simple texto.

Asimismo, a esta combinación de elementos se la conoce también con el nombre de hipermedia y se la define como “una organización de información textual, visual, gráfica y sonora a través de vínculos que crean asociaciones entre información relacionada dentro del sistema”⁵⁸. Es decir, si un vínculo une dos informaciones textuales lo llamamos hipertexto. En el caso de que vinculara dos sonoras, lo podríamos denominar hiperaudio. Lo mismo ocurre con la información visual, que pasaría a llamarse hipervisual. En el caso de que estos tres elementos estén interrelacionados, hecho que ocurre de Internet, nos hallaríamos ante un fenómeno de hipermedia.

Según citan Caridad y Moscoso⁵⁹, Podemos determinar tres factores fundamentales que han contribuido en el desarrollo del hipermedia. El primero es la arquitectura de sistemas, seguidamente la gestión de los proyectos multimedia y en tercer lugar la evolución de las comunicaciones. Un cuarto factor sería la aparición de soportes de almacenamiento masivo como el disco compacto o CD (años 80's). Este soporte inventado por la firma holandesa Philips se destinó, en un principio, a almacenar música. No tardó mucho en desplazar al disco de vinilo debido a la gran cantidad de ventajas que ofrece el CD frente a este último: menor tamaño, mayor

⁵⁸ Caridad, M.; Moscoso, P. : *Los sistemas de hipertexto e hipermedios: una nueva aplicación en informática documental*, Madrid, Salamanca, Fundación Germán Sánchez Ruipérez; Madrid, Pirámide, 1991. pp 48.

⁵⁹ Caridad, M.; Moscoso, P.: *Los sistemas de hipertexto e hipermedios: una nueva aplicación en informática documental*, Madrid, Salamanca, Fundación Germán Sánchez Ruipérez; Madrid, Pirámide, 1991. pp 51.

capacidad, una producción más barata y menos contaminante, y una mayor calidad y fiabilidad ya que se trata de un soporte digital (ceros y unos). Si el CD contenía datos digitales sobre música, nada impedía que en este soporte se almacenara otro tipo de información. La aparición del CD-ROM (Compact Disc-Read Only Memory), o disco compacto de sólo lectura permite almacenar datos multimedia o hipermedia: texto, sonido, imágenes y vídeo, e interrelacionarlas entre sí.

En referencia al concepto de Multimedia, Andreas Whittam Smith, presidente de Newspaper Publishing PLC, empresa editora del diario británico *The Independent*, “la información contenida en un único disco supera en mucho lo que puede caber entre las tapas de un libro normal (...). Es un multimedio , en el sentido de que combina diferentes medios en una forma nueva. Es interactivo; es decir, le permite a uno preguntar al material. (...). La tecnología interactiva y de multimedios sólo se convertirá en un medio verdaderamente nuevo cuando encuentre su propia voz”.

Como dice Jean Cloutier respecto a la integración multimedia, ante este nuevo medio se abre la posibilidad a un lenguaje "polisintético e integrador". Polisintético porque aglutina diversos elementos, que toman sentido sólo dentro del conjunto. Es sincrético, y proporciona acceso a las tres dimensiones espaciales y a la temporal⁶⁰.

Por su parte, Alain Chaptal⁶¹ ofrece una definición de las principales características del medio multimedia que merece la pena reseñar. Se trataría pues de un: “agrupamiento sobre un soporte o según un modo de acceso único de imágenes fijas o animadas, de sonidos, de textos y de datos, almacenados de forma digital, que pueden ser utilizados de forma interactiva en el sitio o a distancia.

El lenguaje HTML

El HTML es un lenguaje de marcas basado en hipertexto (HyperText Markup Language) que se usa para formatear documentos que puedan ser interpretados por cualquier navegador compatible con el protocolo HTTP. Su origen está en la necesidad

⁶⁰ Cloutier, Jean (1994): *L'audiovisuel et le multimedia*, Communication et langages, n° 99, París: Edicions Retz, pp.42-53

⁶¹ Chaptal, Alain (1993). - *Les habits neufs du multimédia*. - Médias Pouvoirs, n°31-32. - pp.211

de disponer de un código que permitiera general ficheros válidos para el Web. El primer objetivo fue conseguir una simplificación del SGML (Standard Generalized Markup Language) un lenguaje muy usado y normalizado para generar hipertexto. Su uso era complejo pero estaba muy extendido en el CERN, por lo que Berners-Lee pensó que realizar una versión simplificada sería sencillo a la vez que facilitaría su uso por parte del personal del CERN que ya estaba acostumbrado a usar el código del SGML.

El HTML fue pensado como una semilla que pudiera desarrollarse en función de la evolución del Web. De esta forma a la vez que el WWW ha ido demandando nuevos servicios, el HTML ha ido evolucionando para incluir nuevas prestaciones. Todo esto manteniendo la compatibilidad con las versiones anteriores del lenguaje y con todo tipo de plataformas informáticas.

El sistema, sin embargo, se veía limitado por dos condiciones. La velocidad de la transmisión y el crecimiento controlado. La primera de estas limitaciones se deriva de la falta de líneas de transmisión adecuadas para la emisión de documentos multimedia. Internet tenía un diseño robusto y adecuado para crecer, pero todavía no había ancho de banda apropiado para poder transmitir por la Red documentos que incluyeran imágenes y sonido. De hecho esta limitación todavía la estamos viviendo en estos momentos, pues la demanda de envío de información multimedia por la Red es superior a las condiciones técnicas que nos ofrece Internet. En cuanto a la segunda de las limitaciones versa sobre cómo hay que evitar que las prestaciones extras que se pueden conseguir a través del uso de determinados navegadores pasen por encima de la idea primordial de que los documentos obtenidos sean compatibles con cualquier sistema operativo que se pueda conectar a Internet. Es decir, los documentos que se incluyen en la web deben de poder ser vistos por cualquier navegador. Evidentemente la simplicidad del HTML permite realizar modificaciones en el código para que en determinados navegadores se pueda obtener una mejor visibilidad de documentos mediante el uso de etiquetas no normalizadas. El uso de estas prestaciones limitan, sin embargo, la universalidad del WWW.

La base de la creación de un documento en formato HTML es el uso de unas marcas especiales llamadas etiquetas. Las etiquetas o *tags* son unos comandos especiales que afectan al formato del documento y se pueden reconocer fácilmente por ir encerradas entre los signos [**<**] y [**>**].

Un documento escrito en código HTML siempre empieza con la etiqueta `<html>` y termina con la etiqueta de cierre `</html>`. Existen diversos tipos de etiquetas que afectan a los aspectos que definen la maquetación del documento. Así pues hay etiquetas para especificar tipos de párrafos, tipografía, colores de fondos, tablas, inclusión de imágenes, forma de señalar los hiperenlaces, etc.

Todo documento HTML se divide en dos partes: Cabecera y cuerpo que se marcan con las etiquetas `<head>` y `<body>`. En la primera de estas partes se incluyen datos técnicos que facilitan la interpretación del documento y su localización por parte de motores de búsqueda. Entre otras cosas en la cabecera se incluye el título del documento y palabras claves para ser utilizadas por los buscadores. La segunda parte del documento, el cuerpo, es la que incluye los elementos que serán visualizados en el navegador.

Este primer código HTML se conoce como HTML estático pues sus páginas no se pueden actualizar de forma automática como sí ocurre con el DHTML o HTML dinámico, que está basado en el uso de bases de datos cuyos campos son los elementos con los que un software especial compone las páginas que serán visualizadas por el navegador. Existen diversas aplicaciones especializadas en la realización de HTML dinámico, las más conocidas son el ASP para uso con servidores de Microsoft, PHP en entornos Linux o JSP en entornos SUN.

Desde sus comienzos Berners-Lee concibió el Web para ser visualizado y editado desde un mismo programa que él definía como navegador/editor. Al igual que pasaba con el SGML era necesario que existiera un editor que evitara tener que tratar directamente con el código fuente, como ocurre con los procesadores de texto que usan un código para permitir el formateo de la página, aunque este código sea invisible para la persona que lo usa. Sin embargo en el caso del HTML ocurrió algo diferente, el uso del nuevo lenguaje era tan simple que se podía leer directamente desde el código fuente. De esta manera el lenguaje se extendió con gran facilidad. Para muchos autores ni siquiera hacía falta usar editor, el código podía escribirse directamente en formato ASCII desde cualquier aplicación que permitiera editar textos y su lectura humana era perfectamente válida.

La necesidad de realizar intercambio de contenidos entre los diferentes sitios web existentes en el mundo ha promovido la revisión del HTML para permitir un

estándar de intercambio que permita la difusión de datos estructurados en campos. Este estándar es el XML (eXtensible Markup Language) un lenguaje cuyas marcas se crean junto con el documento a la vez que se envían. El uso de estos documentos facilita que una agencia de noticias pueda estar actualizando continuamente a sus abonados datos en constante cambio, como pueden ser titulares, cotizaciones de bolsa y novedades de todo tipo. Los datos XML pueden ser directamente interpretados en una página web o alimentar a una base de datos que se actualice de forma automática.

e) Información documentada

Otro aspecto importante a tener en cuenta es la ilimitada cantidad de documentación que puede acompañar a una noticia "on-line", por lo que su contextualización está asegurada. Este acto de documentar se lleva a cabo a través del establecimiento de links que llevan al archivo del periódico, a otras bases de datos especializadas en el tema en cuestión, o a otras páginas simples especializadas en el tema. El usuario tendrá a su disposición todas las fuentes fiables (ya que han de haber sido comprobadas por los periodistas del medio en cuestión) sobre ese tema disponibles en Internet.

f) Información actualizada

El periódico electrónico permitirá a su vez publicar información con una periodicidad mucho mayor que en los medios tradicionales escritos. Del mismo modo, se podrán subsanar los errores en cuanto sean detectados. Ya no se hablará de un único cierre ya que las noticias serán continuamente renovadas.

Hay que tener en cuenta que estamos ante un nuevo medio de comunicación que permite la integración de elementos multimedia, junto con el texto, como audio, video, gráficos interactivos ... etc. El único problema actual para la utilización de estos elementos integrados en una noticia, o de aquellos que el periodista cree que puede ayudar al lector a un mejor entendimiento de la información, es el ancho de banda, problema que a veces se soluciona a través de formatos como "Schockwave".

g) Nueva dimensión del diseño

Asimismo, el diseño de las publicaciones "online" es parte esencial de la misma. Se trata de incitar al usuario a interactuar con la publicación. Para ello hay que ir más allá de la belleza puramente estética, y preocuparse de la funcionalidad de la página, dando preferencia a los elementos gráficos. La navegación dentro de la publicación debe ser sencilla y fácil de seguir. Hay que evitar en todo momento que el usuario se pierda. A la hora de diseñar una publicación en la Red, hay que tener en cuenta (al menos hasta que se cubran las necesidades técnicas actuales) el tiempo de descarga de dicha información, lo cual es en el actualidad uno de los mayores motivos por los que el navegante suele desechar las páginas web con muchos elementos gráficos y multimedia. Aunque este último sea nuestro caso, hay que intentar que, nada más conectar, el usuario tenga texto que leer mientras se cargan las imágenes y otros elementos más complejos.

2.4 LA PERSONALIZACIÓN DE LA INFORMACIÓN

Merece la pena una mención aparte otra inestimable aportación de Internet al proceso de la comunicación: la personalización de la información. Se trata de que el usuario obtenga únicamente aquella información que le interesa, de la forma que él crea más conveniente, de acuerdo a sus propios gustos. Para alcanzar este objetivo, se suele incluir en la publicación una página-registro en la que el lector tiene la oportunidad de determinar los contenidos y el diseño del periódico, según sus intereses. Sin embargo, hay que tener en cuenta que aunque sea el usuario quien supuestamente selecciona las noticias que le llegan, siempre necesitará la ayuda de los profesionales de la información para orientar al lector ante la gran cantidad de datos, muchos de ellos poco fiables, que se encuentran disponibles en la Red.

Con la llegada de los medios electrónicos a través de Internet, la estructura y presentación de los periódicos y revistas digitales podrán "adaptarse" a los gustos, intereses y saberes del lector final, construyendo así no una sino millones de ediciones

personalizadas, las cuales responderán a ciertos parámetros suministrados previamente por el usuario. Si un estudiante venezolano está haciendo su postgrado en Alemania y le interesa el deporte local de su país, la edición digital de su periódico vendrá con los principales titulares e informaciones referidas a las actividades deportivas de su nación.

Este novedoso esquema de comunicación masiva se ha identificado como "narrowcasting", el cual, a diferencia del "broadcasting" (modo tradicional de transmisión de información), está definido por cada uno de los destinatarios finales y no por el emisor.

Esta transformación representa un reto para el "nuevo periodista", pues el hecho de que los medios se adapten a cada usuario determina una hipersegmentación del mensaje comunicacional. Ya no se transmite una noticia, un reportaje o una crónica a través del medio a una masa informe e indeterminada de lectores. Aunque suene imposible, ahora el periodista deberá diseñar la información para cada usuario en particular. El lector ahora manda.

Esta posibilidad de personalización que ofrece la Red no siempre es observada desde un punto de vista positivo, como sería el caso de Walter Bender, Director de la investigación denominada *News in the future* del Instituto Tecnológico de Massachusetts (MIT), y creador del proyecto Daily-Me. Según Stan Lepeak, analista de la Red en la compañía de análisis de mercados Metagroup, un Daily – Me puede meter al no informado de la realidad, en una coraza. En la actualidad, alguien que encienda la televisión para ver el pronóstico del tiempo tiene la oportunidad de ver accidentalmente, los resultados de las elecciones en Rusia y que resulte de su interés en ese momento. Un Daily-Me no se lo hubiese permitido. Hubiese saltado esa información en el supuesto de que no estuviese esa temática dentro de su perfil de intereses⁶².

En referencia a este periódico personalizado, Nicholas Negroponte define al lector del futuro como un editor, porque será él quien irá editando su propio periódico en función de sus necesidades. El usuario del medio en cuestión dispondrá de

⁶² Reproducido por Torres Sabaté, Armengol: *El Acceso a la Información Electrónica Difundida y Publicada, según Perfiles de Usuario Definidos*:
http://www.ictnet.es/+atorres/docs/filtering/capitulo_vi.htm

herramientas que le permitirán realizar esta labor de forma muy automatizada, por lo que la labor de personas que realizan manualmente la búsqueda de información, podrá ser sustituida por sistemas expertos que rastrearán las fuentes de información y reunirán cada día la información que cada uno precisa.

Según exponen M.J. Pérez-Luque y M. Perea en su artículo “El actual periodismo online”⁶³, los servicios informativos que ofrecen actualmente contenidos personalizados, suelen recurrir a alguna de estas tres estrategias:

- a) **Cuestionarios.** Consiste en establecer una página-registro en la primera conexión que realice el usuario al servicio informativo en cuestión. Hoy día puede hablarse de dos tipos de cuestionarios. Por una parte, aquellos que sólo pretenden actuar de control del número de accesos que tiene el periódico, de ahí que tan sólo se exija rellenar unas casillas pidiendo un nombre de acceso y una clave. Por otro, aquellos que específicamente preguntan por los temas o las secciones predilectas del lector. Entre estos últimos pueden citarse los servicios informativos de suscripción y los servicios informativos que respondan al tipo “Daily-Me”.
- b) **Seguimiento de los pasos del usuario cuando navega a través del servicio informativo.** Esta información se utiliza posteriormente para proponer a la audiencia recibir aquellos contenidos que única y exclusivamente son de su interés.
- c) **Agentes inteligentes.**

En la actualidad, las modalidades más comunes de presentación de estos contenidos personalizados son las siguientes⁶⁴:

⁶³ M.J. Pérez-Luque y M. Perea: *El actual periodismo online*:
<http://www.hottopos.com/MirLibro/intrduc.htm>

⁶⁴ M.J. Pérez-Luque y M. Perea: *El actual periodismo online*:
<http://www.hottopos.com/MirLibro/intrduc.htm>

- a) Presentar un menú con aquellas informaciones que se corresponden con la selección que hizo el usuario. Un claro ejemplo sería CRAYON⁶⁵, *Create your Own Newspaper*.

Una variante de esta posibilidad consiste en enviar al usuario la información que él ha solicitado tomándola de diferentes medios de comunicación, y no sólo de noticias elaboradas por el servicio informativo. Esta estrategia está siendo ofrecida en The Gate⁶⁶.

- b) Utilizar las tecnologías PUSH (esta aplicación es utilizada por Point Cast), las cuales remiten al disco duro del ordenador del usuario la información que solicitó en su primer conexión al servicio informativo online. Esta nueva tecnología permite, además, actualizar constantemente los contenidos que se envían al lector.

Incluso los soportes de la información se adaptarán a las comodidades de los lectores-internautas, tal y como afirma Negroponte en su libro "El Mundo Digital". Actualmente ya se está desarrollando la llamada **tableta de Fidler**, como uno de los ejemplos de estos grades avances de la comunicación. Algunos investigadores consideran la opción de que llegará un día en que cada persona lleve pantallas electrónicas ligeras, en lugar de periódicos, o libros de texto. Los estudiantes podrían comprar el material escolar en un formato de tarjeta de memoria, adaptable a la pantalla, o bajárselo directamente de Internet. Roger Fidler, profesor de Periodismo y Comunicación de Masas del estado de Kent (EEUU), comenzó a trabajar en 1992 con una tableta electrónica o "flat paner,, denominada "Portable Document Viewer" o PDV, con la empresa KnightRidder, posteriormente continuó su proyecto en la Universidad. La tableta en cuestión tendría las siguientes características:

- 20cm. x 30 cm.
- 1 cm. de espesor
- 1 kg. de peso
- Permitiría la conexión en línea (terrestre o aérea)

⁶⁵ CRAYON: <http://crayon.es>

⁶⁶ The Gate: <http://www.sfgate.com>

- Permitiría realizar operaciones bancarias, reservas, correo, compras.
- Podrían enviarse documentos sonoros, infográficos y textuales.
- La interacción se realizaría mediante lápiz electrónico y reconocedores de voz.
- Sólo para periódicos

Nuestra opinión es que la personalización de la información será la clave del nuevo periodismo, un periodismo que acercará al lector a sus propios gustos y aficiones; un periodismo que bajo la vigilancia de los profesionales de la información que seguirán actuando como "gatekeepers" comprobando la fiabilidad de las fuentes a las que acceden libremente sus lectores, ofrecerá al usuario la posibilidad de obtener toda la información disponible, en tiempo real, a nivel mundial sobre una cuestión particular. Ya no será el periodista quien, aplicando el filtro de su profesionalidad, seleccionará lo que él considera más importante para el lector; será este último quien disponga para su periódico personalizado, aquellas noticias que aunque no sean de interés general, lo sean del suyo propio.

Web de The Gate (www. sfgate.com/)

SF Gate
San Francisco Chronicle **PM** **KRON 4**


Thursday, May 31, 2001
Updated: 12:07 AM PDT

Jump to:

▶▶ **GO**

[SF Gate Home](#)
[Today's News](#)
[Sports](#)
[Entertainment](#)
[Business](#)
[Technology](#)
[Traffic](#)
[Weather](#)
[Live Views](#)
[Ad Catcher](#)
[Classifieds:](#)

Pugnaciousness And Potatoes
"We are a breed of people who say, This is our home and we are entitled to prote it."—The people of Bonner County, Idaho can't help but a feel a bit proud six children are holding o Sheriff deputies trying to take them into custody.

Catch A Wave

Chronicle / Craig Lee
There's not a soul alive in San Francisco who can remember a May day as hot as this one. When the mercury hit 101 downtown today, the city beat a record set way back in 1887.

Blazes Stop BART
Three small fires halt service just as A game gets out; ride home affected for hours Wednesday afternoon. Chronicle
• [Ammonia scare over in Berkeley](#)

Grid Sweats It Out
Stage 2 warning in effect, but rolling blackouts averted on scorcher of a day
AP
• [Get blackout alerts by e-mail](#)
• [PG&E's bankruptcy bonus bonanza](#)
• [Guy's canny use of media](#)
• [Energy Crisis special report](#)

El periódico inteligente de Negroponte

"El transmisor determina todo y el receptor simplemente toma lo que recibe". Para Nicholas Negroponte, director del Laboratorio de Medios del Instituto de Tecnología de Massachusetts (el célebre MIT), este paradigma comunicacional de los medios masivos tradicionales es el primero que debe modificarse con el advenimiento de las Nuevas Tecnologías.

Para Negroponte, los diarios comunes son producidos "con toda la inteligencia en el punto de transmisión", estableciendo un supuesto orden coherente de acuerdo a géneros periodísticos, fuentes y tipo de información. Sin embargo, el especialista asegura que la mayoría de los lectores de diarios en el mundo entero jerarquiza, clasifica y revisa la edición que tiene en sus manos a su exclusivo parecer, lo cual sugiere una "inteligencia" determinante y concluyente por parte del receptor del mensaje.

En un comunicado⁶⁷ del 1 de julio del 1993, Negroponte decía que el programa que estaba desarrollando en el MIT denominado *News of the Future*⁶⁸ tenía el desafío de desarrollar sistemas para:

- La comprensión por ordenador del contenido de las noticias.
- La comprensión por ordenador de las necesidades específicas e intereses de una persona.

Estos dos conceptos debían de estar reunidos en un sistema personalizado de noticias. Los objetivos a cumplir eran “mejorar la eficiencia de producción, los tiempos de entrega, la conveniencia de presentación, y la relevancia del contenido editorial y de publicidad para el consumidor”.

⁶⁷ Reproducido por Torres Sabaté, Armengol: *El Acceso a la Información Electrónica Difundida y Publicada, según Perfiles de Usuario Definido* :

http://www.ictnet.es/+atorres/docs/filtering/capitulo_vi.htm

⁶⁸ News of the future: <http://nif.www.media.mit.edu/index.html>

Este proyecto se concentra en tres áreas principales:

- **Comprensión de contenidos por el ordenador:** incluye la generación de índices y representaciones del texto basados en el significado de los mismos, no sólo en sus palabras.
- **Modelado por ordenador de procesos e individuos:** sistema que deberá seguir los cambios de intereses sobre las noticias para un usuario, manteniendo un balance entre “especialización” y “exploración” de los artículos. Se desarrollarán para este propósito algoritmos de aprendizaje. Otra parte se concentra en la personalización del contenido a través del análisis de contenidos y modelado del usuario. El objetivo aquí es el de complementar el contenido orientado al usuario con selecciones breves y titulares de historias desarrolladas por los editores de hoy. Otra parte se concentrará en el tema de la publicidad, tal vez, en la forma de “anuncios personalizados” y de “anuncios a medida que los necesitamos”. Una cuarta línea de estudio está en los sistemas “storyteller” o “cuentacuentos”, es decir, programas con conocimientos detallados sobre determinados temas y capaces de producir una narrativa a medida de lo que conoce, y aprende, del entorno del usuario, con sus preferencias e intereses. La última línea es sobre las noticias como herramientas de aprendizaje. Los niños son aquí el objetivo principal de este trabajo dedicado a encontrar el qué y el cómo aprenden ellos a medida que se involucran en las actividades típicas relacionadas con las noticias: selección –captura- análisis y composición de noticias.
- **Colección Automática y Presentación:** comprende investigación en objetos media (texto, audio, vídeo) capaces de ser adaptados en su entorno de aplicación. Por ejemplo, deberían ser capaces de alterar su contenido y modo de presentación en respuesta a los cambios en las preferencias del consumidor. Una línea de trabajo se dedica al “media transcoding”, es decir, al proceso de empaquetar las noticias en un formato diferente al que fueron originalmente capturadas o creadas. Esto es útil para ajustar la presentación a diferentes contextos mientras se preserva al contenido literal. Una tercera línea se enfoca

en la presentación de audio las noticias de manera que se evita recitarlas linealmente. Se pretende que el oyente pueda tener un acceso fluido empleando procedimientos conversacionales y de resúmenes. Una cuarta línea de trabajo se dirige a desarrollar un repertorio de herramientas gráficas de expresión que permitan asistencia informática en la expresión de un diseño dinámico e interactivo. El propósito específico es construir un esquema automático para las noticias en un marco inteligente , espacial, temporal y relacional. La línea final de investigación trata con el “feeling” del papel. Explorará los mecanismos táctiles del usuario para generar señales de entrada/salida útiles para el sistema de información (tales como pasar las páginas o encontrar señaladores), material de pantalla híbrido activo/pasivo para mayor portabilidad , y la posibilidad de utilizar papel reciclado para impresión en el hogar.

Como ejemplo de lo que debería de ser un periódico personalizado, el MIT ha elaborado un periódico con estas características y en continuo desarrollo denominado Fishwrap. Este “periódico”, que comenzó a desarrollarse en otoño de 1993, es un servicio de noticias que lo que hace es ir cambiando el orden de prioridad de sus noticias dependiendo de cuánta gente de la universidad la está leyendo. Cuántos más lectores atraiga una historia , más rápidamente se mueve hacia la cabeza de la lista . Fishwrap, que proviene del dicho periodístico “las exclusivas de hoy envolverán el pescado de mañana”, intenta equilibrar así el deseo individual por la especialización con la necesidad de participar y de conocer el mundo. Ofrece a sus lectores una “ventana egocéntrica” sobre asuntos mundiales, permitiéndoles explorar temas de interés personal seleccionados previamente.

2.5 LA REDACCIÓN *ONLINE*

En la actualidad, Internet se ha convertido en un medio de comunicación de masas, en el que intentan estar presentes los antiguos medios tradicionales. Cada vez es mayor el número de periódicos que decide sacar una edición Online. Incluso la radio y la televisión parecen haber descubierto las grandes cualidades y posibilidades que la Red

de redes les ofrece. Estos medios han creado su "imagen" en Internet, aprovechando así su capacidad y facilidad de difusión. Sin embargo, hay que tener en cuenta la esencia de Internet como medio de comunicación en sí mismo. No se trata simplemente de un soporte en el que adaptar los medios tradicionales. Se trata de una nueva forma de informar y de distribuir dicha información.

A la hora de redactar en un documento cuyo objetivo es su publicación en Internet, merece la pena reflexionar sobre la importancia del soporte en la expresión de un acto lingüístico. Es interesante a este respecto comparar el medio oral, con el medio escrito y el medio electrónico. Para ello, merece la pena reseñar una ilustrativa cita de Umberto Eco, acerca de la sorpresa de un nativo americano ante un mensaje transportado en "papel". La cita corresponde a su Introducción a *El límite de la interpretación*, en la que se relata la historia contada por el explorador y reportero John Wilkins en su viaje a Norteamérica en 1641:

“Hasta qué punto debió parecer extraña el Arte de la Escritura ya en su primera invención, lo podemos comprender por los americanos descubiertos recientemente, que se sorprenden al ver que los hombres conversan con los libros y se esfuerzan por creer que el papel pueda hablar (...)

Hay un bello relato sobre esto, que se refiere a un esclavo indio, que fue enviado por su amo con una cesta de higos y una carta. Durante el camino se comió una gran parte de la carga, y solo entregó la parte restante al destinatario, quien, después de leer la carta y no encontrar la cantidad de higos que debía haber recibido, acusó al esclavo de habérselos comido, poniendo como prueba lo que la carta decía. Pero el indio (a pesar de esta prueba) negaba ingenuamente el hecho, maldiciendo el papel como un testigo falso y mentiroso.

Al poco tiempo de esto, fue enviado con un cargamento igual y una carta en la que se indicaba el número exacto de higos que debían serle entregados al destinatario. El esclavo, de nuevo, como había hecho la otra vez, se comió una buena parte de los higos durante el camino. Pero

en esta ocasión, antes de tocarlos, con el propósito de evitar todas las posibles acusaciones, cogió la carta y la escondió bajo una enorme piedra, pensando que de esa manera, si no le veía comérselos, no podría denunciarle. Pero como fue reprendido con más firmeza que la vez anterior, confesó su falta y admiró la divinidad del papel, prometiendo cumplir en el futuro todas sus encomendaciones con fidelidad”⁶⁹.

El descubrimiento de un nuevo medio de comunicación suele suponer varias reacciones distintas por parte de la persona que lo utiliza: admiración, simple aceptación, desconfianza o rechazo. Pero lo que está claro es que para todos ellos (incluso para los que no lo quieren aceptar) supone un cambio, no sólo para muchos periodistas que a partir de entonces tienen que asumir las características del nuevo medio y adaptarse a él, sino también para la sociedad en general, que gracias a esta evolución verá cada vez más satisfechos sus ideales de información.

Otras ventajas y desventajas de este nuevo medio con respecto a los anteriores, las añade Màrius Serra, escritor y colaborador habitual de *La Vanguardia*, en su artículo “Escribir en tiempos de Internet”⁷⁰, en el que recoge, además, diversas formas de escritura que podemos encontrar hoy en la Red:

“Internet implica un aumento notabilísimo en la producción de textos. El proceso de lectoescritura, ausente en las conversaciones telefónicas, reaparece en el correo electrónico. Esto es innegable. Pero ¿qué tipo de escritura se da en la Red? Creo que podemos hablar de tres grandes paradojas en la escritura internáutica:

- 1) Un retorno entre romántico y chapucero a la escritura apresurada al lado de una práctica obsesiva de la reescritura cercana al cánón clásico.
- 2) Un intento real de abolir la linealidad de la lectoescritura al lado de una simplificación flagrante del cuerpo narrativo que entroniza al esquematismo.

⁶⁹ John Wilkins. 1641. *Mercury or The Secret and Swift Messenger*. 3ª edición, Nicholson, 1707, pp. 3-4, Londres.

⁷⁰ Serra, Màrius: *Escribir en tiempos de Internet*: <http://www.vilanova.org/tintec/ponencia/serra.htm>

3) Una revisión sincera del concepto de autoría al lado de una lucha feroz por conseguir una mayor incidencia del autor en un mercado presuntamente universal.

El primer punto nos lleva a establecer una serie de consideraciones sobre el acto físico de escribir, el segundo al de leer y el tercero nos acerca al ser”.

Con respecto a la escritura apresurada que se practica cuando se escribe en la Red, Serra la define con el anacronismo que contiene la locución catalana "escriure a rajaploma" (escribir a chorro, se entiende que de tinta). Los mejores ejemplos de esta escritura no sólo remiten al espíritu romántico de la inspiración repentina, sino que también nos acercaría a la "liberación de los dedos" de la que hablaban las vanguardias al practicar una escritura presuntamente automática.

En su artículo Serra señala que “la escritura de manos abiertas y "tiempo real" se acerca en sus mejores casos a esa otra convención que denominamos el "habla" y evidencia en los peores las grandes carencias estructurales de la lengua que exhiben impudicamente muchos internautas. En el otro extremo, el fenómeno opuesto de la ultrareescritura se debe fundamentalmente a la facilidad de los mecanismos de corrección. El escritor relee mientras rescribe, utiliza funciones como Busca y Reemplaza, borra sin manías y sólo llena papeleras virtuales. La literatura asistida por ordenador podría llevarnos hacia un nuevo barroco. Pero no es nada seguro que esto vaya a suceder. A pesar de todo, escribir en tiempos de Internet es, sencillamente, escribir”.

En este mismo texto significativo, Serra explica cómo en Internet han aparecido unos curiosos entornos virtuales basados exclusivamente en el lenguaje verbal llamados "moo":

“Los "moo" toman su denominación onomatopéyica de las primera líneas del "Retrato del artista adolescente" de Joyce. Tal como los "yahoos" de Swift han prestado su nombre a uno de los buscadores más efectivos de la Red o los "quarks" de "Finnegan's Wake" han irrumpido en los manuales de física, la loca "moocow" (múvaca) joyceana es el étimo de estos entornos hipertextuales. El usuario transita por un texto discontinuo

que su lectura puede modificar. Uno de los emisores del texto inicial ejerce de mantenedor (oficio parecido al de webmaster o sysop), ocupándose de mantener activo el servidor. Los tecnoeufóricos hablan de cambio de paradigma. David Bolter, por ejemplo, asegura que "los escritores requerirán de un nuevo concepto de estructura para producir textos de ficción en este espacio electrónico cambiante. En vez de una estructura cerrada y unitaria, deberán aprender a concebir su texto como una estructura de estructuras posible. El escritor debe practicar un tipo de escritura de segundo orden, crear líneas coherentes que el lector pueda descubrir sin cerrar las posibilidades de manera prematura o arbitraria. La escritura de este segundo orden será la contribución especial del medio electrónico a la historia de la literatura".

Con respecto a cómo debe variar la redacción en Internet en comparación con los medios tradicionales Ramón Salaverría, director del Laboratorio de Comunicación Multimedia de la Universidad de Navarra, opina lo siguiente:

“Hay problemas de diverso tipo:

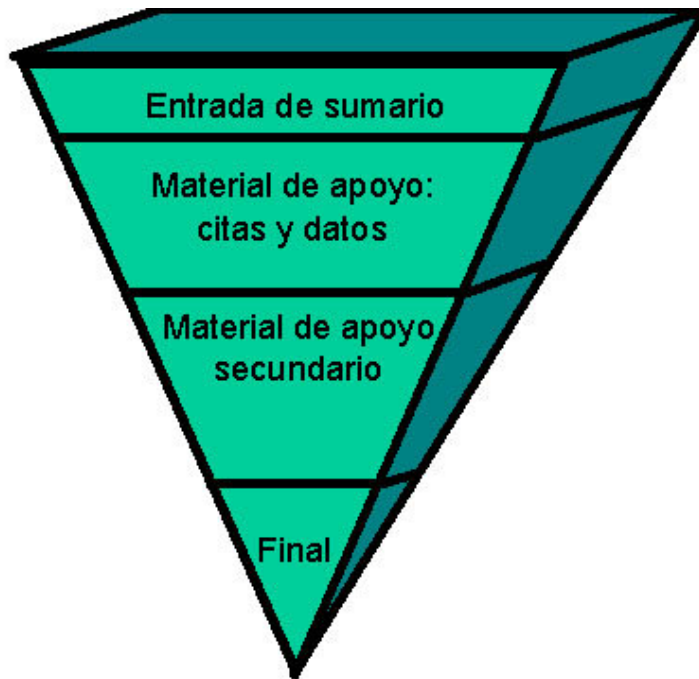
- Para empezar, de extensión: los artículos pensados para la edición en papel cuando se trasladan sin más al ordenador resultan por lo general demasiado extensos y obligan al lector a la engorrosa tarea de avanzar en el texto a lo largo de varias pantallas; además, no le permiten conocer de un vistazo la extensión total del texto, por lo que obligan a una lectura desorientada, sin ofrecer una idea aproximada de la cantidad de información que se proporciona. Esto implica que el lector, sin advertirlo, puede dejar de leer detalles sustanciales del texto informativo porque supone que ya se le ha ofrecido la información relevante en los primeros párrafos y apenas le resta texto por leer.
- También resultan comunes los problemas referidos a la datación de los textos. En los periódicos de papel existe una tendencia a suprimir la referencia temporal en el cuerpo del texto de las noticias pues los redactores presuponen que, aun si no se aporta ese dato, el lector inferirá que el acontecimiento del que se le informa ocurrió la víspera. Sin embargo, cuando esos textos pasan a publicarse en la edición digital esta práctica acarrea confusión. Y es que el concepto del "hoy" se difumina en un entorno como Internet que puede ser consultado

simultáneamente desde los veinticuatro husos horarios del mundo, y a cuyos contenidos, gracias a los archivos en línea, puede accederse varios días, meses o años después de la publicación original. La necesidad de una exacta referencia temporal dentro del texto se hace, de hecho, mucho más obligada en la medida en que los diarios digitales tienden hacia la actualización constante de las informaciones; las referencias temporales generales —"hoy", "ayer", "esta mañana"— se tornan insuficientes cuando las noticias se actualizan con ritmos de horas o, incluso, de minutos.

- Por citar un problema más, mencionaremos la falta de un criterio claro en el uso de los enlaces hipertextuales insertos en el texto. Es un problema que recuerda al mito de la caja de Pandora: cuando, en mitad de la lectura de una noticia, el lector topa con uno de estos enlaces, desconoce qué tipo de información se desplegará en su pantalla si pulsa en él; el enlace puede dar paso a otro artículo de la misma edición, a textos documentales del archivo, a páginas ajenas a la publicación o, quién sabe, incluso a publicidad indeseada. En cualquier caso, supone una incógnita para el lector que no parece acorde con la claridad informativa que debe presidir todo producto periodístico⁷¹.

Salaverría se plantea también en el artículo titulado “De la pirámide invertida al hipertexto”, las posibles inconvenientes de seguir conservando la estructura tradicional de la noticia, basada en el formato de Pirámide invertida.

⁷¹ Salaverría, Ramón: “De la pirámide invertida al hipertexto”. Accesible en: <http://www.unav.es/fcom/mmlab/mmlab/investig/piram.htm>



Su principal conclusión, es que seguir redactando de un modo tradicional supondría prescindir de un nuevo recurso clave en el periodismo: el hipertexto. En este artículo Salaverría afirma lo siguiente:

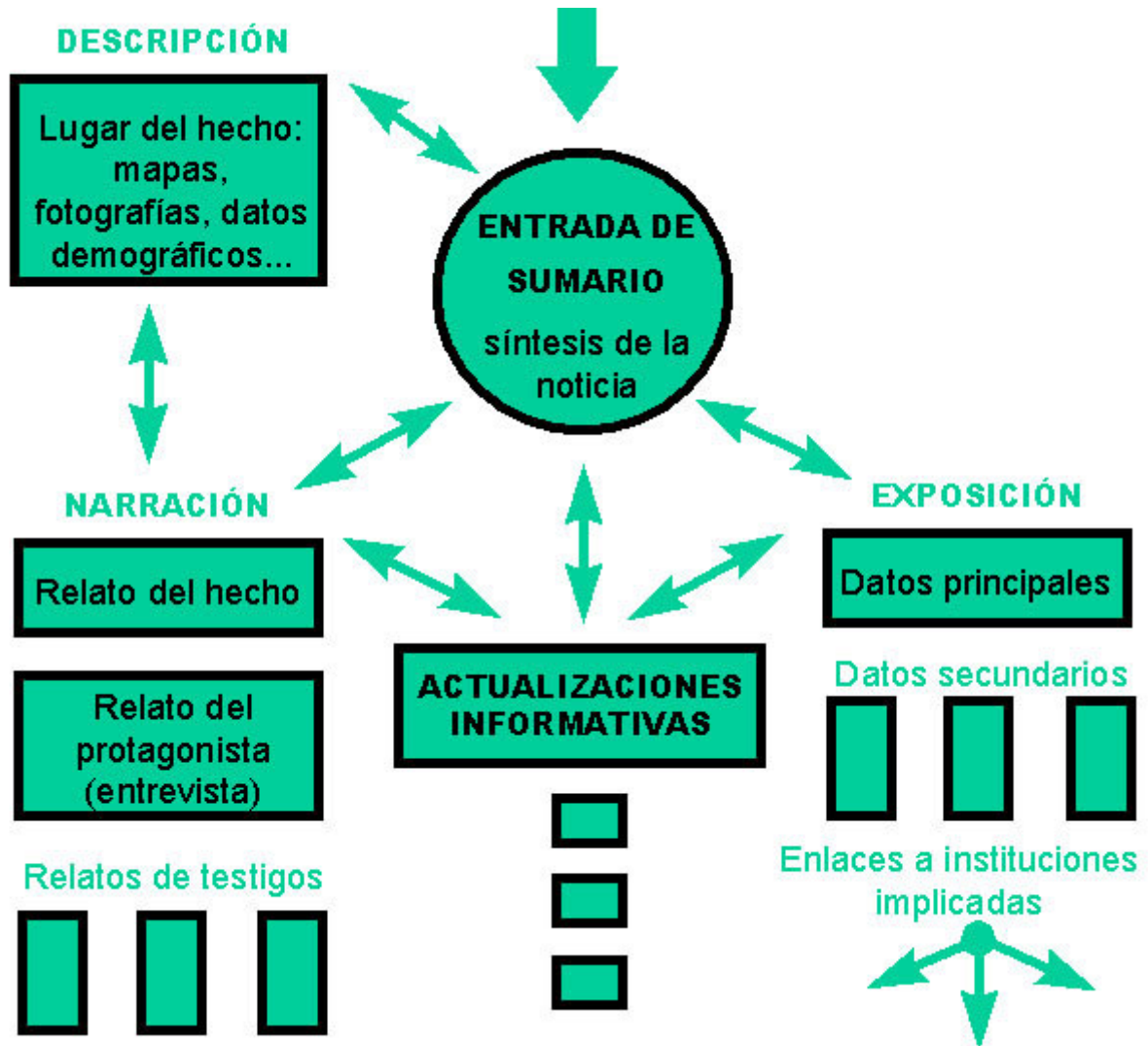
“Gracias al empleo adecuado del hipertexto como recurso redaccional, la noticia en el periódico digital puede superar una limitación y ganar en algo de lo que carece. Puede librarse por un lado de la redundancia y previsibilidad que acarrea la pirámide invertida —pues con ella la noticia se configura como una sucesiva amplificación de datos, sin lugar para la intriga—, y puede adquirir, por el contrario, el desarrollo informativo y estructural del que en buena medida adolece en la prensa de papel.

En virtud del hipertexto, por ejemplo, ya no es obligatorio incorporar en el cuerpo de la noticia pasajes documentales (por tales entendemos aquí desde una simple aclaración de unas siglas hasta, por ejemplo, pasajes biográficos de personas mencionadas en el texto o piezas contextualizadoras con los episodios pasados de la noticia que se comunica). Un enlace hipertextual a terceras páginas permite ampliar, aclarar o relacionar cualquier información, bien mediante nuevos textos o bien mediante recursos infográficos (mapas, recreaciones virtuales...) o audiovisuales (declaraciones, vídeos...). En definitiva, el hipertexto pone por primera vez en manos del lector —no del periodista—

la posibilidad de ampliar hasta donde desee la contextualización documental de cada información y, al mismo tiempo, le libera de leer pasajes documentales indeseados que ralentizan y oscurecen la lectura. En lo que afecta a los redactores, la asimilación de la escritura hipertextual exige una renovación en los modos de estructurar y escribir la información. La novedad principal que acarrea el hipertexto es que, a la hora de redactar la noticia, el periodista debe ser mucho más consciente de los elementos informativos que la conforman. Debe discernir mejor entre la información de última hora y la información documental contextualizadora; debe diferenciar entre la explicación de datos, la descripción de lugares y el relato de acontecimientos; debe identificar con qué soporte —textual, gráfico o sonoro— se informa mejor sobre cada uno de los aspectos de la noticia. En definitiva, la organización hipertextual de la información obliga al periodista a realizar un mayor esfuerzo de selección y jerarquía de los elementos de la información, tareas radicalmente periodísticas.”

Como se puede ver en el cuadro siguiente elaborado por Salaverría, el esquema de la noticia digital no cabría representarlo como una estructura cerrada, sino como un conjunto de elementos interconectados mediante el hipertexto y cuyas partes vendrían dictadas por el acontecimiento informativo comunicado en cada caso.

La célula informativa según Ramón Salaverría



Sin embargo, la aparentemente bien argumentada teoría de Salaverría se enfrenta a la elaborada por Jacob Nielsen quien afirma en su artículo "Inverted Pirámids in Cyberspace"⁷², que los periodistas, a la hora de escribir una noticia, suelen hacerlo en forma de pirámide invertida, comenzando por la conclusión, siguiendo por la

⁷² Nielsen, Jacob: *Inverted Pyramids in Cyberspace*. Junio 1996: <http://www.useit.com/alertbox/9606.html>

información más relevante y concluyendo con documentación en base a contextualizar dicha noticia. La escritura de la noticia en forma de pirámide invertida es útil para los periódicos porque los lectores pueden parar de leer el artículo en cualquier momento y aún así seguir enterándose de la parte fundamental de la noticia. Nielsen señala que en el caso de escribir para la Web, la pirámide invertida es aún más importante desde que se conoce, según bastantes estudios de los usuarios de Internet, que normalmente dichos usuarios no se desplazan en la página (es decir, no hacen *scroll*), por lo que muy frecuentemente dejarán la lectura en la parte superior del artículo. Sin embargo los lectores muy interesados sí que se desplazarán para leer la noticia completa, con todos sus detalles, y terminar así la pirámide invertida. Además de apoyar la estructura de la pirámide invertida, este autor hace también hincapié en la función fundamental del hipertexto a la hora de escribir en Internet. Para él son necesarios los enlaces para dar a los usuarios la posibilidad de acceder a las fuentes primarias de la noticia, o simplemente para una mayor contextualización o ampliación de información de la misma.

Aunque las nuevas formas de trabajo del periodista digital incluidos sus cambios de hábito de escritura para adaptarse al nuevo medio, serán tratados en el capítulo pertinente dedicado a los retos del periodista digital, merece la pena destacar la referencia a los modelos de redacción online que María José Pérez Luque establece , en su libro “El periodismo y las nuevas tecnologías⁷³”. Son las siguientes:

- a) Las redacciones “automatizadas”.
- b) Las redacciones online.

a) Respecto a las redacciones automatizadas, Pérez Luque las define del modo siguiente:

⁷³ Pérez-Luque, María José: *El periodismo y las nuevas tecnologías*. Newbook Ediciones. 1998. pp: 157-164

“Modelo organizativo típico de los inicios de los medios online. Consiste en organizar todo el trabajo de creación de contenidos online en torno a una estructura totalmente automatizada que facilita la conversión de los textos impresos a la versión online”.

Sus rasgos característicos son los siguientes:

- Proceso meramente tecnológico que consiste en la transformación de los contenidos de la versión impresa a lenguaje html mediante un programa informático.
- Se limitan a volcar los contenidos de la versión impresa en la edición online.
- El equipo de trabajo está formado principalmente por técnicos. En algunos casos, algún antiguo miembro de la redacción tradicional se encarga de supervisar que no se hayan producido errores ortográficos, que los enlaces funcionen correctamente, etc.
- Ni siquiera existe un *staff* esditorial online.
- A corto plazo cuentan con lectores fieles (básicamente, porque son compradores habituales del periódico de papel), sin embargo a largo plazo y con su estructura actual están destinados a desaparecer.

Pérez Luque establece una serie de carencias de este método:

- No es posible elaborar contenidos originales, y que por tanto diferencien a ambos productos, por la ausencia de periodistas online.
- No se aprovechan las posibilidades que ofrece el medio de actualización, personalización, contextualización, funcionalidad, etc.
- No ofrecen ningún valor añadido a sus lectores.

Con respecto a este método aquí citado cabe decir que, en mi opinión, los medios de comunicación que lo han desarrollado han sido con el conocimiento de que iba a ser temporal. Por este motivo pienso que caerá en desuso a medida que las características definitorias del periodismo electrónico se vayan afianzando.

b)Respecto a las redacciones online, Pérez Luque las define del modo siguiente:

“Estructura organizativa que caracteriza a los servicios informativos online. Se compone de un equipo formado por periodistas y apoyados por técnicos y/o programadores, cuyo principal objetivo es aprovechar las posibilidades que ofrece el nuevo medio con el fin de ofrecer un producto verdaderamente original y único”.

Según Pérez Luque sus rasgos característicos son los siguientes:

- Surge en el momento en que las empresas de información online se percatan de que para aprovechar las posibilidades de Internet es necesario contar con una redacción online propiamente dicha.
- Actualmente, no existe una modalidad única y universal para crear redacciones online. Los tipos de redacciones online son tan diversos como los medios que pueden consultarse en la Red:

-Variantes en función del grado de implicación

- a) Dedicados exclusivamente a la producción de noticias online.
- b) Dedicados parcialmente a la edición online.
- c) Presencia de un *staff* editorial online que trabaja en coordinación con el *staff* de la versión impresa. Los redactores trabajan indistintamente para una u otra área.

-Variantes en función del número de personas que la integren. Puede oscilar entre una persona hasta el centenar de redactores. La decisión depende fundamentalmente de criterios económicos y no por tanto de la calidad periodística que se desea que tenga el producto ofrecido.

- Los comunicadores online se encargan de elaborar contenidos originales o cuando menos, intentan mejorar las noticias publicadas en la versión impresa aprovechando las características del medio (fundamentalmente la contextualización, las posibilidades de hipertexto, y, en menor medida, la actualización).

Las carencias que se registran en este caso son las siguientes:

- No se observa una integración de conocimientos con el fin de realizar verdadero periodismo multimedia. Para superar esta carencia es necesario que accedan expertos de otros campos.
- Se sigue trabajando individualmente, de ahí que no se creen noticias auténticamente integradas. Es preciso impulsar modos de trabajo que impulsen la colaboración entre todas las personas que trabajan en la redacción y que permitan realizar creaciones conjuntas, inventar entre todos los expertos de las diferentes áreas. Se trata de lograr un modelo más parecido a la organización de los medios audiovisuales.
- Continúa vigente el sistema de un único cierre a última hora del día. Por lo tanto, es necesario establecer la cadencia de actualizaciones más adecuada al ritmo informativo.
- No se ha impulsado suficientemente el establecer y mantener el contacto con los lectores, de forma que se está perdiendo una ocasión única de adecuar los contenidos a las demandas de los usuarios.

Como hemos especificado en el comentario del modelo anterior, estos modelos se mantienen en la actualidad, sin embargo no sin vistas a un prototipo ideal de redacción en Internet que permita aprovechar todas las posibilidades que ofrece este nuevo medio.

2.6 INCONVENIENTES ACTUALES DEL PERIODISMO ELECTRÓNICO

Entre todas las ventajas que ofrece Internet para el periodista cabe tratar también ciertos problemas que el profesional de la información tiene que afrontar⁷⁴.

⁷⁴ Díaz Noci, Javier: *El nuevo periódico electrónico: redefinición del mensaje periodístico tradicional como producto interactivo y multimedia*. Mayo 2000.

- a) La transmisión de imágenes (sobre todo en movimiento) y sonidos, es mucho más lenta que la del mero texto. Hasta que las anunciadas autopistas de la información no sean una realidad, entrado el nuevo milenio, la línea telefónica parece una vía muy insuficiente para transmitir el cada vez mayor caudal informativo.
- b) La legibilidad de las pantallas y su definición es aún muy desigual si la comparamos con el tradicional soporte de papel. La portabilidad de un periódico de papel es otra de las ventajas que las grandes pantallas no pueden igualar, al menos hasta que Knight Ridder y el Massachusetts Institute of Technology den a conocer su “tableta electrónica”, aún en fase experimental⁷⁵. Tampoco ha aparecido, la anunciada “news box” del británico *Evening Standard*⁷⁶. La solución de transmitir el periódico no a un ordenador personal, sino a un PDA (“personal assistant”, una máquina de tamaño y peso pequeños, a caballo entre el ordenador y la agenda), que ya se ha intentado, conlleva el problema de que la pantalla es pequeña, en blanco y negro, de cristal líquido, y por tanto las imágenes son de baja calidad. Por otro lado, aún queda por dilucidar si en el futuro el ordenador y la televisión sobrevivirán como hasta ahora, como dos electrodomésticos independientes, o bien uno se impondrá al otro, o, en una última alternativa, se desarrollará un nuevo aparato que incorpore características y prestaciones de ambos.
- c) Sistemas operativos diversos, que, aunque hace poco tiempo parecían tender a la unificación, continúan sin llegar a una convergencia real: en el mundo PC (liderado por IBM), al tradicional DOS se ha sumado Windows, cuya nueva versión es ya un sistema operativo en sí, mientras que la propia IBM intenta por todos los medios introducir su sistema orientado a objetos, el OS/2; en el mundo Apple Macintosh, el sistema 7; y hay que sumar los sistemas propios de grandes ordenadores y otros, como Unis, cuyo potencial conviene no desdeñar.

⁷⁵ *El Mundo*, 22 Octubre 1993; declaraciones del presidente del proyecto sobre periódicos del futuro en el MIT, Jerome Rubin.

⁷⁶ Dickman, Steven: *Mañana, periódicos electrónicos*. *El País*, suplemento Futuro, 20 de octubre de 1993.

- d) Diversas redes de comunicación (Internet, America Online, Prodigy), que dificultan el acceso a los diversos medios de comunicación.
- e) Indefinición del modelo de periódico electrónico: junto con versiones *online* del periódico tradicional, hallamos formas que, como en el origen del periodismo tradicional, pertenecen a la prehistoria del mismo (almanaques y anuarios, que están presentes en el inicio de la historia del periodismo, se hallan también en la incipiente andadura del periodismo electrónico, sobre todo en versión CD-ROM).

2.7. PERIÓDICOS EN INTERNET EN ESPAÑA ¿GRATUITOS O DE PAGO?

En un intento de hacer frente a las crecientes pérdidas económicas, y una vez vistas las pioneras experiencias internacionales, los principales medios de comunicación en España con presencia en la Red han decidido renunciar a la gratuidad de la información en Internet y apostar por distintos métodos de financiación que impulsen los débiles ingresos publicitarios.

El hecho de que tan sólo un 20% de la población española tenga acceso a Internet y que muchos usuarios aún teman el uso de la tarjeta de crédito para pagar por un producto en la Red no ha amilanado a las principales empresas periodísticas españolas que han visto en el pago por contenido una opción para intentar salir de unas pérdidas cada vez más acusadas.

En este sentido, el director de El Mundo, Pedro J Ramírez y el responsable del Grupo Correo, Miguel Larrea, se muestran partidarios de combinar los contenidos gratuitos con los de pago para preservar la continuidad y rentabilidad del negocio. Pedro J. Ramírez advierte que “Internet se va a consolidar como un medio de comunicación masivo y muy rentable si se produce una bifurcación entre contenidos de acceso gratuito y de pago obligatorio, en especial los que aportan valor añadido a la información diaria, como la consulta de archivos”. Larrea opina que lo que primará será la calidad de la

información “la cultura de la barra libre en Internet se tiene que acabar para que no deteriore la calidad de los contenidos.

En esta línea, el 15 de octubre de 2001 el diario El Mundo le puso precio a su edición digital de la tarde. La oferta de suscripción ofrecía tres opciones: anual, semestral o bimestral y conllevaba meses de regalo según la duración. Por 50 euros anuales el internauta recibe la edición vespertina, de lunes a viernes, a partir de las 18.30. Un año después este diario anunciaba que había decidido comenzar a cobrar por contenidos de la versión impresa del periódico “en consonancia con la tendencia seguida por los más prestigiosos medios internacionales y con intención de mejorar aún más sus contenidos en Internet”. El diario ofrece diversas ofertas. Una de las propuestas permite descartarse la edición nacional en formato PDF a partir de media noche con su aspecto tipográfico tradicional.

La versión del periódico económico Expansión en Internet también ha optado por el cobro por contenidos. Expansión Premium ofrece acceso, sin restricciones, a todas las noticias del diario y a una sección de análisis reforzada, un innovador gestor de carteras, cotizaciones en tiempo real y un servicio de documentación mejorado. La suscripción mensual cuesta 16,95 euros. Asimismo, la edición en Internet de La Gaceta de los Negocios es íntegra de pago y en su página de inicio únicamente ofrece los titulares y los sumarios de algunas informaciones.

Medios de referencia

Mientras que en España la prensa digital está comenzando a dar sus primeros pasos en el ya conocido como ‘pay per read’ (pago por leer), diarios estadounidenses como el Wall Street Journal ya cobran por sus noticias, análisis y reportajes. A pesar de que este diario económico llegó a perder el 90% de sus lectores en Internet cuando instauró el sistema de cuotas, esta publicación digital está considerada como un éxito de los servicios de pago por suscripción. Además, está sirviendo de referencia de muchos otros medios digitales que también han llegado a la conclusión de que la gratuidad de la información no es el camino a seguir si se quiere salir de los ‘números rojos’. El

wsj.com ofrece a sus más de 650.000 abonados noticias de calidad, datos en tiempo real de todas las bolsas, gráficos, datos bursátiles de años anteriores y acceso a una base de datos con años de cobertura de la economía estadounidense y mundial. La sencillez en la navegación es otro de sus puntos a favor. El precio por suscripción anual es de 59 dólares.

Las mayores compañías periodísticas de Estados Unidos sufren, o han sufrido una grave crisis en sus respectivas divisiones digitales. El diario The New York Times tuvo que replantearse a principios de 2001 su estrategia de expansión en Internet. New York Times Co, editora de The New York Times y The Boston Globe constató que las expectativas de crecimiento no se habían visto confirmadas. Este hecho dio lugar al despido de 69 personas de una plantilla de 400 dedicadas exclusivamente a la sección digital. De este modo se pretendía sacar de la crisis a un diario digital que arrastraba unas pérdidas de 3.000 millones de pesetas. Una de sus iniciativas de financiación, fue la de ofrecer al internauta la posibilidad de imprimir la versión impresa del The New York Times desde el ordenador en formato PDF. De este modo se obtiene un periódico más barato. Además se le ofrece a este lector la posibilidad de utilizar un buscador para hallar los temas o términos que más le interesen. Esta misma iniciativa ha seguido el diario francés Le Monde para satisfacer los intereses de aquellos lectores habituados al diseño del soporte papel.

Por su parte, la edición en Internet de Financial Times también ha comenzado a cobrar por contenidos mediante suscripción. FT.com también obtiene ingresos sustanciales a partir de su gran base de usuarios del servicio gratuito. Su objetivo es maximizar su base de lectores para explotar el crecimiento de los ingresos publicitarios cuando se recupere el mercado.

Asimismo, diarios estadounidenses como el Star Tribune de Minneapolis elaboran páginas de Internet para empresas locales; The Washington Post ha creado un centro comercial online dentro de su dominio a través de alianzas con distintas empresas dedicadas al comercio electrónico; y otros periódicos digitales venden servicios de acceso a la Red. Cualquier cosa es válida si con ello se consigue obtener ingresos.

Métodos de financiación

Según una encuesta de Online Publishers Association, que agrupa a unos 20 editores de prensa, como New York Times Digital y The Wall Street Journal Online, cada vez hay un número cada vez mayor de usuarios que muestra su buena disposición a pagar por los contenidos y servicios online, lo cual sugiere que poco a poco se está cambiando la percepción del "todo gratis" al hablar de Internet. Sin embargo, este mismo estudio subraya que aún son muy pocos los medios que se benefician del pago por contenidos y que la publicidad sigue siendo la principal fuente de ingresos. Para obtener financiación por publicidad los medios digitales utilizan dos sistemas: la tarifa plana y la tarifa de coste por impresión (CPI). La tarifa plana es similar a las tarifas publicitarias de los periódicos tradicionales y un banner (anuncio o promoción virtual) tiene un coste fijo mensual. Suele ser la más usada ya que permite al medio obtener mayores ingresos con una limitada difusión. Cuando el medio se consolida y el número de páginas aumenta, el medio puede incrementar sus ingresos cambiando sus tarifas al tipo CPI. Ésta fue la estrategia seguida por El País Digital. El inconveniente de los banners, según un sondeo elaborado en mayo de 2001, es que el 75% de los internautas no repara en ellos y aquellos que lo hacen dicen no encontrar información de utilidad.

Un estudio desarrollado en 1999 por Editor & Publisher sobre 700 medios cibernéticos de todo el mundo, reveló que una cuarta parte de los diarios de Internet había conseguido salir de la crisis gracias a la combinación de diversas vías de financiación, fundamentalmente publicidad online y comercio electrónico de productos y servicios. A esta misma conclusión, basada en la diversificación de ingresos, llegaron el pasado 18 de julio en Valencia, responsables de diversos medios entre los que se encontraban ElPaís.es, ElMundo.es y Marca.com. Entre éstos, El director de Marca.com expuso la posibilidad de combinar publicidad, venta de contenidos a terceros, comercialización de información a través de teléfonos móviles y comercio electrónico.

Otros periódicos como el Financial Times o el Segunda Mano han optado, entre sus métodos de financiación, por la exigencia para acceder a sus contenidos de un registro previo, con lo que van conformando una base de datos con perfiles de usuarios muy beneficiosa para el posible anunciante.

Tanto diversos estudios, como la evidencia, señalan que mientras haya sitios gratuitos con una calidad semejante en sus contenidos a otros de pago, los usuarios optarán por los primeros. Es por esto, por lo que la calidad de la información y los contenidos de valor añadido que marquen una diferencia apreciable, serán las claves que impulsen al usuario a cambiar su visión, hasta ahora arraigada, del ‘todo gratis’ en la Red.

CAPITULO III: EL PAPEL DEL PERIODISTA DIGITAL EN LA SOCIEDAD ACTUAL

*Para nadie es un cuento de ciencia ficción eso que ahora llaman "periodismo electrónico". En Estados Unidos y otros países del mundo ya es común hablar del "proveedor de contenido" en sustitución del tradicional reportero. Las calles que antes se "pateaban" en búsqueda de la noticia se han convertido en "estaciones abastecedoras de datos" ubicadas a lo largo de la superautopista de la información. Pero ¿cuáles son los desafíos de ese nuevo periodista digital?*⁷⁷

3.1 INTRODUCCION

El auge de los medios tradicionales en Internet ha desatado todo un cúmulo de reflexiones sobre cómo presentar de manera más efectiva e idónea la información periodística. Se trata de averiguar la forma de hacer "más atractiva" *la interfaz* gráfica, cómo mantener la atención del lector-internauta y otras especificaciones referidas a aspectos meramente técnicos. En este capítulo nos centraremos en la persona encargada de recopilar y proveer la información, es decir, de las nuevas funciones del periodista.

¿Cómo se puede definir al periodista digital cuando todavía no se puede especificar con exactitud el concepto de periodista en sí mismo?. Está claro que es una pregunta difícil de responder. No nos vamos a aventurar por tanto en ofrecer tal definición sino en desentrañar sus funciones, diferencias y retos con respecto al periodista tradicional.

En un medio como Internet en el que todo aquel que posea un ordenador conectado a una línea telefónica tiene acceso, cualquiera se puede convertir en informador potencial sin que nadie se lo impida. Por lo tanto, está claro que las funciones del periodista online van a ser diferentes a las que tendría en los medios tradicionales.

⁷⁷ Núñez Aldasoro, Antonio J. : es tesista de la carrera de Comunicación Social de la Universidad Central de Venezuela.

Ha habido muchos autores que se han aventurado en definir los perfiles profesionales del periodista digital. En 1995 Vicent Partal veía al nuevo periodista como un guardia urbano que dirige el tráfico de la información, dando preferencia a la circulación de determinados carriles, siendo un punto central en la ordenación del tráfico en las autopistas de la información.

Dos años más tarde, Luis Ángel Fernández Hermana y Alfons Cornella propusieron dos visiones diferentes. Fernández Hermana comparó al nuevo periodista con un cartógrafo, afirmando que en el medio digital las carreteras y autopistas están por construir, y que posiblemente nunca llegarían a estar establecidas. Para este autor, la función del periodista digital es la siguiente⁷⁸ : “ Consiste en darle sentido a un mapa informativo desordenado que pide un diseño comprensible para la intervención del internauta. Los accidentes geográficos –el entramado en el caso del guardia- deben ocupar el lugar que les corresponde a partir de la acción de este nuevo operador de la Red. Para ello, como sucede en el caso del explorador y del cartógrafo, debe servirse de las herramientas apropiadas en cada caso y en función de los objetivos, independientemente de las diferencias propias entre publicaciones por razón de sus finalidades. Son éstas las que determinarán las proporciones entre ayudas a la navegación, contenidos audiovisuales, relaciones entre sistemas informativos, creación de comunidades de interés, calidad de su información o el uso que de ella hagan los usuarios (...). El cartógrafo de la Red requerirá aunar, por lo menos, conocimientos y experiencias que proceden de diferentes campos del conocimiento, como la ingeniería, las telecomunicaciones, el diseño, los flujos de la información y la comunicación, las necesidades coyunturales de los usuarios (demanda y oferta) y la interacción con contenidos de diverso origen cuya síntesis generen productos nuevos”.

Asimismo, Alfons Cornella⁷⁹ nos definió al periodista digital como “infonomista”, un profesional que estudiaba las estructuras comunicacionales y diseñaba economías y flujos óptimos de información.

⁷⁸ Fernández Hermana, L.A. *De exploradores y cartógrafos*. Editorial de En.red.ando (1997). <http://enredando.com/cas/cgi-bin/editorial/plantilla.pl?ident=252>

⁷⁹ Cornella, A. *Eureka: es la infonomía, y somos infonomistas*. Artículo de Extra!Net (1997) <http://www.extra-net.net/articulos/en971108.htm>

En países donde el periodismo electrónico es una fuerte y estable fuente de trabajo, la imagen del reportero-periodista ha dado paso al denominado "proveedor de contenido", una especie de Blade Runner de la información que toma su pauta, busca los datos, hace las entrevistas, toma las fotografías y asiste a los eventos, todo a través de Internet⁸⁰.

Quim Gil en su artículo *Diseñando al periodista digital*⁸¹, señala una serie de parámetros de trabajo del periodista digital que definen muy bien las funciones del profesional de la información al afrontar este nuevo medio. Esta clasificación servirá de base para el desarrollo del capítulo. Son los siguientes:

Rutinas con respecto a las fuentes

Rutinas con respecto a los destinatarios de la información

Rutinas con respecto al medio

Rutinas con respecto a los contenidos

3.2 EL PERIODISTA DIGITAL FRENTE A LAS FUENTES DE INFORMACIÓN

En el periodismo tradicional el periodista tenía que ir en busca de la información. En pocas ocasiones las noticias venían a él. Sin embargo, en este nuevo medio de comunicación bajo el soporte de Internet, el profesional de la información se encuentra justo con el problema opuesto. En cuanto se conecta a la Red tiene a su disposición un mar de información que debe seleccionar y verificar. Debido a esta gran cantidad de información de la que se dispone, el valor del periodista digital se encuentra en su propia valoración de cada fuente.

El uso del correo electrónico y de otras técnicas innovadoras de comunicación en tiempo real, permite entrar en contacto con fuentes de información (y con el propio receptor que, a su vez, también se puede convertir en fuente de información), a nivel

⁸⁰ Núñez Aldasoro, Antonio J: *Los retos del periodista digital*. <http://www.saladeprensa.org>

⁸¹ Gil, Quim: *Diseñando al periodista digital*: <http://andalucia24horas.com/caracola/doca003.htm>

mundial es una de las grandes ventajas que este nuevo medio nos ofrece. Para ilustrar este hecho, cabe destacar el sistema ProfNet (ProfNet)⁸², elaborado en 1993 por la Universidad de Nueva York, que ofrece la posibilidad de tener línea directa con los mejores científicos y académicos del mundo. Actualmente, este sistema cuenta con 750 expertos mundiales que cubren todos los campos y que serían de gran ayuda al periodista en busca de fuentes de información. El informador tan sólo tiene que contactar con ProfNet por cualquier medio electrónico o convencional, describir su objetivo y el tipo de fuente que necesita. Su petición es enviada inmediatamente a toda una red de expertos y en un breve plazo se recibe la respuesta completamente documentada. Es posible además que el periodista entreviste a algún experto en concreto para ampliar el tema. El creador de ProfNet, Dan Forbush, afirma que se reciben unas 200 peticiones semanales⁸³. Prácticamente la mitad de ellas proceden de diarios y agencias informativas; una cuarta parte de emisoras de radio y televisión y la otra cuarta parte, de revistas semanales y mensuales. Sirve información únicamente a medios periodísticos y emplea mecanismos de comprobación como pedir datos personales, el medio para el que trabaja, hora tope para el cierre del texto...etc. Asimismo, asegura la confidencialidad de las informaciones que se intercambian entre la organización y el periodista. La prioridad de atención va en función del orden en que llegan las consultas.

A la hora de buscar información en Internet, los periodistas cuentan, además, con unas guías que facilitan en gran medida su trabajo. Estas guías, entre las que podemos encontrar la “Guía de Internet para periodistas” elaborada por el Laboratorio de Medios de la Universidad de Navarra⁸⁴” o la “Guía de recursos⁸⁵” para periodistas del diario *El País*, consisten en directorios de recursos útiles para usar Internet como herramienta periodística.

Dado el propio carácter de la Red, en el todo el mundo puede publicar, cada vez es más importante que el periodista digital compruebe que las fuentes son en realidad

⁸² ProfNet Network (ProfNet): <http://profnet.com>

⁸³ Armañanzas, Emy: *El periodismo electrónico, resultado de una globalidad económica y social*. Bilbao 1996.

⁸⁴ Guía de Internet para periodistas: <http://www.unav.es/fcom/guia/>

⁸⁵ Guía de recursos: <http://www.elpais.es/p/d/especial/recursos/buscadores.htm>

quiénes dicen ser. Su selección de fuentes será una garantía más para el receptor de la información.

Sin embargo, el entorno global que la Red ofrece, no debe cegar al periodista de las fuentes tradicionales de la información. La recopilación de información no digital y su posterior digitalización e introducción en la Red puede acabar siendo una de las características de los periodistas digitales.

Otra de las funciones del profesional de la información será ofrecer a su audiencia el acceso a todas las fuentes originales que muestren públicamente sus contenidos, permitiéndoles de este modo, que amplíen la información si lo desean. En este nuevo medio no hay limitaciones de tiempo y/o espacio que nos lo impidan. De este mismo modo, y teniendo en cuenta la personalización que permite el medio, el periodista digital podrá dar al receptor la oportunidad de realizar consultas a las fuentes a través de él.

Este apartado será desarrollado ampliamente en el capítulo: "Las fuentes de información del periodista digital".

3.3 EL PERIODISTA DIGITAL FRENTE AL RECEPTOR DE LA INFORMACIÓN

"El transmisor determina todo y el receptor simplemente toma lo que recibe". Para Nicholas Negroponte, director del Laboratorio de Medios del Instituto de Tecnología de Massachusetts (MIT), este paradigma comunicacional de los medios masivos tradicionales es el primero que debe modificarse con el advenimiento de las Nuevas Tecnologías.

En los medios tradicionales es el periodista el que determina lo que el lector, oyente o televidente tiene que ver, pero, gracias a la interactividad que permite Internet, el poder de decisión de cuál es el producto informativo pasa a manos del lector. Ya no es el periodista el que determina cuál es el orden en importancia del producto informativo,

sino que es el usuario de la información quien decide cuáles son en cada momento los elementos informativos que más le interesan.

Además, las características de este nuevo medio ofrecen la posibilidad al periodista de personalizar contenidos dirigidos a individuos o colectivos, satisfaciendo de esta manera sus verdaderas necesidades de información. De esta forma, y en orden a amortizar su publicación, el informador podría ofrecer una primera versión básica de la noticia o documento de forma gratuita, y otra más amplia que le aportara ingresos económicos.

Hay que tener en cuenta ante todo, que los clientes del informador son ahora también sus fuentes. Estos usuarios tienen la posibilidad de informar al medio de una forma fácil y efectiva de noticias que ocurren a nivel local, difíciles de saber por los periodistas.

El profesional de la información en Internet, se dedica a la identificación de fuentes, la extracción de información, el procesamiento de contenidos y su publicación. Pero también debe dedicar buena parte de su tiempo a interactuar con sus lectores y ver lo que éstos le pueden aportar. Esta interactividad que permite la Red produce que el público usuario comience a considerarse miembro de la cibercomunidad que se crea en torno al servicio informativo *online*⁸⁶, promovida fundamentalmente, no por pertenencia al mismo ámbito geográfico, sino por un mismo ámbito de interés.

El usuario que se conecta a Internet, busca ante todo interactividad. Compartir sus experiencias y preocupaciones. Lo que menos importa es dónde estén esas personas físicamente. Por este motivo, son los medios digitales que contienen información personalizada, juegos interactivos, foros de discusión, etc, los que más éxito tienen entre los usuarios. De esta forma, estos usuarios dejan de ser un colectivo anónimo para pasar a ser los destinatarios de una información elaborada según sus gustos e intereses, es decir, un público completamente activo en el proceso de la información.

Asimismo, herramientas como el correo electrónico, permiten al receptor ponerse en contacto directo con el redactor de una determinada información para discutir sobre la noticia, o simplemente para sugerirle algún tema para un artículo. De este mismo modo, el periodista puede ponerse en contacto con los usuarios para que

⁸⁶ Pérez-Luque, María José: "El reto de crear noticias online".
<http://www.ucm.es/info/multidoc/revista/cuad6-7/noticias.htm>

éstos envíen preguntas a un futuro entrevistado por el medio. Este hecho es de gran interés para el medio, ya que permite un mayor conocimiento del público receptor, ajustar la oferta y potenciar los argumentos de venta de los espacios publicitarios⁸⁷ al disponer de información mucho más contrastada de los intereses de la audiencia y de su perfil a la hora de definir sus gustos por un determinado tipo de producto o servicio.

3.4 EL PERIODISTA DIGITAL FRENTE AL NUEVO MEDIO

A mediados de 1998 se reunió en Nueva York un seminario organizado por Editor & Publisher para tratar las consecuencias para la prensa de la aparición de los nuevos medios electrónicos. Las principales conclusiones que se sacaron entonces pueden resumirse en las siguientes⁸⁸:

- Los periódicos no sólo no van a desaparecer, sino que seguirán siendo el medio dominante en el siglo XXI; sin embargo, el entorno está cambiando a gran velocidad.
- Los ingresos publicitarios son amenazados por la aparición de nuevos medios y por la segmentación de la audiencia.
- Conquistar nuevos lectores es una prioridad absoluta; la vieja fórmula del 85% de publicidad y 15% de circulación ya no sirve.
- Se acentúan los cambios en la estructura de propiedad de la prensa: un puñado de grupos “multimedia” controlará los periódicos.

El periodista digital es considerado “más que un redactor”. Internet es un medio de comunicación “multimedia”, es decir, tiene la capacidad de integrar texto, fotografía, vídeo y audio; por lo que los informadores deben concebir su trabajo desde una perspectiva multimedia, atendiendo a elementos de diseño del interfaz y de la estructura

⁸⁷ Pérez-Luque, M.J. y Perea, M: “El actual Periodismo Online”. *Mirandum Magazine*, IV, Ed. Madruv.á. Brasil, 1997. <http://www.hottopos.com/MirLibro/index.htm>

⁸⁸ Bullón, Piedad: *Periodistas, nuevo medio viejas reglas*. Revista *iWorld* N°12. Enero 1999

de contenidos. Por este motivo, Las nuevas generaciones de periodistas que trabajan en este nuevo medio, se caracterizan por una mayor versatilidad, ya que han de ser capaces de integrar textos, fotos, gráficos, audio y vídeo. Esto conlleva un trabajo en equipo y en red mucho más integrado del que se había conseguido en las redacciones de los medios tradicionales.

Según explica José Luis Dader⁸⁹, profesor de la asignatura “Periodismo de precisión” en la Facultad de Ciencias de la Información de la Universidad Complutense de Madrid, comparando la formación del periodista tradicional con el actual: “Hay un cierto mito de la tecnología. Por un lado, en las facultades de periodismo hay familiaridad con la tecnología, pero no se trabaja lo suficiente sobre la mentalidad del futuro profesional, que debería adquirirse a través de otro tipo de enseñanza. Por ejemplo: cómo redactar las noticias en un medio digital, cómo integrar gráficos y vínculos, son aspectos que no se han incorporado a la formación de los periodistas. La información electrónica es transversal. Los criterios que valen para un medio clásico, no siempre sirven para Internet, un medio en el que no basta con reproducir una información de texto que ha dado el periódico o la radio. ¿Por qué no abrir el acceso a bancos de fotos, a información temática exhaustiva? Todo lo cual exigiría concebir de otro modo la selección de los materiales y los criterios de relevancia. El aprendizaje de la tecnología es lo de menos; lo otro requiere un cambio de perspectiva, que a su vez exige un poso cultural que muchos alumnos no tienen ni se les ofrece”.

Asimismo, Ramón Salaverría⁹⁰, director del Laboratorio de Comunicación Multimedia (MMLab) de la Universidad de Navarra, ve el perfil del periodista del futuro de la siguiente manera: “Es un perfil ligado no sólo como hasta ahora a la creación de contenidos, sino, cada vez más, a la gestión de la información. Por eso, uno de los cambios principales consistirá probablemente en mudar del modelo tradicional de enseñanza, consistente en formar a los futuros periodistas en destrezas profesionales dirigidas a enfrentar la escasez de información, hacia otro modelo en el que se forme a los periodistas para enfrentarse a la superabundancia de la información. Con Internet, el

⁸⁹ Bullón, Piedad: *Periodistas, nuevo medio viejas reglas*. Revista iWorld Nº12. Enero 1999

⁹⁰ Salaverría, Ramón: *Criterios para la formación de periodistas en la era digital*:
<http://web8.cti.unav.es/fcom/mmlab/mmlab/investig/crite.htm>

problema ya no es encontrar información, sino distinguir entre lo significativo y lo irrelevante”.

Como se está observando, el proceso de comunicación en este nuevo medio exige a los periodistas una terminología específica, un nuevo lenguaje y nuevas reglas que, según Virginia Luzón Fernández, en su artículo: “Periodista digital, de MacLuhan a Negroponte⁹¹”, se pueden resumir en las siguientes:

- **Gatekeeping mixto:**

En la actualidad, los medios de comunicación electrónicos intentan ganarse la fidelidad de sus consumidores ofreciendo servicios individuales, como la opción de personalizar la pantalla del diario electrónico o de recibir sólo las noticias que interesan al receptor, es el llamado “periodismo a la carta”. Este tipo de información se basa en el denominado “gatekeeping mixto⁹²” donde el receptor selecciona los temas de la actualidad diaria que son de su interés, pero sigue siendo el periodista quien elabora la información: la redacción, la inclusión de imágenes o gráficas, incluso cortes de voz de los protagonistas.

- **Multiedición**

El periodista tiene la posibilidad de actualizar constantemente y en tiempo real la información. No hay un medio físico que impida la modificación de la noticia, cada nuevo dato puede ser ofrecido al cliente, el cual tendrá la ventaja de varias ediciones diarias de un mismo producto.

- **Doble feed-back**

Con el medio electrónico podemos hablar de un doble feed-back en la información. El periodista es a la vez emisor y receptor de la información disponible en la Red, está en el centro de dos diagramas diferenciados de la teoría de la comunicación de Jakobson, de manera que recibe y emite la información, actuando primero como

⁹¹ Luzón Fernández, Virginia: *Periodista digital: de MacLuhan a Negroponte*. Revista Latina de Comunicación Social nº 34. Octubre de 2000. La Laguna (Tenerife): <http://www.ull.es/publicaciones/latina/aa2000kj/w34oc/49s6luzon.htm>

⁹² Pérez Luque, M^o José y Perea, Maiden: *El actual periodismo online*. Laboratorio de Comunicación Multimedia de la Universidad de Navarra. Navarra 1999: <http://www.hottopos.com/MirLibro/index.htm>

receptor y luego como emisor, con la particularidad que los dos procesos de comunicación son simultáneos y tienen lugar a través del mismo canal de comunicación, la Red.

- **Redacción virtual**

El teletrabajo y la teleenseñanza tienen como particularidad común el hecho de que ni el trabajador ni el alumno necesitan desplazarse de su domicilio, el futuro de las redacciones de los medios de comunicación es virtual. No podremos hablar de un centro físico común a los diferentes periodistas de un diario electrónico, cada uno de ellos trabajará de forma independiente, poniendo en común ideas y proyectos mediante el correo electrónico o consejos de redacción basados en la videoconferencia.

- **Narrativa hipertextual**

Los medios electrónicos son no lineales, la navegación es un valor de cambio y una nueva forma de diferenciación respecto a los medios tradicionales. La narrativa hipertextual es la nueva forma de comunicar los hechos. El periodista digital ofrece al receptor de su mensaje los enlaces a las páginas de sus fuentes, a elementos de contextualización, a datos complementarios, a las páginas de los protagonistas, incluso ofrece la opción de que los lectores interesados puedan aportar datos.

Otra función fundamental del periodista no sólo será certificar las fuentes, sino también deberá certificar su propia identidad frente a sus lectores. Asimismo, hay que tener en cuenta la interactividad de la Red, por lo que el informador debe buena parte de su existencia a su eficiencia ante esta nueva “capacidad” de los lectores, inviable en el periodismo tradicional.

Además, la globalidad de la Red, comporta posibles audiencias multiculturales y multilingües. Cada medio digital, opta por trabajar con una audiencia determinada, y a veces se ve limitado por la barrera lingüística. Esto es todavía un reto, aunque ya existen diccionarios multimedia multilingües que ayudan al navegante a traducir por completo las páginas que encuentra en Internet a su idioma materno.

La Red tiene un componente de caos continuo. Hasta el momento, este caos se puede considerar de forma positiva ya que permite unos grados de libertad de expresión desconocidos hasta la fecha. El periodista digital debe gestionar este caos y ofrecer su criterio de ordenación a los lectores que se adentran en una Internet virtualmente sin límites. La ventaja de ser hasta ahora “ilimitada” debe ser aprovechada por el informador a la hora de contextualizar información, utilizar archivos y hacer público el acceso a bases de datos en estado puro, es decir, poner en marcha su vertiente documentalista, enseñando a los lectores a sacar la información que les sea de utilidad de estas bases de datos.

3.5 EL PERIODISTA DIGITAL FRENTE A LOS CONTENIDOS

Según los datos del Congreso de Internet' 99 organizado por la Asociación de Usuarios de Internet, en 1999 la Red contaba con aproximadamente unos 800 millones de páginas web; un año después esta cifra ya se había duplicado y continúa creciendo exponencialmente. Como hemos visto hasta ahora, Internet está considerado como un medio de comunicación más, y como tal, sería lógico pensar que sus contenidos deberían estar redactados o al menos publicados por profesionales de la información. Sin embargo el hecho de que este nuevo medio esté abierto a cualquier persona, a nivel mundial, que tenga un ordenador conectado a Internet, descarta las reglas del periodismo tradicional en este sentido. Internet es un medio, sí; pero no un medio como los demás, sino uno completamente abierto a cualquier profesional, estudiante o ama de casa, en definitiva, a cualquier persona.

Hoy por hoy cuando alguien quiere publicar algo en la Red no tiene que dar cuenta de la veracidad o la ética de los datos que va a difundir. Éste es el motivo por el que el usuario que busca algo en Internet puede sentir cierta desconfianza de lo allí publicado. Nadie le garantiza su autenticidad.

Como estamos viendo, no hace falta ser periodista para publicar en Internet. Las secciones de anuncios de empleo de los periódicos rebosan de ofertas de trabajo para informáticos e ingenieros de telecomunicaciones, sin embargo la demanda de periodistas es mínima. M^a Teresa Sandoval Martín, investigadora de la Universidad de La Laguna (Tenerife), en su artículo “Los periodistas en el entorno digital: hacia el periodista multimedia⁹³” apunta, a este respecto, la teoría siguiente: “Cuando las empresas hayan montado y organizado su infraestructura tecnológica precisarán periodistas que llenen sus páginas de contenidos veraces, creíbles y atractivos. Y, en efecto, algunas de estas empresas, nacidas bajo el paraguas de grandes emporios tecnológicos ya han empezado a contratar a profesionales del sector de los media que destacan del resto por estar familiarizados con los programas informáticos más utilizados en la edición de páginas *web*, la mayoría provenientes del periodismo digital. Esperemos que cuando estas empresas empiecen a poner como condición imprescindible ser periodista para trabajar con los contenidos informativos de sus páginas, esta práctica se extienda de forma generalizada. De esta forma, la captación y elaboración de informaciones se realizará con todos los componentes de objetividad, honestidad, credibilidad, etcétera que pueden aportar los periodistas”.

Uno de los aspectos fundamentales que debe tener en cuenta el periodista digital es la necesaria utilización de un lenguaje universal. Las noticias que publique en la Red podrán ser observadas, no solamente por personas de un determinado ámbito geográfico local, sino por cualquier usuario a nivel mundial. Hilda García Villa en su artículo *el nuevo periodismo en Internet* opina, a pesar de estas distancias geográficas entre unos usuarios y otros, lo siguiente⁹⁴: “ (...) hay informaciones que adquieren una relevancia especial por su interés común. Pero a nivel temático, también se debe considerar que a la gente le interesa lo que le afecta de manera directa. Por esa razón, los diarios y espacios informativos en el ciberespacio se han organizado como portales que le dan fuerza a la noticia local e invitan al usuario a abrir su propio espacio de discusión sobre temas más específicos por segmento profesional, temático o de interés particular”.

⁹³ Sandoval Martín, María Teresa: *Los periodistas en el entorno digital: hacia el periodista multimedia*. <http://www.saladeprensa.org>. Madrid, mayo 2000.

⁹⁴ García Villa, Hilda: *El nuevo periodismo en Internet*: <http://www.saladeprensa.org>.

Otro de los aspectos a tratar a la hora de publicar en Internet, es la importancia de la vigencia y la actualidad. La capacidad de memoria de la Red puede ejercer efectos contraproducentes en la información digital. Cuando compramos un medio impreso en el quiosco o cuando escuchamos las noticias de la radio y la televisión sabemos que la información que estamos obteniendo es reciente. Internet constituye un flujo continuo de información, no precisa esperar al día siguiente para ver publicados los contenidos. Pero a diferencia de la radio y la televisión, que es substitutivo de las informaciones previas, el flujo de la Red es acumulativo, es decir, publicar una información no implica necesariamente la eliminación de la información desfasada o su clasificación en un archivo aparte como base documental. Estos contenidos deben estar relacionados con su fecha de vigencia.

Un buen medio digital debe ofrecer contenidos plenamente actuales, pero lo que siempre debe hacer es mantenerlos vigentes. Una información puede datar del año pasado pero mantener su vigencia. Por ejemplo: “Aznar ha elegido nuevo Gobierno”. Pero cuando el Gobierno cambie habrá perdido su vigencia.

Los medios digitales informativos tienden a estar configurados sobre bases de datos interrelacionados, en las que sea posible realizar modificaciones múltiples con una sola instrucción. El diseño y mantenimiento de estas bases de datos debe correr a cargo de los periodistas digitales. En este sentido la labor del periodista tendrá un elevado porcentaje de documentalista⁹⁵.

Asimismo, el diseño de las informaciones digitales tiende a formatos abiertos, en los que sea posible modificar y ampliar la información. El periodista digital debe de ser capaz de trabajar con este tipo de informaciones abiertas. Continúa siendo apreciada la obtención de la exclusiva, pero en un medio como la Red la cotización de estas exclusivas se deprecia a las pocas horas. Un día después de su publicación, debería estar contextualizado, enlazado, comentado e interpretado. Es tarea del periodista digital organizar esta convergencia de información, opinión, feedback, interpretación y contextualización.

⁹⁵ Edo, Concha: *Los columnistas y la interactividad con los lectores*: <http://www.saladeprensa.org>

M^a Teresa Sandoval Martín en su artículo “Los periodistas en el entorno digital: hacia el periodista multimedia⁹⁶”, señala que en las redacciones actuales tiende a haber dos tipos de profesionales de la información:

1. Aquellos que trabajan con contenidos existentes en la Red; es decir, los profesionales que utilizan Internet para elaborar las noticias de los medios tradicionales, de sus apéndices digitales o para el periodismo que tiene únicamente presencia digital.
2. Los que bien buscan, procesan, analizan o redactan contenidos para espacios de la Red; es decir, los profesionales que desarrollan diferentes tareas con la extraordinaria cantidad de información que está disponible a través de la *World Wide Web*.

Pero Sandoval añade un tercero que considero de vital importancia entre las nuevas facetas de trabajo del periodista digital: el “broker de información”. Este profesional se encarga de buscar información para suministrársela a un tercero. Los buscadores que un usuario común conoce a veces no son suficiente para encontrar una información determinada sobre un tema especializado. El hecho de que esté en la Red es prácticamente seguro, sin embargo su localización puede ser complicada. Aquí se abre por tanto un nuevo campo de trabajo para los profesionales de la información. Esta nueva faceta comienza a surgir en 1996 cuando empiezan a salir al mercado sistemas destinados a enviar noticias que van apareciendo en distintos medios directamente al buzón de correo electrónico del cliente. Tal es el caso de sistemas como el Business Alert, de *Reuters*; el sistema Tracker, del servicio *Online Lexis-Nexis*, o el servicio Discovey, del *Financial Times*". Un ejemplo de este tipo de empresas en España, en la

⁹⁶ Sandoval Martín, María Teresa: *Los periodistas en el entorno digital: hacia el periodista multimedia*: <http://www.saladeprensa.org>. Madrid, mayo 2000.

cual trabajan documentalistas, técnicos y periodistas, es la empresa Documentación y Análisis (D+A), perteneciente a Sofres Audiencia de Medios⁹⁷.

Sandoval en su artículo “Los periodistas en el entorno digital: hacia el periodista multimedia” describe así mismo cuáles son las funciones de este “broker de la información”: “El trabajo que desempeñan estos *brokers* tiene cierta similitud con la línea que siguen los denominados periódicos especializados digitales que a partir de un conjunto de informaciones disponibles en un momento determinado se la envían a sus lectores, aunque se ha visto que estos "periódicos a la carta" han fracasado porque muchas personas no están interesadas en temas específicos y demandan una 'digestión y selección' de los temas. Este servicio de noticias a la carta surgió del periódico *The San José Mercury*, de Estados Unidos, y en la actualidad tiene como prototipo en España *My News*. Resulta evidente, por tanto, que ante la avalancha de información disponible en la red muchas empresas demandarán los servicios de estas empresas o, si no, empezarán a contratar a un tipo de profesional *junior* que es el "buscador de información". Pero estas personas deberán ser expertas no sólo en el manejo de las herramientas para buscar información, sino que deberán investigar las fuentes, seleccionar los documentos, analizarlos, cotejarlos si es preciso y, probablemente, llegará el momento -si no ha llegado ya-, en que tendrán que ser capaces de extraer lo más relevante de cada uno de ellos”.

Una nueva figura surge de la mano de Alfons Cornellá⁹⁸: “el responsable el conocimiento”. “Es la persona bajo cuya responsabilidad recae la acumulación y diseminación de la información y que podría además asumir el papel de filtrador de la información que llega a las empresas, facilitando únicamente la que es relevante para éstas. Esta última actividad ya empieza a darse en los periódicos digitales donde un "asistente de información", que puede ser un periodista recién licenciado o un redactor con poca experiencia, se encarga de leer los mensajes enviados por correo electrónico a la redacción, de seleccionarlos y distribuirlos y de contestar a los usuarios, una figura que nació en el periódico americano *The Chicago Tribune*” afirma Cornellá.

⁹⁷ Sandoval Martín, María Teresa: *Los periodistas en el entorno digital: hacia el periodista multimedia*: <http://www.saladeprensa.org>. Madrid, mayo 2000.

⁹⁸ Cornellá, Alfons: *Nuevas profesiones en la información*, en Revista *Extra!-Net*, 26 de febrero de 1999: <http://www.extra-net.net/>

Un tema que aún no ha sido tratado y que también es de gran importancia en los medios digitales, es el del nuevo papel del columnista. En su artículo titulado “Estudios sobre el mensaje periodístico”, Concha Edo, docente del Departamento de Periodismo I de la Facultad de Ciencias de la Información de la Universidad Complutense de Madrid, ha extraído a este respecto las siguientes conclusiones⁹⁹:

“Se pueden establecer algunas pautas comunes y es posible hacer alguna previsión de futuro:

- Las columnas de opinión no se publican teniendo en cuenta la tecnología disponible en la Red, ni alcanzan la importancia que tienen en la prensa diaria convencional,
- No existe un criterio específico ni en cuanto a los temas -hasta ahora todos se han limitado a reproducir los textos de la edición en papel o a publicar otros semejantes en los diarios exclusivamente digitales-, ni en cuanto al diseño adecuado.
- En casi todos los medios estudiados se comprueba que estos artículos no son los más valorados en cuanto a número de visitas.
- Como no se aprovechan las posibilidades gráficas ni las que ofrece el uso del hipertexto y leer en la pantalla es incómodo -hay que buscar cada artículo y recorrer con el cursor un bloque uniforme que sólo en casos contados incorpora una foto del autor o alguna letra capital en negrita- estos textos están en clara desventaja frente a los convencionales y sólo aumentará el interés por los contenidos si se presentan de una forma más atractiva e interactiva,
- Se puede contar con una facilidad adicional que es el acceso a los artículos atrasados
- Es previsible el aumento del interés de los lectores cuando puedan hablar más fácilmente y en tiempo real con los columnistas y manifestar su

⁹⁹ Edo, Concha: *Los columnistas y la interactividad con los lectores*: <http://www.saladeprensa.org>

opinión respecto a lo que dicen: hasta ahora sólo una mínima parte de los que escriben en los medios ofrecen su correo electrónico.

Hay que tener en cuenta que aún queda mucho que avanzar en todos los aspectos del periodismo *online*. Pero la base ya está forjada: la interactividad, es decir, la relación directa y en tiempo real de los profesionales de la información con el público usuario.

Como hemos podido comprobar a lo largo de este capítulo, la mayoría de las actividades relacionadas con la información que ofrece la Red, son actualmente o serán realizados en un futuro próximo por periodistas. Son los profesionales de la información aquellos que pueden de forma más efectiva sacar el máximo rendimiento de la base de datos mundial que es Internet. Pero para que esto sea posible, según apunta Sandoval¹⁰⁰, hay que recordar que las empresas han de tomar conciencia de los beneficios que les va a reportar en cuanto a eficiencia comunicacional. Con respecto a los periodistas especializados en este nuevo medio, éstos se tienen que preocupar de adquirir las destrezas necesarias, y las facultades de Ciencias de la Información han de ir adecuando sus planes de estudio o los profesores ir introduciendo en sus temarios los conocimientos necesarios para poder ofrecerle al mercado el perfil de profesional de la información que vaya demandando. Una característica esencial del perfil del periodista digital, será por tanto su capacidad de adaptación y la flexibilidad a la hora de realizar su trabajo.

¹⁰⁰ Sandoval Martín, María Teresa: *Los periodistas en el entorno digital: hacia el periodista multimedia*: <http://www.saladeprensa.org>. Madrid, mayo 2000.

3.6. INTERNET COMO HERRAMIENTA DE TRABAJO PARA EL PERIODISTA ACTUAL

La inmensa mayoría de los periodistas consideran la Red como una herramienta imprescindible de su trabajo. Así lo revela el Estudio¹⁰¹ sobre el uso de la Red en los Medios de Comunicación, elaborado por AccesoGroup en colaboración con la firma Deloitte & Touche, que además señala que el 93% de los periodistas encuestados no pueden pasar sin Internet en su profesión.

El informe, que ha encuestado a 616 periodistas de 300 medios de comunicación, revela que estos niveles se alcanzan fundamentalmente en Madrid, Cataluña, Andalucía, País Vasco y Comunidad Valenciana. En el caso concreto de esta última, el porcentaje de periodistas que consideran la Red como imprescindible llega a ser del 100%.

El estudio revela que la gran mayoría de los periodistas encuestados creen que las salas de prensa *on line* son indispensables para su labor profesional. Mediante estas salas de prensa, las organizaciones ponen a su disposición todos los recursos (dossier informativos, material gráfico, notas de prensa, etc.) necesarios para una sencilla búsqueda de información.

Todos los periodistas entrevistados en esta encuesta afirman que Internet es tanto una fuente de información como una fuente de documentación. La información más valorada es el material gráfico, además de los dossiers informativos y las notas de prensa, mientras que los archivos de audio o vídeo comunicados no tienen tanta importancia para el 30% de los encuestados, a excepción de los profesionales de la radio.

Los periodistas afirman, en líneas generales, basarse en las notas de prensa para redactar sus noticias. Sin embargo, las secciones de Internacional, con un 24%, seguido

¹⁰¹ Estudio sobre el uso de la Red en los medios de comunicación. Elaborado por Deloitte & Touche y AccesoGroup (octubre 2002) Accesible en: http://www.deloitte.es/estudio_periodistas/estudio_completo.htm

por Local, con un 10% y Tecnología, con un 8%, son las secciones que menos se apoyan en este tipo de informaciones.

Correo electrónico o teléfono

En su relación con empresas e instituciones, un 34% de los periodistas prefieren el correo electrónico antes que el teléfono. De hecho, un 64% lo valoran como el mejor canal para recibir información por parte de los departamentos y agencias de comunicación. El 50% de los periodistas encuestados reciben entre 15 y 50 *e-mails* diarios, mientras que el 31% de éstos reciben más de 50. Una gran mayoría (87%) piensan que el fax es una herramienta obsoleta. También cabe resaltar que más de la mitad de los encuestados no conceden aún importancia a la mensajería electrónica, o al chat, como herramienta de comunicación.

El estudio concluye que sólo los periodistas con edades comprendidas entre 30 y 50 años estiman imprescindible la utilización de Internet para el desarrollo de su trabajo. Sin embargo, los profesionales más jóvenes, así como los que superan los 50 años, no conceden tanta importancia a esta herramienta. Casi la totalidad de los encuestados consideran fiable la información aparecida en la Red, aunque necesitan cierto grado de credibilidad en la fuente.

El periodista “digital” del futuro será como un “hombre orquesta”

En una entrevista¹⁰² realizada por Juan Foniales en Navegante.com, el profesor de la Facultad de Periodismo de la Universidad de Columbia (Nueva York) Steven S. Ross señaló que en un futuro inmediato, el periodista se tendrá que convertir en un

¹⁰² Foniales, Juan: Entrevista a Steven S. Ross: *El periodista enviará imagen, voz y texto* (publicado el 15 de octubre de 2001 en Navegante.com)

profesional polivalente. Un solo redactor tendrá que enviar imágenes, voz y texto. Y, además, hacerlo con inmediatez. En directo. Apuntó, además, que lo importante va a ser, y es en la actualidad, trabajar en equipo y especializarse en un área concreta.

Preguntado por la posible desaparición de la prensa escrita con la llegada del soporte Internet, Ross afirmó que en 25 años, los diarios desaparecerán por razones medioambientales, pero las empresas informativas se mantendrán. “Las nuevas tecnologías no cambian las necesidades. Cambia el negocio, no el propósito”, señaló. En cuanto al proceso a seguir, apuntó que a distribución de los ejemplares se hará en otro soporte, a través de Internet: “En EEUU el 90% se reparte por suscripción. Es una buena forma de abaratar los costes”.

2º PARTE:

2ª PARTE: LA BÚSQUEDA DE INFORMACIÓN EN INTERNET

1. INTRODUCCIÓN

Ramón Salaverría, director del Laboratorio de Comunicación Multimedia (Universidad de Navarra), suele emplear la siguiente metáfora a la hora de definir los buscadores en Internet, que ilustra acertadamente la tarea de los mismos¹⁰³:

“Imaginemos un pajar del tamaño de un gran estadio de fútbol, repleto de pequeñas briznas de heno hasta diez metros de altura, y lancemos desde el aire una aguja de coser. A continuación, pidamos a alguien que encuentre esa aguja. Ese desventurado individuo estará perdido si pretende hallar la aguja revolviendo desordenadamente la paja. Invertirá horas y horas en vano, hasta la desesperación. Por el contrario, si emplea el utillaje adecuado —un gran imán, por ejemplo— encontrar la aguja puede ser cuestión de segundos”.

Internet es, en la actualidad, una fuente esencial para el trabajo de los periodistas dado el volumen de información que contiene y genera diariamente, además de por su utilidad potencial para facilitar el contacto con fuentes personales (como pudieran ser expertos o testigos de una información) a nivel mundial.

La llamada World Wide Web contiene aproximadamente mil millones de documentos y crece a un ritmo diario de 1,5 millones de páginas¹⁰⁴. Los motores de búsqueda más avanzados no llegan a alcanzar más que una ínfima parte del contenido de esta Web. Específicamente cabe reseñar, que los buscadores que más páginas tienen indexadas son Fast (<http://www.alltheweb.com>) 300 millones de documentos y Northern Light (<http://www.northernlight.com>) con 218 millones. Además hay que tener en cuenta que la World Wide Web, o la Web accesible mediante buscadores, es solo una parte de Internet. Existe también la llamada Web profunda o Internet Invisible, a la que no se puede acceder mediante los buscadores, que es aproximadamente 500 veces mayor que la visible y sobre todo, de una mayor calidad.

¹⁰³ Salaverría, Ramón: *Internet para periodistas*. Escuela de Medios de *La Voz de Galicia*. A Coruña, 20-22 de septiembre de 2000.

¹⁰⁴ *The Deep Web: Surfacing Hidden Value*. Informe de BrightPlanet.com: http://www.completeplanet.com/tutorials/deepweb/introduction_howsearch07.asp.

A la hora de enfrentarse a este mar de información disponible en Internet, el periodista debe de tener un buen conocimiento de los distintos recursos de búsqueda que le ayudarán a alcanzar su propósito y de cómo realizar dicha búsqueda.

Según el periodista Miguel Ángel Díaz Ferreira de la Facultad de Ciencias de la Información de la Universidad Complutense de Madrid, existen cinco formas diferentes de localizar recursos por Internet¹⁰⁵:

- f) Por medio de **publicaciones impresas** que recogen direcciones de la Red. Puede tratarse de trabajos de investigación, artículos de publicaciones especializadas o incluso de periódicos o revistas de información general y, sobre todo, de libros escritos *ad hoc* para las búsquedas. Pueden ser manuales sobre la Red o las “Paginas Amarillas” de Internet. Sus dos principales desventajas son las siguientes:
- Su temprana obsolescencia.
 - La necesidad de introducir los datos necesarios a mano (por tratarse de publicaciones impresas), sin aprovechar la utilidad del hipertexto para moverse por sus páginas.
- g) A través de los **directorios de recursos y clasificaciones temáticas**, que recogen y organizan los recursos de la Red sobre la base de un ordenamiento previo. Su problema es que no siempre están lo suficientemente actualizados y que no recogen todos los recursos de la Red, sino tan sólo los que los autores de la información hayan dado de alta en el servicio. Su ventaja es su propia organización que ayuda a localizar las direcciones con mucha facilidad y de forma muy esquemática.
- h) Utilizando los **motores de búsqueda e indización automáticos**, que pueden localizar cualquier recurso de la Red. Éstos tienen la desventaja de que, si no acotamos mucho nuestra búsqueda, nos pueden proporcionar demasiados enlaces, la mayoría de ellos irrelevantes para nuestra investigación. Su gran ventaja consiste en que la información no necesita ser dada de alta para que el

¹⁰⁵ Citado en: Pérez Luque, María José: *El periodismo y las nuevas tecnologías*. Newbook Ediciones Navarra, 1998. pp: 81-82

motor la localice, por lo que están más actualizados y suelen ser más completos que los directorios y clasificaciones temáticas.

- i) **Consultando los enlaces que ponen a disposición del internauta** otras páginas relacionadas con el tema de nuestra búsqueda. Una vez localizada cualquier página relativa a la materia que se pretende encontrar, bastará con conocer esa dirección, porque es casi seguro que esa página contará con interesantes enlaces que llevarán a los puntos más importantes donde localizar toda la información necesaria. El problema es que no todas las paginas cuentan con esos enlaces.
- j) **Preguntando a los propios usuarios de la Red.** Ellos son, sin duda, los que están mejor informados, los que conocen las ultimas páginas aparecidas no accesibles mediante buscadores. Su información es siempre la más actualizada y suele ser pertinente. El éxito de la búsqueda es precisamente encontrar a algún internauta dispuesto a ayudarnos y que además esté interesado o conozca el asunto sobre el que buscamos información. Hay diversos métodos para localizar a entablar conversación entre los internautas: listas de distribución, grupos de noticias, IRC, Chats, etc.

En la World Wide Web los buscadores los podemos dividir en buscadores automáticos, o motores de búsqueda, y directorios o índices temáticos. Dentro de los motores de búsqueda podemos encontrar, a su vez, buscadores automáticos generales y especializados por campos del conocimiento, y lo mismo ocurre con los directorios (generales y especializados). Asimismo, existen además *webs* de motores de búsqueda que han incorporado directorios en su página y al revés. El principal objetivo es facilitar la búsqueda de información al usuario, aunque como hemos visto, todavía queda mucho por hacer.

A lo largo de este capítulo intentaremos ofrecer una visión global de los recursos que debe conocer el periodista a la hora de buscar eficazmente información por Internet.

4.2. LA BÚSQUEDA DE INFORMACIÓN EN INTERNET

A la hora de utilizar Internet profesionalmente para buscar información hay que tener en cuenta que esta búsqueda de información no es navegar con descontrol por toda la información disponible intentando encontrar algo que te interesa. La búsqueda de información profesional requiere un claro propósito y por consiguiente un conocimiento extendido de los recursos a los que puedes acudir en cada caso.

El verdadero valor del periodista que busca información en Internet, es directamente proporcional al número de fuentes a las que puede llegar más rápidamente, y a la habilidad de dicho periodista para escoger las palabras adecuadas para la búsqueda.

Según el libro *The Information Broker's Handbook*¹⁰⁶, aquellos que se dedican profesionalmente a la búsqueda de información en Internet no deben considerarse capaces de dar soluciones porque para lo que realmente pueden ser útiles y lo que les diferencia del resto son las propias fuentes. Sue Rugge y Alfred Grossbrenner, autores de este libro, señalan que la clave está en ofrecer al cliente el tipo de información que necesita y posibilita que estos individuos puedan resolver sus problemas. Lo que está claro, es que una persona que se dedica a esta profesión nunca puede llegar a ser experto en todas las materias y tampoco debe pretenderlo, tan sólo lo ha de llegar a serlo en la búsqueda de información.

Planteamiento correcto de la búsqueda

Como hemos señalado con anterioridad es primordial a la hora de encontrar específicamente lo que buscas el hecho de saber plantear correctamente la pregunta. Si el término que introduces en un motor de búsqueda es muy vago o general, la respuesta puede ser demasiado amplia como para ser útil. En el artículo de David Novak¹⁰⁷ sobre la búsqueda de Información en Internet, éste apunta que el primer paso a seguir en estos casos, es visualizar aquello que se desea consultar antes de comenzar. De esta manera podemos limitar la búsqueda en cuanto a la amplitud del documento que

¹⁰⁶ Rugge, Sue and Grossbrenner, Alfred: *The Information Broker's Handbook*

¹⁰⁷ Novak David: *The Information Research*. <http://cn.net.au/faq.htm>

buscamos, los autores, la calidad que se necesita, la posibilidad de buscar a un experto en la materia y las herramientas o recursos disponibles en Internet para esa búsqueda concreta. Esto puede marcar la diferencia entre obtener cientos de artículos generales o tres concretos que respondan directamente a tu pregunta. Novak pone el ejemplo de un padre que se pregunta qué puede hacer él por su hijo superdotado.

- En este caso, el profesional de la búsqueda de información debe replantearse la pregunta de la forma siguiente: ¿Qué recursos pueden ayudar a un niño superdotado? o ¿qué suelen hacer los padres que tienen hijos superdotados? o ¿Quién puede aconsejarme sobre la forma de ayudar a un niño superdotado?.
- Seguidamente se debe precisar la pregunta a realizar para obtener respuestas específicas: ¿Dónde puedo encontrar una lista de asociaciones? o ¿Quién conoce asociaciones para niños superdotados? o ¿Qué artículos o libros especializadas existen sobre niños superdotados?.

Conocer los métodos de búsqueda y la tecnología

El profesional de la búsqueda de información en Internet debe tener un conocimiento avanzado de los métodos de búsqueda en cada caso y de la propia tecnología. Las nociones de fuentes específicas sobre una materia también pueden ser de gran ayuda, así como el conocimiento de los metabuscadores. Puede ser de gran ayuda a la hora de mejorar la técnica en este sentido, realizar búsquedas complejas en bases de datos electrónicas, en las que se pone a prueba los conocimientos del experto sobre operadores booleanos, factores de proximidad y campos.

Preguntar cuando sea necesario

Es importante que el profesional de la información conozca sus límites de conocimiento ya que este campo es demasiado amplio y en constante evolución como para conocer absolutamente todos sus aspectos y posibilidades. En este caso no se debe dudar en extender la pregunta a otros expertos en campos más específicos. Es importante en este caso saber a quién preguntar en esta situación y no tener nunca la sensación de no saber,

de ninguna forma, cómo hallar la respuesta a una pregunta. Un buen recurso, en muchos casos, son los bibliotecarios, ya que ellos pueden ayudar a ampliar los conocimientos del informador sobre otros métodos o estilos de ordenar la información que pueden ser de gran utilidad para poder realizar distintos tipos de búsqueda sobre una misma pregunta y así agotar todas las posibilidades.

El formato de Internet

Antes de la aparición de Internet, solíamos buscar la información únicamente en libros o artículos de periódicos o revistas. Sabíamos que cada uno de estos formatos tiene sus características particulares. Por ejemplo, los libros suelen ser densos, normalmente fáciles de comprender y poco actuales, mientras que los artículos suelen ser más escuetos y es más probable que sean actuales.

De igual forma, podemos separar los recursos que encontramos en Internet según su formato, teniendo en cuenta que los “libros electrónicos” tienen prácticamente el mismo formato que los de papel e igual pasa con las noticias cuyo formato es básicamente el mismo se encuentren en el medio que se encuentren.

Una forma de evaluar la importancia de una página web es a través del análisis de los enlaces. En teoría, cuantos más enlaces apunten a una página en cuestión, más calidad tiene dicha página. De esta forma, aquellas páginas que tienen al menos 100 enlaces apuntando a ella pueden ser consideradas fuentes calidad. Mientras que una web sin ningún enlace que apunte a ella es supuestamente poco fiable, ya que ningún otro usuario la ha tenido en cuenta para ponerle un enlace directo en su página.

Sin embargo esta teoría que suelen seguir los motores de búsqueda a la hora de organizar sus bases de datos no es del todo fiable, sobre todo teniendo en cuenta que hay muchas webs de calidad que han sido colgadas recientemente y que no han tenido tiempo de ganarse los enlaces. Otro aspecto a tener en cuenta importante, es que la técnica de comparación entre el número de enlaces y la calidad es bastante conocida y en muchas ocasiones, el ranking en el que se encuentra una página web es producto del capital de la empresa que se haya detrás.

Internet ha cambiado la forma de buscar información en dos sentidos:

- En primer lugar, la página web es un nuevo formato al que tenemos que habituarnos.
- Internet es además una vía a otras muchas herramientas de búsqueda que existían con anterioridad, como son los libros, las noticias o las entrevistas.

Mediante una conexión a Internet, podemos acceder a una gran variedad de bases de datos de gran calidad. Asimismo, también podemos acceder al catálogo de la mayor parte de las bibliotecas de todo el mundo. Realmente estas no son fuentes nuevas, sino un nuevo medio de llegar hasta ellas. El formato que estas fuentes tienen en Internet no dista mucho de cómo es en su formato tradicional. Por ejemplo, las noticias continúan siendo escritas por un periodista del medio en cuestión, después son publicadas en un periódico (o en una página web en nuestro caso) y seguidamente llegan al público. Esto mismo ocurre con los libros, ya que tanto si éstos se publican en la Red como en papel, la cadena hasta que llega al público receptor es parecida: autor-editor-publicación-librerías (virtuales o no) y por último encontramos al lector. Lo que cambia, es que sin movernos de casa o del lugar de trabajo, podemos acceder a bases de datos de librerías como Barnes and Noble, Amazon y The Internet Bookshop, además de a útiles publicaciones estatales como la MOCAT (Estados Unidos), AGIP (Australia), la Stationery Office Online Catalogue (Gran Bretaña) y los catálogos de la Biblioteca del Congreso (LOCOC) y la Biblioteca Británica.

Aunque la industria del libro y de la comunicación no ha cambiado radicalmente, Internet nos ha ofrecido la preciada posibilidad de establecer grupos de discusión a nivel mundial, de comunicarte de forma fácil e inmediata y sin intermediarios con los propios autores de los libros o las noticias, o también acceder de forma inmediata a cualquier otra fuente de información. Esta inmediatez es una de las mayores ventajas que ofrece Internet a sus usuarios.

A pesar de que la mayor parte de la información en Internet es gratuita, y que entre ella se encuentran datos de gran valor, se debe tener en cuenta el valor superior que suele tener una información a la que únicamente se puede acceder mediante pago y que en muchas ocasiones no sería comparable a los resultados de unas cuantas horas navegando por Internet.

A la hora de analizar cómo está organizada la información en la Red de redes hay que partir del hecho de que dicha información no se encuentra de la forma en que

está aleatoriamente, sino más bien lo contrario. Las líneas de orden de estos datos se encuentran perfectamente definidas. Algunas de estas líneas son específicas de cada formato: documento de texto, página web, mensaje de correo electrónico, artículo, etc. Otros patrones se ajustan a la forma en la que la información se nos da a conocer, como las listas de distribución, las preguntas frecuentes o FAQ y los periódicos *online*.

David Novak propone tres definiciones de Internet que se ajustan bastante al hecho que estamos describiendo y que ayuda a una mayor comprensión. Son las siguientes:

1. Internet es una red física; más de un millón de ordenadores continuamente intercambiando información. Internet nos permite intercambiar información en todo el mundo.
2. Internet es un mundo de información disponible en prácticamente cada uno de los tópicos imaginables. Dicha información como una masa casi caótica pero sujeta entre sí a través de claros patrones.
3. Internet es una comunidad creciente de más de 500 millones de personas reales que interactúan, discuten y comparten información *online*. En muchas ocasiones es más útil preguntar a estos usuarios que buscar la información en el ciberespacio.

La forma en que la información es agrupada en Internet (y fuera de Internet) tiene mucho que ver con su contenido, calidad y uso específico de dicha información. Las páginas web, los documentos de texto, las bases de datos, el software y el correo electrónico, tienen cada uno sus características particulares. La mayoría de los libros que encontramos habitualmente en una librería son densos y normalmente poco actuales; mientras que los artículos suelen ser cortos y más recientes. Esto mismo ocurre en Internet. Hay páginas web que son bastante gráficas y en muchas ocasiones no se actualizan periódicamente, otras son sólo texto y por tanto más fáciles de mantener (lo cual no significa que se actualicen con regularidad); las listas de distribución son textos normalmente cortos que forman parte de una discusión sobre un tema concreto. Asimismo, los grupos de noticias están caracterizados por su bajo coste, su vulnerabilidad para contener “contenido basura” (textos sin ningún sentido) ya que

carece de un seguimiento o control y el correo electrónico se basa en el intercambio de mensajes.

Siempre que se realiza una búsqueda en Internet se ha de tener en cuenta qué tipo de información se trata de encontrar, y en base a esto, determinar en qué tipo de “formato” debemos buscar.

4.3 ¿QUÉ CONTIENE LA WORLD WIDE WEB?

En la actualidad la World Wide Web contiene cerca de 5 millones de sitios web, que haría un total de 1.000 millones de páginas que aumentan cada día más. La mayoría de ellas, aproximadamente un 86%, se encuentran escritas en inglés¹⁰⁸.

Pero esto no es todo, la World Wide Web es únicamente una ínfima parte de los contenidos que hallan y que se vuelcan realmente en Internet a través de bases de datos, listas de distribución (news groups), correo electrónico, etc, que pueden ser hasta 500 veces mayor que el contenido de la Web.

Los documentos “web” que se van creando son accesibles a cualquier usuario de la Red gracias a las llamadas “arañas”, una tecnología de los buscadores cuyo cometido es leer cada página que encuentran en su camino, y mandar las palabras que contiene, junto con la información de su ubicación a enormes bases de datos. A continuación, estas arañas siguen los enlaces que se encuentran en esta página hasta llegar a un nuevo sitio, y así sucesivamente. Sin embargo, según afirma *Search Engine Watch*¹⁰⁹, ningún buscador indiza más del 50% de la Red. Según un estudio de la revista *Nature*¹¹⁰ de febrero de 1999, entre todos los buscadores no se cubría más que el 42% de la Red. A pesar de que la tecnología avanza y de que cada vez se amplía más esta cifra, continúa habiendo una cierta tendencia a indizar sobre todo sitios de EEUU, principalmente los más visitados y los sitios comerciales más que relacionados con la educación. También

¹⁰⁸ *The Deep Web: Surfacing Hidden Value*. Informe de BrightPlanet.com:
http://www.completeplanet.com/tutorials/deepweb/introduction_howsearch07.asp.

¹⁰⁹ Search Engine Watch: <http://www.searchenginewatch.com/reports/sizes.html>

¹¹⁰ Lawrence, Steve y Giles, C. Lee, del NEC Research Institute. Citado en *Nature* del 08/07/1999 (resumen accesible en <http://www.wwwmetrics.com>).

son poco privilegiados en este aspecto los documentos escritos en lenguas minoritarias que podemos encontrar en la Red en una baja proporción.

Según la empresa *Wordtracker*¹¹¹, los usuarios realizan diariamente un total de 250 millones de consultas a los buscadores. La palabra más buscada ha sido hasta hace poco la palabra “sex” desbancada por el formato de compresión musical “mp3”. Gracias a estos buscadores no es tan difícil encontrar algo en la Web, tan sólo hay que saber cómo utilizarlos adecuadamente, usar los llamados operadores booleanos, combinando

palabras -y, o, no- para definir la búsqueda e intentar que la palabra clave a utilizar no pueda hallarse en otro campo.

A la hora de analizar el contenido disponible en la Red debemos tener también en cuenta lo efímeros que son algunos enlaces. La media de cualquier enlace, hace algunos años, era de 44 días. En febrero del 2000, según un estudio de Inktomi¹¹², frente a 4,2 millones de sitios accesibles que se examinaron, 0.7 millones eran inalcanzables debido a la desaparición de su servidor o a su traslado. Con el propósito de preservar la mayor parte de la Web posible, incluyendo a los grupos de discusión, en 1995 Brewster Kahle comenzó a desarrollar el proyecto “The Internet Archive¹¹³”. Hasta el momento lleva almacenados 15, 5 terabytes. La mayoría de las webs que contiene son obtenidas mediante arañas, aunque también se aceptan páginas cedidas por sus propietarios.

Para facilitar aún más la búsqueda, según el artículo de Juan José Millán “El libro de medio billón de páginas¹¹⁴”, se ha creado cierto software que permite ampliar el campo morfológico y semántico de las búsquedas; por ejemplo preguntando por dirigir, se accederá también a dirigido, dirigiendo y a guiar, conducir, etc. Del mismo modo, existen también buscadores que rompen la frontera de la lengua incorporando a la búsqueda sus traducciones en otros idiomas (conducir, drive, conduire).

Por lo tanto a la pregunta de qué contiene la World Wide Web, la respuesta es una infinidad de información que el usuario tiene que aprender a encontrar. Que no se

¹¹¹ Wordtracker: <http://www.wordtracker.com/>

¹¹² Inktomi: <http://www.inktomi.com/webmap/>

¹¹³ The Internet Archive: <http://www.archive.org/>

¹¹⁴ Millán, Juan José: *El libro de medio billón de páginas*: <http://jamillan.com/ecoling.htm>

encuentre lo que se busca, no significa necesariamente que no esté ahí. En este mismo artículo, Juan José Millán identifica a los buscadores como intermediarios, es decir, como el filtro por el que los contenidos de las *webs* llegan al público. “Imaginemos que la única vía de acceso a todas las publicaciones mundiales fueran los ficheros de una decena de bibliotecas. En la Web estamos así, con la diferencia de que en los buscadores no ha habido bibliotecarios que apliquen sus saberes clasificadores: sus programas actúan ciegamente, analizando el código de la página, y clasificando los sitios según criterios formales. Este último aspecto es básico: cuando puede haber cientos o miles de sitios que responden a una determinada búsqueda, figurar en los primeros puestos de la lista de respuestas de un buscador puede ser la diferencia entre

existir o no. No extrañará, por tanto, que haya quien lo intente por todos los medios”, afirma.

4.4. LA CREDIBILIDAD Y LA VALORACIÓN DE LA INFORMACIÓN HALLADA EN LA WEB

Como hemos señalado en varias ocasiones con anterioridad, la Red es un gran mar de información completamente abierto y accesible para que cualquier persona que posea un ordenador con conexión a Internet pueda hacer su aportación. Por ahora no existe ningún filtro capaz de valorar la veracidad o ética de las miles de páginas que se vuelcan diariamente en la Web.

En la actualidad todo son facilidades a la hora de crear una página en Internet. Los portales llegan a ofrecer de 25 a más Megabites de forma gratuita por darse de alta en sus servidores. Un usuario, una clave y un espacio completamente libre de ataduras y censuras para publicar. Por ahora todo parece bien. Sin embargo, esta misma frase que simula una situación “perfecta”, conduce a un cierto problema. La Red se encuentra repleta de información útil para sus usuarios, pero del mismo modo también se pueden encontrar, con la misma facilidad, documentos falsos, con una falta absoluta de ética o incluso que rayan el delito.

Como hemos afirmado con anterioridad, no existe ningún filtro que, en un principio impida publicar en la Red. Lo que sí es posible es utilizar un *software* (ej: censorware o cibernanny) para restringir el acceso a ciertos sitios que podrían resultar perjudiciales, sobre todo para los niños.

En el caso que nos ocupa, es decir, el de los periodistas que utilizan Internet como medio para buscar sus fuentes, éstos deben comprobar cualquier información hallada en la Red ya que cualquiera puede haberla publicado bajo nombre falso. Del mismo modo, a la hora de publicar en la Web, el periodista debe regirse por la misma ética de veracidad y exactitud que tendría en un medio tradicional.

4.5. LOS COMIENZOS DE LA NAVEGACIÓN POR LA RED

La World Wide Web tuvo un éxito inmediato gracias a su naturaleza descentralizada. Aunque los primeros servidores estuvieran alojados en el CERN, ni siquiera Berners-Lee ejercía ningún tipo de control sobre quien decidía establecer un nuevo servidor en cualquier sitio de Internet. Lo único que se requería para esta tarea era enlazar con otros servidores e informar a otros usuarios de la Red sobre los nuevos servidores de manera que éstos pudieran crear a su vez enlaces de hipertexto de vuelta.

Sin embargo esta naturaleza descentralizada también creó un problema. Aunque era sencillo para el usuario navegar de un servidor a otro siguiendo los enlaces de hipertexto, la navegación comenzó a resultar más difícil a medida que la Red iba creciendo. No había una autoridad central que creara y mantuviera organizados el creciente número de documentos disponibles. Para facilitar la comunicación entre estos primeros usuarios de la Red de redes, Berners Lee estableció una lista¹¹⁵ de servidores de la Red a la que se podía tener acceso a través de los enlaces de hipertexto, creando, de esta manera, el primer directorio de la Red.

Durante los comienzos de la World Wide Web, la creación de una nueva página web tenía que ser notificada al resto de los usuarios a través del correo electrónico o a través de los foros de discusión. En esta época, algunos usuarios de la Red comenzaron a crear listas de enlaces a sus webs favoritas. Algunas de las listas más populares eran las mantenidas por autores como: John Makulowich, Joel Jones, Justin Hall y el equipo de la compañía de publicidad de O'Really & Associates. Estas listas se convirtieron en un medio para anunciar el lanzamiento de nuevas páginas web. Sin embargo, el proceso de darse a conocer en estas páginas requería facilitar cierta información sobre cada uno de los nuevos sitio web, una ralentización del proceso no compatible con el imparable crecimiento de la Red, ya que era imposible mantener las listas actualizadas.

Lo que era realmente necesario, en este momento de auge de la World Wide Web, era la automatización del descubrimiento y organización de nuevas páginas. Las

¹¹⁵ Accesible en: <http://www.w3.org/History/19921103-hypertext/hypertext/DataSources/bySubject/Overview.html>.

técnicas de búsqueda de información habían estado disponibles desde 1960 pero solamente eran efectivas en búsquedas en sistemas cerrados. De esta forma se establecieron nuevas herramientas que permitieran la búsqueda en servidores intercomunicados entre sí.

CAPÍTULO 5-RECURSOS DE BÚSQUEDA

5.1. INTRODUCCIÓN

A la hora de realizar una búsqueda en Internet, el usuario tiene a su disposición una serie de recursos que le facilitan encontrar lo que desea. Es importante conocerlos para no perderse en el mar de información, cada vez mayor, disponible para el usuario.

Principalmente son los siguientes:

- **Directorios**
- **Motores de búsqueda**
- **Anillos Web**
- **Portales (Horizontales y verticales)**
- **Bases de datos**

Tal y como se ha explicado con anterioridad, a parte de estos recursos principales existen otros muchos como preguntar a otros usuarios de la Red a través del e-mail o de los foros de discusión. En esta ocasión vamos a centrarnos en los cuatro primeros cuyo conocimiento por parte del profesional de la información consideramos fundamental. El quinto, las bases de datos, lo analizaremos en profundidad en la tercera parte de esta tesis, dedicada a la Red profunda o Internet invisible.

A continuación exponemos una breve historia de la evolución de las herramientas de búsqueda en Internet, para dar paso, a continuación, a una explicación detallada y con ejemplos de los recursos de búsqueda disponibles en la Red de redes.

5.2 LAS PRIMERAS HERRAMIENTAS DE BÚSQUEDA

Se puede afirmar que las primeras herramientas de búsqueda y recuperación de información datan de finales de la década de 1950, principios de la de 1960; sin embargo estas técnicas eran utilizadas en sistemas cerrados. Las primeras herramientas de búsqueda en Internet no tenían ni siquiera las habilidades más básicas debido principalmente a que estas técnicas que se utilizaban de forma eficaz en sistemas

cerrados no tenían la misma utilidad cuando intentaban adaptarse a una estructura abierta¹¹⁶.

En los primeros tiempos de Internet, cuando se trataba de acceder a un archivo en Internet había que completar un proceso de dos partes: en primer lugar se necesitaba establecer conexión con el ordenador remoto donde se encontraba el archivo en cuestión usando un programa de emulación de terminal llamado Telnet. A continuación se necesitaba usar otro programa denominado File Transfer Protocol (FTP) para poder descargarte el documento en tu propio ordenador.

Durante algunos años el acceso a los distintos archivos disponibles en Internet estuvo condicionado al conocimiento, por parte del usuario, tanto de la dirección IP del ordenador donde se encontraba el documento, como la de la localización exacta y nombre del documento que se estaba buscando. En este momento de la historia de las herramientas de búsqueda normalmente cuando se intentaba buscar algo se utilizaba el correo electrónico o los foros de discusión para preguntar entre los pocos “internautas” que entonces había por las direcciones en cuestión. Esta situación mejoró considerablemente gracias a la introducción de los servidores FTP anónimos, es decir, en los que no eran necesaria la introducción de una contraseña, que era un servicio de archivos centralizado específicamente diseñado para facilitar el intercambio de archivos. Los archivos en los servidores FTP estaban organizados en directorios específicos. Esta estructura hacía más fácil el acceso a la lista de archivos que se encontraban en el servidor, sin embargo aún era necesario el conocimiento previo de los contenidos que había en el servidor.

El primer motor de búsqueda de archivos incluidos en servidores FTP fue “Archie”, creado en 1990 por un pequeño equipo de administradores de sistemas y estudiantes de la universidad McGill en Montreal (Canadá). Éste es considerado el prototipo de los motores de búsqueda actuales. Archie recorría Internet buscando archivos disponibles en servidores anónimos FTP; una vez encontrados, descargaba las listas de los directorios de cada uno de estos servidores FTP anónimos. Las listas resultantes eran archivadas en una base de datos central, que permitía la búsqueda

¹¹⁶ Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

interna, llamada “Internet Archives Database” (Base de Datos de Archivos de Internet) en la universidad McGill y era actualizada de forma mensual.

A pesar de haber significado un avance significativo, Archie limitaba las búsquedas a un archivo específico, además de poder utilizarse únicamente en ordenadores con funciones específicas. Sin embargo, esto no impidió que casi el 50% de los usuarios de Internet en Montreal a principios de la década de los años 90, usara este motor de búsqueda, según el jefe del equipo creador de Archie, Peter Deutsch. En 1995, este equipo ya había distribuido 63 licencias de instalación de la red Archie en todo el mundo.

En 1991, Mark McCahill y su equipo crearon en la universidad de Minnesota una alternativa a Archie, a la que llamaron “Gopher”, ya que este era el nombre de la mascota de la universidad (Golden Gopher). Gopher era básicamente una combinación de Telnet y protocolos FTP que permitía a los usuarios pulsar en los enlaces de un menú para acceder a la información solicitada. Permitía, además, la búsqueda por categorías y el acceso al texto completo de documentos, gráficos y archivos de música. Esta nueva herramienta supuso un importante avance a la hora de facilitar la búsqueda de información en Internet.

La apariencia de Gopher fue precursora de lo que más tarde evolucionarían a ser los directorios de la Red de redes (como Yahoo!), ya que le ofrecía al usuario una visión general y estructurada de toda la información de la que disponía, daba la oportunidad de pulsar en enlaces de hipertexto para moverse a áreas más especializadas y encontrar, de esta forma los documentos que se precisaban. Uno de los inconvenientes de Gopher era que fue diseñado para proporcionar listados de archivos disponibles en ordenadores con una localización específica. Mientras que los servidores Gopher permitían la búsqueda en ellos, no había un directorio centralizado que permitiera la búsqueda en otros ordenadores que estuvieran usando Gopher a Internet o al “Gopherspace” (espacio Gopher) como era llamado entonces. En noviembre de 1992, Fred Barrie y Steven Foster del Servicio de Sistemas de Ordenadores de la universidad de Nevada resolvió este problema creando un programa llamado “Verónica” (acrónimo de “Very Easy Rodent-Oriented Net-Wide Index to Computerized Archives”) que fue concebida como una herramienta de búsqueda centralizada para archivos Gopher.

En 1993, otro programa llamado “Jughead” (acrónimo de “Jonzy’s Universal Gopher Hierarchy excavation and Display”) incorporó la búsqueda por palabras clave y los operadores booleanos que permitían la búsqueda en Gopher.

Un tercer protocolo de búsqueda desarrollado en esta fecha fue “Wide Area Information Servers (WAIS), creado por Brewster Kahle, que era muy parecido a los metabuscadores actuales. El programa WAIS era ejecutado en el ordenador local y permitía la búsqueda de información en otros servidores de Internet usando el lenguaje natural, en lugar de los comandos del ordenador. Los propios servidores eran los responsables de interpretar la búsqueda y devolver los resultados apropiados, liberando al usuario de la necesidad de aprender un lenguaje específico de búsqueda de cada servidor.

WAIS usaba una extensión del protocolo estándar Z39.50 que era muy usado en estos años. Principalmente su función era proporcionar un solo protocolo entre un ordenador y otro para la búsqueda de información (texto, fotografías, archivos sonoros y documentos en otros formatos). La calidad de los resultados de búsqueda era el resultado directo de la efectividad con que cada servidor interpretaba la búsqueda de WAIS.

Los primeros protocolos de búsqueda en Internet representan un gran avance con respecto a las toscas herramientas de acceso que proporcionaban Telnet y FTP. Sin embargo aún carecían de la habilidad para establecer conexiones (texto, sonido, imágenes, etc) y de esta forma establecer enlaces que transformaran los datos en información útil para el usuario. En definitiva, aunque cada vez los métodos de búsqueda eran más sofisticados aún no eran ni conocidos, ni utilizados por una comunidad masiva; prácticamente su uso estaba limitado a científicos, agencias de gobiernos y centros académicos.

En este mismo tiempo un ingeniero de programas de Suiza llamado Tim Berners Lee comenzó a desarrollar un programa llamado “Enquire Upon Everything” que daría origen a la llamada World Wide Web (WWW) .

5.3. NAVEGAR Y BUSCAR

Llegados a este punto, y antes de comenzar a describir los distintos tipos de buscadores que nos pueden ayudar a localizar información en Internet, merece la pena aclarar las diferencias entre dos términos que se confunden en numerosas ocasiones y que son básicos a la hora de comprender correctamente el tema que nos ocupa.

Hay fundamentalmente dos métodos para localizar información en Internet: navegar y buscar¹¹⁷. Podemos definir navegar como el proceso de seguir los enlaces de hipertexto creados por otros usuarios. Un enlace de hipertexto redirecciona a otro documento, a otra imagen u a otro objeto en la Red. La palabra o palabras que componen el enlace son el título o descripción del documento que se obtiene cuando pulsas en dicho enlace. Por lo tanto, el hecho de navegar por la Red lo podemos considerar bastante simple e intuitivo.

Sin embargo, buscar depende de un programa que se encarga de hacer coincidir las palabras clave que el usuario especifica con los documentos más relevantes que existen en la Red. La búsqueda efectiva requiere aprender a usar las herramientas de búsqueda, así como desarrollar, a base de práctica, la habilidad para obtener resultados satisfactorios.

Cuando la Red era aún pequeña, navegar era suficiente para localizar la información que se necesitaba, sin embargo la amplitud que ha adquirido en la actualidad hace imposible esta tarea y el método más eficiente suele ser la búsqueda.

Aunque se analizarán más adelante los diferentes métodos de las herramientas de búsqueda, merece la pena citarlas en este apartado para explicar mejor las diferencias entre los dos términos que nos ocupan. Uno de los métodos es el “directorio” creado a partir de primeras herramientas de Internet como Archie y Gopher. El otro método, el motor de búsqueda, está basado en las técnicas clásicas de búsqueda de información usadas en las bases de datos cerradas (es decir, sin acceso a Internet) tradicionales. Normalmente los directorios poseen una estructura adecuada para poder “navegar” por

¹¹⁷ Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

ellos mientras que los motores de búsqueda permite únicamente la búsqueda por palabras claves o frases. Mientras que los motores de búsqueda trabajan con una herramienta automática que se suele utilizar una vez al mes (aunque depende del buscador) con el objetivo de indexar las nuevas páginas creadas, los directorios son organizados por un equipo de personas que se dedican a seleccionar los enlaces que incluyen en este directorio, y normalmente los clasifican por materias.

A pesar de esta clara diferenciación, hay que tener en cuenta que cada vez está habiendo una mayor proliferación de alianzas entre motores de búsquedas y directorios, de manera que tanto uno como otro poseen los dos métodos de búsqueda.



Imagen del motor de búsqueda de Google (<http://www.google.com>)



Imagen del directorio de Google (<http://www.google.com/dirhp?hl=esom>)

Tabla comparativa de los motores de búsqueda y los directorios

Directorios	Motores de búsqueda
Normalmente de escaso tamaño debido a que su mantenimiento es manual.	No tiene restricciones de tamaño a causa de su alimentación artificial.
Los enlaces son seleccionados en base a su calidad.	No hay un control de calidad en cuanto a los enlaces que incorpora.
Podría resultar pobre para búsquedas exhaustivas.	Recomendado para las búsquedas exhaustivas ya que la mayoría de ellos permite, además, el uso de herramientas de búsqueda avanzadas.
Normalmente los enlaces de un directorio apuntan al nivel más alto de una web, es decir, a la "home page" o página de inicio de dicha web.	Suelen indexar el texto completo de gran parte, sino todas, de las páginas en cada sitio web.

5.4. BUSCADORES GENERALES DE INTERNET

Son los buscadores que contienen en sus bases de datos todo tipo de *webs* y contenidos, y los predominantes hasta el momento. Existen dos tipos en función de cómo se genera la búsqueda de información: los **directorios**, se accede a los resultados de la búsqueda a través de un directorio formado por categorías y subcategorías o sobre unas palabras clave previamente introducidas en la base de datos, y, los **motores de búsqueda**, que van creando sus bases de datos rastreando Internet y reconociendo los programas "meta tags" de los sitesy/o palabras contenidas en los documentos de cada web. Los buscadores que más se suelen utilizar son: Yahoo, Altavista, Google, Lycos, Infoseek y Excite. Estos grandes buscadores generales, tienden a especializarse. La tendencia entre ellos está siendo la especialización geográfica por países. Del mismo modo, hay también una tendencia a que los directorios incorporen sus propios motores de búsqueda, y que estos últimos incluyan en su web directorios con el fin de facilitar lo más posible la búsqueda al usuario.

Aunque, como hemos podido observar, todos los buscadores tienen una serie de características comunes, cada uno de ellos tiene ciertas particularidades que conviene conocer a la hora de lograr el resultado más preciso. La mayor parte de ellos se puede buscar por palabras clave; otros permiten buscar por preguntas, por ejemplo "¿Cuántos presidentes ha tenido EEUU?", el resultado será una selección de direcciones sobre los presidentes de EEUU.

Para precisar aún más la búsqueda se pueden asociar las palabras clave mediante los llamados operadores que son herramientas que permiten acotar o ampliar la búsqueda.. Aunque hay algunos comunes a todos, cada motor de búsqueda puede tener otros específicos.

Fundamentalmente, existen tres tipos de operadores:

- a) **Operadores lógicos o booleanos** (nombre proveniente del matemático inglés George Boole).
- b) **Operadores posicionales**
- c) **Operadores de exactitud o truncamiento**

a) **OPERADORES LÓGICOS O BOOLEANOS**

Son los siguientes:

AND: Une dos palabras clave y devuelve todos los documentos que contengan ambas palabras aunque estas se encuentren en partes muy distantes del documento. Algunos buscadores utilizan para este fin, el signo "+" o el signo "&"

OR: Une dos palabras clave, pero devuelve todos los documentos que contengan al menos una de las dos palabras. Es útil para introducir sinónimos o palabras en otros idiomas. Algunos buscadores utilizan con este fin, el signo "|"

XOR: Une dos palabras clave, y devuelve aquellos documentos que contengan una u otra, pero no los documentos que contengan las dos

NOT: Permite excluir una palabra clave. Permite restringir la búsqueda. Algunos buscadores utilizan para este fin, el signo "-" o el signo "!"

Estos operadores se pueden combinar de múltiples formas, haciendo uso del paréntesis

b) **OPERADORES POSICIONALES**

Permiten medir la distancia entre las palabras que se intentan buscar en el documento.

Son los siguientes:

- **Near:** El operador de proximidad permite buscar palabras que suelen ir juntas o separadas por pocos términos. Por ejemplo una búsqueda sobre tecnologías de la información y la comunicación. Queremos encontrar páginas sobre periodismo electrónico. Ej: periodismo near (electrónico or digital). Esta consulta me permitiría encontrar documentos sobre periodismo electrónico o digital, siempre que no haya mucha separación entre los dos términos.
- **FAR:** Este operador es lo contrario de NEAR. Al utilizarlo se localizan documentos en los que los términos de la búsqueda aparecen con 25 palabras o más de distancia al menos en un caso. Como esto no elimina la posibilidad de que los términos aparezcan también más cerca dentro de la misma página, FAR es más útil cuando se usa en conjunción con otras expresiones.
- **After** (después): Encuentra palabras en un orden relativo. Por ejemplo: "social" before "seguridad".
- **Before** (antes): Funciona como el operador Y, siendo su única diferencia que los términos deben aparecer en el orden que se especifique, pero pueden encontrarse a cualquier distancia en el mismo documento. Si no está seguro de cómo se escribe Gettysburg, se puede hallar el texto del famoso discurso del presidente de Estados Unidos Abraham Lincoln en aquella ciudad escribiendo lo siguiente: *fourscore **BEFORE** liberty*.
- **Within:** En este caso, el buscador encontrará palabras comprendidas en un radio particular. Por ejemplo: (mujeres) within 5 (violencia) recupera registros que contienen la palabra “mujeres” y “violencia” en cualquier orden y dentro de un

radio de 5 palabras de separación entre ellas. Cualquier número puede ser usado para determinar el radio de proximidad.

- **ADJ.** Se debe utilizar este operador cuando se desee encontrar documentos en los que aparezcan los términos juntos, sea en el orden que sea. Ej: carreras ADJ coches sabiendo que tal búsqueda localizará tanto carreras de coches como coches de carreras.

Las expresiones que incluyen varias palabras adyacentes como por ejemplo: “Primera Guerra Mundial”, deben escribirse entre comillas. Asimismo, algunos motores ofrecen la posibilidad de utilizar lo que se conoce como “comodines” o “Wildcards”, que se simboliza mediante el signo *. Éstos permiten escribir sólo una parte del término deseado (normalmente sus primeras letras). Por ejemplo, realizando la siguiente búsqueda: enfer*, el motor buscará todas las palabras que empiecen por estas letras: enfermo, enfermera, enfermería, etc.

c) **OPERADORES DE EXACTITUD O TRUNCAMIENTO**

Estos operadores suelen usarse como comodines e indican al buscador que, en el espacio que ocupa falta cualquier cadena de caracteres. Ejemplo: car*, nos devolverá: cara, carabina, caramelo, carátula, etc.

También se puede utilizar el asterisco en medio de una palabra. Ejemplo: tor*o, nos devolverá: toro, torso, torero, tornadizo, etc. Algunos buscadores utilizan el signo "\$" .

El signo “?” sustituye a una sola letra. Es útil para buscar palabras en masculino y femenino. Ejemplo: cociner?

Otras herramientas de búsqueda avanzada son las siguientes:

- **Funciones especiales para búsquedas en la Web**

anchor:texto. Busca páginas que tienen un hipere enlace, cuyo texto contiene la palabra o frase especificada.

applet:clase. Busca páginas que contienen la aplicación (applet) de Java especificada

domain:nombre de dominio. Busca páginas en el dominio especificado

host:nombre. Busca páginas en el ordenador especificado

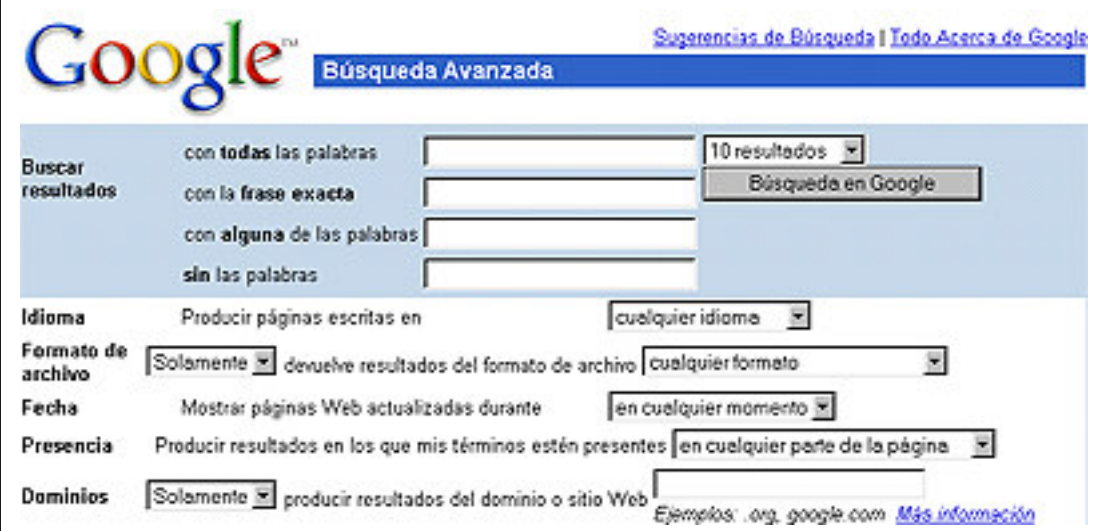
image:nombre de archivo. Busca páginas que contienen imágenes con el nombre de archivo especificado.

link:direccion. Busca páginas con enlaces a la dirección URL especificada

text:texto. Busca páginas que contengan el texto especificado en cualquier parte de la página excepto en su dirección URL, en un hipere enlace o en una imagen tag.

title:texto. Busca páginas cuyo título de página contenga la palabra o frase especificada.

url:texto. Busca páginas que contengan la palabra o frase especificada en su dirección URL.



The image shows the Google Advanced Search interface. At the top left is the Google logo. To its right is the text "Sugerencias de Búsqueda | Todo Acerca de Google". Below this is a blue header bar with the text "Búsqueda Avanzada". The main search area is a light blue box containing several options for searching results: "con todas las palabras", "con la frase exacta", "con alguna de las palabras", and "sin las palabras". Each option has a corresponding input field. To the right of these options is a dropdown menu showing "10 resultados" and a button labeled "Búsqueda en Google". Below the search options are several filter sections: "Idioma" (Producir páginas escritas en cualquier idioma), "Formato de archivo" (Solamente devuelve resultados del formato de archivo cualquier formato), "Fecha" (Mostrar páginas Web actualizadas durante en cualquier momento), "Presencia" (Producir resultados en los que mis términos estén presentes en cualquier parte de la página), and "Dominios" (Solamente producir resultados del dominio o sitio Web). At the bottom right of the filter section, there are examples ".org, google.com" and a link "Más información".

Imagen de la búsqueda avanzada del motor de búsqueda Google (http://www.google.com/advanced_search?hl=es).

5.5. MOTORES DE BÚSQUEDA

5.5.1. FUNCIONAMIENTO DE LOS MOTORES DE BÚSQUEDA

Los motores de búsqueda son bases de datos que contienen el texto completo de páginas web indexado. Cuando un usuario utiliza un buscador, no está buscando en toda la World Wide Web sino en la base de datos específica del motor. Esto permite que se obtengan rápidos resultados, hecho que no sería posible si cada vez que se realiza una consulta se tuvieran que rastrear en tiempo real los millones y millones de documentos que se encuentran disponibles en la Web.

A la hora de crear estas bases de datos, los motores de búsqueda, con una periodicidad normalmente mensual (aunque depende del motor de búsqueda en cuestión) utilizan las llamadas “arañas” o robots inteligentes que “saltan” de una web a otra a través de los enlaces de hipertexto registrando la información allí disponible. Esta base de datos también se alimenta de las *webs* que ceden sus propietarios, los cuales indican al motor de búsqueda, mediante un formulario, la existencia de su servicio. Otro mecanismo de búsqueda que poseen los motores es el que activan los usuarios de Internet. Su función es recuperar información de la base de datos que previamente se ha elaborado mediante el rastreo de la “araña”.

Los motores de búsqueda contienen varias partes diferenciadas¹¹⁸:

- a) El **robot** o **araña** que localiza las páginas web.
- b) El **indexador**, que indexa cada palabra en cada página y lista las palabras indexadas en una enorme base de datos.
- c) El **procesador de búsqueda** que compara la búsqueda del usuario con la lista de palabras indexadas y recomienda los documentos más afines.

¹¹⁸ Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

Los robots o arañas con los que funcionan los motores de búsqueda, son agentes que atraviesan de forma automática la World Wide Web con distintos fines, como pueden ser los siguientes:

- Mantener la estructura hipertextual del Web.
- Realizar análisis estadísticos del Web.
- Descubrir nuevas aportaciones
- Crear de forma automática bases de datos textuales a partir de los documentos HTML distribuidos por los distintos servidores.

Estos robots parten de una lista inicial de URLs. Una vez en esas páginas, van siguiendo los enlaces que en ellas se encuentran, almacenando a su vez en una base de datos parte o la totalidad del texto disponible en las webs que encuentra a su paso, y así sucesivamente. El SGBD o Sistema de Gestión de Bases de Datos, es parecido a las bases de datos documentales. Cuando el robot localiza una página, ésta recibe como entrada el texto localizado y produce como salida un índice invertido. En dicho índice suelen almacenarse las raíces de las palabras y excluirse los términos vacíos. Un ejemplo de índice invertido de las frases “la vida es buena”, “malo o bueno”, “buen amor” y “amar la vida” podría contener identificadores para cada frase (numeradas del 1 al cuatro), y la posición de cada palabra dentro de la frase:

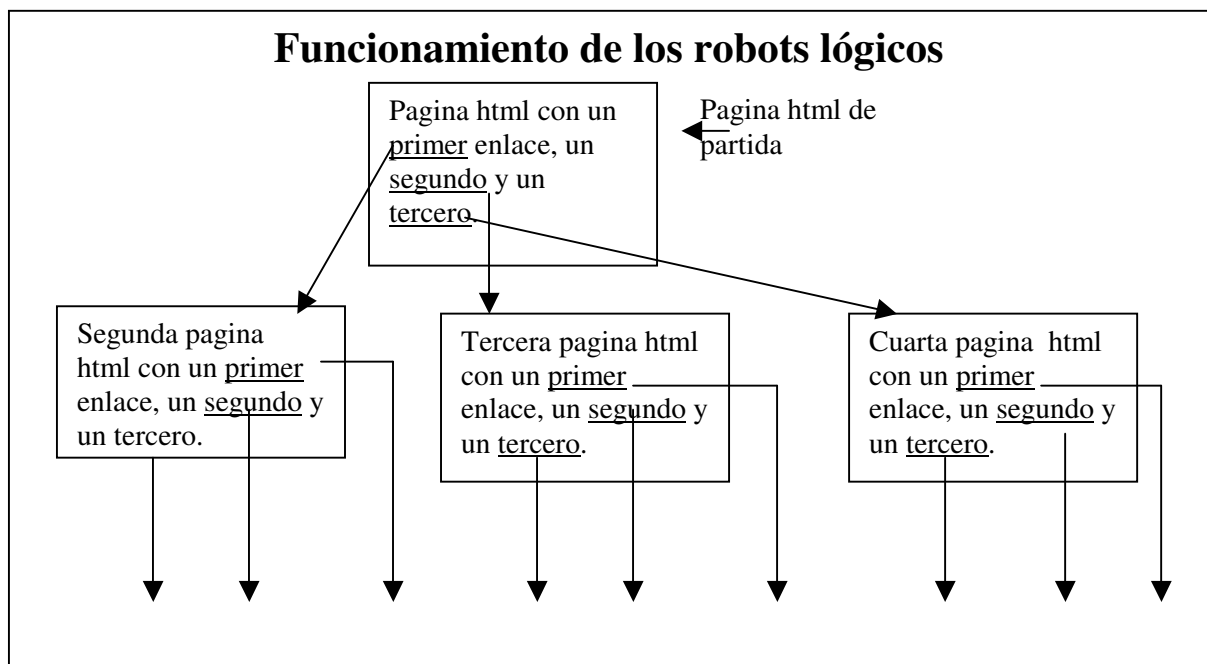
malo	(2,1)		
bueno	(1,3)	(2,3)	(3,1)
es	(1,2)		
vida	(1,1)	(4,3)	
amar	(3,2)	(4,1)	
la	(4,2)		
o	(2,2)		

Para facilitar la búsqueda, algunos motores de búsqueda eliminan las palabras denominadas “stop words” o palabras muertas como pueden ser: “es”, “y” “en”, etc. Asimismo, el índice puede eliminar también la puntuación, los espacios múltiples y convertir todas las letras a minúscula. Algunos motores ahorran espacio en sus índices

cortando palabras hasta que alcanzan su raíz, confiando en que el procesador de búsqueda expanda dicha búsqueda añadiendo sufijos a la raíz de los términos en cuestión.

La indización del texto completo de las páginas web permite al motor llegar más allá de la mera coincidencia de palabras clave. Si se graba la localización de cada palabra, los operadores booleanos como “NEAR”, “AND, y “OR” pueden usarse para limitar las búsquedas. Normalmente el motor puede localizar la página en cuestión buscando por frases completas. Y en caso de que este motor indexe también el código HTML, además del texto de la página, las búsquedas podrán ser delimitadas a campos específicos de la página como pueden ser el título, la URL y el cuerpo.

Hay que tener en cuenta que las bases de datos no contienen los documentos originales, sino únicamente las direcciones de los documentos; por lo que el robot tarda bastante en volver a indexar una página puede ser que ésta esté desactualizada cuando después de un tiempo sea resultado de una búsqueda de un usuario. En este caso, lo más aconsejable es volver a copiar la url (dirección web) en un nuevo navegador. De esta forma se podrá acceder a la página en cuestión sin intermediarios.



La mayoría de los motores de búsqueda pueden encontrar páginas, además de a través de las arañas, mediante una opción de “añadir URL” que permite al autor de una Web notificar al motor de búsqueda la dirección de esta página.

El procesador de búsqueda es la parte más compleja del motor. Está compuesto por varias partes que incluyen: el formulario de búsqueda, la máquina que evalúa una búsqueda y la hace coincidir con los documentos más relevantes en la base de datos en la que están indexadas las páginas, y los resultados de dicha búsqueda.

El formulario y los resultados varían poco de un motor a otro. Todos tienen modelos de búsqueda simples y avanzados, y la mayor parte de los resultados son similares con algunos añadidos como pueden ser webs relacionadas con las páginas encontradas, las búsquedas más populares, etc.

Lo que realmente es lo más significativo de un motor es la forma en que éste calcula la relevancia. Algunos de ellos se basan en análisis estadísticos del texto y desarrollan sofisticados métodos de comparaciones en su tarea de localizar los documentos más relevantes a una búsqueda. Otros analizan el enlace intentando usar la totalidad de la web como medio para conseguir la mayor relevancia posible. El método que utilizan es encontrar los documentos más citados por otros autores de webs para una determinada búsqueda.

Estas fórmulas son el secreto mejor guardado de las compañías dueñas de las máquinas de búsqueda, y sus algoritmos son actualizados constantemente para mejorar la calidad o incorporando las últimas tecnologías contra los “spammers” o aquellos que se dedican a incluir “webs basura” en los buscadores, que entorpece el buen funcionamiento de éstos. Es tarea del profesional de la búsqueda de información en Internet, el conocimiento de las peculiaridades de cada uno de los motores que suele utilizar con el objetivo de seleccionar el tipo de motor más adecuado en cada búsqueda.

La calidad de los robots o motores de búsqueda depende de los siguientes aspectos¹¹⁹:

- La eficacia para descubrir nuevos documentos y mantener la lista de direcciones.
- La cantidad de información que guarda el fichero invertido sobre cada documento.
- La potencia del lenguaje de consulta. Ej: Lycos es uno de los más exhaustivos en número de sedes Web pero el que tiene peor lenguaje de consulta y no indexa el texto completo de los documentos. Web Crawler es el que menor número de documentos indexa pero tiene uno de los lenguajes de interrogación más flexibles.

Los problemas más comunes de los motores de búsqueda pueden ser técnicos, es decir, derivados de una sobrecarga y saturación en los servidores web o por una sobrecarga en la infraestructura de la red de comunicaciones. En segundo lugar pueden ser debidos a la indexación del hipertexto. En este sentido puede deberse a que las unidades de información son más pequeñas y peor definidas; a que es imposible fijar los límites de un documento o a que la heterogeneidad de la información dificulta la creación de índices uniformes. Y en tercer lugar, estos problemas pueden derivarse de la calidad de la entrada de datos debida al mal uso de etiquetas HTML, a una terminología incorrecta en los títulos y encabezamientos y a que no se excluyen documentos de poca importancia.

Aunque su funcionamiento es sencillo, los motores de búsqueda deben estar bien programados para afrontar retos. La mayoría de los motores envían peticiones simultáneas de miles de páginas en las que deben examinar y comparar la nueva URL con la ya existente en sus índices. Debido a lo caro que resulta este proceso la mayor parte de estos motores limitan el número de páginas que visitarán de cada sitio Web. Este método de búsqueda siguiendo los enlaces de hipertexto es muy eficaz según un estudio¹²⁰ de Altavista, Compaq e IBM elaborado en mayo de 2000, en el que se concluían los siguientes puntos:

¹¹⁹ Olmeda Gómez, Carlos. Universidad Carlos III (Madrid)

¹²⁰ Broder, Andrei, et al (2000): *Graph Structure in the Web*. Paper presented at Ninth International World Wide Web Conference. Amsterdam. (Accesible en: <http://www.almaden.ibm.com/cs/k53/www9.final/>).

- Para cualquier fuente y página de destino elegida al azar, la probabilidad de que exista un enlace directo de esta fuente a su página de destino es sólo del 24%.
- Si existe un enlace directo entre las páginas elegidas al azar, es decir, si hay una página web que corresponda perfectamente a la búsqueda con un enlace directo desde otra página, el número de enlaces de media que hay que pulsar hasta llegar hasta ella es 16.
- Si existe un camino indirecto, es decir, que los enlaces pueden seguirse hacia atrás o hacia delante (técnica disponible para las arañas de los motores de búsqueda pero no para una persona que utiliza un buscador), la media para conseguir la página deseada son 6 enlaces de separación.
- Más del 90% de todas las páginas en la Red son localizables entre ellas siguiendo los enlaces hacia delante o hacia atrás, lo cual sugiere que una búsqueda eficiente por parte de los motores de búsqueda puede cubrir la mayor parte de la Red visible, sin embargo en numerosas ocasiones muchas de las páginas que son perfectamente accesibles por las arañas no son añadidas a los índices por razones económicas, temporales (páginas en continuo cambio), etc.

Los motores de búsqueda tienen una serie de métodos propios que los definen, en cuanto a funcionamiento, que pueden ser técnicos o debidos a la elección de los arquitectos o ingenieros que mantienen estos motores. Entre éstos los más importantes son los siguientes¹²¹:

- **Coste de rastreo de webs:**

Durante el rastreo de las webs, el proveedor del motor de búsqueda debe mantener los ordenadores en su estado óptimo para que puedan soportar la cada vez mayor expansión de la Red. Cada vez que la araña localiza una página y la almacena

¹²¹ Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

en la base de datos cuesta dinero. Asimismo, hay también gastos asociados con el proceso de búsqueda, pero generalmente el rastreo de Internet es la parte más cara del mantenimiento de un motor de búsqueda.

Dado el hecho de que, en la actualidad, los recursos de todos los motores de búsqueda son limitados, los responsables de dicho motor han de tomar decisiones para que el coste resultante del rastreo no sea excesivo. Algunos motores limitan el número total de páginas que indexan, descargando las páginas anticuadas cuando se encuentran con las nuevas. Otros limitan la frecuencia del rastreo, por lo que su base de datos es muy probable que se encuentre anticuada. Otros limitan el rastreo a ciertas partes o dominios de la Red intentando no indexar páginas duplicadas.

Cuando un motor de búsqueda decide limitar el rastreo, significa que páginas que potencialmente podían ser incluidas en su índice, no lo están. Podría decirse que estas páginas son parte de la “Red profunda” o “Internet Invisible” que analizaremos en la tercera parte de esta tesis; sin embargo, estas páginas son perfectamente indexables, lo que ocurre es que los motores de búsqueda no las registran a propósito. Los profesionales encargados de este proceso deben establecer una estrategia de búsqueda y ser conscientes de lo que implica limitar el rastreo.

Otro de los problemas con los que se encuentran los motores de búsqueda al realizar los rastreos es una gran cantidad de páginas duplicadas y de basura en la Red. La acción de excluir las “páginas basura” y asegurar la entrada en el índice de todas las páginas relevantes es una tarea tan difícil que se podría llegar a afirmar que pasará bastante tiempo antes de que algún motor de búsqueda logre totalmente tener un índice perfectamente compuesto y comprensible de Red de redes.

- **Rastreadores torpes**

Los rastreadores o arañas son programas muy simples en su nivel más básico. Éstos han sido diseñados sin otra función que encontrar y captar páginas web. Para

descubrir páginas que no han sido indexadas las arañas utilizan los enlaces que localizan en cada página. Si una página web no tiene ningún enlace que apunte hacia ella, el motor de búsqueda no podrá encontrarla a no ser que alguien (el autor o cualquier otra persona) haga saber al motor de su existencia a través del formulario que se encuentra en todas las webs de los buscadores de “Añadir URL”.

Otro problema con el que se encuentran los motores de búsqueda relativo a las arañas, es que a éstas les llevaría demasiado tiempo rastrear la World Wide Web en su totalidad, a pesar de que logran alcanzar millones de páginas al día. El poco tiempo que tienen las arañas para rastrear es también una cuestión importante a tener en cuenta que puede repercutir positiva o negativamente en los resultados de la búsqueda. En primer lugar, normalmente siempre suele haber cierto tiempo entre que la página es publicada hasta que la araña la descubre. En segundo lugar, también suele pasar tiempo desde que la araña indexa por primera vez la página hasta que vuelve a rastrearla buscando nuevo contenido.

Algunas arañas de última generación poseen la propiedad de determinar, en cierta medida, la calidad de las páginas web, o si es una web que cambia mucho por lo que debe ser rastreada con frecuencia.

- **Habilidades y expectativas del usuario**

En numerosas ocasiones, los usuarios tienen unas expectativas poco realistas de lo que se puede pedir a un motor de búsqueda y de lo que éstos pueden llegar a encontrar. El hecho de obtener los documentos que más relevantes para nosotros en una búsqueda entre millones o billones de páginas, únicamente usando unas pocas palabras clave, es una tarea bastante difícil.

Para resolver este problema, algunos motores crean una serie de resultados programados con anterioridad que se adaptan a las búsquedas más populares. Por ejemplo: si un cantante saca una canción muy esperada por el público, el motor puede ser pre-programado para responder a los resultados anteponiendo enlaces a webs sobre su biografía, su discografía y otras webs relativas a ese tipo de música en particular.

Otro método es ajustar los resultados, de manera que los más populares, es decir, las páginas normalmente más buscadas referentes al tema objeto de la búsqueda, aparecen en primer lugar.

- **Respuesta rápida frente a resultados elaborados**

Al igual que en el anterior punto, el usuario también suele tener expectativas creadas en cuanto a la inmediatez del listado de resultados cuando se realiza una búsqueda en cualquier motor. Para poder satisfacer esta demanda de rapidez, muy pocas veces los motores realizan un análisis exhaustivo que les permitiría ofrecer unos resultados más precisos, en el caso de dispusieran de más tiempo. Normalmente se toma el camino más corto, se trunca el total de los resultados y posiblemente son omitidos documentos importantes.

Afortunadamente, gracias a la rápida evolución de la tecnología que permite una cada vez mayor capacidad de procesamiento y al considerable incremento del ancho de banda, los motores pueden usar mejor su técnica y, por tanto, sus resultados, sin tener que sacrificar la rapidez. A pesar de esto, el cada vez mayor crecimiento de la Red, va incluso a más velocidad que la capacidad de los ordenadores y el ancho de banda debido a que a medida que la Red crece, el espacio de búsqueda necesario para recorrerla completamente también se incrementa.

- **Tendencia a registrar sólo texto:**

La mayor parte de los motores de búsqueda de última generación se encuentran altamente optimizados para indexar texto. Si no hay texto en una página, es decir, si sólo hay una imagen o un archivo de audio, el motor no lo indexará. Este tipo de objetos el motor puede recoger algunos detalles como el nombre del archivo y su localización, aunque no mucho más. Sin embargo en la actualidad se está estudiando el método para que comiencen a reconocer este tipo de objetos no textuales.

- **Palabras clave**

Tener claro el concepto de palabra clave, es fundamental a la hora de proceder a realizar una búsqueda.

Los buscadores y algunos directorios de búsqueda utilizan las palabras clave para localizar los enlaces dentro de su base de datos y presentarlos como resultados.

Estos resultados aparecen listados en páginas que contienen 10, 15 o 20 de estos enlaces. Esta forma de presentación varía de acuerdo a cada Buscador.

El quedar colocado entre los primeros significa una correcta utilización de las etiquetas (metatags) por parte del responsable de la web.

El título de la página es la carta de presentación de la misma, es lo primero que ven los posibles visitantes del Web. Las palabras clave que definen el contenido de la página deben aparecer en dicho título. Los Buscadores con robot utilizan el título como elemento básico de indexación. Este título es el que se halla dentro del código HTML entre las instrucciones <TITLE> y </TITLE>, que causa que aparezca como título en la ventana del navegador al cargar la página.

Para hacer figurar el título en diversos buscadores, aparte de la existencia de la etiqueta <TITLE>, algunos motores de búsqueda prefieren emplear el contenido de la metaetiqueta *title*.

Si el buscador encuentra dos o más títulos igual de coincidentes con la palabra o frase buscada, los ordena alfabéticamente, por lo que si el título comienza por alguno de los primeros caracteres de la codificación ASCII tiene muchas posibilidades de aparecer en los primeros lugares de las búsquedas.

El orden alfabético ASCII es:

" # \$ % & ' () * + , - . / 0 1 2 3 4 5 6 7 8 9 : ; < = > ? @ A B C D E F G H I J K L M N O
P Q R S T U V W X Y Z [\] ^ _ ` a b c d e f g h i j k l m n o p q r s t u v w x y z { | } ~

- **La importancia de la relevancia**

Se denomina "relevancia" a la correspondencia e importancia que una página tiene respecto al criterio de búsqueda introducido en un buscador. Los grandes motores de búsqueda, se encaminan hacia mostrar los resultados ordenados por ese factor, tratando que en primer lugar en los resultados aparezcan las páginas más relevantes.

Por "popularidad", se entiende la cantidad de visitas que recibe un web, y en su defecto, al no ser un dato manejable apenas por los motores de búsqueda, se relaciona con la cantidad de enlaces que otros webmasters hacen a un sitio. La popularidad que de un sitio web pueda tener es aún más importancia para muchos robots de búsqueda que las palabras clave.

En definitiva, cuantos más enlaces tenga un sitio web, más posibilidades tendrá de conseguir las primeras posiciones dentro del tipo al que la web pertenezca.

La determinación de la relevancia de una página respecto a una palabra clave en los buscadores, es algo que está en constante evolución. La tendencia es primar la popularidad, teniendo en cuenta a su vez la relevancia de las páginas que enlazan a un determinado web referente a las mismas palabras o frases clave.

Antes, bastaba con incluir las URL o direcciones web en páginas que se dedicaban simplemente a recopilar enlaces para aumentar la popularidad del web. Este sistema está perdiendo eficacia, y cada vez se valora más el origen de los enlaces que su cantidad para determinar la relevancia y con ello la posición en que mostrar una página en los resultados de una búsqueda.

El hecho de que otros sitios web hagan enlaces a una determinada web depende de varios factores:

- En primer lugar, se trata crear un sitio interesante y útil.
 - En segundo lugar, los enlaces no se consiguen en poco tiempo. La clave está en solicitar las altas correctas en los principales buscadores y directorios.
1. Además, también es muy importante el conocimiento del código HTML. El uso de editores web automáticos, aunque ayudan a construir una página "aceptable" afectan el rendimiento que puede alcanzar al dar altas, ya que la mayoría incluye gran cantidad de código "basura" o inservible y en algunos casos hasta impiden que sea indexado correctamente por los motores de búsqueda.

Aunque la forma en la que los buscadores evalúan la relevancia de los documentos que encuentran suele ser uno de los secretos mejor guardados por las empresas propietarias, la tecnología que utiliza el motor de búsqueda Google, es bastante conocida, y nos servirá para mostrar un ejemplo de cómo se decide la relevancia de los documentos:

Google utiliza como sistema de búsqueda y de organización la tecnología PageRank. Esta tecnología aprovecha la naturaleza estructural de la Web, es decir, el modo en que cada página web puede vincularse con cualquier otra de forma inmediata. El buscador evalúa la importancia de una página por los votos que ésta recibe. Google analiza además la página que emite el voto. Los votos que son emitidos por páginas que son "importantes" tienen mayor ponderación. Estos resultados de alta calidad para el usuario, reciben PageRank superior y son colocados más arriba al ordenar los resultados. PageRank es, por tanto el indicador general de importancia de Google y no depende de una consulta específica sino de las características de una página, basada en datos de la Web que el buscador analiza utilizando algoritmos complejos que evalúan la estructura de vínculo.

A continuación exponemos una descripción de los principales motores de búsqueda generales:

- **ALTAVISTA**

(<http://www.altavista.magallanes.net>)

El proyecto de Altavista comienza en el verano de 1995 en los laboratorios de Digital en Palo Alto (California), comenzando a funcionar el 15 de diciembre de 1995. Tras tres semanas de funcionamiento ya recibía dos millones de peticiones diarias. En mayo de 1996 tenía indexadas 30 millones de *webs* y recibía 12 millones de consultas diarias.

Este motor proporciona la posibilidad de efectuar dos tipos de búsqueda: una simple (Simple Search) y una avanzada (Advance Search).

- a) La búsqueda simple se ofrece en la página de bienvenida y se presenta con un formulario simple que permite introducir las palabras clave. El primer menú desplegable permite efectuar una búsqueda o por páginas web o por archivos de los foros Usenet. El segundo menú (Display the Results) permite obtener los resultados de la búsqueda bajo distintos aspectos. A la hora de introducir la palabra clave, Altavista tiene varias posibilidades. En cuanto al formato de las palabras clave, hay que tener en cuenta que las minúsculas y las mayúsculas son importantes. Si escribimos todo con minúscula, Altavista buscará por todas las palabras sin importarle su estilo. Sin embargo, si escribimos alguna de las letras con mayúscula, lo que buscará será la copia exacta de la palabra que se ha escrito. Con los acentos ocurre lo mismo que con las mayúsculas y minúsculas. El comodín * puede utilizarse para reemplazar una o mas letras, pero debe utilizarse después de la tercera letra como mínimo, ej: med* (medico, medicina, etc). Este comodín solo reemplaza hasta cinco letras y no tiene en cuenta ni las mayúsculas, ni las cifras. Por último, si el motor encuentra demasiados resultados para la palabra que contiene el comodín, se ignorará la búsqueda.

Con respecto al formato de interrogaciones, se utiliza, en primer lugar para buscar palabras juntas; en este caso hay que ponerlas entre comillas, ej: "Primera Guerra Mundial". En el caso de que se escriban palabras distintas separadas por espacios, Altavista realiza un búsqueda y

proporciona un resultado según su cálculo de pertinencia. Los operadores booleanos funcionan aquí de la siguiente forma:

- Cuando se realiza una búsqueda enlazada por Y, hay que utilizar el signo +.

- Si lo que se pretende es excluir una palabra de una búsqueda se utilizará el signo -. Ej: Africa +habitantes-geografía.

Altavista permite asimismo realizar búsquedas específicas por partes de los documentos indexados: url, texto, enlaces, imágenes, etc.

- Si lo que se desea es efectuar una búsqueda por un enlace de hipertexto, se utilizará: **anchor:page**. En este caso buscará la palabra *page* sólo si se halla dentro de un enlace de hipertexto.

- Si se busca *applet Java*, se utilizará lo siguiente: **applet: javaclass10**. La búsqueda se concentrará en las páginas que contengan la clase *javaclass10* en los enlaces que lanzan un *applet Java*.

-Con respecto a la búsqueda de una imagen, el término a utilizar es **image:charlot.jpg**. Buscará las páginas que contengan una imagen llamada *charlot.jpg*.

-La búsqueda en la url se realiza de la siguiente forma, **link: elpais.es**. En este caso buscará todas las páginas que contengan enlaces hacia otra página remota con la expresión *elpais.es*.

-Si lo que se pretende es buscar únicamente en el texto, se utiliza **text:Lenin**. Buscará la palabra Lenin sólo en el texto de las páginas (no en los enlaces, títulos, etc)

-La búsqueda por título de página se realiza de la siguiente manera: **title:hola**. Buscará la palabra *hola* en los títulos de los documentos.

Por último, se puede buscar en el interior de un url: **url:casa.html**. En este caso buscará la expresión *casa.html* en el url completo de las páginas fuente.

b) En la búsqueda avanzada de Altavista, se puede utilizar todo lo que hemos especificado en la búsqueda simple, además de los operadores booleanos explicados anteriormente: AND, OR, NEAR, NOT. La zona de “Results Ranking Criteria” permite definir el orden de visualización de datos en la página de resultados.



- **GOOGLE!**

(<http://www.google.com>)

GOOGLE nació en 1998. Sergie Brin y Larry Page, de la universidad de Stanford (EE UU) estaban trabajando en un proyecto de clase para intentar identificar patrones en la estructura de enlaces de la Red. Fue este estudio lo que les dio pie a intentar diseñar un motor de búsqueda basado la estructura de los enlaces.

Originalmente el motor de búsqueda creado fue llamado “Googol” haciendo referencia al número del 10 elevado a 100, lo cual representa el número infinito de documentos de búsqueda en la Red. Tras presentar el proyecto a los inversores, éstos recibieron un cheque a favor de “Google”, por lo que decidieron cambiarle el nombre.

En la actualidad, Google es uno de los mejores motores de búsqueda que hay en la actualidad. Existen versiones en los siguientes idiomas: Alemán, Chino (simplificado), Chino (tradicional), Coreano, Danés, Español, Finlandés, Francés, Holandés, Inglés Italiano, Japonés, Noruego, Portugués y Sueco. Además, reconoce e indexa documentos en formato: PDF, RTF, PostScript, Word, Excel, PowerPoint, entre otros. Cada 28 días este motor indexa 3 billones de documentos web y actualiza más de 3 millones de páginas web importantes cada día.

A continuación se expone una breve descripción de algunos de los campos más relevantes este buscador:

- **Campo de búsqueda**

Para introducir una consulta en Google, hay que escribir algunas palabras clave descriptivas y pulsar el botón Búsqueda en Google para obtener su lista de resultados pertinentes. También se puede buscar una frase usando comillas. Las palabras entre comillas dobles ("como esta") aparecerán juntas en todos los documentos devueltos. Las búsquedas de frases con comillas son útiles cuando se buscan expresiones famosas o nombres específicos.

- **Botón de búsqueda**

Este botón sirve para enviar otra consulta de búsqueda. También puede enviar una consulta presionando la tecla Entrar.

- **Me siento afortunado**

El botón "Me siento afortunado" lleva de forma automática y directa a la primera página Web que Google devolvió en la consulta. Se debe utilizar cuando se está

prácticamente seguro que con la palabra o palabras clave utilizadas se obtendrá una web que es relevante sobre las demás y que es la que se está buscando.

- **Menú desplegable de resultados**

Este buscador permite ajustar el número de resultados que se ve en cada página. Sus opciones son 10, 30 o 100 y el valor predeterminado es 10.

- **Título de la página**

La primera línea del resultado es el título de la página Web encontrada. A veces, en lugar de un título hay una dirección URL. Esto significa que aún no ha indizado la página Web que devuelve Google, de modo que no conoce su título. También podría significar que el autor de la página no le ha puesto un título. Sin embargo, si el texto asociado a estos vínculos coincide con su consulta, podría devolver la página como resultado, aún cuando no se haya indizado.

- **Google y RealNames**

La marca RN (RealNames) que aparece algunas veces en la página de resultados de búsqueda en Google indica una nueva asociación destinada específicamente a mejorar la búsqueda, manteniendo al mismo tiempo la integridad de ésta.

Las palabras clave son nombres de marcas, productos, servicios o compañías que dirigen específicamente al sitio Web del propietario. Por ejemplo, si se escribe "Microsoft", el primer resultado que se obtendrá será el sitio oficial de esta compañía. Cuando la palabra clave coincide con el primer resultado de Google, esta palabra clave y la marca RN aparecerán al final del título. Cuando la palabra clave difiere del primer resultado de Google, aparecerá sobre los resultados de Google.

- **Texto bajo el título**

Este texto es un resumen de la página de resultados devuelta. El texto devuelto no necesariamente es el primer texto de la página, sino un extracto de la página que incluye uno o más de los términos de consulta y aparecerán resaltados en negrita.

- **Línea de estadísticas**

Esta línea indica el número de resultados devueltos para su consulta y la cantidad de tiempo que tardó en realizarse su búsqueda.

- **Resultado con sangría**

Cuando este buscador encuentra múltiples resultados del mismo sitio Web, el resultado más pertinente se indica primero en la lista y las demás páginas pertinentes de ese mismo sitio aparecen con sangría.

- **Dirección URL de resultado**

Esta es la dirección Web del resultado devuelto.

- **Enlace en caché**

Después de la dirección URL aparece un vínculo en caché. Si se pulsa en este enlace, se ve el contenido de la página Web en el momento en que se indexó. Para sitios Web con contenidos que cambian con rapidez, tales como los sitios de noticias, el contenido actual de una página Web puede que no sea una coincidencia adecuada para su consulta. El contenido en caché es el contenido que Google usó para juzgar si esta página es una buena coincidencia para la consulta. Este enlace estará ausente en los sitios que no se hayan indexados y en aquellos cuyos propietarios hayan solicitado quitar el contenido en caché. Al pulsar en este enlace,

, la página resultante tendrá cierta información de encabezado HTTP en la parte superior. Las fechas indican la antigüedad del contenido en caché, pero lo más significativo es que recuerda que esta es una versión en caché de la página y no la página misma. Esta opción es bastante importante ya que si por alguna razón, el enlace del sitio no lo conecta a la página actual, aún podrá recuperar la versión en caché y encontrar ahí la información que necesita.

- **Tamaño**

El siguiente número de la línea es el tamaño de la parte del texto de la página Web. Se omite para los sitios que Google aún no ha indexado.

- **GoogleScout**

Cuando se pulsa en un enlace de GoogleScout para un resultado en particular, Google automáticamente busca en la Web las páginas que se relacionan con este resultado. Generalmente, GoogleScout encontrará aproximadamente una docena de páginas de alta calidad para usted. Cuando GoogleScout busca las páginas relacionadas, intenta encontrar páginas en el mismo nivel de generalidad. Por ejemplo, si la página de inicio es la página principal de una universidad, GoogleScout devolverá las páginas principales de otras universidades. Pero si la página de inicio es el departamento de ciencias de la computación, GoogleScout encontrará departamentos de ciencias de la computación relacionados y no universidades relacionadas. GoogleScout se puede usar para muchos propósitos. Si al usuario le interesa el contenido de un sitio en particular, pero desearía que tuviera más información, GoogleScout puede encontrar sitios con contenido similar con los cuales probablemente no se esté familiarizado. Si está interesado en investigar sobre un campo en particular, a través de esta opción se puede encontrar un gran número de recursos en forma bastante rápida, despreocupándose de las palabras clave que se pueden usar en esos sitios.

Búsqueda avanzada en el buscador Google

La mayoría de los motores de búsqueda tienen opciones de búsqueda avanzada. Aunque muchas de ellas son comunes a todos, existen otras específicas a cada uno de ellos y que merece la pena resaltar. En el buscador Google son las siguientes:

- **Búsquedas de frases**

Con Google, puede buscar frases agregando comillas. Las palabras entre comillas dobles ("casa") aparecerán juntas en todos los documentos devueltos (a menos que sean palabras vacías, palabras comunes especiales que requieren un signo "+"). Las búsquedas de frases con comillas son útiles cuando se buscan expresiones famosas o nombres específicos.

Ciertos caracteres sirven como conectores de frases. Google reconoce guiones, barras, puntos, signos igual y apóstrofes como conectores. Los conectores de frases funcionan como comillas; por ejemplo, madre-política se trata como una frase, aunque las tres palabras no estén entre comillas.

- **Palabras vacías**

Google ignora las palabras y caracteres comunes, conocidos como palabras vacías. Automáticamente no considera términos como "http" y "com," al igual que ciertos dígitos únicos y letras únicas, porque estos términos rara vez ayudan a limitar una búsqueda y pueden disminuir significativamente la velocidad de ésta.

Se debe usar el signo "+" para incluir palabras vacías en la búsqueda. También se puede incluir el signo "+" en las búsquedas de frases.

- **Cómo excluir palabras de su búsqueda**

A veces es útil optar por excluir ciertas palabras o frases, cuando el usuario necesita todos los resultados pertinentes excepto los que contienen cierta palabra o frase. Google admite esta funcionalidad "negativa" con el signo menos ("-"). Se debe usar este signo para excluir en forma determinada un término de la búsqueda. De esta forma Google ignorará todas las páginas que contengan esa palabra.

- **Funciones especiales de búsqueda**

Algunas palabras cuando están seguidas de dos puntos, tienen significados especiales para Google. Actualmente, este buscador admite uno de estos operadores especiales:

La consulta `link: <URL>` muestra todos los vínculos de retorno para una dirección URL determinada; es decir, las páginas que señalan esa dirección URL. Por ejemplo, `link:www.google.com` mostrará todas las páginas que señalan la página principal de Google. No se puede combinar una búsqueda `link:` con una búsqueda de palabra clave normal.

Tecnología PageRank

El buscador Google utiliza como sistema de búsqueda y de organización la tecnología PageRank. Esta tecnología aprovecha la naturaleza estructural de la Web, es decir, el modo en que cada página web puede vincularse con cualquier otra de forma inmediata. El buscador evalúa la importancia de una página por los votos que ésta recibe. Google analiza además la página que emite el voto. Los votos que son emitidos por páginas que son "importantes" tienen mayor ponderación. Estos resultados de alta calidad para el usuario, reciben PageRank superior y son colocados más arriba al ordenar los resultados. PageRank es, por tanto el indicador general de importancia de Google y no depende de una consulta específica sino de las características de una página, basada en

datos de la Web que el buscador analiza utilizando algoritmos complejos que evalúan la estructura de vínculo.

- **FAST**

(<http://www.alltheweb.com>)

Fast, que es desarrollado por la empresa noruega Fast Search a finales de 1999, destaca por su rapidez y por su sistema de entrega de resultados altamente relevantes, basados en algoritmos, mediante los cuales ha logrado construir un banco de datos de 300 millones de páginas, fundamentado en el análisis y preselección de más de 700 millones de páginas visitadas. A su amplitud y rapidez hay que añadir la facilidad de manejo y sus amplias posibilidades de búsqueda avanzada.

Web del buscador Fast (www.alltheweb.com)



The image shows the homepage of the FAST search engine. At the top left, the slogan "All the Web, All the Time™" is displayed. To the right is the "fast" logo, where the letters are stylized and colored (orange, blue, black). Below the slogan and logo is a search interface consisting of a dropdown menu set to "all of the words", a search input field, and a "FAST Search" button. Underneath the search bar are three links: "Help", "Customize", and "Advanced Search". In the center, there are two news snippets: "FAST Mobile Search index grows to more than 10 million wireless web pages" and "FAST launches World's Largest Multimedia Search. Try it Now...". At the bottom, a navigation bar lists various search services: "WEB SEARCH | MOBILE SEARCH | FTP SEARCH | MP3 SEARCH | MULTIMEDIA SEARCH | LISTENING ROOM".

- **NORTHERN LIGHT**

(<http://www.northernlight.com>)

NorthernLight nace en 1995 en Cambridge, Massachusetts. Contiene un amplio índice con más de 200 millones de direcciones. Para el ordenamiento de las páginas tienen en cuenta diversos aspectos. El más importante es el número de veces que aparecen los términos por los que se ha buscado. También son importantes el título de la página, y el número de enlaces que reciba la página. Ofrece la posibilidad de clasificar automáticamente los resultados en carpetas, así como consultar colecciones de artículos de periódicos norteamericanos y algunos archivos y bases de datos públicas de Estados Unidos.

Web del buscador Northernlight (www.northernlight.com)

The screenshot displays the Northern Light search engine interface. At the top left is the 'divine' logo and the Northern Light logo. The main search area includes a search bar with a dropdown menu showing 'Special Collection', 'Invested Reports', and 'MarkIntel Market Research'. Navigation links include 'Power Search', 'Business Search', 'Search News', and 'Invested Search'. A sidebar on the left contains 'My Account', 'My Alerts', and 'My Help'. A central section titled 'Northern Light Content Solutions' lists various sources like 'Scripts', 'Advertising Age', and 'AFX'. A right sidebar features an 'Important Notice' about the acquisition by Divine, Inc. and a 'Special Editions' section with links to 'Banking Industry', 'Computer/Videos', etc.

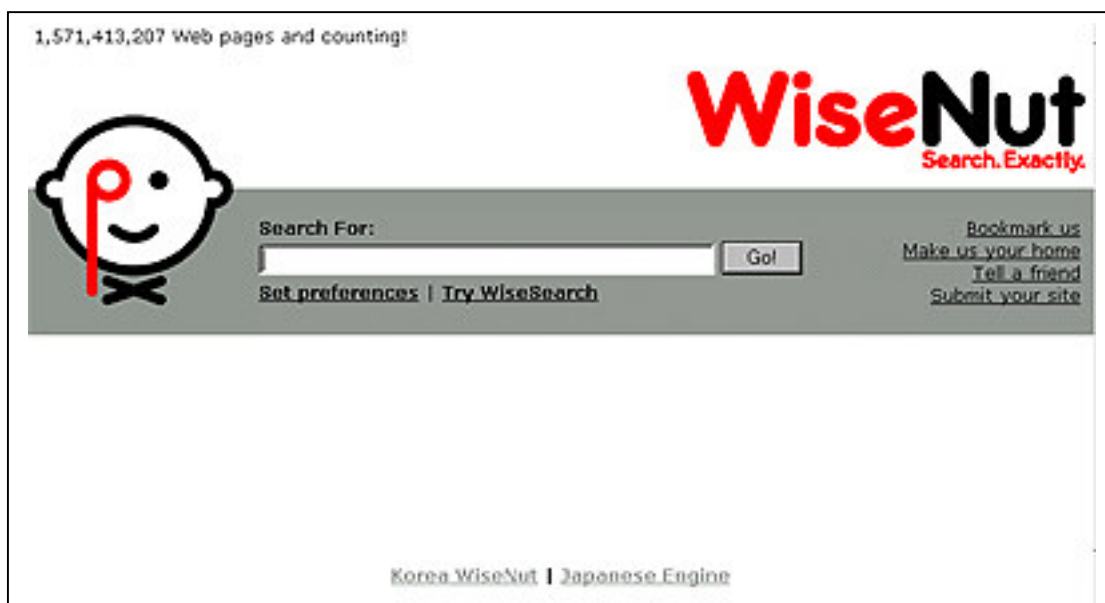
- **WISENUT**

Wisnut fue fundado por Yogirl Yun en 1999 y está ubicado físicamente en Santa Clara, California. Entre sus inversores se encuentran Samsung y Hyundai.

Este motor de búsqueda, incorpora en su base de datos más de mil quinientos millones de páginas indexadas, destacando en la eficacia de los resultados suministrados. Además, muestra hasta tres imágenes de cómo es el sitio de la búsqueda sin necesidad de salir de este buscador. Cuando se muestran los resultados, un enlace final titulado "Sneak-a-Peek" muestra una imagen tomada del sitio al que hace referencia. Asimismo, el enlace "See more 3 pages from this site", te permite ver tres páginas del sitio con sus correspondientes capturas de pantalla. Esta característica es de gran ayuda para saber cómo es un web, sin necesidad de tener que esperar a que se cargue la página.

Wisnut, incorpora un algoritmo inteligente de ordenación de páginas sensible al contexto, evaluando los contenidos, y clasificándolos por la importancia de los mismos.

Web del buscador Wisnut (Accesible en: <http://www.wisnut.com>)

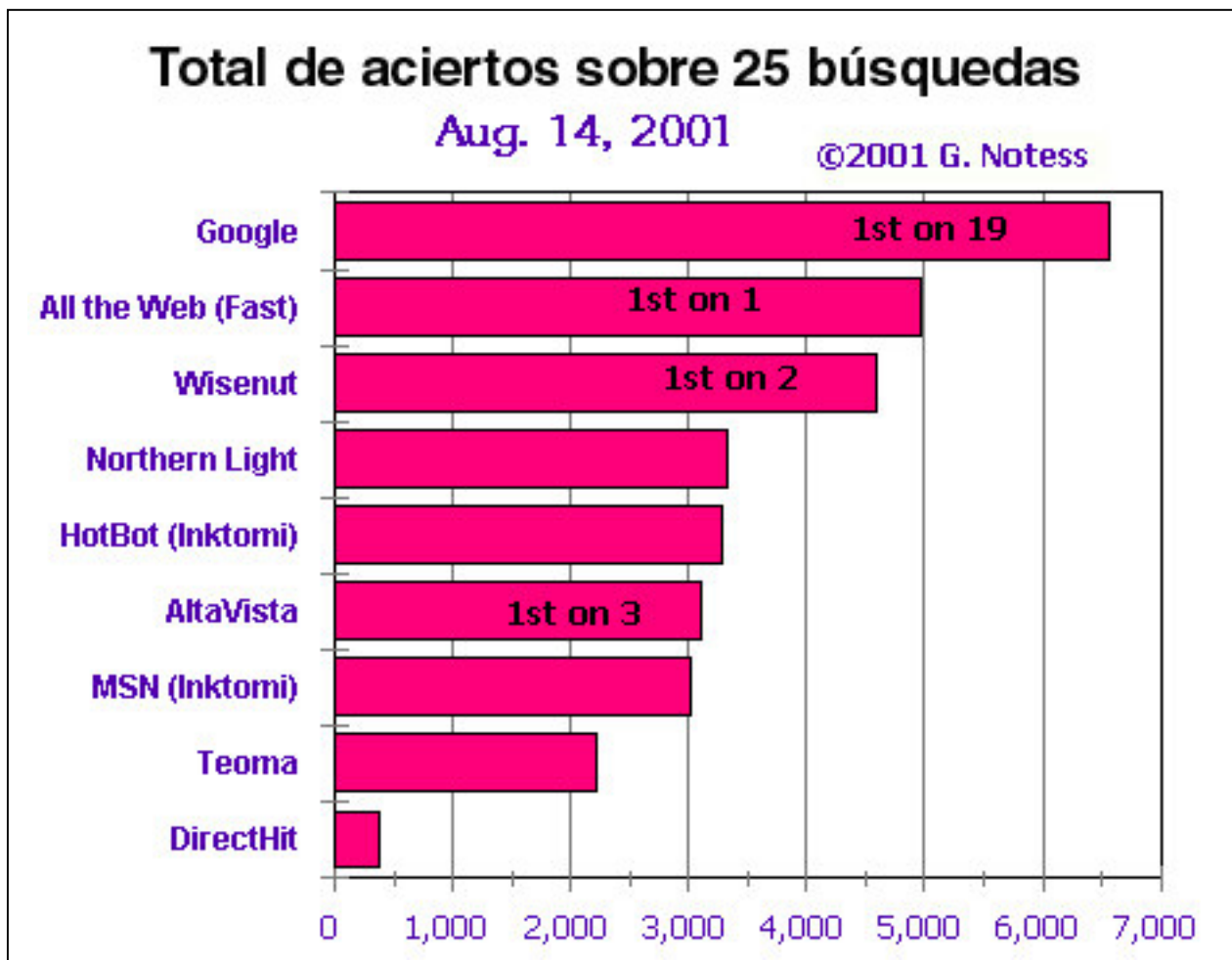


5.5.2. LOS BUSCADORES GENERALES Y EL TAMAÑO DE LA RED

Un estudio de Notess¹²², elaborado el 14 de agosto de 2001, junto con Inktomi, MSN Search y Hotbot, comparaba la capacidad de nueve motores de búsqueda. En este análisis se usaron para realizar la búsqueda 25 palabras simples. Google fue el que obtuvo un mayor número de aciertos. Asimismo, 19 de las 25 búsquedas logró colocar el primero el resultado más óptimo.

La base de datos de Fast, que representa a All The Web, obtuvo el segundo lugar pero tan solo logró que uno de los resultados más óptimos se colocara en primer lugar de la lista.

Wisnut se coloca en tercer lugar. Northern Light en el cuarto y Hotbot en quinto. Por su parte AltaVista obtuvo el sexto lugar, aunque consiguió situar los primeros los resultados más relevantes en dos ocasiones de las 25 búsquedas.



Cuando se analizan estos datos partiendo del total de resultados verificados de las 25 búsquedas, la base de datos de Google logra el primer lugar. En la tabla expuesta a continuación se muestra el número total de aciertos de cada motor de búsqueda:

Google	6,567
All the Web	4,969
WISEnut	4,587
Northern Light	3,321
HotBot	3,277
AltaVista	3,112
MSN Search	3,005
Teoma	2,219
Direct Hit	381

En el siguiente gráfico podemos ver como el motor Google lidera el número de aciertos. En segundo lugar encontramos a Fast y en tercero a Wisenut.

5.5.3. METABUSCADORES

Los “metabuscadores” son aquellos permiten realizar una búsqueda en varios buscadores a la vez. Uno de sus inconvenientes, además de un mayor tiempo de espera, es que no es posible precisar la búsqueda ya que cada uno de los motores que engloba tienen sus propias peculiaridades de búsqueda. Su método de funcionamiento es el siguiente: cuando el usuario realiza una búsqueda, el metabuscador las dirige a sus motores asociados componiendo una lista de aciertos que representan, teóricamente, las mejores respuestas a la pregunta. Posteriormente algunos ofrecen la posibilidad de

ordenar por relevancia (entendiendo por relevancia el grado con que la web resultante concuerda con la búsqueda realizada por el usuario o con la información que necesita. El cálculo de la relevancia que realiza el motor de búsqueda implica varios factores, normalmente es expresada en un porcentaje al lado de cada enlace proporcionado por el motor), formatear los resultados de forma consistente, verificar la accesibilidad o eliminar enlaces muertos. Podemos encontrar distintos tipos de buscadores:

-Aquellos que no agrupan los resultados. Se debe revisar el listado resultante de cada uno de los motores en los que se ha realizado la búsqueda. En este caso hay muchas posibilidades de que existan webs duplicadas. Un ejemplo de este caso es Dogpile (<http://www.dogpile.com>).

-Aquellos que agrupan los resultados. Suelen ser los más numerosos. Poseen la ventaja de que eliminan los duplicados. Un ejemplo es Metacrawler (<http://www.metacrawler.com>).

Los más importantes, a nuestro juicio, son los siguientes:

- **COPERNIC**

(<http://www.copernic.com>)

Copernic, producto de la empresa canadiense Copernic Technologies, es un buscador múltiple que transfiere una ecuación de búsqueda a un conjunto de buscadores de manera simultánea, recupera las referencias pertinentes y las ordena según el grado medio de relevancia obtenido de cada uno de los buscadores. La empresa creadora de *Copernic* se constituyó en 1996 con el propósito de desarrollar soluciones innovadoras en tecnologías de agentes de búsqueda para conseguir un acceso y una gestión eficientes de la gran masa de información presente en Internet. El programa se desarrolló el año siguiente en una versión experimental, y se puso al alcance del público en noviembre de 1998. Desde entonces se han lanzado versiones con elementos nuevos, la mayoría de los cuales, no obstante, sólo están disponibles en las ediciones comerciales. La edición

gratuita ofrece los servicios básicos de consulta y, con la intención de persuadir a los posibles compradores de las ediciones comerciales, da una idea bastante aproximada de las posibilidades del programa completo. Los tipos de Copernic comercializados son los siguientes:

Copernic Basic, la edición gratuita, permite consultar simultáneamente cerca de 80 buscadores importantes agrupados en siete categorías: “El web” más una categoría opcional relacionada con un idioma o un país, “Grupos de noticias”, “Direcciones de correo electrónico”, “Compra de libros”, “Compra de hardware” y “Compra de software”. A pesar de ser la edición reducida de un programa comercial, ofrece tantas o más posibilidades que las versiones completas de otros buscadores múltiples como *Lexibot*, *NetAttaché Pro* o *Internet EZ Search*.

Copernic Plus, la edición comercial más económica, permite acceder a más de 1.000 fuentes de información agrupadas en 90 categorías de búsqueda especializada, algunas de las cuales son bastante irrelevantes; he aquí algunos ejemplos: “Enciclopedias”, “Salud”, “Multimedia”, “Ciencias”, “Negocios y finanzas”, “Descarga de software”, “Cine”, “Artículos sobre las tecnologías”, “Música”, etc.

Copernic Pro, la edición comercial más completa, ofrece, además, otras prestaciones interesantes: actualización de búsquedas programadas, servicio de alerta y verificador ortográfico de las búsquedas.

- **DOGPILE**

(<http://www.dogpile.com>)

Permite realizar las búsquedas en veintitrés buscadores a la vez con bastante rapidez. Además permite buscar específicamente en grupos de noticias, servidores FTP, etc. Fue

desarrollado por Aarón Flin. Dogpile admite operadores booleanos a la hora de refinar la búsqueda.

- **METACRAWLER**

(<http://metacrawler.com>)

Es uno de los metabuscadores más completos. Permite buscar en los siguientes buscadores: Opentext, Lycos, WebCrawler, Infoseek, Excite, Inktomi, Galaxi y Yahoo.

- **METAINDICE**

(<http://www.metaindice.com>)

Es el primer metabuscador hispano que activa 23 buscadores de recursos exclusivamente en español. Uno de sus principales inconvenientes es que no permite buscar frases exactas.

- **C4**

(<http://www.c4.com>)

Permite búsquedas booleanas, por proximidad y opciones personalizadas. Ordenación por ponderación.

- **IXQUICK METASEARCH**

(<http://ixquick.com>)

Busca en directories, motores y multimedia. Ordena por relevancia. Múltiples tipos de búsquedas.

- **PROFUSIÓN**

(www.profusion.com)

Busca en 9 máquinas o en las tres más rápidas. Permite seleccionar ponderación.

Web del metabuscador Dogpile (www.dogpile.com)

DÖGPİLE®
All results, no mess.

Local Search: Fetch

Select: The Web

[Yellow Pages](#) [White Pages](#) [Classifieds](#)

Doggy Daily
[Looking for warmth this winter? Check the weather.](#)

Joke of the Day
[Joke of the Day](#)

Shopping
[Go Shopping](#)

Stores
[AbtElectronics.com](#)
[Amazon.com](#)
[gap.com](#)
[Spiegel](#)

Departments
[Apparel & Accessories](#)
[Electronics](#)
[Gifts & Flowers](#)
[Sporting Goods](#)

Father's Day

Dogpile References

Favorite Fetches

Autos	Dating	Insurance Quotes	New Domains
Boats	Gambling/Casino	Legal	Psychic Readings
Car Rental	Health	Loans	Travel
Data Recovery	Help Desk Software	Mortgages	Further Fetches...

Search Tools

Maps	Jobs/Careers	City Guides	Find Friends
Weather	Horoscopes	Health Center	Classifieds
Stock Quotes	Singles Pics ♥	Gov't Listings	Health Insurance

Web Directory

Auto	Reference	Entertainment	Travel
Food & Wine	Sports	Lifestyle	Work & Money
Health	Computing	Personal	
Home	Connecting	Shopping	

Download cool FREE content ...

5.5.4. SOLAPAMIENTO DE LOS MOTORES DE BÚSQUEDA

La Red de redes suele ser descrita como un enorme pajar de información en el que los usuarios de Internet intentan encontrar una pequeña aguja. El doctor Matthew Koll, pionero en el estudio de la Red profunda o Internet invisible que analizaremos en la siguiente parte de esta tesis, señalaba que hay varias maneras de que el motor de

búsqueda pueda hallar la aguja dentro del pajar. Se han de considerar los siguientes aspectos¹²³:

- Una aguja conocida en un pajar conocido.
- Una aguja conocida en un pajar desconocido.
- Una aguja desconocida en un pajar desconocido.
- Cualquier aguja en un pajar.
- La aguja más afilada en un pajar
- Muchas de las agujas afiladas en el pajar.
- Cosas como agujas en cualquier pajar.
- Todas las agujas en un pajar.
- Afirmación de que no hay agujas en el pajar.
- Hazme saber cuando aparece una nueva aguja.
- ¿Dónde se encuentran los pajares?
- Agujas, pajares... cualquier cosa.

Como se ha señalado con anterioridad, los motores de búsqueda no indexan la Red en su totalidad. Se podría pensar que si se combina de los mayores motores se obtendría se obtendría un mapa de la Web mucho más completo; sin embargo esto no sería cierto según Greg Notess¹²⁴

¹²³ Koll, Matthew (1998): Major trends and Issues in the Information Industry. ASIDIC Newsletter. (<http://www.asidic.org/techsumf99.html>).

¹²⁴ Notess, Greg (2000): Search Engine Statistics: Database Overlap. (Accesible en: <http://searchengineshowdown.com/stats/overlap.shtml>).

6.5. DIRECTORIOS

Los directorios temáticos generales se diferencian de los motores de búsqueda en que los primeros son elaborados y organizados por personas. Ésta es la razón por la que ofrecen su base de datos por temas clasificados jerárquicamente. No precisan de motor de búsqueda ya que descendiendo por categorías y subcategorías se puede llegar a la información deseada. Permite a los responsables de un sitio web registrarse mediante un formulario.

Entre las ventajas de los directorios se encuentra, además de la calidad de las *webs* indexadas, la posibilidad de hacerse una idea de los sitios mas relevantes sobre un determinado tema. Dos de sus principales inconvenientes es su lentitud y su reducido catalogo, si los comparamos con un motor de búsqueda. Además, en la mayor parte de los directorios hay un alto índice de páginas que ya no existen y que continúan siendo listadas porque no han sido dadas de baja. A la hora de buscar en un directorio hay que tener también en cuenta que en numerosas ocasiones, el orden en que están colocados los enlaces responden a un criterio comercial. Hay empresas que pagan porque su web ocupe los primeros lugares, ya que está demostrado que de esta forma aumenta su tráfico.

A continuación exponemos una explicación de los directorios más usados:

- **YAHOO!**

(<http://www.yahoo.com>)

Yahoo se creó en abril de 1994 por iniciativa de dos estudiantes de la universidad de Standford, David Filo y Jerry Yang. Yahoo! significa: Yet Another Hierarchical Officious Oracle. A comienzos de 1995 Marc Andreessen, cofundador de Netscape

ofrece a los creadores de Yahoo la posibilidad de instalar su motor de búsqueda en los ordenadores de su empresa.

Los empleados de Yahoo examinan páginas web y recursos de todo el mundo y las incluyen en esta “guía temática universal”. Una vez examinado el material, incorporan cada una de las páginas en una categoría predeterminada, hacen un pequeño resumen de su contenido y lo publican en el catálogo general de más de 20.000 categorías y cientos de miles de páginas documentadas. Por tanto Yahoo no almacena páginas, sino que proporciona los enlaces precisos. Normalmente cataloga de forma prioritaria los documentos de preguntas frecuentes (FAQ) y los recursos que contienen índices de cualquier tipo. Ofrece además algunos detalles de interés para el usuario, como por ejemplo, las nuevas incorporaciones a la base de datos aparecen con un icono destacado (“New”) durante unas semanas. También se suele utilizar otro icono (unas gafas rojas, “Cool”) si el sitio es especialmente interesante.

Además de la posibilidad de navegación por categorías y subcategorías, Yahoo ofrece la posibilidad de buscar información por palabras clave. Sus ventajas principales son su sencilla y atractiva interfaz y su facilidad de uso.

- **OPEN DIRECTORY**

(<http://www.dmoz.org>)

Directorio de recursos de Internet donde prima la calidad sobre la cantidad. No es propiamente un buscador, pues no busca páginas automáticamente, sino que se nutre de las direcciones de recursos enviadas por los propios autores de las páginas. Su mayor virtud es que las páginas, antes de ser incluidas en el catálogo, son revisadas por un editor experto en el tema.

- **GALAXY**

(<http://galaxy.Einet.net/galaxy.html>)

Galaxy es un directorio temático general al que se puede recurrir para localizar información de referencia y enlaces relevantes. Si no se logra encontrar lo que se busca por temas, es posible buscar también por palabras clave.

Web del directorio Yahoo (www.yahoo.com)

Calendar Messenger Check Email **YAHOO!** What's New My Personalize Help

[claim-your-name.com](#) claim it before it's gone Enter to win the Ultimate Outdoors Trip. [Weight Control Problems?](#)

Search [advanced search](#)

Y! Shopping Depts: [Books](#), [CDs](#), [DVDs](#), [Flowers](#) Stores: [Ciraque](#), [Avon](#), [Barnes&Noble](#), and more

Shop [Auctions](#) · [Classifieds](#) · [PayDirect](#) · [Shopping](#) · [Travel](#) · [Yellow Pages](#) · [Maps](#) **Media** [Finance](#) / [Quotes](#) · [News](#) · [Sports](#) · [Weather](#)

Connect [Careers](#) · [Chat](#) · [Clubs](#) · [Experts](#) · [GeoCities](#) · [Greetings](#) · [Mail](#) · [Members](#) · [Messenger](#) · [Mobile](#) · [Personalize](#) · [People Search](#)

Personal [AddrBook](#) · [Briefcase](#) · [Calendar](#) · [My Yahoo!](#) · [Photos](#) **Fun** [Games](#) · [Kids](#) · [Movies](#) · [Music](#) · [Radio](#) · [TV](#) [more...](#)

Yahoo! Auctions - Bid, buy, or sell anything!

Categories	Items
· Antiques	· Computers
· Cameras	· Electronics
· Coins	· Sports Cards
· Comic Books	· Stamps
· Barry Bonds	
· Tiger Woods	
· PlayStation 2	
· MP3 Players	
· My Auctions	
· Morgan Dollies	
· Longaberger	
· Palm Pilots	
· Baseball Cards - McOwire , A-Rod , Jeter , Bonds , Sosa , Griffey Jr. , Ichiro	

In the News

- [McVeigh to seek execution delay](#)
- [Palestinian minister Hussein dies of heart attack](#)
- [CDC: HIV-AIDS spreading rapidly among young, gay men](#)
- [FBI: Violent crimes up slightly](#)

[more...](#)

Marketplace

- [Y! Autos](#) - buy, sell, maintain & worship
- [Y! Store](#) - build an online store in

Arts & Humanities
[Literature](#), [Photography](#)...

Business & Economy

News & Media
[Full Coverage](#), [Newspapers](#), [TV](#)...

Recreation & Sports

5.7. ANILLOS WEB

Los *webRing* o anillos web son grupos de webs relacionadas sobre un mismo tema. Cada Web inserta en sus páginas un código de programación gracias al cual aparecen agrupadas distintas webs por temas. Existen anillos sobre, Informática, Marketing, Medicina, Vídeo Juegos, Alpinismo, y un largo etcétera.

El WebRing fue creado por Sage Weil, a partir del proyecto EUROPA¹²⁵ (Expanding Unidirectional Ring of Pages), que era un conjunto de páginas enlazadas a modo de anillo. Este proyecto había sido ideado por Denis Howe, en el Imperial College de Londres.

Pero el concepto de WebRing, nació cuando Weil mejoró la idea implantada en EUROPA. Lo que hizo fue enlazar las páginas en un anillo de web's utilizando un script CGI centralizado. De esta forma, bastaba asignar en cada una de las Web's de cada anillo, una dirección al WebRing central, y ésta servía de enlace a la siguiente Web, con lo que se eliminaba la necesidad de actualizar los tags con cada nueva incorporación.

El script CGI comenzó a funcionar en Agosto de 1.995, y un mes más tarde, después de pequeñas modificaciones, se fueron creando más Anillos. Poco más tarde apareció el primer anillo temático, el cual estaba dedicado a la enseñanza del Inglés. La web¹²⁶ de Weil no tardó en hacerse famosa. Más tarde sería vendida a Starseed (<http://www.starseed.com>).

¹²⁵ Accesible en: <http://www.wombat.doc.ic.ac.uk/europa.html>

¹²⁶ WebRing – Accesible en: <http://www.webring.com>

WebRing News

- [News April 1, 2002](#)
- [Logo Design Contest](#)

Featured Rings

Plan a Spring trip with [Small Hotels Around the World](#), or meander [US Highway/Route 66](#).

Still too Wintery? Find past generations at [AAA Genealogy](#), or work with the next in [Parenting](#).

WebRing Numbers

View system traffic [here](#).

Registered with WebRing?

User ID

WebRing: Your place on the web

- Are you **looking** for something? You can easily visit and explore concentrations of web sites. Simply search or browse the WebRing directory below.
- Use WebRing to participate in highly focused and rich online communities.
- Do you have a web site of your own you'd like to share with others?
 - Try [joining](#) a Ring and increase traffic to your site!
 - Or, [create](#) a Ring of your own and develop an online community.

[Business & Finance](#)
Investments, Industry

[Hobbies & Crafts](#)
Antiques, Collecting

[Computers & Internet](#)
Network, New Access

[Music](#)
Rock & Pop, Rap, Punk

Top Rings

- [Yahoo! Messenger Skins Ring](#)
- [Midi Ring](#)
- [AraAmerican Web Ring](#)
- [Mazda Rotary WebRing](#)
- [The PalmPilot Web Ring](#)
- [The Live WebCam Ring](#)
- [Anime Music Videos](#)
- [Ring of Cartoon Dolls](#)
- [The KISS DollRing](#)
- [The PICmicro ring](#)
- [more...](#)

Hot Rings

- [Yahoo! Messenger Skins Ring](#)
- [Midi Ring](#)
- [The Live WebCam Ring](#)
- [Anime Music Videos](#)
- [Ring of Cartoon Dolls](#)
- [... Always Greatest Websites...](#)
- [The KISS DollRing](#)
- [Mazda Rotary WebRing](#)
- [The PalmPilot Web Ring](#)
- [The PICmicro ring](#)
- [more...](#)

Top NEW Rings

- [Apple Anton Ghoo](#)
- [Photographers on Line](#)

Página web: WebRing (<http://www.webring.com>)

Un WebRing, no deja de ser una alternativa a los Motores de Búsqueda o Buscadores. Cuando se accede a uno de ellos, por lo general es que se desea obtener webs sobre un tema muy concreto. Al realizar la búsqueda, el "ruido" que hay en la base de datos del Buscador, ofrece una lista más o menos larga de Web's relacionadas con el tema motivo de nuestra consulta. El problema es que por lo general esta lista es demasiado larga y por lo general un porcentaje relativamente elevado de las direcciones ni tan siquiera corresponden al tema concreto de la búsqueda ya que dichas Web's han sido optimizadas al ser dadas de alta para aparecer en la mayor cantidad posibles de búsquedas.

Sin embargo, en un WebRing o Anillo Web, tan sólo están aquellas páginas relacionadas con el tema del Anillo, por lo que si se accede a un Anillo relacionado con el deporte, obtendremos sólo información de Web's que tratan de deporte.

Asimismo, otro beneficio igualmente importante lo obtienen los responsables de las Web's que forman parte de cada uno de los Anillos, ya que el Tráfico que se genera desde el Anillo al que pertenecen, es un Tráfico claramente interesado en el tema de su Web. Con ello lo que se consigue una Audiencia definida.

A la hora de decidir cual es el mejor método de búsqueda a utilizar, debemos tener en cuenta las limitaciones de cada uno de ellos. En concreto, las de los Anillos web son las siguientes:

- Es necesario darse de alta en un Anillo para formar parte del mismo.
- Si no se está dado de alta una web en cuestión nunca podrá ser localizada en las bases de datos de los WebRing's.
- Dado que no todas las páginas están conectadas o se conectarán a un WebRing, para rastreos o búsquedas exhaustivas es necesario recurrir a los Motores de Búsqueda.
- Si lo que se desea buscar, es un Web en concreto, los Anillos no son útiles.
- Sólo se pueden buscar términos genéricos o específicos o definidos, pero no por nombres o por dominios.

Otras webs que ofrecen este método de búsqueda son las siguientes:

- Web Ring - <http://www.webring.com> - En Inglés.
- The Rail - <http://www.therail.com> - En Inglés.
- Looplink - <http://www.looplink.com> - En Inglés.
- Web Tower - <http://www.salamander.com/~hexagon/tower> - En Inglés.
- Ring Surf - <http://www.ringsurf.com> - En Inglés.
- Spanish Ring - <http://www.spanishring.com> - En Castellano.

5.8.PORTALES HORIZONTALES, VERTICALES Y CANALES TEMÁTICOS

Los portales son sitios Web que actúan como distribuidores a otros sitios vinculados a un tema o campo. Se distinguen dos tipos: los portales horizontales y los verticales. Los primeros ofrecen información y servicios a un universo amplio. Sin embargo, los portales verticales o vortales, ofrecen información específica y seleccionada sobre un tema.

Por su parte, los canales temáticos se centran en materias específicas que suelen interesar a un amplio número de usuarios. Se diferencian de los portales verticales principalmente en que estos últimos tienen una dirección común a varios de ellos, que se suelen integrar en un portal generalista, mientras que los canales temáticos son independientes empresarialmente hablando.

A continuación se expone una muestra como ejemplo:

- **YAHOO!**

(<http://www.yahoo.com>)

Portal horizontal (generalista)

- **TERRA**

(<http://www.terra.es>)

Portal horizontal (generalista)

- **INVERTIA.COM**

(<http://www.invertia.com>)

Es el portal vertical especializado en economía de Terra (<http://www.terra.es>). Posee además una gran variedad de utilidades.

- **CANAL SALUD**

(<http://www.canalsalud.com>)

Portal vertical del grupo PRISA “Inicia”, dedicado a la salud.

- **NOSOTRAS.COM**

(<http://www.nosotras.com>)

Es un canal temático dirigido a las mujeres.

5.9. OTROS RECURSOS DE BÚSQUEDA

5.9.1. GUIAS TEMÁTICAS ESPECIALIZADAS

Cuando el objetivo de la búsqueda es muy concreto, más que recurrir a motores de búsqueda o directorios, es recomendable acudir a consultar las guías temáticas especializadas en el tema concreto de la búsqueda. Estas guías suelen ser elaboradas por expertos por lo que suelen ser bastante precisas y tienen una buena actualización y categorización de los recursos.

Algunas de estas guías especializadas son las siguientes:

- **A COLLECTION OF SPECIAL SEARCH ENGINES**

(<http://www.leidenuniv.nl/ub/biv/specials.htm>)

Página de la biblioteca de la Universidad de Leiden (Holanda) con cientos de enlaces a buscadores, índices y guías generales y especializadas.

- **NEWSPAPER ARCHIVES ON THE WEB**

(<http://metalab.unc.edu/slanews/internet/archives.html>)

Esta guía elaborada desde una universidad norteamericana, reúne varios centenar de enlaces directos a archivos en línea de periódicos tanto de Estados Unidos como del resto el mundo.

- **RECURSOS DE MEDICINA EN INTERNET**

(<http://est.unileon.es/medicina>)

Guía especializada en medicina. Incluye cientos de direcciones, bien estructuradas, de páginas *web* y otros servicios de Internet sobre medicina.

5.9.2. BUSCADORES ESPECIALIZADOS

Los buscadores especializados, como su propia palabra indica, se centran en un tema concreto. Son muy útiles ya que tienen la ventaja de que la búsqueda se realiza sobre un tema específico sin necesidad de que el usuario lo indique. Por ejemplo: si se está buscando algo sobre el Sol en un buscador especializado en Astronomía, el usuario en cuestión tiene la seguridad de que los enlaces resultantes serán *webs* sobre nuestra estrella relacionadas íntimamente con la Astronomía. Asimismo, algunos buscadores especializados se centran en otro tipo de contenidos, como pueden ser las noticias, biografías, empresas, o expertos en una materia determinada.

Aunque estos buscadores tienden a ser bastante rigurosos, puede ser también que no sea así. Para asegurar la búsqueda se debe consultar también los buscadores generales.

A continuación se expone una clasificación con ejemplos ilustrativos de los buscadores especializados más comunes y que más pueden ayudar al profesional de la información a la hora de realizar su trabajo.

➤ **BUSCADORES ESPECIALIZADOS POR ÁREAS DEL CONOCIMIENTO**

Permiten realizar búsquedas por temas específicos, lo que permite precisar la búsqueda a un campo del conocimiento. Son de gran utilidad para los periodistas especializados. A continuación se ofrecen algunos ejemplos:

- **SCIRUS**

(<http://www.scirus.com/>)

Potente herramienta de búsqueda de recursos relacionados con la ciencia y la investigación científica. Divide los resultados entre páginas de acceso libre y las de pago, que en su mayoría pertenecen a publicaciones periódicas.

- **BIOFORUM**

(<http://www.bioforum.net/links/>)

Buscador y directorio temático hispano especializado en sitios sobre naturaleza, medio ambiente y ecología.

- **BUSCAFINANZAS**

(<http://www.buscafinanzas.com>)

Buscador hispano de recursos sobre finanzas y negocios en general. Posee también un directorio temático de recursos.

- **BUSCADEPORTE**

(<http://www.buscadeporte.com/>)

Buscador temático español de deportes: deportes de invierno, acuáticos, ciclismo, motor, gimnasia, atletismo, baloncesto, fútbol, etc.

Buscador especializado en ciencia e investigación científica Scirus

(www.scirus.com)

for scientific information only | [Customize Scirus](#) | [About Scirus](#) | [Submit URL](#) | [Contact Scirus](#) | [Help](#)

SCIRUS Search [Advanced Search](#)

New: [MEDLINE](#) citations now available on Scirus

All of the words All Sources [Search](#)

Subject Areas

<input checked="" type="checkbox"/> All	<input type="checkbox"/> Earth and Planetary Sciences	<input type="checkbox"/> Mathematics
<input type="checkbox"/> Agricultural and Biological Sciences	<input type="checkbox"/> Economics, Business and Management	<input type="checkbox"/> Medicine
<input type="checkbox"/> Astronomy	<input type="checkbox"/> Engineering, Energy and Technology	<input type="checkbox"/> Neuroscience
<input type="checkbox"/> Biosciences	<input type="checkbox"/> Environmental Sciences	<input type="checkbox"/> Pharmacology
<input type="checkbox"/> Chemistry and Chemical Engineering	<input type="checkbox"/> Life Sciences	<input type="checkbox"/> Physics
<input type="checkbox"/> Computer Science	<input type="checkbox"/> Materials Science	<input type="checkbox"/> Social and Behavioral Sciences

➤ **BUSCADORES DE NOTICIAS**

Son buscadores automáticos que indexan diariamente noticias publicadas por medios de comunicación en Internet. Gracias a este tipo de buscadores, es posible buscar todas las noticias que se publican en los medios de comunicación digitales sobre un mismo tema.

Algunos de ellos son los siguientes:

- **TITULARES.COM**

(<http://titulares.com>)

A través de palabras clave, localiza titulares de noticias del día en más de 70 ediciones digitales de periódicos en español.

- **PRENSA DIGITAL**

(<http://www.prensadigital.com>)

Selecciona a diario noticias de interés general y artículos publicados en periódicos digitales españoles.

- **INFOJUMP**

(<http://www.infojump.com>)

Buscador de noticias de periódicos y artículos de revistas. Afirma poseer una base de datos de 5 millones de artículos procedentes de más de 4.000 publicaciones.

➤ **BUSCADORES DE EMPRESAS**

Existen algunos buscadores de empresas que pueden facilitar el trabajo del periodista a la hora de localizarlas en Internet. A continuación se exponen algunos ejemplos:

- **TELEFÓNICA: PÁGINAS AMARILLAS**

(<http://www.paginas-amarillas.es>)

Desde esta página se puede acceder a la base de datos de las páginas amarillas de todas las provincias de España.

- **SPAIN INDUSTRY: BUSCADOR DE EMPRESAS ESPAÑOLAS**

(<http://www.spainindustry.com>)

Posee un catálogo de unas 75.000 empresas, que se pueden buscar por los productos importados o exportados, así como por las actividades de producción.

- **INTERPYME**

(<http://www.interpyme.com>)

Índice de empresas españolas que poseen página propia en Internet. Contiene más de 3.000 direcciones.

- **EUROPAGES: THE EUROPEAN BUSINESS DIRECTORY**

(<http://www.europages.com>)

Directorio de más de 500.000 empresas europeas, establecidas en 30 países.

- **COMPANY SLEUTH**

(<http://www.companysleuth.com>)

Buscador internacional de empresas que permite encontrar información específica sobre miles de empresas.

➤ **BUSCADORES DE EXPERTOS**

Los buscadores de expertos son directorios organizados de especialistas a nivel mundial quienes ofrecen sus servicios para asesorar a los periodistas sobre temas complejos.

Normalmente esta asistencia suele ser gratuita, aunque puede darse el caso de que alguno de los expertos cobre por su aportación a un artículo.

- **PROFNET**

(<http://www.profnet.com>)

Professor Network es una base de datos de expertos norteamericanos especializados en diversos temas.

- **EXPERTS.COM**

(<http://www.experts.com>)

Permite hallar especialistas internacionales sobre materias muy diversas. El sistema de búsqueda es sencillo: por tema, palabra clave o categoría.

- **GUÍA DE EXPERTOS DE LA UNIVERSIDAD DE NAVARRA**

(<http://www.unav.es/un/expertos/>)

Esta guía permite hallar profesores de la Universidad de Navarra expertos en cientos de temas científicos y humanísticos dispuestos a asesorar a los periodistas sobre los temas específicos.

Web del buscador de expertos Profnet (www.profnet.com)

ProfNet Global™

The Shortest Distance Between A Journalist And A Source

Linking the expert resources of:

[ProfNet](#)

[ARCES \(France\)](#)

[Universities UK Expertnet](#)

[Informationsdienst Wissenschaft \(Germany\)](#)

Good sources are hard to find.

That's why we created ProfNet, a collaborative of information officers linked by Internet to give journalists free, convenient access to expert sources.

We represent organizations of every type in North America and Europe -- colleges, universities, corporations, think tanks, laboratories, medical centers, non-profit organizations, government agencies, and PR agencies. There's no limit to the subjects we cover.



For Journalists



For Information Officers

➤ **BUSCADORES DE BIOGRAFÍAS**

Los buscadores de biografías pueden resultar de gran utilidad a la hora de documentar noticias, reportajes o entrevistas. Aunque la mayoría de medios de comunicación suelen disponer de un archivo de documentación propio, merece la pena tener presente y conocer este recurso para poder acceder a él en caso necesario.

Algunos de los más interesantes son los siguientes:

- **BUSCABIOGRAFÍAS**

(<http://www.buscabiografias.com>)

Esta base de datos permite obtener de forma gratuita biografías en español de personajes célebres españoles e internacionales. A finales del año 2000, contaba con casi un millar de biografías.

- **BIOGRAPHICAL DICTIONARY**

(<http://www.s9.com/biography/>)

Biografías en inglés de más de 27.000 personas de todo el mundo.

- **DIRECTORIO DE BASES DE DATOS BIOGRÁFICAS**

(<http://www.library.upenn.edu/resources/reference/general/biography.html>)

Elaborado por la biblioteca de la Universidad de Pennsylvania, este directorio norteamericano da acceso a más de 20 bases de datos biográficas.

➤ **BUSCADORES DE BUSCADORES**

Los buscadores de buscadores, como su propia palabra indica, se utilizan para buscar buscadores especializados sobre una materia concreta. Algunos de los más utilizados son los siguientes:

- **BUSCOPIO**

(<http://www.buscopio.com>)

El más completo y mejor organizado directorio de buscadores en español. Tiene catalogados más de 3.000 buscadores nacionales e internacionales.

- **SEARCHPOWER**

(<http://www.searchpower.com>)

Directorio con más de 2.000 buscadores generalistas y especializados, clasificados en 80 categorías temáticas.

- **SEARCHABILITY**

(<http://www.searchability.com>) Directorio de buscadores temáticos especializados y metabuscadores.

5.9.3. BUSCADORES AUTOMÁTICOS DE SERVICIOS DE INTERNET

En Internet no solo podemos encontrar información documental sino también un servicio aún mas valioso para el periodista que es la posibilidad de obtener las direcciones personales de sus fuentes, además de entrar en contacto con especialistas en diversas materias a nivel mundial a través de foros.

Los buscadores de servicios de mayor utilidad periodística son los siguientes:

➤ **BUSCADORES DE NÚMEROS DE TELÉFONO**

Los buscadores de número de teléfono son de bastante utilidad para el periodista a la hora de buscar números de teléfono y datos precisos de sus fuentes de información:

- **TELEFÓNICA: PÁGINAS BLANCAS**

(<http://www.telefonica.es/index/comun/pagblancas.html>)

Da acceso a la base de datos íntegra de todos los teléfonos de España.

- **TELEPHONE DIRECTORIES ON THE WEB**

(<http://www.teldir.com>)

Sencillo directorio telefónico y páginas amarillas de la mayoría de los países del mundo.

- **EUROCALL WORLD DIRECTORIES**

(<http://www.eurocall.net/world/world.htm>)

Directorio de páginas amarillas y guías telefónicas de todo el mundo.

➤ **BUSCADORES DE DIRECCIONES DE CORREO ELECTRÓNICO**

En la actualidad, prácticamente cualquier persona que puede utilizar con una cierta asiduidad un ordenador con conexión a Internet tiene cuenta de correo electrónico. Este servicio que posibilita Internet es de la máxima utilidad para los periodistas, ya que nos permite, de una forma sencilla, entrar en contacto directo e inmediato con fuentes de

primera mano de todo el mundo. Si el profesional de la información no conoce la dirección e-mail de su posible fuente, puede buscarla en los buscadores especializados en direcciones de correo electrónico. Algunos de los más utilizados son los siguientes:

- **WHO WHERE**

(<http://www.whowhere.com>)

Buscador de personas de todo el mundo. Existe una versión en español en la siguiente dirección: <http://www.spanish.whowhere.com>.

- **E-MAIL ADDRESSES**

(http://www.emailaddresses.com/email_find.htm)

Metabuscador internacional de direcciones de correo electrónico.

- **INTERNET ADDRESS FINDER**

(<http://www.iaf.net>)

Buscador internacional de direcciones de correo electrónico.

➤ **BUSCADORES DE LISTAS DE DISTRIBUCIÓN**

Las listas de distribución son de gran utilidad periodística ya que permiten realizar foros de discusión sobre temas específicos con expertos u otros interesados en la materia a nivel mundial, a través del correo electrónico. Los buscadores de listas de distribución facilitan la búsqueda de la lista específica en la que está interesado el usuario. Algunos de ellos son:

- **E-LISTAS**

(<http://www.e-listas.com>)

Buscador de listas de correo en español. Ofrece un buscador convencional, una posibilidad de búsqueda avanzada y un directorio temático de listas.

- **REDIRIS: BUSCÓN**

(<http://www.rediris.es/list/buscon.es>)

Buscador de listas académicas en español. A través de este buscador de listas se pueden encontrar las más de 300 listas de correo en español dependientes de RedIRIS, la red académica de Internet en España.

- **CATALIST**

(<http://seagate.sunet.se/lists/listref-eng.html>)

Buscador internacional de listas de correo. Da acceso a decenas de miles de listas de acceso público.

➤ **BUSCADORES DE GRUPOS DE NOTICIAS**

Los grupos de noticias, también llamados *newsgroups*, son foros en los que los usuarios intercambian mensajes sobre un tema específico. Principalmente se diferencian de las listas de distribución en que para participar en los grupos de noticias no hace falta una suscripción previa y tampoco tiene moderador, lo que a veces suele suponer poca credibilidad. Sin embargo es aquí donde con frecuencia suelen surgir los primeros rumores sobre personajes y temas de actualidad por lo que el periodista debe tener este recurso en cuenta, siempre que compruebe los datos obtenidos.

Los buscadores de grupos de noticias permiten encontrar este tipo de servicio, además de hallar mensajes en sus archivos. Algunos de los más conocidos son los siguientes:

- **GNEWS**

(<http://www.gnews.com>)

Según sus autores, este buscador de grupos de noticias da acceso a una base de datos con más de 10 millones de mensajes publicados en unos 33.000 grupos de noticias.

- **REMARQ**

(<http://www.remarq.com>)

Potente y efectivo buscador internacional de grupos de noticias.

➤ **BUSCADORES DE BASES DE DATOS**

Gracias a este tipo de buscadores, podemos encontrar bases de datos de información estadística, demográfica, bibliográfica, etc, de gran utilidad periodística. Los mas relevantes son los siguientes:

- **THE BIG HUB**

(<http://www.thebighub.com>)

Este buscador permite buscar directamente en unas 1.500 bases de datos en Internet. Para facilitar la búsqueda, los accesos a las bases de datos están organizados de forma temática, ofreciendo además una breve descripción de cada una.

- **DIRECT SEARCH**

(<http://gwis2.circ.gwu.edu/~gprice/direct.htm>)

Directorio con más de 2.000 bases de datos clasificadas por el documentalista norteamericano Gary Price, de la George Washington University.

- **THE INVISIBLE WEB**

(<http://www.invisibleweb.com>)

Cataloga más de 10.000 bases de datos, buscadores y archivos de todo tipo. Posee un índice temático de recursos de búsqueda y permite realizar búsquedas avanzadas.

5.9.4. BUSCADORES POR SOPORTES DE INFORMACIÓN

Además de buscar información en Internet u otros servicios que le ayuden a fundamentar sus informaciones, el periodista a veces precisa de recursos gráficos, audio

o vídeo para ilustrar su noticia o reportaje, y de esta forma ofrecer una información más completa y atractiva al usuario.

Para encontrar estos recursos la Red dispone también de buscadores especializados que permiten localizar elementos gráficos, audio y vídeo. Sin embargo, a la hora de utilizar estos recursos para informaciones propias, el periodista debe tener en cuenta los derechos de autor de estos elementos aunque, a veces, estén libres de ellos. Normalmente se suele encontrar estos buscadores como servicio gratuito en los grandes buscadores automáticos y directorios generales de la Red. La mayor parte de estos servicios gratuitos suelen corresponder a entidades públicas.

También podemos localizar mediante estos buscadores, documentos en un formato distinto al html. Normalmente, estos documentos completos de gran extensión, se publicaron originalmente en un soporte físico y no fueron pasadas al formato html, sino que se publican en la Red como reproducciones facsímil de los originales en formato PDF. Este formato se puede visualizar con el programa de la marca Adobe “Acrobat Reader”.

Este sistema de publicación en la Red ha sido patentado por Adobe por lo que el buscador de este tipo de documentos también ha sido desarrollado por esta empresa.

A continuación se expone una muestra de los principales buscadores por soportes de información:

- **SEARCH PDF - ADOBE ONLINE**

(<http://searchpdf.adobe.com>)

Buscador específico de documentos en formato PDF que permite rastrear entre más de un millón de estos informes y documentos existentes en servidores de Internet de empresas, universidades e instituciones públicas repartidas por todo el mundo.

- **LYCOS**

(<http://www.lycos.com>)

Para buscar fotografías, sonidos y vídeos: <http://richmedia.lycos.com/>

Para buscar documentos en servidores FTP: <http://ftpsearch.lycos.com/>

Para buscar músicas en formato MP3: <http://mp3.lycos.com/>

- **ALTAVISTA**

(<http://www.altavista.com>)

Según el tipo de documentos que se desee, se puede elegir alguno de estos tres buscadores: AltaVista Images, AltaVista MP3/Audio o AltaVista Video.

Posee además buscadores de imágenes especializados por temas.

- **NATIONAL ARCHIVES AND RECORDS ADMINISTRATION**

(<http://www.nara.gov/nara/nail.html>)

Da acceso a la base de datos NAIL del gobierno de Estados Unidos, que posee más de 124.000 fotografías, mapas, sonidos y documentos históricos digitalizados de uso libre. Es de gran utilidad a la hora de hallar fotografías históricas.

- **THE PICTURE COLLECTION**

(<http://www.thepicturecollection.com>)

Permite adquirir, previo pago, imágenes del archivo fotográfico de Time Inc. (con revistas como *People*, *Fortune*, *Sports Illustrated*, *Money*, *Entertainment Weekly*...).

Posee un avanzado sistema de búsqueda de fotografías.

5.9.5. FUENTES PERIODÍSTICAS EN INTERNET

Además de lo citado hasta ahora, en Internet podemos encontrar gran variedad de publicaciones especializadas en periodismo, así como listas de recursos para periodistas, organizaciones, foros, etc, que pueden ayudar al profesional de la información a estar al tanto de su profesión, además de ser útil para su trabajo.

Algunas de estas fuentes son las siguientes:

➤ **PUBLICACIONES DIGITALES ESPECIALIZADAS**

En la actualidad existen en Internet gran variedad de revistas, diarios y boletines, en su mayoría gratuitos, especializados en periodismo, en su mayor parte, en su faceta digital. Aunque muchos de ellos comenzaron siendo meras copias de su versión en papel, a medida que avanza el desarrollo de Internet (y de las propias empresas periodísticas), estas publicaciones digitales se han ido independizando de su formato original. Asimismo han surgido otras específicamente pensadas para ser difundidas a través de la Red.

Algunas de estas publicaciones son las siguientes:

- **NOTICIAS DE LA COMUNICACIÓN**

(<http://www.noticom.es>)

Edición en Internet de la revista mensual que analiza el panorama empresarial y editorial de la prensa periódica y los medios audiovisuales en España.

- **EDITOR & PUBLISHER**

(<http://www.mediainfo.com>)

Incluye diversas secciones y columnas periódicas de gran calidad.

- **AMERICAN JOURNALISM REVIEW**

(<http://ajr.newslink.org>)

Artículos sobre las últimas tendencias de la prensa digital.

- **ONLINE JOURNALISM REVIEW**

(<http://ojr.usc.edu/>)

Publicación estadounidense de análisis y debate sobre las novedades del periodismo digital.

- **EUROPEMEDIA**

(<http://www.europemedia.com>)

Semanario en inglés que informa sobre la actualidad de las publicaciones digitales europeas.

- **MEDIA CENTRAL**

(<http://www.mediacentral.com>)

Ofrece un directorio de enlaces muy variado y amplio sobre la actualidad editorial internacional.

➤ **GUÍAS PARA PERIODISTAS**

Estas guías son de gran utilidad sobre todo para los periodistas que no están muy familiarizados con la búsqueda de información por Internet. La mayor parte de ellas suelen ofrecer enlaces comprobados y clasificados por áreas del conocimiento. Su objetivo es facilitar en la medida de lo posible el trabajo de búsqueda de datos que realiza el periodista a través de Internet.

Algunas de las principales guías son las siguientes:

- **SALA DE PRENSA**

(<http://www.saladeprensa.org/>)

Esta web, elaborada desde México, posee enlaces académicos y profesionales sobre diversos ámbitos del periodismo en España y América Latina. También dispone de gran variedad de artículos especializados sobre diversos temas (sobre todo sobre periodismo digital) que son actualizados con una cierta periodicidad.

- **JOURNALISMNET**

(<http://www.journalismnet.com/>)

Journalismnet reúne múltiples buscadores internacionales generalistas y especializados.

- **JOURNALISTEXPRESS**

(<http://www.journalistexpress.com>)

Ofrece un índice de las mejores publicaciones de la Red y a un estructurado catálogo de recursos orientados a la investigación periodística.

- **GUÍA DE INTERNET PARA PERIODISTAS**

(<http://www.unav.es/fcom/guia/>)

Guía elaborada por Ramón Salaverría, director del Laboratorio de Medios de la Universidad de Navarra. En él se catalogan por secciones gran variedad de recursos para periodistas, así como un registro actualizado de medios hispanos e internacionales en Internet, recursos académicos sobre periodismo, bases de datos, etc.

- **CIENCIAS DE LA INFORMACIÓN EN INTERNET**

(<http://www.ucm.es/BUCM/inf/0500.htm>)

Amplia y bien estructurada selección de recursos de Internet sobre periodismo, elaborada en la Biblioteca de la Facultad de Ciencias de la Información de la Universidad Complutense de Madrid.

- **THE EUROPEAN JOURNALISM PAGE**

(<http://www.demon.co.uk/eurojournalism/general.html>)

Selección de recursos europeos de gran utilidad para el periodista.

- **WWW VIRTUAL LIBRARY: JOURNALISM**

(<http://209.8.151.142/vlj.html>)

Página elaborada por el documentalista John Makulowich con un completo y bien estructurado listado de enlaces sobre diferentes facetas del periodismo en Internet.

Guía de Internet para periodistas de la Universidad de Navarra

(<http://www.unav.es/fcom/guia/>)



➤ LISTAS DE DISTRIBUCIÓN

Las listas de distribución son foros en los que diversos profesionales, en este caso del periodismo, se intercambian, a través del correo electrónico, informaciones para un mejor uso profesional de la Red como herramienta periodística. Estas listas son de una gran utilidad periodística ya que permite poner en contacto directo a nivel mundial a periodistas interesados en un mismo tema. De esta forma pueden discutir sobre temas de actualidad o realizar consultas específicas para una información. Normalmente suelen estar moderadas, aunque a veces esto no impide que algunos mensajes no concuerden con el tema de la lista. Éste es uno de sus inconvenientes.

Algunas de las listas que más se suelen utilizar son las siguientes:

- **PERIODISMO**

(<http://www.rediris.es/list/info/periodismo.html>)

Lista de correo incluida en la red académica española, RedIRIS, que reúne a casi 400 profesionales y académicos hispanohablantes. Los debates suelen estar muy ligados a la actualidad de la profesión periodística, sobre todo en España.

- **EN.MEDI@**

(<http://enredando.com/cas/enmedi@.htm>)

Lista dedicada a debatir sobre el panorama de la comunicación en la Sociedad de la Información.

- **FIDEL (Fuentes de Información y Documentación en Línea)**

(<http://exlibris.usal.es/merlo/fidel/index.htm>)

Lista de correo incluida en RedIRIS sobre fuentes documentales y recursos de búsqueda de información en la Red con más de un millar de suscriptores.

- **CARR-L**

(listserv@ulkyvm.louisville.edu)

Computer-Assisted Reporting and Research List (CARR-L) es una de las listas más interesantes para aquellos periodistas que deseen emplear Internet como herramienta

para sus reportajes. Se centra en el periodismo asistido por ordenador, pero también aborda asuntos variados sobre periodismo. Está moderada por el profesor Elliot Parker, de la American University.

- **IRE-L**

(listserv@mizzou1.missouri.edu)

Investigative Reporters and Editors List (IRE-L) es el foro de discusión profesional de numerosos periodistas de investigación, principalmente estadounidenses.

➤ **RECURSOS ACADÉMICOS**

Cada vez hay un mayor número de universidades y centros de enseñanza que apoyan y desarrollan la investigación sobre periodismo digital. Muchas de ellas suelen difundir a través de su web algunos de sus nuevos avances, por lo que puede resultar útil para el periodista acceder a estos sitios con cierta periodicidad.

Algunos de estos centros son los siguientes:

- **LABORATORIO DE PERIODISMO ELECTRÓNICO – EHU/UPV**

(<http://www.pd.lp.ehu.es/laboratorio/>)

Centro de investigación sobre periodismo digital creado desde el Departamento de Periodismo II de la Universidad del País Vasco. Ofrece acceso a artículos de sus profesores, reseñas de actividades y selección de enlaces, entre otros contenidos.

- **LABORATORIO DE COMUNICACIÓN MULTIMEDIA (MMLab)**

(<http://www.unav.es/fcom/mmlab/>)

Centro de investigación básica y aplicada sobre la comunicación digital, creado por la Facultad de Comunicación de la Universidad de Navarra en 1995. Su sitio da acceso, entre otros contenidos, a las páginas de las asignaturas y cursos coordinados desde allí, a artículos a texto completo publicados por los profesores que lo componen y a varios recursos de Red desarrollados en el propio Laboratorio.

- **EUROPEAN JOURNALISM CENTRE**

(<http://www.ejc.nl/index.html>)

Situada en Maastricht (Holanda), esta institución independiente promueve la actualización profesional de los periodistas europeos a través de seminarios de formación, publicaciones, etc.

- **COLUMBIA UNIVERSITY: GRADUATE SCHOOL OF JOURNALISM**

(<http://www.jrn.columbia.edu>)

Esta escuela de Nueva York se trata, probablemente, de la más prestigiosa del mundo en cuanto al mundo del periodismo se refiere. En concreto, su Center for New Media se

encuentra entre los centros de investigación y docencia más avanzados en el campo de la comunicación digital.

5.9.6. OTROS RECURSOS DE REFERENCIA

Como podemos ver a lo largo de este capítulo, Internet es en la actualidad una fuente esencial de gran utilidad para el periodista actual. Además de los recursos que hemos reseñado hasta ahora, merece la pena hacer una mención a otro tipo de recursos que pueden servir de referencia al profesional de la información.

➤ **PUBLICACIONES DIGITALES**

El periodista que utiliza Internet como fuente debe de tener en cuenta las publicaciones digitales ya que son un recurso de gran valor. Gracias a este medio, podemos acceder a un amplio abanico de publicaciones, la mayoría de ellas gratuitas, de forma instantánea y a nivel mundial.

Las publicaciones digitales las podemos encontrar generales y especializadas. Las **generales** suelen ser en su mayoría las ediciones en Internet de los diarios tradicionales. Como muestra exponemos algunos ejemplos de los diarios digitales más prestigiosos.

- **EL PAÍS** (Madrid)
(<http://www.elpais.es>)

- **EL MUNDO** (Madrid)
(<http://www.elmundo.es>)

- **ABC** (Madrid)
(<http://www.abc.es>)

- **EL CORREO** (Bilbao)
(<http://www.elcorreodigital.com>)

- **LA VANGUARDIA** (Barcelona)
(<http://www.lavanguardia.es>)

- **CORRIERE DELLA SERA** (Milán, Italia)
(<http://www.corriere.it>)

- **LA NACIÓN** (Buenos Aires, Argentina)
(<http://www.lanacion.com.ar>)

- **LE MONDE** (París, Francia)
(<http://www.lemonde.fr>)

- **SÜDDEUTSCHE ZEITUNG** (Munich, Alemania)
(<http://www.sueddeutsche.de>)

- **THE NEW YORK TIMES** (Nueva York, Estados Unidos)
(<http://www.nytimes.com>)

- **THE GUARDIAN** (Londres, Reino Unido)
<http://www.guardian.co.uk>

Con respecto a las **publicaciones digitales especializadas**, en la Red hay una gran variedad especializadas en todo tipo de temas. Algunas de ellas (suelen ser las más prestigiosas) son versiones digitales de los originales de papel. Otras, la mayoría, han surgido específicamente creadas para la Red.

A continuación se exponen varios directorios en los que se pueden encontrar publicaciones especializadas.

- **MUNDO LATINO - PRENSA**

(<http://www.mundolatino.org/prensa/>)

Directorio de medios de comunicación de España y América Latina, organizado por países y por tipo de medios.

- **EDITOR & PUBLISHER INTERACTIVE: MEDIA LINK**

(<http://www.mediainfo.com/emedial/>)

Uno de los más completos directorios mundiales de medios de comunicación en Internet.

- **NEWSCENTER OF GARY PRICE**

(<http://gwis2.circ.gwu.edu/~gprice/newscenter.htm>)

Directorio de medios de comunicación en Internet elaborado por Gary Price, documentalista de la George Washington University.

➤ **ORGANIZACIONES**

A través de las *webs* de las organizaciones se puede acceder a boletines y noticias, informes sobre el sector, artículos sobre investigaciones, bases de datos y otras informaciones de gran interés para el periodista. Algunas de las más interesantes para el profesional de la información son las siguientes:

- **LA MONCLOA**

(<http://www.la-moncloa.es>)

A través de esta dirección se puede acceder a los sitios en Internet de todas las instituciones públicas del Estado (gobierno, ministerios, tribunales, etc.) y autonómicos de España, así como a los organismos públicos dependientes.

- **ASSOCIATION CENTRAL**

(<http://www.associationcentral.com>)

Portal vertical que permite hallar asociaciones de todo tipo, especialmente en el entorno anglosajón.

- **ACADEMICINFO.NET**

(<http://www.academicinfo.net>)

Directorio de recursos académicos de calidad en Internet, a escala mundial.

➤ **BIBLIOTECAS VIRTUALES**

En las bibliotecas virtuales el periodista puede acceder, normalmente de forma gratuita, a grandes archivos bibliográficos digitalizados a texto completo de multitud de obras.

Este servicio suele estar promovido por instituciones que pretenden difundir la cultura de un determinado país.

Algunas de las principales son las siguientes:

- **BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES**

(<http://www.cervantesvirtual.com>)

Esta biblioteca tiene previsto dar acceso gratuito a más de 30.000 títulos en español para el año 2003. Por ahora la mayoría de los títulos que tiene indexados son grandes clásicos de la literatura en lengua castellana.

- **PROJECT GUTENBERG**

(<http://www.gutenberg.net>)

Es el proyecto de biblioteca virtual más antiguo.

- **FOROS**

Cuando el periodista usa los foros como fuente de información tiene que tener en cuenta que no suelen tener ningún filtro por lo que los comentarios que allí se exponen distan mucho de ser creíbles. Sin embargo, estos foros son una herramienta útil sobre todo para los medios de comunicación, ya que ofrecen la posibilidad a los usuarios que navegan por el medio en cuestión, de participar sobre diversos temas de actualidad, propiciando de esta forma la interactividad entre el medio y el lector.

- **PÁGINAS PERSONALES**

En la actualidad cada vez más portales y sitios de Internet ofrecen la posibilidad de crear páginas personales en sus servidores de forma gratuita o de pago (dependiendo del espacio que vaya a ocupar dicha página). Ésta es una de las grandes ventajas o inconvenientes de Internet, (depende de cómo se mire) cualquiera puede publicar de forma inmediata en la Red. La mayoría de estas páginas no presentan, en una primera instancia, una utilidad específica para el periodista. Aunque puede ser que en ciertos casos las *webs* personales supongan fuentes de primera mano. Como se ha dicho ya en varias ocasiones a lo largo de este capítulo, el periodista siempre debe comprobar sus fuentes, y más en estos casos

Uno de los directorios más recomendables de páginas personales es el siguiente:

- **PERSONALES**

(<http://www.personales.com/listado/>)

Esta web tiene registrados varios miles de páginas personales en español en más de cuarenta categorías temáticas.

5.9.7. CÓMO ELEGIR UN MOTOR DE BÚSQUEDA O UN DIRECTORIO

El listado que se expresa a continuación es una recomendación de qué herramienta de búsqueda se debe usar en cada caso¹²⁷:

BÚSQUEDA POR CAMPOS:

- **Audio/Música:**

- **akoo.com:** <http://www.akoo.com/index.cfm>
- **AllTheWeb:** <http://www.alltheweb.com/?cat=mp3&query=>
- **AltaVista:** <http://www.altavista.com/sites/search/advaudio>
- **CNET MP3 Search:** <http://home.cnet.com/category/1-4004.html?dd.cn.txt.0520.09>
- **Dogpile:** <http://www.dogpile.com/>
- **FindSounds.com :** <http://www.findsounds.com/>
- **FtpFind:** <http://www.ftpfind.com/>
- **Genie Knows:** <http://www.genieknows.com/>
- **HotBot SuperSearch:**
<http://hotbot.lycos.com/?MT=&SM=MC&DV=0&LG=any&DC=10&DE=2&AM1=MC&act.super.x=134&act.super.y=6>
- **Ithaki:** <http://ithaki.net/>
- **Ixquick Metasearch:** <http://ixquick.com/>
- **Lycos Music Downloads:** <http://music.lycos.com/downloads/>
- **Mamma:** <http://www.mamma.com/>
- **Lycos Multimedia Search:** <http://richmedia.lycos.com/>
- **MSN Search Advanced Search:**
<http://search.msn.com/advanced.asp?q=>
- **MetaMission:** <http://metamission.com/>

¹²⁷ Fuente: Cohen, Laura: *How to Cose a Search Engine or Directory*. Universidad de Albany (EE UU). Accesible en: <http://library.albany.edu/internet/choose.html>

- **Researchville: Audio:** <http://www.researchville.com/1025/>
- **SearchTurtle.com:** <http://www.searchturtle.com/>
- **Singingfish:** <http://www.singingfish.com/>
- **StreamSearch:** <http://www.streamsearch.com/>

- **Fecha de la última modificación:**

- **AltaVista Advanced Search:** <http://www.altavista.com/cgi-bin/query?pg=aq&styp=stext>
- **AltaVista Search Assistant:**
<http://www.altavista.com/sites/search/power>
- **Google Advanced Search:** http://www.google.com/advanced_search
- **HotBot:** <http://hotbot.lycos.com/>
- **MSN Search Advanced Search:**
<http://search.msn.com/advanced.asp?q=>

- **Dominio/Sede/URL:**

- **AllTheWeb Advanced Search:**
<http://www.alltheweb.com/advanced?t=all&c=web>
- **AltaVista Search Assistant:**
<http://www.altavista.com/sites/search/power>
- **Direct Hit Advanced Google Advanced Search:**
http://www.google.com/advanced_search
- **HotBot SuperSearch:**
http://hotbot.lycos.com/?MT=&SM=MC&DV=0&RG=all&RD=RG&DC=10&DE=2&act.super=More+Search+Options&_v=2&OPs=MDRTP
- **MSN Search Advanced Search:** <http://search.msn.com/advanced.asp>
- **Namedroppers.com:** <http://www.namedroppers.com/>
- **SearchEdu.com:** <http://www.searchedu.com/>

➤ **Virtual Learning Resources Center:** <http://www.virtuallrc.com/>

- **Localización geográfica**

➤ **AltaVista Search Assistant:**

<http://www.altavista.com/sites/search/power>

➤ **Clickey:** <http://www.clickey.com/>

➤ **Fossick:** <http://fossick.com/Search.htm>

➤ **HotBot SuperSearch:**

http://hotbot.lycos.com/?MT=&SM=MC&DV=0&RG=all&RD=RG&DC=10&DE=2&act.super=More+Search+Options&_v=2&OPs=MDRTP

➤ **Mama Power Search:** <http://www.mamma.com/mammapower.html>

➤ **MetaCrawler Power Search:**

http://www.go2net.com/index_power.html

➤ **MSN Search Advanced Search:** <http://search.msn.com/advanced.asp>

➤ **Northern Light Geosearch:**

<http://www.northernlight.com/geosearch.html>

- **IMÁGENES**

➤ **AllTheWeb:** <http://www.alltheweb.com/?cat=img&query=>

➤ **AltaVista:** <http://www.altavista.com/cgi-bin/query?mmdo=1&stype=simage>

➤ **The Amazing Picture Machine:** <http://www.ncrtec.org/picture.htm>

➤ **Cobion:** <http://www.visoo.com/default.htm>

➤ **Ditto:** <http://www.ditto.com/>

➤ **Dogpile:** <http://www.dogpile.com/>

➤ **FtpFind:** <http://www.ftpfind.com/>

➤ **Genie Knows:** <http://www.genieknows.com/>

- **Google Image Search:** <http://images.google.com/>
- **HotBot:** <http://hotbot.lycos.com/>
- **Ithaki:** <http://ithaki.net/>
- **Ixquick Metasearch:** <http://ixquick.com/>
- **Lycos Multimedia Search:** <http://richmedia.lycos.com/>
- **Mamma:** <http://www.mamma.com/>
- **MSN Search Advanced Search :** <http://search.msn.com/advanced.asp>
- **Picsearch :** <http://www.picsearch.com/>
- **Researchville: Images:** <http://www.researchville.com/1023/>
- **SearchTurtle.com:** <http://www.searchturtle.com/>
- **Yahoo! News Image Gallery :**
<http://dailynews.yahoo.com/headlines/g/ts/>

- **IDIOMA:**

- **AllTheWeb:** <http://www.alltheweb.com/>
- **AltaVista:** <http://www.altavista.com/>
- **Google Language Tools:** http://www.google.com/language_tools
- **HotBot:** <http://hotbot.lycos.com/>
- **MSN Search Advanced Search:** <http://search.msn.com/advanced.asp>

- **MULTIMEDIA Y VIDEO**

- **akoo.com:** <http://www.akoo.com/index.cfm>
- **All TheWeb:** <http://www.alltheweb.com/?cat=vid&lang=any&query=>
- **AltaVista:** <http://www.altavista.com/>
- **Dogpile:** <http://www.dogpile.com/>
- **FtpFind:** <http://www.ftpfind.com/>
- **Genie Knows:** <http://www.genieknows.com/>
- **HotBot:** <http://hotbot.lycos.com/>

- **Lycos Multimedia Search:** <http://richmedia.lycos.com/>
 - **Mamma:** <http://www.mamma.com/>
 - **Researchville: Video:** <http://www.researchville.com/1026/>
 - **SearchTurtle.com :** <http://www.searchturtle.com/>
 - **Singingfish:** <http://www.singingfish.com/>
 - **StreamSearch:** <http://www.streamsearch.com/>
- **LENGUAJE DE PROGRAMACIÓN**
 - **HotBot SuperSearch:**
http://hotbot.lycos.com/?MT=&SM=MC&DV=0&RG=all&RD=RG&DC=10&DE=2&act.super=More+Search+Options&_v=2&OPs=MDRTP
 - **MSN Search Advanced Search:** <http://search.msn.com/advanced.asp?q=>
- **NOMBRES PROPIOS**
 - **AltaVista:** <http://www.altavista.com/>
 - **HotBot:** <http://hotbot.lycos.com/>
- **TÍTULO**
 - **AllTheWeb Advanced Search:**
<http://www.alltheweb.com/advanced?t=all&c=web>
 - **AltaVista:** <http://www.altavista.com/>
 - **AltaVista Search Assistant:** <http://www.altavista.com/sites/search/power>
 - **Direct Hit Advanced:**
<http://www.directhit.com/advancedsearch.asp?alias=advsrch&qry=>
 - **Google Advanced Search:** http://www.google.com/advanced_search

- **HotBot:** <http://hotbot.lycos.com/>
- **iLOR:** <http://www.ilor.com/>
- **Mamma:** <http://www.mamma.com/>
- **Virtual Learning Resources Center:** <http://www.virtuallrc.com/>

LÓGICA DE BÚSQUEDA:

- **Operadores booleanos**

- **AltaVista Advanced Search:** <http://www.altavista.com/cgi-bin/query?pg=aq&what=web>
- **Dogpile:** <http://www.dogpile.com/>
- **HotBot:** <http://hotbot.lycos.com/>
- **Ixquick Metasearch:** <http://ixquick.com/>
- **metaEUREKA:** <http://www.metaeureka.com/>
- **ProFusion:** <http://www.profusion.com/>

- **Lógica booleana completa con paréntesis (ejemplo: comportamiento and (gatos o felinos)).**

- **AltaVista Advanced Search:** <http://www.altavista.com/cgi-bin/query?pg=aq&what=web>
- **HotBot:** <http://hotbot.lycos.com/>
- **Ixquick Metasearch:** <http://ixquick.com/>
- **MSN Search Advanced Search:** <http://search.msn.com/advanced.asp>

- **Lógica booleana mediante formularios**

➤ **AllTheWeb Advanced Search:**

<http://www.alltheweb.com/advanced?t=all&c=web>

➤ **AltaVista Search Assistant:** <http://www.altavista.com/sites/search/power>

➤ **AOL.COM Search Options:** <http://search.aol.com/refine.adp>

➤ **Google Advanced Search:** http://www.google.com/advanced_search

➤ **HotBot:** <http://hotbot.lycos.com/>

➤ **iLOR Advanced Search:**

http://www.ilor.com/searchilor.lor?&start=0&pg=aq&as_q=

➤ **Lycos Pro:** <http://lycospro.lycos.com/>

➤ **Metor:** <http://www.metor.com/>

➤ **MonkeySweat:** <http://www.monkeysweat.com/>

➤ **MSN Search Advanced Search:** <http://search.msn.com/advanced.asp>

➤ **ProFusion Advanced:** <http://search.msn.com/advanced.asp>

➤ **Snoopa Advanced Search:** <http://www.snoopa.com/AdvSearch.jsp>

• **Operadores de proximidad**

➤ **AltaVista Advanced Search:** <http://www.altavista.com/cgi-bin/query?pg=aq>

➤ **Google** (por defecto) : <http://www.google.com/>

➤ **Ixquick Metasearch:** <http://ixquick.com/>

OPCIONES DE BÚSQUEDA:

• **Términos alternativos de búsqueda**

➤ **Applied Semantics:** <http://www.oingo.com/>

➤ **SurfWax:** <http://www.oingo.com/>

➤ **Turbo10:** <http://turbo10.com/>

- **Un formulario**

- **AltaVista Search Assitant:** <http://www.altavista.com/sites/search/power>
- **Google Advanced Search:** http://www.google.com/advanced_search
- **HotBot:** <http://hotbot.lycos.com/>
- **Lycos Pro:** <http://lycospro.lycos.com/>
- **metaEUREKA:** <http://www.metaeureka.com/>
- **Metor:** <http://www.metor.com/>
- **MSN Search Advanced Search:** <http://search.msn.com/advanced.asp>
- **ProFusion Advanced:**
<http://www.profusion.com/CatNav.asp?ID=1&AGTID=1&queryterm=>
- **Snoopa Advanced Search:** <http://www.snoopa.com/AdvSearch.jsp>
- **Virtual Learning Resources Center:** <http://www.virtuallrc.com/>

- **Diferenciar mayúsculas y minúsculas**

- **AltaVista:** <http://www.altavista.com/>
- **Go.com:** <http://www.go.com/>
- **Ixquick Metasearch:** <http://ixquick.com/>

- **Buscar entre los resultados obtenidos**

- **AltaVista:** <http://www.altavista.com/>
- **HotBot:** <http://hotbot.lycos.com/>
- **Lycos:** <http://www.lycos.com/>
- **MetaMission:** <http://metamission.com/>

- **Buscar documentos similares**

- **AltaVista:** <http://www.altavista.com/>
- **Google:** <http://www.google.com/>

- **iLOR:** <http://www.ilor.com/>
- **MetaMission:** <http://metamission.com/>
- **ProFusion:** <http://www.profusion.com/>

- **Buscar términos relacionados**
 - **Direct Hit:** <http://www.directhit.com/>
 - **HotBot:** <http://www.hotbot.com/>
 - **Ixquick Metasearch:** <http://ixquick.com/>
 - **Lycos:** <http://www.lycos.com/>

- **Truncamiento**
 - **AltaVista:** <http://www.altavista.com/>
 - **Ixquick Metasearch:** <http://ixquick.com/>

- **Truncamiento automático**
 - **HotBot:** <http://hotbot.lycos.com/>
 - **Lycos:** <http://www.lycos.com/>

- **Lenguaje natural**
 - **Albert:** <http://albert.com/demo.php>
 - **AltaVista:** <http://www.altavista.com/>
 - **Ask Jeeves:** <http://www.ask.com/>
 - **Ithaki :** <http://ithaki.net/>
 - **Ixquick Metasearch:** <http://ixquick.com/>

- **Traducción automática**
 - **Babel Fish:** <http://world.altavista.com/>

- **Frases entre comillas**

- La mayoría de los motores de búsqueda ofrecen esta opción

- **Frase exacta a través de formularios**

- **AllTheWeb Advanced Search:**

<http://www.alltheweb.com/advanced?t=all&c=web>

- **Fossick:** <http://fossick.com/Search.htm>

- **Google Advanced Search:** http://www.google.com/advanced_search

- **HotBot :** <http://hotbot.lycos.com/>

- **MetaMission:** <http://metamission.com/>

- **MonkeySweat:** <http://www.monkeysweat.com/>

- **MSN Search Advanced Search:**

<http://search.msn.com/advanced.asp?q=&RS=CHECKED&co=20>

PRESENTACIÓN DE RESULTADOS

- **Resultados agrupados por conceptos**

- **Google:** <http://www.google.com/>

- **Google Web Directory:** <http://directory.google.com/>

- **iLOR:** <http://www.ilor.com/>

- **Teoma:** <http://www.teoma.com/>

- **Resultados agrupados por conceptos y/o sedes**

- **AOL.COM Search:** <http://search.aol.com/>

- **Direct Hit:** <http://www.directhit.com/>

- **Ixquick Metasearch:** <http://ixquick.com/>

- **Resultados como series de preguntas para ayudar a localizar las sedes más relevantes**
 - **Ask Jeeves:** <http://www.ask.com/>
 - **Atomica:** <http://www.atomica.com/>
 - **Guidebeam:** <http://guidebeam.com/>
 - **Oingo:** <http://www.oingo.com/>
 - **Query Server:** <http://www.queryserver.com/>
 - **SurfWax:** <http://www.surfwax.com/>
 - **Teoma :** <http://www.teoma.com/>
 - **Vivisimo:** <http://vivisimo.com/>

- **Resultados basados en enlaces. Ej: número de enlaces de las páginas más relevantes**
 - **Guidebeam:** <http://guidebeam.com/>
 - **Query Server:** <http://www.queryserver.com/>
 - **Teoma:** <http://www.teoma.com/>
 - **Turbo10:** <http://turbo10.com/>
 - **Vivisimo:** <http://vivisimo.com/>

- **Resultados basados en la popularidad, seleccionados por otros buscadores**
 - **AllTheWeb:** <http://www.alltheweb.com/>
 - **AltaVista:** <http://www.altavista.com/>
 - **HotBot:** <http://hotbot.lycos.com/>
 - **WISEnut:** <http://www.wisenut.com/>

- **Agrupar los resultados por sedes web con la opción de ver todos**
 - **Karnak:** <http://www.karnak.com/>
 - **Northern Light Search Alerts:**
http://www.northernlight.com/docs/alerts_help_about.html
 - **TracerLock:** <http://www.tracerlock.com/>

- **Almacenar las búsquedas en el servidor y que éste te avise de nuevos resultados**
 - **Copernic:** <http://www.copernic.com/>

CALIDAD DE LA BASE DE DATOS, TAMAÑO Y TIPO

- **La Red invisible (no cubierta por buscadores convencionales)**
 - **Atomica:** <http://www.atomica.com/>
 - **Complete Planet:** <http://www.completeplanet.com/>
 - **Direct Search:** <http://gwis2.circ.gwu.edu/~gprice/direct.htm>
 - **Google:** <http://www.google.com/>
 - **The InvisibleWeb:** <http://www.invisibleweb.com/>
 - **Invisible-web.net:** <http://www.invisible-web.net>
 - **Lycos Directory: Searchable Databases :**
http://dir.lycos.com/Reference/Searchable_Databases/
 - **MetaIQ :** <http://www.metaiq.com/>
 - **MetaMission:** <http://metamission.com/>
 - **ProFusion:** <http://www.profusion.com/>
 - **Search.Com:** <http://www.search.com/>
 - **Search Engines and News:** <http://www.internets.com/>

- **Metabuscadores con eliminación de duplicados**
 - **C4:** <http://www.c4.com/>
 - **Chubba:** <http://www.chubba.com/>
 - **Copernic:** <http://www.copernic.com/>
 - **Fossick:** <http://fossick.com/Search.htm>
 - **Ithaki:** <http://ithaki.net/>
 - **Ixquick Metasearch:** <http://ixquick.com/>
 - **Mamma:** <http://www.mamma.com/>
 - **MetaCrawler:** <http://www.metacrawler.com/>

- **MetaMission:** <http://metamission.com/>
- **ProFusion:** <http://www.profusion.com/>
- **Query Server:** <http://www.queryserver.com/>
- **SearchTurtle.com:** <http://www.searchturtle.com/>
- **SurfWax :** <http://www.surfwax.com/>

- **Usenet (grupos de noticias)**
 - **Chubba:** <http://www.chubba.com/>
 - **Gripe.com:** <http://gripe.com/>
 - **Ithaki:** <http://ithaki.net/>
 - **MetaCrawler:** <http://www.metacrawler.com/>
 - **metaEUREKA:** <http://www.metaeureka.com/>
 - **MetaMission:** <http://metamission.com/>
 - **Google Groups:** <http://groups.google.com/>
 - **Yahoo! Search Options:** <http://search.yahoo.com/search/options>

- **Un directorio temático pequeño pero evaluado**
 - **About:** <http://home.about.com/index.htm>
 - **Academic Info:** <http://www.academicinfo.net/>
 - **Argus Clearinghouse :** <http://www.clearinghouse.net/>
 - **BUBL Link:** <http://bubl.ac.uk/link/>
 - **INFOMINE:** <http://infomine.ucr.edu/>
 - **Librarians' Index to the Internet :** <http://www.lii.org/>
 - **Scout Report Archives :** <http://scout.cs.wisc.edu/archives/>
 - **Virtual Learning Resources Center:** <http://www.virtuallrc.com/>
 - **WWW Virtual Library:** <http://www.vlib.org/>

- **Un directorio comercial pero sin evaluar**
 - **Yahoo! :** <http://www.yahoo.com/>

- **Fuentes de referencia** (Diccionarios, enciclopedias, búsqueda de correos electrónicos, enciclopedias, noticias, estadísticas, etc)

- **Reference Collection** (Biblioteca de la Universidad de Albany) :
<http://library.albany.edu/reference/>

5.10. EL FUTURO DE LOS BUSCADORES

La evolución previsible del mercado de búsqueda de información en Internet la situamos en tres factores, según Jorge Juan García Alonso¹²⁸:

- **Sofisticación** de los motores de búsqueda y bases de datos: en la actualidad estamos en la "prehistoria" de los buscadores.
- **Especialización**: la mayoría de los buscadores en zonas geográficas o temas o ambos simultáneamente. Internet es un medio global, pero en un porcentaje muy alto funcionará lo "local".
- **Concentración** de los grandes buscadores genéricos en muy pocas empresas, que darán entrada a diferentes escalones de información, probablemente proveídos por otras empresas más especializadas en un tema en una región del mundo.

Estos tres puntos, señalan una parte real de lo que será el futuro de las herramientas de búsqueda en Internet, sin embargo es difícil saber cuál será el camino real de las nuevas tecnologías relacionadas con los motores de búsqueda. Lo que es claro es que es una evolución imparabile y constante, y que de una forma u otra beneficiará al periodista

¹²⁸ García Alonso, Jorge Juan: *El futuro de la búsqueda de información*. 1999. (Acesible en <http://www.webred.com>)

(en el caso del tema de esta tesis) en cuanto a que facilitará la búsqueda precisa de información.

En el Congreso Internacional sobre Estrategias de los Motores de Búsqueda de 2001¹²⁹, en referencia al futuro de este método de búsqueda de información en Internet, Craig Silverstein, director de tecnología de Google eligió como tema de discusión ilustrativo el futuro de la búsqueda en los años 2260-2264, en el contexto de las series de televisión de Star Trek. Silverstein señalaba que dentro de 260 años el motor de búsqueda “Star Trek” tendrá resuelto el problema de la búsqueda. Entenderá búsquedas verbales y será capaz de superar la imprecisión comunicativa humana concibiendo el contexto correcto de dichas búsquedas. Será capaz de buscar en bases de datos con información infinita de forma inmediata, y dará los resultados en una cálida voz imitando el sonido humano.

El director de tecnología de Google reconoció en este congreso que probablemente será en el año 2260 cuando esta tecnología realmente llegue a existir , pero también señaló que tanto Google como otras empresas dedicadas a los motores de búsqueda están desarrollando continuamente nuevas tecnologías con este objetivo concreto, cuyos primeros pasos se verán en los próximos años.

Para lograr una mejor comunicación, se está investigando tecnologías de inteligencia artificial que darán lugar a una mayor internacionalización de las búsquedas de manera que los motores serán capaces de operar con múltiples lenguajes, hecho que ya se está comenzando a observar.

Asimismo, dando prioridad a la rapidez de devolución de resultados, se continuará mejorando la actualización de la índices de las bases de datos, mostrando especial atención a técnicas que permitan una mejor elección de qué páginas son necesarias ser rastreadas más con mayor frecuencia.

Silverstein apuntaba en esta conferencia, que la mayoría de las compañías de los motores de búsqueda están trabajando en intentar conseguir que se puedan rastrear

¹²⁹ The Search Engine Strategies 2001 conference (SES). Agosto 2001. San Francisco.

páginas que no sean de texto o creadas con el lenguaje HTML, como contenido multimedia, animaciones, mensajes de Usenet y mapas.

Por su parte, el jefe estratégico de Inktomi, Andrew Littlefield visionaba las mejoras en los motores de búsqueda desde el punto de vista de la diferenciación internacional, poniendo como ejemplo un motor específico que proporciona a los canadienses franceses resultados tanto de Francia como de Canadá, pero con la peculiaridad de que dicho motor cuenta con la “sabiduría” de que los ciudadanos franceses no están interesados en los sitios web franceses-canadienses por lo que los filtra para aquellos usuarios que se encuentren en Francia.

Asimismo, Littlefield opina que uno de los grandes retos que deberán afrontar los motores de búsqueda en un futuro, será el de “entender” qué es lo que el usuario quiere buscar realmente, presentando más tipos de información potencialmente relevante para que el usuario sea capaz de escoger lo que verdaderamente le interesa de forma más fácil.

Desde un punto de vista mucho más práctico se posicionaron Bill Bliss, director general de MSN Search; John Gentry, directivo de GoTo y Scott Gtz, director general de búsqueda del directorio Yahoo!. Sus opiniones se centraron en señalar que en el futuro los usuarios que más paguen por realizar una búsqueda serán los que obtengan resultados más precisos. Para ellos se trata de vislumbrar nuevos planes de negocio no de mirar en una bola de cristal.

En definitiva, el rápido avance de las nuevas tecnologías nos permitirán en los próximos años comenzar a vislumbrar nuevos descubrimientos en el campo de los motores y de los recursos de búsqueda de información en general. De cualquier forma, y pase lo que pase, el objetivo final a alcanzar será facilitar la búsqueda del usuario de Internet y ofrecerle los resultados que se ajusten de forma más precisa a su consulta.

3ª PARTE: LA RED PROFUNDA Y SU UTILIDAD PARA EL PERIODISTA

CAPÍTULO 6-LA RED PROFUNDA

CAPÍTULO 6- LA RED PROFUNDA

6.1. INTRODUCCIÓN

El contenido que se encuentra disponible en Internet es sensiblemente mayor de lo que aparentemente se suele concebir y que se suele identificar con lo que se halla en la World Wide Web (protocolo http). También se puede encontrar información a través de otros protocolos ajenos a la Red: el FTP (protocolo de transferencia de ficheros), el e-mail, las news, Telnet, Gopher, etc. Del mismo modo, incluso dentro de la Red, la mayoría de los usuarios únicamente conocen la existencia de los contenidos que les presentan buscadores como Altavista, Google, Excite, o Northern Light, o directorios como Yahoo, About.com, etc, y muchos de ellos acaban frustrados tras realizar una búsqueda y no encontrar lo que específicamente buscaban, sino un mar de documentos aparentemente relacionados pero sin utilidad precisa para el usuario.

Según Michael K. Bergman, autor de un informe de la compañía Bright Planet¹³⁰ (julio 2000), existe una Web invisible que tendría una extensión de aproximadamente 500 veces mayor que la visible (que es la única que detectan los buscadores convencionales). Y su tamaño total sería de 550.000 millones —medio billón— de documentos. Su utilidad, y por tanto su conocimiento por parte del periodista, es incalculable debido a la gran cantidad de información que posee.

La primera imagen que se ofrece a continuación ilustraría la llamada World Wide Web o Red accesible mediante buscadores, mientras que el segundo gráfico ilustraría la totalidad de Internet.

¹³⁰ *The Deep Web: Surfacing Hidden Value*. Informe de BrightPlanet.com (Accesible en: http://www.completeplanet.com/tutorials/deepweb/introduction_howsearch07.asp).

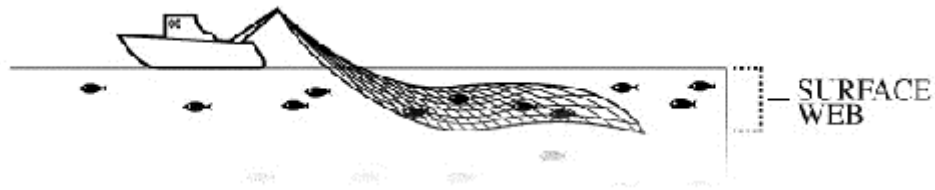


Grafico 1. La Web superficial, la que controlan los buscadores. Imagen tomada de "The Deep Web: Surfacing Hidden Value"¹³¹

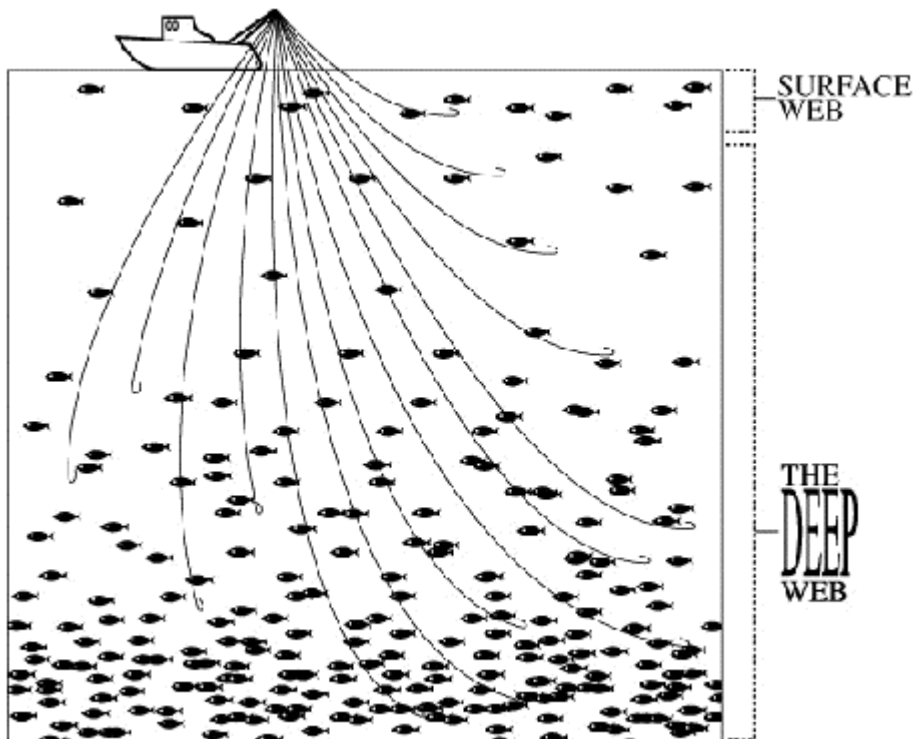


Gráfico 2. La Web profunda, la que presuntamente controla el software de BrightPlanet. Imagen tomada de "The Deep Web: Surfacing Hidden Value"¹³².

¹³¹ White Paper. *The Deep Web: Surfacing Hidden Value*, por Michael K. Bergman con Mark Smither y Will Bushee, © 2000 BrightPlanet.com LLC, se puede bajar en versión Word en <http://www.completeplanet.com/Tutorials/DeepWeb/index.asp>

6. 2- EL TAMAÑO DE LA RED

Analizar la cantidad de información disponible en la Red es un tema complejo dada su continua evolución. Actualmente existen muy pocos estudios que aludan a este tipo de análisis debido a su constante dinamismo. Como paso previo a nuestra explicación, merece la pena reseñar la cantidad de información producida y almacenada en el mundo según el soporte aludiendo a los datos ofrecidos por el School of Information Management and Systems (SIMS) de la Universidad de Berkeley.

Estimaciones de la producción mundial de contenidos originales almacenados digitalmente usando métodos de compresión, en terabytes en 1999¹³³.

Método de almacenamiento	Tipo de contenido	Terabytes/año. Estimación superior.	Terabytes/año. Estimación mínima.	Tasa de crecimiento anual
Papel				
	Libros	8	1	2%
	Diarios	20	2	-2%
	Periódicos	12	1	2%
	Documentos oficiales	195	19	2%
	Sub total:	240	23	2%
Film				
	Fotografías	410.000	41.000	5%
	Cine	16	16	3%
	Rayos X	17.200	17.200	2%
	Subtotal	427.216	58.216	4%
Ópticos				
	CD's música	58	6	3%
	CD's datos	3	3	2%
	DVD's	22	22	100%
	Subtotal	83	31	70%
Magnéticos				
	Cintas	300.000	300.000	5%
	Discos duros de PC	766.000	7.660	100%
	Servidores departamentales	460.000	161.000	100%
	Servidores de	167.000	108.550	100%

¹³² White Paper. "The Deep Web: Surfacing Hidden Value", por Michael K. Bergman con Mark Smither y Will Bushee, © 2000 BrightPlanet.com LLC, se puede bajar en versión Word en <http://www.completeplanet.com/Tutorials/DeepWeb/index.asp>

¹³³ School of Information Management and Systems (SIMS), Universidad de Berkeley, 2000.

	empresas			
	Subtotal	1.693.000	577.210	55%
Total		2.120.539	635.480	50%

Es interesante observar la evolución de estos medios. Así, algunos medios ópticos (CD audio, CD Rom y DVD) y magnéticos como soportes de información se incrementan a tasas que pueden ser consideradas como notables de hasta el 100% anual, otras encuentran crecimientos vegetativos en los denominados “soportes tradicionales” como el papel y los filmes.

El mundo produce actualmente entre 1 y 2 exabites anuales de información, equivalentes a 250 megabites por cada habitante del planeta. Un exabite es un billón de gigabites o 10^{18} bites. Los documentos impresos comprenden el 0,03% del total¹³⁴. El almacenamiento magnético es el medio de guardar información que más rápidamente crece, con una capacidad de almacenaje que se duplica cada año.

Las actuales capacidades de compresión informática reducen la necesidad de almacenamiento, bajando sin preocupación el tamaño normalmente requerido en uno a dos órdenes de magnitud.

Las principales tendencias observadas en este campo, según el informe del Dr. Ricardo Petrisans¹³⁵ sobre la Extensión de la Red son las siguientes:

- **La dominación de lo digital:** el 93% de la información actualmente producida se almacena bajo forma digital. Los discos duros de las PC's contienen un 55% del almacenamiento realizado cada año (último dato disponible de 1999). Así, la mayor parte de la información producida es almacenada en formato digital, esto es, además “nace” en formato digital, producida por los individuos y las organizaciones a partir de los ordenadores.
- **La migración magnética:** Los contenidos de vídeo, audio e impresos se están movilizandando rápidamente hacia los soportes y almacenamientos magnéticos y ópticos.

¹³⁴ Estudio de la School of Information Management and Systems (SIMS), Universidad de Berkeley.

¹³⁵ Petrisans Aguilar, Ricardo: *La extensión de la Red y la información*. Montevideo, enero 2002

- **La transición de la cinta:** el almacenamiento en cinta supera en 10 veces el almacenamiento en disco, pero es usado exclusivamente como archivo. El almacenamiento en discos es mucho más atractivo, inclusive para archivos y está en constante crecimiento particularmente por la rápida declinación de costos que se indica en los cuadros siguientes y resulta mucho más sencillo manejar los datos almacenados de esta manera.
- **Los impresos:** el material publicado en el mundo, si fuera expresado en caracteres ASCII, ocuparía un espacio menor a los 5 Terabytes. Si bien no se encuentra un descenso significativo en la producción informática de material impreso, aún cuando sí una tendencia al estancamiento con migración hacia el medio magnético.
- **Cantidad de imágenes:** 80 millones de fotografías son tomadas cada año, requiriendo algo más de 400 petabytes para su almacenamiento, equivalente a ochenta millones de veces los requerimientos para el almacenamiento de texto.

En términos de producción de información, la posición de los Estados Unidos en este mundo digital indica que ese país produce el 35% del total del material impreso, el 40% de las imágenes y algo más del 50% del material digital almacenado y producido en el mundo en el año.

Una correlación que lleva a conclusiones de gran interés, se relaciona con la comparación expuesta en los cuadros siguientes, que permiten la comparación entre producción de información digital, capacidad de almacenamiento de la misma, costo del almacenamiento y consumo, lo que tiene un peso trascendente en el fenómeno de migración a Internet

En primer lugar, la producción anual de información, medida en terabytes, tomando como base el año 2001, muestra lo siguiente:

Producción anual de información.

Tipo de documento	Terabytes
Documentos de oficina	198
Noticias	25
Revistas	23
Libros	5
Diarios	1

Otra comparación realizada en el estudio de Berkeley consiste en observar la producción anual de información individual, no considerada como “información pública” y sin comparación con ésta:

Producción anual de información individual:

Ítem	Cantidad	Terabytes
Fotografías	80 billones de imágenes	410.000
Videos domésticos	1,4 millones de cintas	300.000
Rayos X	2 billones de imágenes	17.200
Discos duros	200 millones de drives instalados	13.760

Cuando se compara la producción de la “información individual” con la cantidad de información publicada, se encuentra que la primera es 600 veces superior a la segunda. Además de la Red, Usenet y el correo electrónico incluyen una gran cantidad de información individual y están omitidos de la tabla por las dificultades de medición que implican, así como la dificultad metodológica de distinguir algunos casos de información “individual” y de “información pública”.

Producción anual de información publicada:

Ítem	Títulos	Terabytes
Libros	968.735	8
Diarios	22.643	25
Semanarios	40.000	2
Magazines	80.000	10
Documentos de oficina	7.500.000.000	195
Cine	4.000	16
CD's musicales	90.000	6
Cd's de datos	1.000	3
DVD-videos	5.000	22
Total		283

Por su parte, la capacidad de almacenamiento en discos duros en el mundo, (expresado en petabytes) se presenta, tomando un periodo entre 1995 y el estimado hacia el 2003. La primera fecha se toma en consideración en razón de ser el primer año efectivo en que se produce en el mundo la popularización de Internet.

De la observación surge una multiplicación clara de esa capacidad, lo que, en primera instancia parece no arrojar cuellos de botella actuales o en el futuro inmediato en materia de capacidad de almacenamiento.

Capacidad de almacenamiento digital en el mundo.

Año	Petabytes
1995	500
1996	600
1997	650
1998	1.000
1999	1.800
2000	2.200
2001 (e)	4.200
2002 (e)	8.100
2003 (e)	14.000

El tercer elemento trascendente en este caso, es el relacionado al costo de almacenamiento digital. El costo por gigabyte de almacenamiento muestra una evolución de baja constante, consistente con la oferta de almacenamiento. La tendencia es, en este sentido, sumamente nítida, acompañando el descenso de los costos de *hardware* y con una proyección a la baja que parece irreversible, donde, el negocio puro de almacenamiento no experimenta un repunte en términos económicos, salvo si se encuentra asociado con la prestación de servicios de valor agregado.

Costo de almacenamiento de la información digitalizada.

Año	Dólares por Gigabyte
1988	11.600
1989	9.200
1990	6.900
1991	5.200
1992	3.000
1993	1.300
1994	600
1995	300
1996	100
1997	650
2000	10
2005 (e)	1

En síntesis, el costo del almacenamiento magnético se desploma rápidamente. En el 2000, el costo del almacenamiento del gigabyte se ubicaba en U\$S 10, después de los más de U\$S 11.000 de 1988 y las previsiones de la Universidad de Berkeley son que para el 2005 es que ese costo se ubique en U\$S 1, lo que en los actuales términos de Internet tampoco podría ser considerado como el umbral definitivo.

Finalmente, resulta de interés poder medir el consumo de información (cifras de 1999) y únicamente medidos para hogares de los Estados Unidos, expresados en horas anuales y traducidos a megabites.

Cuadro N° 12:**Consumo de información en los Estados Unidos (1992/2000).**

Ítem	Horas 1992	Horas 2000 (e)	Megabites en 2000	% de cambio 2000/1992
Televisión	1510	1571	3.142.000	4
Radio	1150	1056	57.800	-8
Música grabada	233	269	13.450	15
Diarios	172	154	11	-10
Libros	100	96	3	-4
Revistas	85	80	6	-6
Videos hogareños	42	55	110.000	30
Video juegos	19	43	21.500	126
Total	3.324	3.380	3.344.783	1,7

La Red superficial y la Red profunda

La autoría de los estudios más relevantes sobre el tamaño de la Red superficial corresponde a Lawrence and Giles del NEC Research Institute en Princeton, Nueva Jersey. Su primer gran estudio sobre este tema publicado en la revista *Science* en 1998, estimaba el tamaño total de la web superficial en 320 millones de documentos¹³⁶. Una actualización de dicho estudio empleando una metodología diferente, fue publicado en la revista *Nature* en 1999¹³⁷. Esta última investigación establecía 800 millones de documentos indexables en la Web.

Algunas de las conclusiones extraídas fueron las siguientes:

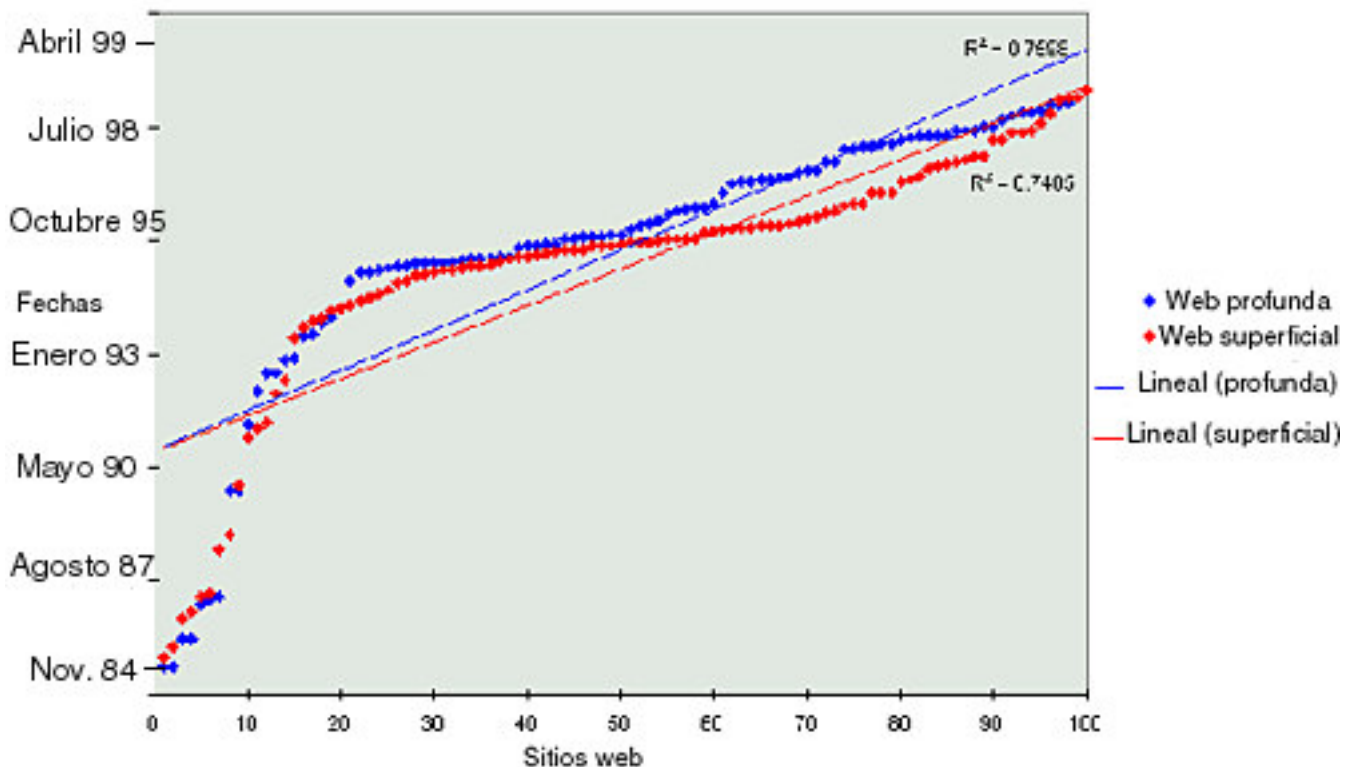
¹³⁶ Lawrence, S y Giles C.L: *Searching the World Wide Web*. *Science* 80. 98-100. 3 de Abril 1998.

¹³⁷ Lawrence, S y Giles C.L: *Searching the World Wide Web*. *Nature* 400. 107-109. 8 de Julio 1999.

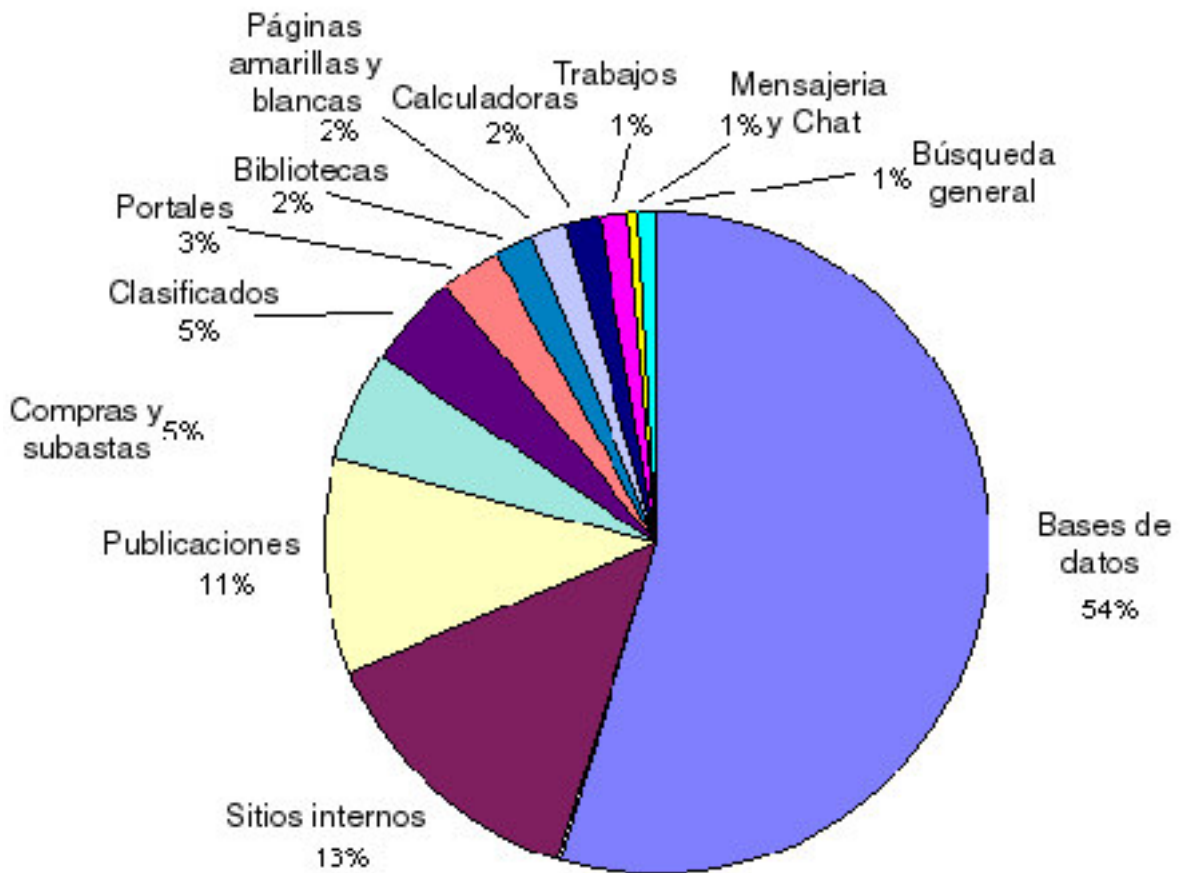
- a) La superficie de la Red cubierta con la mayoría de los motores de búsqueda ha caído de un máximo de un 32% en 1998 a 16% en 1999, siendo Northernlight el buscador que más páginas tiene indexadas.
- b) Los metabuscadores, que pueden usar varios motores de búsqueda a la vez, pueden incrementar aproximadamente en un 3, 5% (con respecto al apartado anterior) la capacidad de búsqueda.
- c) Los documentos web más populares, es decir aquellos que tienen muchos enlaces (*links*) de referencia a otros documentos, tienen ocho veces más posibilidades de ser indexados que aquellos que no los tienen.

Junto con la empresa Inktomi¹³⁸, NEC logró estimar a principios del año 2000 mil millones de documentos.

Gráfico comparativo entre las tasas de crecimiento de la Red profunda y la superficial¹³⁹



¹³⁹ White Paper. "The Deep Web: Surfacing Hidden Value", por Michael K. Bergman con Mark Smither y Will Bushee, © 2000 BrightPlanet.com



En realidad no hay una frontera clara que separe los contenidos en Internet. Hay veces que documentos de la Web profunda pueden aparecer en la superficial y al revés.

Como se ha explicado con anterioridad, los contenidos “superficiales” se encuentran en páginas estáticas fácilmente accesibles por los motores de búsqueda, mientras que los documentos en la Web profunda sólo presentan dinamismo en respuesta a una pregunta directa. Sin embargo, una vez que hemos obtenido esta respuesta, a ésta se le asocia una URL o dirección web (que normalmente contiene el número que le corresponde en la base de datos) por la que posteriormente se podrá obtener este mismo documento. Es decir, si se elabora otra página web y se coloca en ella la URL obtenida de una búsqueda en una base de datos, cualquier usuario podrá

pulsar en ella y obtener esa misma información. Además, en este caso ese documento de la base de datos pasará a formar parte de la web superficial ya que al estar su enlace de acceso en una página estática, puede ser encontrado por los motores de búsqueda.

Debemos tener en cuenta, para concluir, que establecer el tamaño real de la Web profunda es prácticamente imposible debido a que al sólo poderse localizar los contenidos de las bases de datos a través de preguntas directas, nunca se podrán hacer todas las preguntas necesarias para que resultaran todos los documentos interiores. Sin embargo que no vayamos conocer su tamaño exacto no significa que el usuario vaya a tener más problemas localizando el documento que busca, ya que éste puede preguntar directamente a la base de datos en cuestión y obtendrá la respuesta precisa. Asimismo también hay que tener presente el dinamismo de las propias bases de datos y de las webs en general que implica un movimiento continuo de documentos, por lo que si alguna vez acabáramos de contabilizarla, tendríamos que volver a empezar.

A medida que avance la tecnología los métodos de búsqueda de información en Internet se irán desarrollando y precisando cada vez mas, hasta que llegue un momento en que todo Internet será transparente y fácilmente accesible a través de nuestras búsquedas. Aún entonces necesitará ser organizada y valorada por profesionales de la información, los cuales deberán conocer de forma precisa todos los recursos de búsqueda disponibles, en aras de sacar el mayor provecho de esta gran fuente de información que es Internet.

6. 3. RED PROFUNDA VERSUS RED SUPERFICIAL

Según explican Chris Sherman y Gary Price en su libro *The Invisible Web*, no siempre es fácil distinguir cuando es recomendable usar los recursos disponibles en la Red profunda, en lugar de los buscadores generales. Estos autores enumeran varias reglas a seguir que indican cuando se deben usar los recursos de la Red profunda¹⁴¹:

- **Cuando se está familiarizado con el tema en cuestión.** Si se conoce bien un determinado tema, normalmente los recursos de la Red profunda ofrecen un tipo de información más precisa. Además, el hecho de conocer el tema sobre el que se está realizando la búsqueda ofrece, a parte de métodos para localizar nuevas fuentes, la ventaja añadida de saber términos precisos que ayudan a obtener mejores resultados cuando se utiliza la herramienta de búsqueda.
- **Cuando se está familiarizado con herramientas específicas de búsqueda.** Algunos de los recursos disponibles en la Red profunda abarcan muchos temas, sin embargo el hecho de que ofrezcan al usuario un sofisticado interfaz supone que se consigan mejores resultados en comparación de los que se obtendrían con las herramientas de búsqueda generalistas. Si se restringe la búsqueda usando operadores booleanos u otras funciones de búsqueda avanzada generalmente se encontrará, de forma más fácil y precisa, el objeto de la búsqueda.
- **Cuando se está buscando una respuesta precisa.** Cuando se está buscando una simple respuesta a una pregunta lo último que necesitamos es una lista de cientos de posibles resultados, como normalmente obtendríamos al utilizar un motor de búsqueda generalista. Bastantes recursos de la Red profunda están diseñados para poder buscar en ellos por cuestiones específicas, como pueden ser: un número de teléfono, un nombre, bibliografía, etc.

¹⁴¹ Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

- **Cuando se quieren resultados acreditados y exhaustivos.** Los motores de búsqueda generalistas nunca podrán ser capaces de ofrecer al usuario el tipo de información acreditada que se puede obtener de la Red profunda. Este tipo de buscadores están caracterizados por una búsqueda menos exhaustiva, la rapidez y la ausencia de filtros de selección, por lo que los resultados tienen mucha probabilidad de contener ruido. Además, la inmensidad de la Red de redes provoca que un cierto número de documentos de gran importancia sean ignorados por estos motores de búsqueda generalistas.
- **Cuando la actualidad de la información que se está buscando es importante.** Los recursos disponibles en la Red profunda normalmente están más actualizados que los que podemos encontrar a través de los motores de búsqueda generalistas o los directorios.

A pesar de que, como podemos ver, la Red profunda es muy útil para los usuarios de Internet, sobre todo para aquellos que buscan información rigurosa y especializada, se debe tener en cuenta que aún existe mucha información y muchos datos que no han sido digitalizados e incluidos en Internet. Son los siguientes:

A pesar de que, como podemos ver, la Red profunda es muy útil para los usuarios de Internet, sobre todo para aquellos que buscan información rigurosa y especializada, se debe tener en cuenta que aún existe mucha información y muchos datos que no han sido digitalizados e incluidos en la Red o que tienen restringido el acceso a determinados usuarios. Son los siguientes¹⁴²:

- Bases de datos privadas y servicios de información. Esto incluye el servicio Thomson's Dialog, LexisNexis y el Dow Jones, que tienen restringido el acceso a sus sistemas de información para suscriptores de pago.
- Muchos documentos públicos y del gobierno. Aunque los gobiernos cada vez tienden más a publicar documentos de interés público en Internet, muchos de

¹⁴² Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

ellos no son incluidos por una gran variedad de razones, entre ellas, la posibilidad de que la publicación de alguno de estos documentos pudiera atentar contra derechos individuales. Algunos servicios privados como KnowX (<http://www.knowx.com>) ofrece acceso limitado a archivos públicos previo pago.

- Periódicos académicos. En la actualidad las publicaciones académicas de las universidades están suponiendo un negocio bastante lucrativo para sus editores, que no están dispuestos a plantearse la posibilidad de poder perder algunos de sus clientes habituales, o a ver disminuido su beneficio, si ponen su producto en la Red. Por este motivo muchos de ellos no se hallan en Internet. Sin embargo hay algunas excepciones. La colección especial de Northernlight, por ejemplo, pone a disposición de los internautas y a un bajo precio, amplios contenidos de estos periódicos que antes sólo estaba disponible mediante una suscripción cara o una licencia expedida por el servicio de información privado en cuestión. Actualmente esta colección especial contiene aproximadamente 25 millones de documentos que se han ido almacenando desde 1995 e incorpora unos 250 nuevos cada mes. Por otro lado, ResearchIndex (<http://researchindex.dlib.vt.edu/cs>) permite al usuario acceder a copias de estas publicaciones que se encuentran alojadas en páginas personales, aunque estas no sean las versiones oficiales. En definitiva, por ahora, las bibliotecas públicas y físicas de las universidades continúan siendo una fuente de información importante a tener en cuenta a la hora de realizar una investigación, ya que muchos de los libros o documentos de los que disponen no se encuentran aún disponibles en la Red.
- Texto completo de muchos periódicos y revistas. Muy pocos periódicos y revistas ofrecen sus archivos íntegros y aquellas publicaciones que lo hacen, únicamente han digitalizado y han puesto a disposición del usuario el archivo de los últimos 10 o 20 años. Hay varias razones para esta forma de actuar. Los editores son conscientes de que el contenido que, a veces, el contenido que publican puede revalorizarse con el tiempo. Los derechos de autor son otra

cuestión a tener en cuenta. Muchos autores tienen los derechos sobre los textos que han publicado en periódicos y revistas. Los editores y los autores deberán llegar a un acuerdo para que estos contenidos, que se encuentran disponibles en los medios tradicionales, puedan incluirse en la Red. Del mismo modo, es importante que los usuarios de la Red comiencen a ser conscientes de que no todo lo que hay en Internet es gratuito. Y no es gratuito por el mismo motivo que no lo es un periódico tradicional. Ambos son negocios que necesitan ser rentables para subsistir. Por este motivo, actualmente están comenzando a aparecer sistemas de micropagos o servicios de suscripción que permiten acceder a contenidos de periódicos y revistas *online* y a otros valores añadidos del medio digital.

Estos aspectos hay que tenerlos muy en cuenta. Cuando se está realizando una investigación sobre un tema específico las fuentes en Internet, tanto las disponibles en la Red profunda, como en la superficial, son recursos importantes a tener en cuenta. Sin embargo, no se debe olvidar de consultar las posibles fuentes en soporte tradicional: libros, periódicos, revistas, informes no digitalizados, etc. Como se ha señalado a lo largo de este capítulo, no todos los datos disponibles en los soportes tradicionales están incluidos en la Red y probablemente nunca lo estarán. Más bien una ínfima parte sí lo están. Cada vez más, pero aún sigue siendo una pequeña parte.

Estos aspectos hay que tenerlos muy en cuenta. Cuando se está realizando una investigación sobre un tema específico las fuentes en Internet, tanto las disponibles en la Red profunda, como en la superficial, son recursos importantes a tener en cuenta. Sin embargo, no se debe olvidar de consultar las posibles fuentes en soporte tradicional: libros, periódicos, revistas, informes no digitalizados, etc. Como se ha señalado a lo largo de este capítulo, no todos los datos disponibles en los soportes tradicionales están incluidos en la Red. Más bien una ínfima parte. Cada vez más, pero aún sigue siendo una pequeña parte.

6. 4- LA RED PROFUNDA O INTERNET INVISIBLE

Fue en 1994 cuando el Dr. Jill Ellsworth acuñó el término “Internet Invisible” para referirse a la información que no podían encontrar los buscadores convencionales. La importancia de que la información contenida en las bases de datos pudiera ser localizada mediante buscadores u algún otro método, llevó a desarrollar la primera utilidad de búsqueda, el llamado AT1, que fue concluido a comienzos de 1997. Sin embargo la empresa PLS, propietaria del AT1, fue absorbida por America Online (AOL) en 1998 y este servicio fue abandonado.

Como hemos apuntado con anterioridad, esta supuesta “invisibilidad” de ciertos sitios de Internet no es sino fruto de la tecnología actual de los motores de búsqueda convencionales. Que no sean accesibles mediante este método de búsqueda, no significa que no lo sean si se utiliza en este empeño una tecnología diferente y más avanzada. Por este motivo algunos autores, (y nosotros en este caso) prefieren definir esta parte de la web como Red Profunda.

Existe una web en Internet denominada “The Invisible Web” (<http://www.invisibleweb.com>) de gran utilidad a la hora de buscar documentos en la Web profunda, ya que tiene catalogados más de 10.000 bases de datos, buscadores y archivos de todo tipo, no accesibles mediante buscadores convencionales.

Asimismo, la compañía BrightPlanet han elaborado una tecnología denominada “LexiBot” que rastrea la Red profunda y encuentra resultados precisos a consultas realizadas. “LexiBot” es definida como un motor de interrogación directo. Otra de sus características es su capacidad para cualificar, clasificar y organizar. Sin embargo es su habilidad para interrogar directa y simultáneamente a múltiples sitios de búsqueda lo que le permite recuperar el contenido de la Web profunda.

Con respecto a esta nueva técnica hay que añadir que al igual que es capaz de rastrear información en la Web profunda, también permite buscar en la Web superficial dado el hecho de que los actuales motores de búsqueda son en si mismos bases de datos que admiten búsqueda.

Desde un punto de vista documental se distinguen cuatro grandes categorías en la Web profunda¹⁴³:

- a) Catálogos de bibliotecas y bases de datos bibliográficas. Los OPAC y las bases de datos de registros bibliográficos accesibles a través de pasarelas web resultan imposibles de interrogar exhaustivamente por los motores convencionales.
- b) Bases de datos. Se incluyen todas las bases de datos no bibliográficas, con un amplio abanico de tipologías, desde las bases alfanuméricas o a texto completo hasta las obras de referencia, tipo diccionario o enciclopedia.
- c) Revistas electrónicas y archivos de documentos. Tanto las de acceso gratuito, pero que exigen registro previo, como las de pago protegidas por palabra clave son invisibles a los motores. Añádanse las que solo o recen la recuperación de sus artículos mediante búsqueda y no utilizan navegación por sumarios de contenidos.
- d) Depósitos y documentos en formatos que no se pueden indexar. Todos aquellos documentos construidos con lenguajes de descripción de páginas (Acrobat, Postscript, etc...).

La Red profunda o Internet invisible está compuesta por aquellas páginas o archivos, normalmente de alta calidad, que los motores de búsqueda generales no pueden (debido a limitaciones técnicas), o no quieren (deliberadamente) añadir a sus índices de páginas web.

Básicamente, los motores de búsqueda son diseñados para indexar páginas web. Como señalamos en la 2ª parte de esta tesis, éstos utilizan programas denominados “arañas” para rastrear y encontrar páginas web en los servidores de todo el mundo. Para un servidor, no existe ninguna diferencia entre que la petición de una determinada página venga de un usuario que está utilizando cualquier navegador, o de una araña de

¹⁴³ Aguillo, Isidro: *Internet invisible o Infranet: definición, clasificación y evaluación*.
<http://dois.mimas.ac.uk/DoIS/data/Papers/julmjoifp4172.html>

un motor de búsqueda. En cualquiera de estos casos, el servidor responderá a la petición enviando la página solicitada al ordenador que la ha requerido.

La principal diferencia entre un usuario que utiliza un navegador y la araña, es que la persona es capaz de escribir manualmente la dirección web (URL) en el navegador y descargarse la página solicitada. Las arañas obviamente carecen de esta posibilidad por lo que su técnica para encontrar otras páginas recae en los enlaces que puede encontrar en las páginas por las que comienza la búsqueda. Si una página no tiene enlaces que apunten hacia ella, la araña no podrá encontrarla. Estas páginas forman la mayor parte de la Red profunda. Sin embargo, estas páginas no conectadas, pueden dejar de ser “invisibles” en el caso de que se cree un enlace en otra web que apunte hacia ellas, o que el autor solicite al motor de búsqueda que su página sea rastreada, a través de los formularios de “añadir URL”.

Los problemas técnicos comienzan cuando la araña encuentra un archivo en el que no hay texto simple. Los motores de búsqueda están diseñados para indexar texto, además de estar altamente cualificados para realizar búsquedas y otras operaciones sobre textos simples, pero si nos salimos de este formato, es fácil que no obtengamos el mismo resultado, al menos con las herramientas de búsqueda disponibles en la actualidad. Éste está siendo uno de los grandes retos de los motores de búsqueda y cada vez más se están empezando a incorporar opciones que permiten buscar ciertos tipos de archivos no textuales, incluyendo imágenes, audio y video, archivos PDF o Postscripts (excepto Google), Flash, Shockwave, programas ejecutables o archivos comprimidos. El problema de estos archivos es que no están programados con HTML. Técnicamente la mayoría de estos formatos pueden ser indexados por los motores eligen no indexarlos por razones de negocio. En primer lugar, hay mucha menos demanda de este tipo de formatos que de textos HTML; y además estos formatos son más difíciles de indexar por lo que requieren mayores recursos del servidor. Este tipo de archivos no suponen una parte muy grande de la Red profunda, sin embargo, a medida que pasa el tiempo se va extendiendo rápidamente su uso, particularmente para documentos de alta calidad, concernientes a los gobiernos o alguna otra autoridad.

En realidad, el mayor reto al que se enfrentan los motores de búsqueda recae en el acceso a la información que se encuentra en las bases de datos. En este sentido, estaríamos hablando de falta de accesibilidad a quizá millones de bases de datos con

información de alta calidad para el usuario. El problema radica en que cada una de las bases de datos es única en diseño, estructura, herramientas y método de búsqueda. A diferencia de los archivos HTML en los que el motor puede rastrear e indexar el contenido, la información que se encuentra en las bases de datos es difícil de acceder por las siguientes razones:

- Normalmente, los motores de búsqueda no tienen dificultad para hallar las “home page” o página de inicio de las bases de datos, ya que se podría decir que son páginas como cualquier otra página web, programadas, normalmente en HTML. Sin embargo, la forma de acceder al contenido al que da acceso este “interface” es incomprendible para el motor.
- Mientras que estos “interfaces” poseen potentes herramientas para el usuario, éstas actúan bloqueando a la araña del motor de búsqueda. Cuando la araña se encuentra con la base de datos, es como si hubiera llegado a la entrada de una gran biblioteca bloqueada con puertas de seguridad. En este caso el buscador puede localizar e indexar la dirección de la biblioteca, pero no puede acceder a los libros o documentos que contiene.

Estas bases de datos componen la mayor parte de la que nosotros llamamos Red profunda. Son accesibles a través de la Red, pero pueden o no estar realmente en la Red. A la hora de buscar dentro de la base de datos, se debe usar los buscadores y otras herramientas de búsqueda que ofrece la propia base de datos. La ventaja de esto, es que el usuario tiene a su disposición herramientas de búsqueda creadas específicamente para obtener el mejor resultado en una base de datos concreta. La desventaja radica en que necesitas encontrar la base de datos en primer lugar, tarea en la que a veces los motores de búsqueda ayudan y otras no.

Hay diferentes tipos de bases de datos a los que se puede acceder a través de la Red. Sin embargo, se debe tener en cuenta que el hecho de que los documentos web estén almacenados en una base de datos no los hace automáticamente parte de la Red profunda. De hecho, algunos sitios web usan las bases de datos, no simplemente por la posibilidad que ofrecen de sofisticadas herramientas de búsqueda, sino porque la arquitectura de las bases de datos permite un mantenimiento más fácil del sitio.

- En primer lugar encontramos aquellas bases de datos diseñadas para albergar contenidos personales de usuarios concretos. Un ejemplo de este caso es: My Yahoo!, Mi ya, etc. Estos sitios utilizan bases de datos que generan páginas HTML para un usuario específico. Por el hecho de que cada página y contenido es creado para cada usuario concreto, no habría necesidad de indexarla en un motor de búsqueda general.
- En segundo lugar, están las bases de datos diseñadas para albergar datos que se actualizan en tiempo real, como pueden ser valores de Bolsa, información del tiempo, horarios de trenes, autobuses o aviones, etc. Estos datos son necesarios guardarlos en bases de datos debido a su volumen y a la gran rapidez con la varía dicha información. Técnicamente, la mayoría de estos datos podrían indexarse porque la información, aunque es tomada de una base de datos, es publicada en una página HTML normal y corriente. Pero el hecho de que cambie con tanta frecuencia y su valor sea de tan limitada duración, es uno de los puntos determinantes para que no merezca la pena indexarse. Además a los motores de búsqueda les cuesta bastante mantener actualizada este tipo de información. Mantener actualizada una web con información en tiempo real consumiría tantos recursos que, aunque tal y como avanza la tecnología en la actualidad no se puede decir que sea imposible, sería muy difícil para un motor de búsqueda convencional.
- La información disponible en el tercer tipo de bases de datos, es recuperada a través de potentes herramientas de búsqueda especializadas y diseñadas para recuperar la información usando el método más rápido y efectivo posible. La única forma de acceder al contenido de estas bases de datos es a través de las herramientas de búsqueda de las propias bases de datos. Éstas forman la mayor parte de la Red profunda.

Tipos de contenido en la Red profunda	¿Por qué es invisible a los buscadores generalistas?
Páginas web sin enlaces	Si no hay enlaces que apunten hacia ellas

	los motores de búsqueda no pueden localizarlas.
Páginas cuyo contenido son imágenes, audio y vídeo.	No suele haber suficiente texto para que el motor de búsqueda entienda el contenido de la página.
Páginas cuyo principal contenido son documentos en formato PDF, Postscript, Flash, Shockwave, programas ejecutables o archivos comprimidos.	Técnicamente se pueden indexar, pero normalmente son ignoradas por motivos de negocio o políticos.
Contenidos en bases de datos relacionadas	Las arañas de los motores de búsqueda no pueden rellenar todos los campos requeridos en sus formularios interactivos.
Contenido en tiempo real	Técnicamente es posible indexarla, pero no merece la pena por su carácter efímero, gran volumen y muy cambiante.
Contenido generado de forma dinámica	Este contenido es irrelevante para la mayor parte de los usuarios.

6. 5. LOS MOTORES DE BÚSQUEDA Y LA RED PROFUNDA

La función principal de los motores de búsqueda es ayudar a los usuarios a localizar en la Red documentos de hipertexto, principalmente programados con el lenguaje Hypertext Markup Lenguaje (HTML). A medida que la Red ha ido evolucionando, se han ido incluyendo documentos o archivos en otros lenguajes y formatos a los que, poco a poco, se están adaptando los motores de búsqueda, aunque, por ahora, la localización de documentos de texto es la opción más segura.

Los documentos HTML se componen de dos partes: una cabeza (head) y un cuerpo (body), separadas claramente en el código fuente. La parte de la cabeza contiene el título y también puede contener información adicional que describa el documento lo cual puede ser usado por el motor de búsqueda para clasificar la página. La parte del cuerpo contiene el documento en sí mismo y, por tanto, lo que el motor de búsqueda tendría que registrar. La simplicidad de este formato permite que a los motores les sea fácil indexar cada palabra del documento e incluirlas en sus enormes bases de datos. Las dificultades comienzan cuando el contenido no consiste en una de estas páginas simples.

El primer paso que un rastreador debe seguir es averiguar si el acceso a unas determinadas páginas en un servidor está restringido. El programador de la página puede usar, principalmente, tres métodos para prevenir que los motores de búsqueda indexen una página. Dos de ellos utilizan específicamente técnicas para bloquear a través de Robots de Protocolo de Exclusión que consisten en una serie de reglas que permite a los programadores especificar qué parte de los servidores están abiertas a los rastreadores de los motores de búsqueda y cuáles están vedadas. El programador crea una lista de archivos y directorios que no deben de ser rastreados o indexados y guarda esta lista en el servidor en un archivo llamado "robots.txt". Este archivo opcional se sitúa en el nivel más alto de un sitio web. En realidad, se trata simplemente de una petición al buscador para que no indexe estas páginas y, normalmente, los motores de búsqueda lo respetan.

Una segunda opción para prevenir que una página sea indexada actúa del mismo modo que los "robots.txt" pero sobre una página específica. Los programadores pueden incluir una etiqueta (meta tag) con la instrucción "noindex" en la parte superior del documento. Ambas instrucciones se usan para bloquear a los motores de búsqueda. La

única diferencia entre las dos órdenes que la etiqueta de “noindex” se utilizaría para una página determinada, mientras que “robots.txt” puede usarse para prevenir que se indexen páginas individuales, grupos de páginas o webs enteras.

Un tercer modo de prevenir que una página sea rastreada e indexada es mediante contraseñas. Este método es más seguro que los dos anteriores ya que establece una barrera técnica mientras que los anteriores se trata de una medida voluntaria. Sin embargo esta opción no suele ser muy usada ya que a las páginas que lo tienen solamente pueden acceder aquellos que conocen la contraseña, mientras que a las que tienen sólo el robot de protocolo de exclusión puede acceder cualquier usuario excepto los motores de búsqueda.

Las páginas que usan alguno de estos tres métodos son también parte de la Red profunda. La mayoría de ellas no tienen una barrera técnica que les impida ser indexadas, simplemente sus programadores piden voluntariamente que no lo sean. Una vez que la araña ha determinado si le es permitido el acceso a una página, el siguiente paso es intentar captarla y dejarla en manos del indexador del motor de búsqueda. Este acto es definitorio para que una página sea visible o invisible para el buscador y, por tanto, para la mayoría de los usuarios de Internet.

A continuación exponemos algunas de las variaciones que los motores de búsqueda encuentran en su tarea de descubrir páginas en la Red¹⁴⁴:

- **Caso 1:** El motor de búsqueda encuentra una página programada en lenguaje HTML que incluye gráficos. Es el más común de los casos y sería perfectamente indexable.
- **Caso 2:** El motor encuentra una página hecha en HTML pero ésta contiene campos de texto, casillas para chequear u otros componentes que requieren la interactividad del usuario. En este caso la página en sí puede ser indexada ya que está hecha con el lenguaje de programación propicio para ello; sin embargo, el contenido que se encuentra dentro de estos formularios que el usuario puede ver tras enviar los datos pedidos, son invisibles al motor de búsqueda. Si se encuentra esto en una página hay dos posibilidades:

¹⁴⁴ Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

- 1) El formulario que se encuentra dentro de la página se usa para seleccionar preferencias de usuario. Otras páginas del sitio están programadas en lenguaje HTML plano, teniendo que en cuenta que hay enlaces en otras webs que apuntan hacia ellas. En este caso tanto el formulario como el contenido que se esconde detrás es visible e indexable por los motores . Un ejemplo de este caso es la web de Hoover's Business Profiles¹⁴⁵ , que incluye un formulario para buscar empresas pero que presenta los datos de estas compañías que páginas HTML que pueden ser indexadas.

- 2) El formulario es usado para recoger información que genera páginas dinámicas cuando los datos son enviados. En este caso, aunque el formulario es visible el contenido detrás de éste es "invisible" a los motores. Debido a que los formularios pueden tener literalmente infinitas variaciones y que entre sus funciones está el acceso a contenido dinámico, los motores de búsqueda serán incapaces de acceder a él. Un buen ejemplo de este caso es la base de datos del World Bank Group's Economics of Tobacco Control Country Data Report¹⁴⁶ , que permite seleccionar cualquier país y elegir una amplia variedad de informes de cada uno de ellos. En el futuro los formularios supondrán menos retos para los motores de búsqueda. En la actualidad se están desarrollando varios proyectos para crear rastreadores más inteligentes que son capaces de rellenar formularios y procesar información. Uno de ellos utiliza herramientas preprogramas para interactuar con los formularios de bases de datos específicas. Otros utilizan una combinación de fuerza bruta con inteligencia artificial para intentar adivinar la forma de rellenar los formularios y traspasarlos para averiguar la información que hay detrás. De todas formas, aunque los buscadores generales adquieran la habilidad de

¹⁴⁵ Hoover's Business Profiles: accessible en: <http://www.hoovers.com>)

¹⁴⁶ Base de datos de World Bank Group's Economics of Tobacco Control Country Data Report. Accesible en <http://www1.worldbank.org/tobacco/database.asp>.

rastrear el contenido de las bases de datos, es probable que los buscadores específicos de cada base de datos continúen siendo la mejor fórmula para interactuar con ellas.

- 3) El rastreador encuentra una página generada dinámicamente. La forma más sencilla de saber si una página ha sido dinámicamente o no es ver si lleva el símbolo “?” en su dirección web o URL. Técnicamente estas páginas son visibles a los buscadores. Los buscadores pueden rastrear cualquier página que puede ser visualizada en un navegador. Un ejemplo de este tipo de sitios web es el motor de búsqueda experimental Compaq’s SpeechBoot¹⁴⁷ que usa contenidos de audio y vídeo por reconocimiento de voz y convierte los archivos de “streaming” a texto. En la actualidad, las páginas generadas dinámicamente suponen un reto para muchos buscadores. Estas páginas son creadas a partir de encriptaciones, un programa de ordenador que realiza una selección a partir de varias opciones que dan lugar a una nueva página. Hasta que este programa o “script” empieza a actuar el buscador no tiene forma de conocer que es lo que realmente éste va a hacer. Realmente el script debería crear simplemente una nueva página web. Sin embargo, algunos programadores “poco éticos” han creado “scripts” millones de páginas similares, aunque no idénticas en un esfuerzo para sobrecargar a los buscadores de páginas inútiles. Estas “trampas” pueden resultar muy dañinas para los buscadores, por lo que muchas de las empresas de buscadores han tomado la decisión de evitar rastrear o indexar URL que generan contenido dinámico. Estas páginas componen una gran proporción de la llamada “web opaca” ya que potencialmente pueden ser indexadas pero no lo son. En este sentido, la empresa Inktomi ofrece la siguiente explicación sobre su buscador “Slurp”¹⁴⁸: “Slurp ofrece la posibilidad

¹⁴⁷ Compaq’s experimental SpeechBoot search engine. Accesible en: <http://www.speechbot.com>.

¹⁴⁸ Buscador de Inktomi “Slurp”: <http://www.inktomi.com/slurp.html>

de rastrear enlaces dinámicos o documentos generados dinámicamente. Sin embargo no los rastrea por defecto. Y hay una serie de buenas razones para esto”. Un par de ellas son que los documentos generados dinámicamente pueden crear infinitas URL’s y que los enlaces y documentos generados dinámicamente pueden ser diferentes cada vez que se recuperan por lo que indexarlos no supone ninguna utilidad”. A medida que la tecnología de los buscadores avanza lo más probable es que parte de este contenido generado dinámicamente sea indexado. Normalmente será el contenido que consiste esencialmente en páginas estáticas que están recogidas en bases de datos por razones específicas. Desde el momento en que los motores de búsqueda aprendan qué sitios con contenidos generados dinámicamente son fiables o que están libres de “trampas”, éstos comenzarán a aparecer en los índices de los motores de búsqueda. Sin embargo, por ahora, continúan formando parte de la Red profunda.

- 4) El rastreador encuentra una página HTML que no contiene nada que pueda ser indexado. Hay miles, sino millones, de páginas que contienen flash, imágenes en formatos .jpg, .gif, gráficos web, archivos de audio u otro contenido no textual sobre marcos HTML básicos, en el cuerpo de la página. Este tipo de páginas son parte de la Red profunda porque no contienen nada que los motores de búsqueda puedan indexar. Los motores especializados en búsqueda multimedia como ditto.com y Webseek son capaces de reconocer algunos archivos no textuales e indexar una mínima información sobre ellos, como puede ser el nombre del archivo y su tamaño, pero esto aún no supone una solución para la búsqueda efectiva de este tipo de arribos por palabras clave.
- 5) El buscador encuentra un sitio que ofrece información dinámica en tiempo real, como puede ser una web que contenga datos de la Bolsa o sobre llegadas y salidas de vuelos. Este tipo de sitios forman también

parte de la Red profunda principalmente porque, desde un punto de vista práctico, no es conveniente indexarlos. A pesar de ser técnicamente posible indexar la mayoría de estos datos que son ofrecidos en tiempo real, su único valor es el de llevar un archivo histórico y la enorme cantidad de datos que supondría acapararía rápidamente la capacidad del motor de búsqueda. Un buen ejemplo de este tipo de sitios es TheTrip.com's Flight tracker¹⁴⁹, que ofrece información en tiempo real sobre la llegada de vuelos de numerosas compañías.

- 6) El buscador encuentra un PDF o un archivo Postscript. Éstos son formatos de texto que guardan la apariencia de un documento y lo muestran de forma idéntica independientemente del tipo de ordenador que el usuario esté usando. Técnicamente la tarea de convertir un archivo PDF o Postscript a texto simple, para que pueda indexarlo el buscador, no es difícil. Sin embargo la mayoría de los motores no los indexan ya que llevan más tiempo y el proceso es costoso. Otra de las razones es que muchos de estos documentos son muy técnicos o académicos, por lo que son útiles para una comunidad pequeña pero irrelevantes para la mayoría de usuarios de los motores. Sin embargo, a raíz de que los gobiernos e instituciones públicas están empezando a adoptar estos formatos para difundir documentos oficiales a través de la Red, la política de los motores está variando y muchos están comenzando a indexar estos archivos. Un motor de búsqueda experimental llamado Researchindex¹⁵⁰, creado por científicos del instituto Netresearch, no solamente indexa archivos PDF y Postscript, sino que también aprovecha la ventaja de las características únicas que aparecen en común en los documentos, usando el formato para mejorar los resultados de la búsqueda. Por ejemplo, los documentos académicos normalmente suelen citar otros documentos e incluyen una

¹⁴⁹ TheTrip.com's Flight tracker. Accesible en: <http://www.trip.com/ft/home/0,2096,1-1,00.shtml>..

¹⁵⁰ Researchindex. Accesible en: <http://www.researchindex.com>

lista de referencia de otros materiales. Además de indexar el texto completo de los documentos, Researchindex también crea un índice de citas que permite localizar estos documentos de referencia más fácilmente. En definitiva, hasta que los buscadores comiencen a indexar regularmente archivos PDF y Postscript, tal y como lo está haciendo Google en la actualidad, la mayor parte de este tipo de documentos seguirán formando parte de la Red profunda.

- 7) Y por último, el buscador encuentra una base de datos que ofrece un interfaz web. Hay miles de bases de datos que contienen información de gran calidad y que está disponible a través de la Red de redes. Sin embargo, los motores no pueden indexar este material ya que a pesar de que el interfaz de dicha base de datos está normalmente creado en un lenguaje de programación HTML, la base de datos en si misma puede haber sido creada con anterioridad al desarrollo HTML por lo que su sistema es incompatible con los protocolos usados por los motores. También es posible que la base de datos requiera un registro previo del usuario, que el acceso esté reservado a usuarios específicos o que sea necesario el pago previo al acceso.

Dificultades en los rastreadores

Aunque como se ha explicado con anterioridad, existen razones técnicas por las que los motores de búsqueda no pueden indexar los recursos disponibles en la Red profunda, también existen otra serie de razones referentes a la validez de la información en cuestión, su autor o su calidad. Teniendo en cuenta el hecho de que cualquier persona que tenga un ordenador conectado a la Red puede publicar en ella todo tipo de datos con gran facilidad y sin ningún filtro, el hecho de realizar una selección y autenticación de las webs que se indexan en determinadas bases de datos rigurosas disponibles en la Red profunda, es un factor importante.

Uno de los problemas más importantes con el que se encuentran los rastreadores de los motores de búsqueda generalistas son webs “poco éticas” programadas para bloquear a estas herramientas de los motores haciéndoles caer en un ciclo en el que indexan millones de “páginas basura”. Muchos de los motores de búsqueda más importantes en la actualidad que se han visto obligados a lidiar con esta tesitura, han desarrollado estrictas guías con las que se intenta evitar este tipo de webs. Sin embargo, en ocasiones, el efecto que se produce, a parte del deseado, es el de dejar de indexar contenido legítimo.

En definitiva, cuando un usuario cualquiera y más el profesional de la información accede a una información, ya sea en la Red profunda como en la superficial, debe de ser crítico y no fiarse a primera vista de lo que hay allí publicado ya que cualquiera puede haberlo escrito sin seguir ningún tipo de rigurosidad o ética. A continuación se exponen una serie de preguntas que se debería hacer el periodista, y en realidad todos aquellos usuarios que busquen información rigurosa en la Red¹⁵¹.

- **¿Quién mantiene el contenido allí disponible?** La primera pregunta que se debe hacer el usuario al tratar de averiguar el rigor de una web es quién es su autor o la persona responsable de su actualización. En este aspecto, el profesional de la información debe de actuar como lo haría en el caso de que estuviera analizando la credibilidad de una fuente que no estuviera en Internet. Los sitios web de las empresas suelen estar dotados de gran cantidad de información sobre ella. Sin embargo estos datos suelen estar sesgados a favor de la propia compañía. Es por esto por lo que se deben comparar con otras informaciones que se pudieran encontrar en otros sitios. Por otro lado, las webs de los gobiernos o de las instituciones académicas suelen ser bastante fiables en sus contenidos, aunque en este caso el usuario también deberá tener en cuenta al autor. En el caso de que se trate de una web creada por un particular el usuario deberá atender a las posibles notas aclaratorias sobre la identidad del autor. A pesar de la aparente calidad que pueda tener la

¹⁵¹ Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

información en ellas disponible, éstas siempre serán las fuentes menos fiables, dada la facilidad con la que se puede acudir al engaño.

- **¿Quién es el proveedor del contenido?** Dado el hecho de que en la Red cualquiera puede publicar, a la hora de atender a la veracidad o rigurosidad de una información allí publicada, es necesario prestar especial atención a la identidad del autor. En el caso de que se trate de un particular, el usuario debe intentar localizar una biografía de éste que cite su posible bibliografía anterior u otras experiencias demostrables en el trabajo en cuestión que prueben la rigurosidad del autor en la materia sobre la que ha escrito. Algunos motores de búsqueda facilitan este trabajo. En el caso de Google, éste intenta identificar al autor examinando la estructura de todos los enlaces del sitio web para averiguar cuantos enlaces de otras páginas apuntan hacia la página web en cuestión. Google también chequea si hay enlaces hacia esa determinada página web desde sitios webs importantes que pudieran otorgarle una mayor credibilidad. Sin embargo esta regla no es rigurosa. Se puede dar el caso de que una página creada por un premio Nobel no aparezca en los primeros lugares del listado de resultados de Google si otras páginas importantes de la Red no tienen enlaces apuntando hacia ella.
- **¿Hay posibilidad de que la información esté sesgada?** La Red está llena de datos no veraces derivados de la facilidad de publicación en la Red. A pesar de que una información en una página pueda parecer objetiva, siempre hay que contrastarla. Es particularmente importante observar si tiene algún tipo de patrocinio de un grupo o de una organización y averiguar a qué tipo de actividades se dedican, sobre todo en informaciones relativas a la salud. Algunas grandes empresas se dedican a difundir fuentes de información para mejorar condiciones de salud específicas o intentan atraer al usuario hacia “inmejorables” estilos de vida que podrían conseguir usando un determinado producto. Indiscutiblemente el usuario o el periodista en este caso, debe de tener en cuenta de que esta información está sesgada a favor de la empresa.

Asimismo, la Red es un medio de gran poder para los grupos activistas que intentan ejercer influencia negativa sobre una determinada compañía. La mayoría de ellos intentan parecer objetivos en la Red presentando lo que parece ser información contrastada y veraz, cuando en realidad se trata de datos procedentes de un solo punto de vista. En definitiva, la clave está en no fiarse de la primera impresión y en contrastar siempre la información.

- Examinar la URL o dirección web
- Examinar los enlaces que contiene la página
- ¿Está la información actualizada?
- Usar el sentido común.

6.6. NIVELES DE PROFUNDIDAD DE LA INFORMACIÓN DIGITAL

Olvidando las razones técnicas, hay otras motivos por las que cierto tipo de información no es indexada por los motores de búsqueda. En realidad, la llamada Red profunda podríamos dividirla en varios niveles dependiendo del grado de “invisibilidad” para el buscador que tenga. Serían los siguientes¹⁵²:

- a) La Red opaca
- b) La Red privada
- c) La Red de propiedad
- d) La verdadera Red profunda

a) **La Red opaca** se compone de archivos que, aunque técnicamente pueden ser indexados, son dejados de lado por los motores. Esta parte de la Red es bastante amplia y suele representar un reto para los usuarios que intentan encontrar algo a este nivel.

¹⁵² Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

Mientras que cierto contenido de la Red profunda es localizable si se conoce cómo encontrarlo, la información en la Red opaca normalmente es más difícil de localizar por las siguientes razones:

1) Rastrear un sitio web e indexar cada una de sus páginas es una operación intensa y costosa para un buscador. Hace unos años, la mayoría de los motores de búsqueda sólo indexaba había unas pocas páginas, a modo de ejemplo, por cada sitio web, en lugar de realizar un rastreo profundo e indexar cada una de las páginas de dicho sitio. De esta forma, también limitaban el coste del proceso. En general, los motores de búsqueda no suelen revelar la profundidad de su rastreo, aunque la tendencia de todos ellos es indexar cuántas más páginas mejor. A medida que el coste del rastreo y de indexar las páginas va disminuyendo y el tamaño de los índices de los motores de búsqueda se ha ido convirtiendo en una “competición” entre todos ellos, la profundidad del rastreo está dejando de ser un problema para los usuarios. Sin embargo, sigue sin haber garantía de que cada página de cada sitio web esté siendo rastreada e indexada. Ésta es una de las principales razones por la que cierto contenido que podría resultar interesante para el usuario puede ser “invisible” para aquellos que no conocen bien todas las posibilidades o recursos de búsqueda en Internet.

2) La Red está en constante cambio ya que continuamente se están añadiendo quitando páginas. Incluso los buscadores más potentes pueden visitar únicamente alrededor de 10 millones de páginas diarias, una pequeña fracción del número total de páginas disponibles en la Red. Por este motivo cada motor debe decidir la mejor forma de desplegar sus rastreadores y crear un horario que determine la frecuencia de rastreo de una determinada página o sitio web.

Un artículo de los expertos en búsqueda de información en la Red Steve Lawrence y Lee Giles publicado en 1999 por la revista *Nature*, afirmaba que indexar páginas nuevas o modificadas por uno de los motores más potentes podía llevar meses. Aunque aparentemente la situación desde entonces ha mejorado, muchos motores aún continúan actualizando sus índices una vez al mes e incluso con menos frecuencia. Hay que tener en cuenta que no es suficiente para un motor visitar una página una sola vez y asumir que dicha página estará disponible en un periodo de tiempo más o menos largo.

Para que no se diera el problema de mantener en el índice páginas que ya no existen los rastreadores deberían volver a las webs con una periodicidad mayor, no solamente para verificar su existencia, sino para realizar una nueva copia del contenido (que puede haber variado) e indexar posibles nuevas páginas que hayan sido añadidas al sitio.

Según un estudio de Koehler¹⁵³, parece que la vida media de una página web es inferior a dos años y la de un sitio web es algo más de dos años, lo cual significa que si un rastreador vuelve a un sitio que ha indexado dos años atrás, puede contener el mismo número de direcciones web (URL's) pero sólo la mitad de páginas originales que han sido reemplazadas.

Asimismo, los nuevos sitios suelen ser los menos propicios para ser encontrados por los rastreadores, ya que, en teoría, habrá pocos sitios en la Red que tengan enlaces apuntando hacia ellos. Hasta el momento en que los motores logran encontrarlos e indexarlos, éstos también forman parte de la Red profunda.

3) Es bastante común que al realizar una búsqueda en algún motor el resultado sea una lista muy larga de webs encontradas. Sin embargo, la mayoría de los motores restringen el número total de resultados desplegados. Normalmente suelen ser entre 200 y 1.000 documentos. Cuando ocurre esto, significa que la mayoría de las páginas que el buscador ha encontrado relevantes son inaccesibles, dado que la lista de resultados ha sido arbitrariamente truncada. Estas páginas también formarían parte de la Red profunda.

Los usuarios que conocen bien todas las posibilidades de los buscadores pueden solucionar este problema usando estrategias de búsqueda y los filtros de búsqueda avanzada que ofrecen la mayoría de los motores.

4) Para que un rastreador de un motor de búsqueda pueda acceder a una página hace falta que el autor de la web utilice la opción de “añadir URL” para pedir que su página sea visitada e indexada, o que el buscador por su cuenta descubra la página

¹⁵³ Koehler, Wallace. Digital libraries and World Wide Web sites and page persistence. 2000 (Accesible en: <http://informationr.net/ir/4-4/paper60.html>)

siguiendo los enlaces que apuntan hacia ella desde otras webs. Cuando los autores no dan a conocer las páginas al buscador, ni tampoco hay enlaces apuntando hacia ellas, a estas páginas se las llama “desconectadas”. No son rastreadas e indexadas simplemente porque el motor no tiene forma de localizarlas.

A menudo, estas páginas no suelen tener barreras técnicas para los motores. Un estudio de IBM, Altavista y Compaq elaborado en Mayo de 2000 sacó a la luz que el número total de páginas desconectadas suponía el 20% de las webs que potencialmente podían ser indexadas.

En definitiva, la Red opaca es grande pero no impenetrable por los motores de búsqueda. Aquellos usuarios que conocen bien los recursos de búsqueda pueden encontrar lo que buscan en esta parte de la Red. Además, los motores están constantemente mejorando sus métodos para intentar localizar e indexar las páginas disponibles en la Red opaca.

b) **La Red privada** está compuesta por aquellas páginas web que aunque técnicamente pueden ser indexadas, son excluidas deliberadamente de los motores de búsqueda. El autor de una página puede prevenir que sea indexada por los siguiente métodos:

- Mediante una contraseña que proteja la página. La araña de un motor de búsqueda no puede atravesar un formulario en el que se requiere un nombre de usuario y una contraseña.
- Usar los archivos “robots.txt” para advertir a la araña que no acceda a la página en cuestión.
- Usar una etiqueta “noindex” en la parte superior de la programación de la página para prevenir que la araña lea el cuerpo de la noticia.

En general la Red privada no es muy usada por los usuarios. Estas páginas simplemente se valen de la Red para ofrecer acceso a ciertos usuarios con permiso para consultarlas.

c) Los motores de búsqueda tampoco pueden indexar las páginas que se hallan en la llamada “**Red de propiedad**”, porque únicamente son accesibles para aquellos usuarios que previamente se han registrado. En la mayoría de los casos este registro es gratuito. A pesar de esto, y por ahora, el motor de búsqueda no puede superar el proceso de registro, por lo que estas páginas le son inaccesibles. Un ejemplo de este caso es el diario estadounidense *The New York Times*.

Hay otro tipo de páginas como *The Wall Street Journal*, contenidas en la “Red de propiedad” que están disponibles por una pequeña cuota por cada página que el usuario desee ver o por suscripción.

d) Hay algunas páginas que son técnicamente inaccesibles por los motores de búsqueda. Éstas componen la **verdadera Red profunda**. Aunque es cierto que debido a la constante evolución de los motores llegará un momento en que estas páginas entren a formar parte de la “Red superficial”. Entre aquellas más susceptibles de permanecer más tiempo inaccesibles están las que usan formatos que los motores actuales no están programados para reconocer, como: PDF, Postscript, Flash Shockwave, programas ejecutables y archivos comprimidos. Principalmente existen dos razones por las que estas páginas no son indexadas. En primer lugar, estos archivos tienen poco texto o directamente no tienen nada por lo que son difíciles de comparar su relevancia con respecto a otros documentos que sí contienen texto. La solución a este problema podría ser añadir información en las etiquetas HTML de la cabecera de la programación de la página. Pero incluso en este caso únicamente se indexaría la descripción de la etiqueta. La segunda razón por la que algunos de archivos no aparecen en los índices de búsqueda, es porque los motores han elegido omitirlos. En realidad pueden ser indexados, pero no lo son. Un buen ejemplo de este caso es el buscador “Research Index”, que permite rastrear e indexar archivos PDF, Postscript y archivos comprimidos en tiempo real, creando una base de datos específica a cada búsqueda. El

motor de búsqueda “Altavista” es capaz de indexar más de 250 formatos pero la parte gratuita y pública de este buscador sólo permite unos pocos.

Las páginas de contenido dinámico suelen ser las más problemáticas. Normalmente tampoco se trata de una dificultad técnica sino más bien la falta de voluntad de indexar este tipo de contenido. La mayoría de estas páginas están generadas en HTML por lo que, en teoría, son fácilmente indexables. Sin embargo, se han dado bastantes casos de rastreadores que se han visto atrapados en las llamadas “trampas de araña” que es un sitio con miles o millones de páginas diseñadas únicamente para bloquear a los motores. Visto esto hay muchos motores que no indexan páginas de contenido dinámico por esta razón.

Finalmente, la información que se encuentra en bases de datos y a la que no se puede acceder más que a través de una búsqueda específica en la propia base de datos, forma realmente parte de la Red profunda. Por ahora, los motores no están programados para entender la estructura de la base de datos o el lenguaje de comandos que usa para extraer la información.

6.7. LOS AGUJEROS NEGROS DE INTERNET

A comienzos de enero de 2002, tres investigadores de Arbor Networks¹⁵⁴ publicaron un informe, donde se indica el análisis de un problema detectado pero escasamente estudiado en la Red. El denominado “agujero negro” o “falla profunda” o “zona oscura” es el conjunto de la Red que permanece absolutamente aislado de cualquier intento de acceso. En este estudio se indica que es equivalente al 5% de la superficie de la Red, medida en términos de sitios web, equivalente a 100 millones de ellos. Esto supone, que un “sitio” desaparezca de la Red o, desde otro punto de vista, que el usuario se encuentre en un callejón sin salida.

Si bien ese 5% no tiene un enorme impacto en el usuario final, pone de manifiesto, en primer lugar, un problema técnico de importancia que se ahonda en la medida en que crece el tamaño y la complejidad de Internet, y que manifiesta la vulnerabilidad de la infraestructura de enrutadores, que son los encargados de asegurar el correcto funcionamiento de la Red. También, dentro de esta zona de inaccesibilidad se registran sitios del Ejército de los Estados Unidos que utilizan antiguas direcciones que los actuales enrutadores ya no reconocen.

La metodología utilizada en este informe se ha enfocado en realizar un análisis del comportamiento de diversos enrutadores centrales de Internet en los Estados Unidos, Europa y Asia, concentrándose en las interrupciones de navegación que se producen en forma más persistente, descartando aquellos que se resuelven en cuestión de minutos o de horas.

Las explicaciones del estudio señalado indican que esa falla se produce tanto por errores de diseño de los enrutadores, por “accidentes técnicos” o bien por la manipulación. De esta manera, el número de sitios que un usuario puede ver depende del lugar o del punto en el que se inicia la navegación y los posibles destinos, se

¹⁵⁴ Labovitz, Craig, Ahuja, Abha y Baley, Michel.

encuentra, en un determinado porcentaje de situaciones, determinados por factores que no controla.

Algunas de las explicaciones ensayadas en el estudio mencionado son:

- Disputas entre proveedores de servicios que pueden provocar que haya “desconexiones” de porciones de la Red a los que la empresa competidora da acceso,
- Configuración errónea de los enrutadores, a lo que se agrega que existen partes de Internet más interconectadas que otras y, algunas secciones de la Red actúan involuntariamente como “porteros” de otras secciones que se ubican más allá de ellas, impidiendo el acceso.
- Acciones de hackers y otros manipuladores informáticos ilegales que utilizan para sus acciones los agujeros que encuentran en el software de la Red.

Frente a estos problemas han comenzado a surgir diversos servicios especializados en la ordenación de la Red profunda, fundamentalmente directorios dedicados a la colección de enlaces a las bases de datos.

Entre ellos se encuentran los siguientes:

- **CompletePlanet**: Servicio creado por Bright Planet para ayudar a localizar recursos y bases de datos. Incluye un directorio que contiene algo más de 20.000 bases de datos accesibles en la Red.
- **Lexibot**: es una herramienta de software creada por Bright Planet para la realización de búsquedas en las bases de datos. Realiza un lento y complejo proceso de búsqueda, por lo que es una aplicación que corre del lado del “cliente”.
- **Alphasearch**: facilita acceso a más de 800 “pasarelas” especializadas.
- **Direct Search**: constituye una enorme colección de enlaces a las bases de datos y sitios de búsqueda, clasificados en forma temática.

- **Infomine**: desarrollado por la biblioteca de la Universidad de California, se especializa en información de tipo académico y científico, con aproximadamente 20.000 recursos de este tipo.
- **Intelliseek**: es una empresa especializada en “*aplicaciones inteligentes*”. Distribuye Bulls Eye, un software que tiene cierta similitud con Lexibot.
- **Invisibleweb.com**: es un sitio creado por Intelliseek, constando de un directorio con algo más de 100.000 enlaces a bases de datos, archivos y motores de búsqueda.
- **Lycos-Searchable Databases**: índice de bases de datos con aproximadamente 7.000 recursos de búsqueda especializados.
- **Open Directory Project**: es un gran directorio temático, con algo más de 2,5 millones de páginas recopiladas y organizadas por editores voluntarios.
- **RefDesk**: colección de enlaces a sitios.
- **Search Engine Watch**: es uno de los sitios pioneros en la búsqueda de referencias dentro de la web profunda.
- **Webdata**: es una guía con bases de datos, con evaluaciones sobre la calidad de las mismas.

6.8. LA BÚSQUEDA EN LA RED EN BASE A LA NECESIDAD DEL USUARIO

Antes de comenzar una búsqueda en Internet, es importante conocer los recursos que se deben utilizar dependiendo el tipo de búsqueda que se desee realizar.

Fundamentalmente, podemos considerar dos tipos de sitios en la Red:

- Aquellos que proveen de contenido
- Aquellos que facilitan la navegación web y el descubrimiento de recursos.

La mayoría de los webs pertenecientes a la Red profunda son proveedores de contenido, y normalmente este contenido se encuentra alojado en una base de datos, por lo que la única forma que hay de acceder a él es usando las herramientas propias del sitio. El punto clave en el que hay que fijarse para ver si este contenido puede ser o no indexado, es la forma en la que dicho contenido de la base de datos está disponible. Si ésta está compuesta simplemente por archivos en HTML y no está protegido por ninguna contraseña, ni restringido el acceso por robots de protocolo de exclusión, el contenido es perfectamente accesible por los motores, y por lo tanto no formaría parte de la Red profunda. Para que se dé este caso, el contenido debe estar almacenado en la base de datos y solamente debe de ser accesible usando las herramientas que facilita la propia base de datos.

En la Red hay algunos sitios, en los que parte de su contenido se puede indexar y otro no. Un ejemplo claro de este tipo mixto es el sitio web de la Biblioteca del Congreso de los Estados Unidos, que está considerado como uno de los que contienen mayor volumen de información de la Red. La mayor parte de su contenido está almacenado en bases de datos inaccesibles para los motores. Sin embargo, tiene también miles de páginas en HTML que sí pueden ser indexadas. Incluso algunos sitios duplican su contenido, manteniendo una de ellas en la base de datos y la otra como página HTML normal. Estas copias suelen ser llamadas sitios “espejo” o sitios “sombra” y suelen servir de contenido alternativo para que los motores puedan indexarlo. Un ejemplo de este último caso es la base de datos de la fuente de información sobre educación Clearinghouse (ERIC¹⁵⁵), en la que el material

¹⁵⁵ The Education Resource Information Clearinghouse (ERIC): accesible en: <http://www.accesseric.org>).

almacenado en la base de datos también aparece en periódicos digitales, libros y otras publicaciones.

También podemos tener algunos casos en los que es difícil saber si un motor sería capaz de indexarlo o no. En el caso de que se produzca esta ambigüedad, y el usuario no tenga claro si debe usar las herramientas propias del sitio o usar buscadores generales, la mejor opción es la primera, ya que los recursos propios del sitio la mayor parte de las veces suelen ofrecer mejor resultado, y sobre todo, suelen evitar el “ruido”, es decir, resultados que no tienen nada que ver con la consulta.

Direcciones web

La mejor forma de determinar si una página pertenece a la Red profunda o a la superficial, es examinando la dirección web o URL. La mayoría de las URLs son referencias directas a una página específica: ej. <http://www.elpais.es>; <http://www.google.com>. Cuando el usuario pulsa en un enlace, el navegador busca la página HTML solicitada y la despliega para el usuario. El rastreador de un motor de búsqueda realiza este mismo proceso, enviando una solicitud a un servidor web para que éste le devuelva una página web específica.

Sin embargo las direcciones web indirectas no apuntan a una página web física específica. En realidad, éstas contienen información que será ejecutada por un código en el servidor, y será este código el que genere la página que el usuario verá. Normalmente los rastreadores no siguen el rastro de las URLs que llevan estas llamadas a códigos.

A veces, las claves por las que un motor de búsqueda decide si unas páginas van a ser o no indexadas son indicadas mediante símbolos o palabras en las direcciones web. El símbolo más común usado para indicar la presencia en una página de contenido dinámico es el signo de interrogación. Sin embargo es conveniente tener en cuenta que aunque normalmente la interrogación es usada para ejecutar códigos que generan páginas de contenido dinámico, a veces también son usadas como un medio para alertar al servidor de que se está pasando información adicional usando otras variables que se añaden después del signo de interrogación. Estas variables normalmente se usan para trazar una ruta dentro de un sitio web, productos específicos dentro de una carta de compra y otros propósitos que no tienen nada que ver con la Red profunda.

Normalmente, las direcciones web que contienen las palabras “cgi-bin” o “javascript”, ejecutan un código para generar una página. Sin embargo, no se debe pensar que una página pertenece a la Red profunda basándonos en una única evidencia. Ejemplos de direcciones web indirectas son las siguientes: <http://us.imdb.com/Name?Newman,+Paul> (apunta a la lista de Paul Newman en la base de datos de películas en Internet (IMDB)).

En el caso de que una URL parezca ser indirecta y susceptible de ejecutar un código, es relativamente fácil para el usuario determinar si esa dirección web va a ser rastreada o no. Tan sólo hay que situar el cursor en dicha dirección a la izquierda del signo de interrogación, y eliminar este signo de interrogación y todo lo que se encuentre a la derecha de él. A continuación se pulsa la tecla de “Enter” para que el navegador intente buscar si hay alguna página que coincide con el fragmento de URL que queda. Si la página es encontrada y es la esperada por el usuario esta dirección web sería directa. En este caso, el signo de interrogación ha sido usado como medio para pasar información adicional al servidor, no para ejecutar un código. Al tratarse de una página en simple HTML el motor no tendría ningún problema en indexarla. En el caso de que la página que aparece no sea la que el usuario pensaba o el resultado es un mensaje de error, significa que la información tras el signo de interrogación es necesaria para completar el código que genera la página dinámica. Sin esta información, el servidor no sabe qué datos coger de la base de datos para crear la página.

A veces es difícil determinar si una dirección web apunta a contenido que es generado de forma dinámica. Muchos navegadores guardan información sobre la página en determinadas variables que están escondidas para el usuario. Pulsando el botón de “actualizar” se reenvía la información usada para construir la página de nuevo al servidor, y se vuelve a elaborar. Sin embargo, puede ser que la página haya sido guardada en la memoria caché del ordenador del usuario. La mejor forma de analizar si una URL pertenece a la Red profunda es abriendo una nueva ventana del navegador, cortar y pegar dicha dirección web en la nueva ventana, y recargar la página.

Es importante analizar la estructura de las direcciones web. Muchos de los sitios que contienen un gran volumen de información pueden parecer a primera vista parte de la Red profunda, cuando simplemente son sitios web especializados. Hay muchos directorios especializados en la Red que comparten características de lo que podría ser

un sitio web perteneciente a la Red profunda, pero, sin embargo, son perfectamente accesibles por los motores de búsqueda. Normalmente estos sitios están estructurados como directorios hieráticos diseñados para que el usuario pueda navegar por categorías o tópicos de información y normalmente ofrecen tanto sofisticadas herramientas de búsqueda como la posibilidad de navegar por la estructura del directorio. Sin embargo, incluso si estos sitios están integrados por cientos o miles de páginas HTML, la mayor parte de ellas no forman parte de la Red profunda ya que los rastreadores no encontrarían ningún problema en indexarlas. De hecho, estos sitios normalmente poseen una gran estructura interna de enlaces que hace que el trabajo de la araña o rastreador, sea aún más sencillo. En realidad, muchos de los sitios que dicen tener grandes colecciones de contenido perteneciente a la Red profunda, incluye servicios de búsqueda especializados que son perfectamente accesibles por los motores. Llegados a este punto merece la pena reseñar que el estudio de Bright Planet sobre Internet Invisible analizado en el primer capítulo de esta parte de la tesis, podría estar equivocado al afirmar que la "Internet Invisible" o "Red profunda" es de 400 a 550 veces mayor. Muchas de las fuentes citadas en este estudio son directorios especializados que el buscador puede indexar. Bright Planet incluye también datos efímeros como puede ser la información meteorológica u otro tipo de datos en su mayor parte inservibles para los usuarios. Según estiman Chris Sherman y Gary Price en su libro "The invisible Web", la Red profunda es entre 2 y 50 veces más grande que la Red superficial.

Robots de protocolo de exclusión

A la mayoría de los webmasters de los sitios web que cambian constantemente no les interesa que sus páginas sean indexadas. Si la página varía diariamente y el rastreador sólo la visita una vez al mes, el resultado es una permanente desactualización de dicha página. Además, algunos de estos sitios tienen libre acceso a sus contenidos durante un cierto periodo de tiempo, y luego éstos pasan a ser de pago. Un buen ejemplo de este caso son las versiones *online* de muchos periódicos estadounidenses, como el Washington Post y el Wall Street Journal. En España, la modalidad de periódicos digitales de pago está comenzando en la actualidad ya que el "pago por contenidos" es,

ahora mismo, el único medio de obtener beneficio para las empresas con este tipo de productos. La publicidad en Internet ha demostrado ser un mal método de financiación, al menos por ahora.

Para impedir el paso de los rastreadores a un sitio web, los webmasters utilizan los Robots de protocolo de exclusión. Se trata de una serie de reglas que permite al webmaster transmitir al rastreador a qué parte del servidor puede y a cual no puede acceder. El encargado del web crea una lista de archivos o directorios que no deberían ser indexados, y la guarda en un archivo llamado robots.txt. Un ejemplo de un archivo robots.txt perteneciente al diario *online Los Angeles Times*, es el siguiente:

```
User-agent: *  
Disallow: /RealMedia  
Disallow: /archives  
Disallow: /wires/  
Disallow: /HOME/  
Disallow: /cgi-bin/  
Disallow: /class/realestate/dataquick/dqsearch.cgi  
Disallow: /search
```

El campo de “User-agent” especifica que la araña o rastreador debe prestar atención a las instrucciones que siguen. El asterisco es una señal que significa que todos los buscadores deben leer y respetar el contenido de este archivo. Cada comando “Disallow” es seguido por el nombre de un directorio específico que se encuentra en el servidor de *Los Angeles Times*, y supone que los rastreadores tienen prohibido el acceso a ellos. También es posible prevenir que el rastreador indexe una página específica incluyendo una etiqueta de “no index” en la cabecera de la programación del documento, de la siguiente forma:

```
<html>  
<head>  
<title>Keep Out, Search Engines!</title>  
<META name= “robots” content= “noindex, nofollow”>  
</head>
```

Tanto los archivos “robots.txt” como las etiquetas de “noindex” pueden usarse para bloquear a los motores de búsqueda. La única diferencia entre los dos es que la etiqueta de “noindex” es para una página específica y el archivo “robots.txt” puede usarse para prevenir indexar páginas individuales, grupos de archivos o incluso sitios webs completos.

Cuando utilizar los recursos de la Red Profunda

A la hora de decidir cuando un usuario debe usar los recursos disponibles en la Red profunda debe tener en cuenta las siguientes normas¹⁵⁶:

- Para que la búsqueda en la Red profunda sea efectiva, el usuario debe conocer bien el tema sobre el que está realizando la búsqueda. La familiaridad con el tema en cuestión ofrece también la ventaja de conocer mejor los términos o palabras clave que permitirán hallar los mejores resultados, así como métodos para localizar nuevas fuentes.
- Es una buena solución buscar en la Red profunda cuando estás familiarizado con herramientas de búsqueda específicas. Algunas de las fuentes disponibles en esta parte de la Red cubren una gran variedad de temas diferentes, sin embargo suelen estar dotados de potentes herramientas de búsqueda que prácticamente con toda probabilidad le harán conseguir al usuario mejores resultados que si usa las herramientas de búsqueda generalistas. Una buena forma de acotar la búsqueda en estos casos es usar los operadores booleanos u otras herramientas de búsqueda avanzada.
- La búsqueda en la Red profunda es útil cuando se está intentando obtener una respuesta precisa. Lo que no se debe hacer, en este caso, es usar los motores de búsqueda generalistas ya que el resultado que se suele obtener son largas listas de

¹⁵⁶ Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

direcciones que, de algún modo, tienen algo que ver con lo que el usuario ha solicitado.

- Cuando se pretende conseguir una respuesta exhaustiva o rigurosa. Los motores de búsqueda generalistas “nunca” (aunque hablando de la evolución de la tecnología no se debería emplear esta palabra), podrán asegurar al usuario la rigurosidad de las webs cuya dirección proporcionan, debido a diversas razones entre las que se encuentran la falta de selección y de tiempo disponible en cada búsqueda. El resultado es una gran cantidad de ruido entre el que puede que se encuentre la respuesta que el usuario ha pedido o puede que no. Sin embargo, la mayor parte de las webs de la Red profunda sí pueden, ya que suelen sufrir un riguroso proceso de selección.
- Por último, en el caso de que la necesidad del usuario sea encontrar datos actualizados, es más probable que los encuentre en la Red profunda que usando las herramientas de búsqueda generalistas y los directorios.

6.9. ANÁLISIS DE LOS MAYORES SITIOS WEB DISPONIBLES EN LA RED PROFUNDA

A continuación exponemos un análisis del informe de la consultora Bright Planet que analiza 60 de los mayores sitios web de la Red profunda¹⁵⁷. En dicho estudio se establecen tres objetivos:

- Estimar el número total de documentos que contiene cada sitio.
- Recoger como ejemplo un mínimo de diez resultados de cada sitio, guardar dicho documento expresado en HTML y estimar su peso en bytes.
- Indexar y analizar cada página para determinar el tema que trata.

A cada uno de estos sitios le han realizado una serie de pruebas, en orden de importancia descendente:

- 1) Fueron enviados correos electrónicos a los contactos correspondientes de cada sitio solicitando verificación del número total de archivos de cada web y su peso. El 13% de estos sitios contribuyó aportando la documentación requerida.
- 2) Inspeccionar cada una de las páginas del sitio web incluyendo las secciones de ayuda, las FAQs, etc.
- 3) Tamaño de los documentos, no solamente los estimados por el sitio web sino también consultando otras fuentes de referencia que pudieran aportar información sobre el sitio en cuestión, con el objetivo de contrastar los datos.
- 4) Análisis del sistema de búsqueda de la base de datos del sitio en cuestión. Algunos sitios ofrecen estadísticas del número total de búsquedas de las que ha sido objeto su base de datos.
- 5) Una vez analizados todos estos sitios, aquel que no cumpla los requisitos mínimos es retirado de la lista.

¹⁵⁷ *The Deep Web: Surfacing Hidden Value*. Informe de BrightPlanet.com (Accesible en: http://www.completeplanet.com/tutorials/deepweb/introduction_howsearch07.asp).

- 1) En base a estos criterios, BrightPlanet seleccionó 60 sitios de la Red profunda lo suficientemente significativos como para darnos una visión general de esta parte de Internet. Son los siguientes:

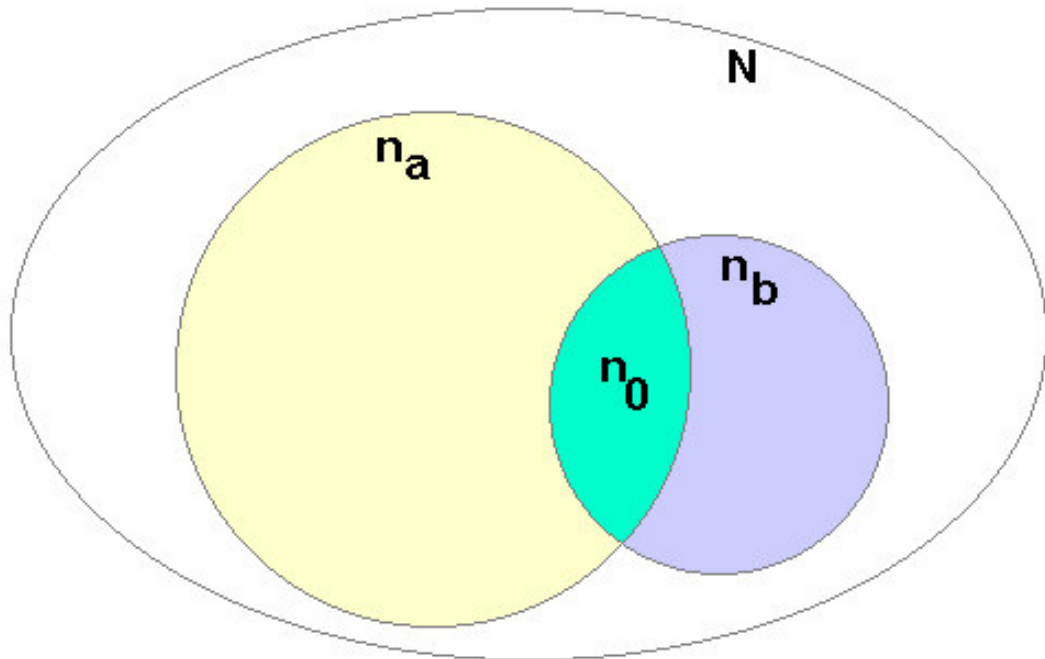
Nombre	Tipo	URL o dirección web	Tamaño (GBs)
National Climatic Data Center (NOAA)	Público	http://www.ncdc.noaa.gov/ol/satellite/satelliteresources.html	366,000
NASA EOSDIS	Público	http://harp.gsfc.nasa.gov/~imswww/pub/imswelcome/plain.html	219,600
National Oceanographic (combined with Geophysical) Data Center (NOAA)	Público/Pago	http://www.nodc.noaa.gov/ , http://www.ngdc.noaa.gov/	32,940
Alexa	Público	http://www.alexa.com/	15,860
Right-to-Know Network (RTK Net)	Público	http://www.rtk.net/	14,640
MP3.com	Público	http://www.mp3.com/	4,300
Terraserver	Público/Pago	http://terraserver.microsoft.com/	4,270
HEASARC (High Energy Astrophysics Science Archive Research Center)	Público	http://heasarc.gsfc.nasa.gov/W3Browse/	2,562
US PTO - Trademarks + Patents	Público	http://www.uspto.gov/tmdb/ , http://www.uspto.gov/patft/	2,440
Informedia (Carnegie Mellon Univ.)	Público	http://www.informedia.cs.cmu.edu/	1,830
Alexandria Digital Library	Público	http://www.alexandria.ucsb.edu/adl.html	1,220
JSTOR Project	Limitado	http://www.jstor.org/	1,220
10K Search Wizard	Público	http://www.tenkwizard.com/	769

UC Berkeley Digital Library Project	Público	http://elib.cs.berkeley.edu/	766
SEC Edgar	Público	http://www.sec.gov/edgarhp.htm	610
US Census	Público	http://factfinder.census.gov	610
NCI CancerNet Database	Público	http://cancernet.nci.nih.gov/	488
Amazon.com	Público	http://www.amazon.com/	461
IBM Patent Center	Público / Privado	http://www.patents.ibm.com/boolquery	345
NASA Image Exchange	Público	http://nix.nasa.gov/	337
InfoUSA.com	Público / Privado	http://www.abii.com/	195
Betterwhois (many similar)	Público	http://betterwhois.com/	152
GPO Access	Público	http://www.access.gpo.gov/	146
Adobe PDF Search	Público	http://searchpdf.adobe.com/	143
Internet Auction List	Público	http://www.internetauctionlist.com/search_products.html	130
Commerce, Inc.	Público	http://search.commerceinc.com/	122
Library of Congress Online Catalog	Público	http://catalog.loc.gov/	116
Sunsite Europe	Público	http://src.doc.ic.ac.uk/	98
Uncover Periodical DB	Público/Pago	http://uncweb.carl.org/	97
Astronomer's Bazaar	Público	http://cdsweb.u-strasbg.fr/Cats.html	94
eBay.com	Público	http://www.ebay.com/	82
REALTOR.com Real Estate Search	Público	http://www.realtor.com/	60
Federal Express	Público	http://www.fedex.com/	53
Integrum	Público	http://www.integrumworld.com/eng_test/index.html	49
NIH PubMed	Público	http://www.ncbi.nlm.nih.gov/PubMed/	41
Visual Woman (NIH)	Público	http://www.nlm.nih.gov/research/visible/visible_human.html	40
AutoTrader.com	Público	http://www.autoconnect.com/index.jtmpl/?LNX=M1DJAROSTEXT	39
UPS	Público	http://www.ups.com/	33
NIH GenBank	Público	http://www.ncbi.nlm.nih.gov/Genbank/index.html	31
AustLi (Australasian)	Público	http://www.austlii.edu.au/austlii/	24

Legal Information Institute)			
Digital Library Program (UVa)	Público	http://www.lva.lib.va.us/	21
Subtotal de sitios públicos y mixtos			673,035
DBT Online	Pago	http://www.dbtonline.com/	30,500
Lexis-Nexis	Pago	http://www.lexis-nexis.com/lnccl/	12,200
Dialog	Pago	http://www.dialog.com/	10,980
Genealogy - ancestry.com	Pago	http://www.ancestry.com/	6,500
ProQuest Direct (incl. Digital Vault)	Pago	http://www.umi.com	3,172
Dun & Bradstreet	Pago	http://www.dnb.com	3,113
Westlaw	Pago	http://www.westlaw.com/	2,684
Dow Jones News Retrieval	Pago	http://dowjones.wsj.com/p/main.html	2,684
infoUSA	Pago / Público	http://www.infousa.com/	1,584
Elsevier Press	Pago	http://www.elsevier.com	570
EBSCO	Pago	http://www.ebsco.com	481
Springer-Verlag	Pago	http://link.springer.de/	221
OVID Technologies	Pago	http://www.ovid.com	191
Investext	Pago	http://www.investext.com/	157
Blackwell Science	Pago	http://www.blackwell-science.com	146
GenServ	Pago	http://gs01.genserv.com/gs/bcc.htm	106
Academic Press IDEAL	Pago	http://www.idealibrary.com	104
Tradecompass	Pago	http://www.tradecompass.com/	61
INSPEC	Pago	http://www.iee.org.uk/publish/inspec/online/online.html	16
Subtotal de sitios de pago			75.469
TOTAL			748,504

Estimación del número total de sitios

El estudio de Bright Planet utiliza estas muestras con distintos objetivos. Uno de ellos es intentar averiguar o aproximarse al número total de sitios web disponibles en la Red profunda. La técnica básica que aplica pasa por un “análisis de solapamiento”. Esta técnica la podemos ver ilustrada en el siguiente diagrama:



El análisis de solapamiento supone la comparación del número de resultados obtenidos de dos recursos de búsqueda diferentes (por ejemplo: Altavista y Google): n_a y n_b y observar el grado de resultados comunes entre ambos (n_0). Asumiendo estas dos listas de resultados al azar, n_a y n_b , puede ser estimado el número total de sitios, es decir, N . Este supuesto estadístico parte del hecho de que la distribución de datos es de tipo normal, es decir, que cada uno de los conjuntos tiene la misma probabilidad de acceder a la información. Sin embargo, esto no quiere decir que esta probabilidad ocurra, sino que estadísticamente están en el mismo nivel. Este hecho tan cuestionable afecta al resto del desarrollo de la investigación, pero es una buena suposición de arranque debido a que actualmente no existen estudios equivalentes previos.

Realmente esta distribución no se da en la práctica porque los algoritmos de búsqueda difieren con cada buscador y, por tanto, los resultados de indexación son distintos. Por este motivo, la distribución no debe de ser normal sino que debe tener algún tipo de sesgo que, dada la cantidad de datos, puede llegar a ser muy importante.

Tratar de analizar el tamaño total de la Red profunda, es complicado, ya que al no haber estudios previos hay que basarse en suposiciones. El objetivo último del estudio de BrightPlanet es vender un producto, por lo que todos estos datos que aporta deben de ser cuestionados. Sin embargo este documento base aporta mucho valor debido a que orienta hacia una solución práctica. Si cualquier otra persona o empresa se planteara este mismo problema, sin contar con recursos ilimitados, tendría que hacer suposiciones similares para simplificar el problema.

Escasez de direcciones IP

A la hora de calcular la profundidad de la Red, debemos de tener en cuenta que la cantidad de direcciones IP superficiales es limitada. De hecho se acabarán pronto. Por lo que, según esta teoría, el número de sitios web también es limitado. Al menos por ahora, y hasta que se amplíe este número.

Originalmente, Internet fue desarrollada como un medio de comunicación para la comunidad científica. Puede decirse, que, hasta 1994, no se había pensado en ella como un marco económico y social contextual, paralelo al mundo físico. Su popularización derivó en un crecimiento exponencial que ha terminado derivando en un problema físico de capacidad. De esta manera, cuando se creó, hace unos veinte años el Protocolo IP, la idea de que Internet fuera a necesitar la cantidad de direcciones que tiene actualmente era poco previsible.

El problema consiste, en este caso, en la creciente escasez de direcciones que se ajustan al protocolo de Internet, las direcciones IP. Cada ordenador tiene una dirección única IP, dentro del actual protocolo de la Red que sigue el Ipv4, esto es, la versión cuarta del protocolo para redes de computadoras, que tiene un funcionamiento bastante análogo al sistema telefónico. Cada dirección IP es única, representada en cuatro conjuntos de números separados por puntos. Cada conjunto puede ir, en la actualidad, del 0 al 255, lo que termina arrojando cerca de 4.000 millones de combinaciones posibles.

La agencia internacional IANA (Internet Assignet Number Authority) es la encargada de asignar a cada servidor de Internet unas IPs únicas para que sus usuarios se puedan identificar en Internet. Su registro en el Continente Americano esta delegado a ARIN (<http://www.arin.net>) , y en Europa y Oriente Próximo a RIPE

(<http://www.ripe.net>), generalmente su registro es llevado a cabo únicamente por los proveedores de Internet más grandes.

Además de la asignación de estos Nodos IP, también existe un depósito central de nodos IP a nombres (De 212.19.215.67 a www.greenpeace.org), aparentemente parece que fuera Network Solutions (<http://www.networksolutions.com>) quien mantiene este depósito, pero este depósito central de nombres a nodos IP se mantiene actualmente en 13 DNS¹⁵⁸ servers raíz ("root servers") en el Mundo. (5 en Estados Unidos, 2 en Sudamerica, 3 en Europa y 3 en Asia). Estos servidores raíz contienen información relativa a los dominios .com, .edu, .org. También existen otros servidores raíz de cada país .mx , .ar , .cl.. y las raíces más recientes como .tv,.aero..etc. La cuota actual que se paga al registrar un dominio es para la manutención de este servicio.

Las IPs pueden ser estáticas o dinámicas. Las estáticas son las que no cambian y suelen otorgarse bajo contrato. Son direcciones fijas de sitios específicos de Internet. Normalmente suelen ser contratadas por empresas. A esta dirección IP se le suele asignar un nombre. Ej: 10.90.40.43 es una dirección IP correspondiente al dominio de 5Dias.com (sitio web del diario de información económica *CincoDias*).

Las IPs dinámicas son las que tienen la mayoría de los usuarios. Cuando un usuario se conecta a Internet el servidor le asigna una dirección IP para que se pueda identificar en la Red. Esta dirección IP es aleatoria y no debe estar siendo usada por otro usuario porque sino podrían entrar en conflicto.

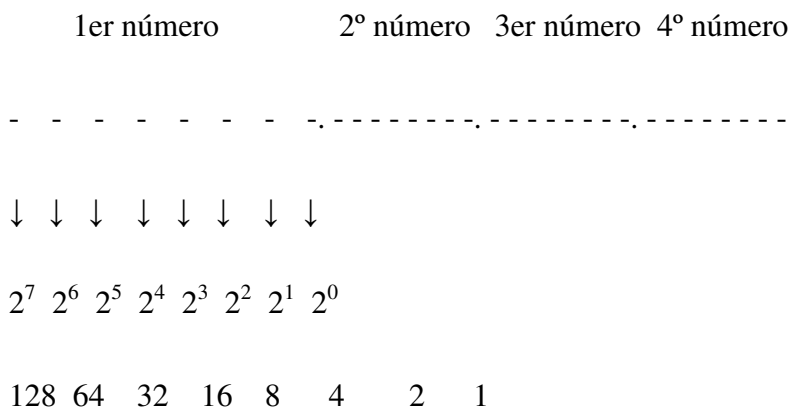
Para entender la limitación actual de las direcciones IP pasamos a explicar en qué consiste dicha dirección.

¹⁵⁸ DNS ("Domain Name System") es otra herramienta primordial para el funcionamiento de Internet. DNS es implementado através del Software llamado BIND ("Berkley Internet Name Domain") y es la base para la resolución de nombres en Internet, de manera que cada vez que alguien busca www.elpais.es, o se envía un mail a redacción@cincodias.com esta resolución se hace en un DNS server.

Una dirección IP consta de 4 números enteros binarios de 8 bites cada uno, separados por puntos, representados en decimal. Al ser cada número de 8 bites, el máximo valor que puede tomar cada número es de 255.

Explicaremos sólo el 1er número, teniendo en cuenta que ocurre lo mismo para los otros tres.

.



Cada raya es un bite, que puede ser un 1 o un 0 (ya que es código binario)

Ej: 10010010.-----.-----.-----

↓ ↓ ↓ ↓

El 1er número de esta IP sería: $1 \times 2^7 + 1 \times 2^4 + 1 \times 2^1 = 128 + 16 + 2 = 146$

Esa IP sería 146. *. *. *

Ejemplo de IP completa: 10001001.10010100.00000001.10111000 = 137. 148. 1. 216

Por lo tanto las direcciones IP son limitadas porque el rango para cada número de la IP comprende desde el 1 al 255 y no se puede superar; por lo que deducimos que el máximo número de IPs sería: $255 \times 255 \times 255 \times 255 = 4.228.250.625$ IPs distintas. Y esta cifra no tardará mucho en alcanzarse.

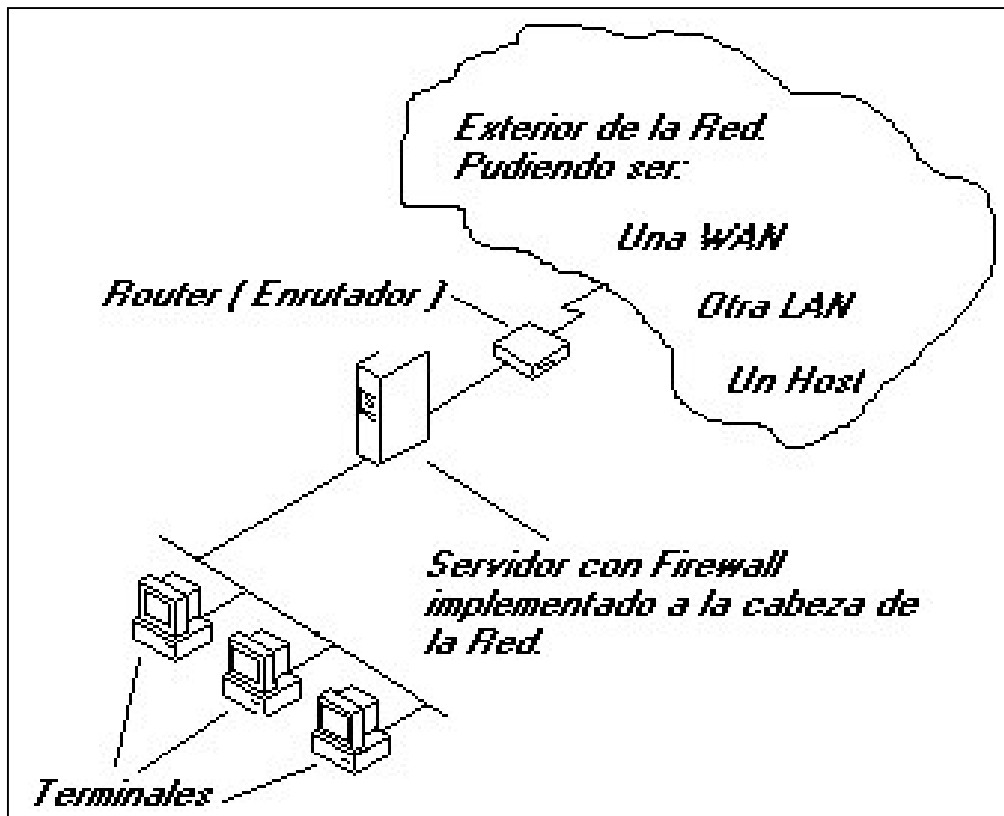
En definitiva, teniendo en cuenta este hecho, sí podríamos calcular el tamaño de la Red superficial. Sin embargo, intentar averiguar el tamaño de la Red profunda es más complicado y más que en análisis técnicos, se basa en suposiciones. En este sentido, es importante remarcar el hecho de que de cada una de estas direcciones IP es posible “colgar” un verdadero universo de información creando redes locales (LAN) y utilizando los llamados “Routers”, que pueden estar protegidos contra usuarios no autorizados mediante un “Firewall” (Cortafuegos). Los Routers (Encaminadores) se sitúan entre dos redes y a la vez que encaminan los paquetes entre una red y otra buscando los recorridos más “rápidos” o más precisos, realizan o pueden realizar un filtrado de paquetes (Por eso también le podríamos llamar Firewall, porque están haciendo ese trabajo), comprobando las IP y los puertos (Cada paquete lleva una franja de información en la cuál se incluyen las IP de salida y destino y el puerto). El Router no es sólo un elemento físico para conectar una LAN (red local) a una internet (WAN), sino que es una muy buena opción a la hora de ampliar una LAN.

Estas redes locales utilizan direcciones IP que no interfieren con las IP globales ya que deben estar (por asignación) entre los siguientes tres rangos:

10.0.0.0 - 10.255.255.255

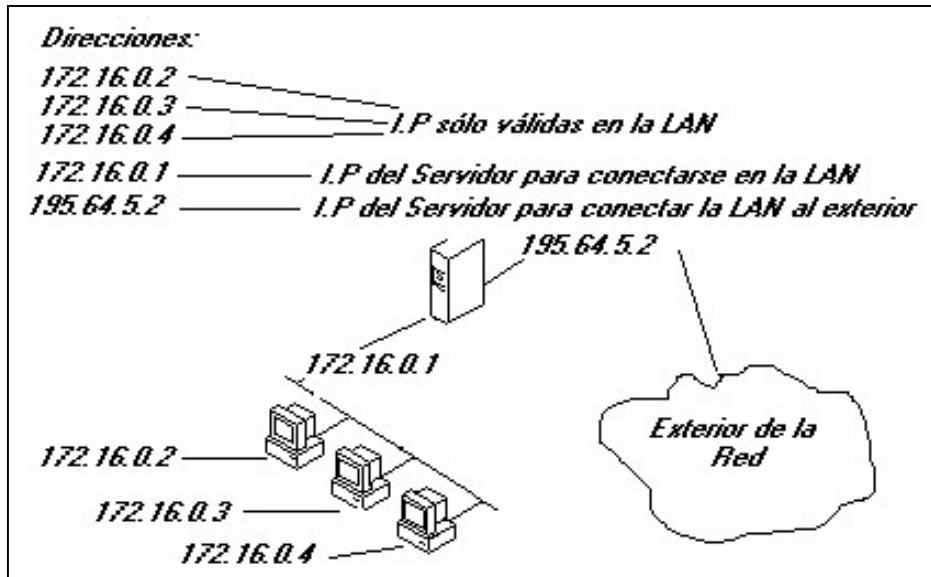
172.16.0.0 - 172.31.255.255

192.168.0.0 - 192.168.255.255



En este gráfico podemos observar 3 terminales en una red con un servidor a la cabeza con Firewall y un Router. Cada terminal de esta LAN, incluido el Servidor tiene una dirección IP personal que la va a identificar en la red local, pero el Firewall tendrá otra

que será la que haga posible una identificación con el exterior. Al instalar el Firewall se dota al ordenador servidor con las dos direcciones IP: una para que se puedan conectar los terminales de la LAN a él y otra real de identificación con el exterior.



El Firewall define la organización de los paquetes de información, y la red sólo podrá acceder a los parámetros que el Firewall tenga permitido o posibilite mediante su configuración. Por ejemplo, si un terminal de la red intenta enviar un paquete a una dirección IP no autorizada, el Firewall rechazará éste envío impidiendo realizar ésta transmisión. Con el Firewall es posible definir el tamaños de los paquetes, IP con las que no interesa comunicación, deshabilitación de envíos o recogida de paquetes por determinados puertos, imposibilitar el uso del comando Finger, etc.

La escasez creciente de direcciones IP proviene de varias circunstancias. Entre ellas está el creciente número de internautas, y de máquinas conectadas a la Red, lo que lleva inclusive, en algunos casos a una combinación de una pluralidad de direcciones IP de un solo usuario. La llegada inminente del sistema de telefonía UMTS (la tercera

generación de telefonía móvil) que requerirá que cada aparato tenga su propia dirección IP, agudiza esta situación.

Otras soluciones

Una solución a este problema de escasez de direcciones IP ha sido la creación de los denominados “NAT”, aparatos que hacen de traductores de direcciones en la Red. Esta tecnología permite a un conjunto de empresas que conecten ordenadores a Internet utilizando una única dirección IP, sin embargo, cuenta con ciertas limitaciones: las direcciones privadas conectadas a la dirección única que aparece en Internet, no tienen acceso a servicios como telefonía o vídeo conferencia.

Asimismo, en la actualidad se está desarrollando el Protocolo Ipv6, el cual, teóricamente podría solucionar el problema, elevando el número de direcciones a números prácticamente ilimitados, donde, de acuerdo a sus creadores, la posibilidad es de tener mil direcciones de Internet por metro cuadrado de superficie terrestre. El nuevo protocolo consiste en ocho conjuntos de cuatro números o letras, separados por puntos. Dentro de las ventajas que se han anunciado destaca, un aumento en la seguridad; la información viaja encriptada desde el origen y da mayores certezas acerca del emisor.

Sin embargo, el cambio de protocolo no es una tarea sencilla ya que la transición debe ser realizada, además, asegurando el funcionamiento ininterrumpido de Internet.

La profundidad de la Red

Siguiendo con el estudio de Bright Planet, para calcular la profundidad de la Red, se parte, de dos servicios de indexación A y B, cada uno con 50 items de entre una población de 100 items posibles. En promedio se esperan 25 items compartidos entre los dos servicios; son 25 porque cada uno tiene el 50% de la población. Lo que Bright Planet propone es que el total de la población se puede calcular como :

$$A/(Intersección/B) = 50/(25/50) = 100.$$

La única forma de relacionar la intersección de estos grupos al total de la población es suponer que los servicios de indexación tienen la misma probabilidad de acceso a la misma información que se va a indexar. Esto implica que:

- 1) Tanto A como B podrían catalogar el mismo documento sin preferencia para ninguno.
- 2) Entre dos servicios de indexación la diferencia sería el tamaño de sus bases de datos (o la potencia de sus equipos) y la cantidad de tiempo que lleven indexando documentos y haciendo crecer sus bases de datos.

En realidad, ninguno de estos dos presupuestos es cierto. Tan sólo es necesario revisar el tamaño de las bases de datos de servicios de indexación y comparar entre ellos. En realidad, por diferencias de desarrollado de tecnología, entre otras cosas, algunos motores de búsqueda tienen una mayor probabilidad de indexar información que otros. En este caso, en el análisis de paridad hay que incluir la probabilidad de indexación y el modelo se vuelve muy complejo. Además, los números propuestos por Bright Planet son demasiado pequeños, aunque ellos mismos sugieren que el resultado debe ser el límite mínimo. Sin embargo, como hemos señalado con anterioridad, este informe es importante dado que no existen estudios previos relevantes relativos a este tema

Conclusiones al estudio de Bright Planet

Este estudio es el primero que trata de describir de forma científica la Red profunda. La conclusión extraída es que esta parte de Internet es del orden de 500 veces mayor que la parte accesible mediante los motores de búsqueda más comunes entre los usuarios, y sobre todo, su contenido es de una mayor calidad. En 2000, año en el que fue elaborado el informe, el número total de sitios web de la Red profunda excedía los 100.000 y su crecimiento era exponencial, mayor que el que presenta la “Red superficial”. La

información que albergan estos sitios es de una gran utilidad tanto para investigadores como para aquellos que buscan datos especializados, precisos y rigurosos. Además, el 95% de esta información “escondida” está disponible para el público general sin ninguna restricción. Tan sólo hay que saber dónde buscar.

El estudio revela que la Red profunda posee unas claras diferencias frente a la superficial. Cada uno de estos sitios web posee una gran diversidad en el número de documentos que contiene. Algunos de ellos llegan a alcanzar los 5 millones de documentos, aunque la media está en 4.950. El tamaño de ciertas bases de datos aquí disponibles es de 74,4 megabytes, mientras que la media se encuentra en 169 kilobytes.

De media, los sitios de la Red profunda reciben un 50% más de tráfico que los sitios superficiales (123.000 al mes vs. 85.000). Además, en porcentaje, los sitios de la Red profunda contienen más enlaces de hipertexto que los de la superficial (66 vs. 83 enlaces). Esto podría sugerir que estos sitios de la Red profunda son muy populares, según la teoría de que aquel sitio que posee un mayor número de enlaces es más popular entre los usuarios (ya que es más fácilmente localizable por los buscadores). Sin embargo, esta parte de Internet no es muy conocida por la mayor parte de los usuarios.

Uno de los resultados de este estudio más valiosos es la apreciación de que el 97,4% de los sitios disponibles en la parte de Internet que estamos analizando está disponible para cualquier usuario sin ningún tipo de restricción; un 1,6% es mixto, es decir, algunos documentos son libremente accesibles para cualquier persona, mientras que otros requieren suscripción o el pago de una tasa; y solamente un 1,1% del total están completamente sujetos a suscripción o a un pago previo. Este hecho es especialmente significativo debido a que las bases de datos más conocidas entre los usuarios, es decir, Dialog, Lexis-Nexis o la edición digital del Wall Street Journal pertenecen a este último grupo; por lo que, en un primer momento, éstos podrían suponer que la mayor parte de este tipo de recursos también son de pago.

Por otro lado, el estudio establece que los sesenta mayores sitios web pertenecientes a la Red profunda estudiados contienen aproximadamente 750 terabytes de datos, lo cual supone un tamaño 40 veces mayor que la disponible en la Red

superficial. También estima que el número total de documentos disponibles en estos sitios es superior a 85.000 millones. Dos tercios de ellos son de libre acceso para cualquier usuario de Internet, lo cual representa el 90% del contenido total disponible en este grupo de 60.

6.10. EL FUTURO DE LA RED PROFUNDA

En el futuro, la Red profunda será, al mismo tiempo, más grande y más pequeña que la actual, según afirman Gary Price y Chris Sherman en su libro *The invisible Web*¹⁵⁹. Esta paradoja se explica teniendo en cuenta el desarrollo de las nuevas tecnologías y métodos para hacer más útil y accesible la información disponible en la Red profunda. La evolución de estas técnicas será directamente proporcional al coste de la potencia de los ordenadores (cuanto más bajo mejor) y al aumento del ancho de banda.

Sin embargo, en el futuro, se mantendrá el problema de incompatibilidad entre la velocidad y la calidad, y facilidad de uso y potencia. Las nuevas tecnologías lograrán abrir al usuario corriente los recursos disponibles en la Red profunda, de forma sencilla, pero esto no es suficiente. A pesar de que la nueva generación de herramientas que utilizarán los buscadores podrán acceder a gran parte de la información disponible en esta “parte oscura” de Internet a través de motores de búsqueda generales, el crecimiento exponencial de la Red profunda volverá a superar a estas herramientas.

Motores de búsqueda inteligentes

En el futuro, los motores de búsqueda habrán evolucionado en gran medida en cuanto a la elaboración de sus índices, que se adaptarán más a la realidad de aquello que está disponible en la Red. En parte, las nuevas tecnologías empleadas estarán basadas en el desarrollo de “arañas” o “robots de búsqueda” inteligentes en cuanto a su modo de actuar cuando encuentra una página web. Es decir, estas herramientas localizan los enlaces en una web, los colocan en fila y después los rastrean en el orden en que han

¹⁵⁹ Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).

sido añadidos a la fila. En teoría, teniendo en cuenta que la interconexión entre las páginas es alta, las arañas tendrían que poder encontrar e indexar la mayoría de las páginas disponibles en la Red. Sin embargo, en la práctica, debido a que la mayor parte de los motores de búsqueda limitan la búsqueda en profundidad de un sitio web, muchos de estos sitios son sólo parcialmente indexados.

A medida que avance la potencia de los ordenadores y su capacidad de almacenaje, además de abaratarse el precio del ancho de banda, los motores de búsqueda irán eliminando sus restricciones en el rastreo de los sitios. Sin embargo, teniendo en cuenta que la Red continúa creciendo de forma exponencial, la araña continuará estando limitada a visitar cada sitio cada cierto periodo de tiempo, lo que significa que las páginas indexadas no siempre serán el fiel reflejo del contenido real del sitio.

Las “arañas inteligentes” intentarán aprender más de cada sitio y automáticamente adaptarán su frecuencia de visita a su popularidad o importancia para asegurar lo más posible la actualidad de su contenido. Las páginas que cambien con frecuencia serán rastreadas cada tiempo, mientras que las páginas estáticas que cambian menos, serán ignoradas por estos robots durante mayores periodos. Estos rastreadores inteligentes, también estarán programados para realizar juicios sobre la relevancia de un determinado documento y la calidad de los enlaces que contiene. Aquellos de poca calidad o de contenido no apropiado no serán indexados. De esta forma el resultado de la búsqueda del usuario será más fiable y de una mayor calidad.

También habrá avances en cuanto a los web que eligen las arañas para ser rastreadas. Cohen, McCallum y Quass, en su artículo “Learning to Understand the Web”¹⁶⁰ prevén el surgimiento de un gran número de “centrales de datos” especializadas en los motores de búsqueda. “Es probable que muchos motores de búsqueda se centren en distintas especialidades temáticas. Estas herramientas especializadas también variarán en su capacidad de rastreo en profundidad. En definitiva, la mayor parte de las búsquedas centradas en tópicos comunes serán respondidas por una de las pocas docenas de bases de datos generales; mientras que las

¹⁶⁰ Cohen, William W., Andrew McCallum, and Dallon Quass (2000): Learning to Understand the Web. IEEE Data Engineering Bulletin 23.3.

búsquedas temáticas más específicas serán respondidas por una de las pocas miles de bases de datos especializadas”.

Metadatos

Los metadatos que aparecen en la cabecera de la programación de las páginas web, son elementos esenciales en el trabajo de recopilación de los motores de búsqueda.

Principalmente, estos metadatos contienen información sobre el documento, incluyendo el nombre del autor, un sumario y palabras clave descriptivas. La forma tradicional de los metadatos es usada desde 1996, pero el hecho de haber sido objeto de abuso por parte de los llamados “spammers” (aquellos que tratan de sabotear el funcionamiento correcto de los motores de búsqueda, añadiendo indicadores “trampa” en los metatags) ha llevado a muchos motores a ignorar estas cabeceras o a usarla, en ocasiones, para el cálculo de la relevancia del documento.

En la actualidad existen gran cantidad de proposiciones para la presentación de los metadatos. Entre ellas, la opción más cercana a convertirse en universal, es decir, en ser aceptada por todos, es el RDF (Resource Description Framework), que usa la sintaxis del lenguaje de programación XML (Extensible Markup Language). Todas estas proposiciones pretenden lograr el objetivo de crear datos comprensibles por los motores de búsqueda. Entre otras cosas, éstas tienen la habilidad para introducir vocabulario controlado en la ecuación de búsqueda que permitirá unir diferentes términos o conceptos. También proveerá de la estructura para la descripción, clasificación y manejo de los documentos web.

RDF permitirá mantener un control mayor sobre la calidad e integridad en el proceso de autorización de los metadatos. Los servicios especializados de búsqueda podrán imponer reglas a los autores del sitio web a los que le será requerido un mayor nivel de consistencia y validación en sus documentos del que existe en la actualidad. Para los servicios de búsqueda generales, esta nueva forma de presentar los metadatos será más problemática, ya que un lenguaje tan rico como el XML ofrece infinidad de nuevas posibilidades de sabotaje para los “spammers”.

Algunos motores de búsqueda, como Northern Light, añaden sus propios metadatos a las páginas a medida que las van indexando. Sin embargo, este proceso es caro y no es probable que sea adoptado por el resto de los motores.

El verdadero potencial de los metadatos surgirá cuando los autores de webs acepten de forma mayoritaria una certificación de autoridad que garantice la confianza y buena voluntad del sitio web en cuestión.

Algo más que texto

Una gran proporción de la Red profunda consiste en objetos no textuales, entre ellos: imágenes, sonidos, vídeo en streaming, y otros archivos en formatos que presentan problemas para unas “arañas” diseñadas para capturar texto. La investigación para recuperar información multimedia está centrada en el diseño de nuevas herramientas que tienen la capacidad para extraer las características particulares de cada objeto multimedia y permite búsquedas basadas en la comparación de estos atributos. Por ejemplo: localiza imágenes de rayos-X que muestran características similares a una fractura de tibia; o localiza las grabaciones de audio del presidente de la Reserva Federal de EEUU, Alan Greenspan, durante la fuerte caída de los mercados en 1987.

En cada búsqueda, el motor hará coincidir cada una de las características específicas que se piden. Estas características pueden haber sido asociadas con imágenes individuales utilizando los metadatos. En caso contrario, pueden ser reconocidas si se usa un programa de búsqueda específicamente entrenado para este propósito.

A medida que avance la tecnología que permite la búsqueda de objetos multimedia, es probable que las búsquedas de texto pasen a formar sólo una parte de las amplias posibilidades que ofrecerán los motores.

Bases de datos

Las bases de datos componen la mayor parte de la Red profunda. Técnicamente no hay razón para que la generación actual de “arañas” o robots de búsqueda no puedan acceder a esta información. Los motores de búsqueda no indexan contenido dinámico porque el

lenguaje de encriptación usado para montar las páginas web de la base de datos, puede también ser usado para formar “trampas” a las arañas, haciéndolas entrar en un círculo sin fin o colapsándolas con miles o millones de páginas inútiles.

Esta situación podría remediarse si los motores de búsqueda pudieran establecer códigos de ética para estos sitios dinámicos que permita indexar estas páginas sin peligro. Una de las maneras de hacerlo, sería hacer que estos sitios web de contenido dinámico pagaran a los motores por rastrear e indexar sus páginas. Inktomi¹⁶¹ está actualmente realizando una investigación en este aspecto. Cobra por este tipo de capturas como parte de una prueba para lograr recoger el contenido dinámico sin peligro, asumiendo que nadie tiene tan poca inteligencia como para pagar por incluir una “URL-trampa” que en cualquier caso será borrada del índice en cuanto sea descubierta.

Otra de las soluciones sería establecer acuerdos individuales entre los sitios web y los motores de búsqueda, especificando qué páginas deberían ser incluidas en la base de datos del motor. Asimismo, los sitios web con dominios completamente fiables, como pueden ser los “.gov” o “.edu” (correspondientes al Gobierno de EE UU y educación), pueden ser abiertos a los motores sin peligro, incluso si contienen páginas dinámicas.

El uso de las nuevas tecnologías también pueden ayudar a solventar este problema. La información puede ser extraída de las bases de datos usando diversas técnicas mediante las cuales un programa prueba a entrar en una base de datos, extraer resultados y examinarlos para determinar qué tipo de acción tomar. Estos agentes que “entienden” como interactuar con los formularios de búsqueda de las bases de datos, pueden también ser usados para introducir los parámetros requeridos en dichos formularios y generar páginas de resultados que pueden ser rastreadas e incluidas en la base de datos del motor.

Actualmente sólo hay unas pocas técnicas que permiten superar las barreras de estas bases de datos. Sin embargo, la creciente evolución de las tecnologías de búsqueda pueden motivar que en un periodo de tiempo más o menos corto, los motores de búsqueda generales puedan comenzar a rastrear la información disponible en las bases de datos. A pesar de esto los mejores resultados siempre se obtendrán utilizando

¹⁶¹ Inktomi: <http://www.inktomi.com>

la metodología de búsqueda de las propias bases de datos. Los motores de búsqueda generales suelen tener la política de que “cualquier respuesta es mejor que ninguna” , en aras de llegar a un mayor número de usuarios que puede que puede que no utilizaran dicho motor, sino les proporcionara siempre algún tipo de resultado. Por el contrario, las bases de datos intentan ofrecer información más específica y limitada dirigida a un menor número de usuarios, teniendo en cuenta que “es preferible que no haya ninguna respuesta antes de que haya muchas incompletas o imprecisas. Estas claras diferencias prueban el hecho de que si el usuario, o en nuestro caso el periodista, está buscando una información específica, una base de datos especializada en un área concreta siempre dará mejores resultados que un motor de búsqueda general.

Lenguajes de búsqueda de hipertexto

Si la Red está considerada como una enorme y desorganizada base de datos, pueden ser creados lenguajes de búsqueda que permitan sofisticados procesos de búsqueda de la información disponible en la World Wide Web. La primera generación de estos lenguajes incluyen W3QL y WebSQL. La siguiente generación, más potente, utiliza lenguajes como STRUQL, FLORID y WebOQL. Estos lenguajes de búsqueda de hipertexto pretenden unificar la Red, permitiendo avanzadas búsquedas similares a las disponibles en las bases de datos actuales. Sin embargo, en lugar de estar limitadas a una o dos bases de datos a un tiempo, estos nuevos lenguajes permiten tratar a toda la Red como una única base de datos.

La búsqueda en tiempo real

Aunque no se puede decir que nada es “imposible” dada la creciente evolución de la tecnología, el hecho de que un usuario introduzca un término o concepto a buscar en un motor y el rastreo se realice en tiempo real en la totalidad de la Red, es muy poco

probable. Sin embargo, en la actualidad hay proyectos como el IBM Fetuccino¹⁶² que no permite la búsqueda en tiempo real, pero comprueba y verifica cada uno de los resultados obtenidos en tiempo real. Asimismo, también aumenta los resultados de la primera búsqueda rastreando en otras direcciones donde pudiera hallar información relevante.

Por otro lado, con el uso de herramientas adecuadas, un usuario podría realizar un rastreo en tiempo real. Si dicho usuario contara con un buscador personalizado, éste podría irlo alimentando con las direcciones web que considerara de su interés. El rastreador personal comenzaría con estas páginas y seguiría sus enlaces buscando nuevos documentos relevantes, teniendo como referencia los sitios web iniciales.

En definitiva, a pesar de la rápida evolución de las nuevas tecnologías de búsqueda, la Red profunda será constituyendo una amplia parte de Internet. El grado de crecimiento de la información es tan grande que los motores de búsqueda generales probablemente nunca puedan abarcarla. Incluso si las bases de datos de la Red profunda consiguen ser rastreadas por los motores de búsqueda generales, las herramientas propias de cada una de las bases de datos específicas siempre serán más efectivas y más rápidas.

¹⁶² IBM Fetuccino Project: <http://www.ibm.com/java/fetuccino/fetuccino-abstract.html>

CUARTA PARTE: ANEXOS, CONCLUSIONES Y BIBLIOGRAFÍA

ANEXO 1

RELACIÓN DE RECURSOS DE LA RED PROFUNDA

DIRECTORIO DE BASES DE DATOS DE LA RED PROFUNDA

Tal y como hemos señalado en la tercera parte de esta tesis, la Red profunda contiene unas 500 veces más información que la disponible en la Red superficial o parte de Internet accesible por los buscadores generalistas. Ante los problemas de acceso a estos contenidos diversas empresas han elaborado una serie de directorios. Los más destacables son los siguientes:

CompletePlanet: Servicio creado por Bright Planet para ayudar a localizar recursos y bases de datos. Incluye un directorio que contiene algo más de 20.000 bases de datos accesibles en la Red.

Lexibot: es una herramienta de software creada por Bright Planet para la realización de búsquedas en las bases de datos. Realiza un lento y complejo proceso de búsqueda, por lo que es una aplicación que corre del lado del “cliente”.

Alphasearch: facilita acceso a más de 800 “pasarelas” especializadas.

Direct Search: constituye una enorme colección de enlaces a las bases de datos y sitios de búsqueda, clasificados en forma temática.

Infomine: desarrollado por la biblioteca de la Universidad de California, se especializa en información de tipo académico y científico, con aproximadamente 20.000 recursos de este tipo.

- **Intelliseek:** es una empresa especializada en “aplicaciones inteligentes”. Distribuye Bulls Eye, un software que tiene cierta similitud con Lexibot.
- **Invisibleweb.com:** es un sitio creado por Intelliseek, constando de un directorio con algo más de 100.000 enlaces a bases de datos, archivos y motores de búsqueda.
- **Lycos-Searchable Databases:** índice de bases de datos con aproximadamente 7.000 recursos de búsqueda especializados.

- **Open Directory Project:** es un gran directorio temático, con algo más de 2,5 millones de páginas recopiladas y organizadas por editores voluntarios.
- **RefDesk:** colección de enlaces a sitios.
- **Search Engine Watch:** es uno de los sitios pioneros en la búsqueda de referencias dentro de la web profunda.
- **Webdata:** es una guía con bases de datos, con evaluaciones sobre la calidad de las mismas.

RECOPIACIÓN DE RECURSOS DISPONIBLES EN LA RED PROFUNDA

A continuación exponemos una recopilación de recursos no accesibles mediante buscadores convencionales, ordenados temáticamente, que pueden ser de una gran utilidad para el periodista, no sólo por el gran volumen de información que contienen sino también por la calidad de dicha información¹⁶³.

ARTE Y ARQUITECTURA

➤ ARQUITECTURA

- **Base de datos de ciudades y edificios:**

<http://content.lib.washington.edu/cities/>

Esta base de datos contiene una colección de imágenes digitalizadas de edificios y ciudades de todo el mundo. Disponible para estudiantes, académicos y profesionales de la información.

- **UN-HABITAT- CitiBase**

<http://www.unhabitat.org/guo/citibase/citibase.asp>

Base de estadísticas de ciudades del mundo. Contiene datos geográficos, meteorológicos, demográficos, sobre el nivel de vida de sus ciudadanos, medio ambiente, infraestructuras, etc. Se puede buscar por país, ciudad y clasificación estadística. En inglés.

- **Base de datos de edificios altos:**

¹⁶³ Fuente: Propia y procedente del libro de Chris Sherman y Gary Price: *The Invisible Web: Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today, Inc. 2001.

<http://hrbd.ce.lehigh.edu/>

Esta Base de datos contiene datos sobre miles de edificios altos: los últimos construidos, estadísticas, imágenes visuales, vídeos y listados de empresas profesionales por categorías de especialidad. Acceso Limitado para los no miembros de la organización.

- **S*P*I*R*O (Diapositivas de arquitectura):**

<http://arch.ced.berkeley.edu/resources/archslides.htm>

Acceso al catálogo de más de 200.000 diapositivas de la Biblioteca de Diapositiva de Arquitectura de la Universidad de California en Berkeley..

➤ **ARTISTAS**

- **Artcyclopedia:**

<http://www.artcyclopedia.com/index.html>

Enciclopedia útil para todo tipo de consultas sobre arte en Internet. La base de datos contiene enlaces a más de 80.000 trabajos de más de 7.000 artistas.

- **Listado de artistas ULAN:**

<http://www.getty.edu/research/institute/vocabulary.html>

Este listado contiene aproximadamente 220.000 nombres e información biográfica de artistas y arquitectos.

- **Quien es quien en el arte americano:**

<http://www.artstar.com/>

Base de datos que contiene información sobre 11.724 artistas de Estados Unidos, Canada y México.

➤ **GALERÍAS EN LA RED**

- **Biblioteca del Congreso de EE UU:**

<http://loc.harpweek.com/default.asp>

Permite el acceso a una de las más importantes colecciones de documentos sobre la política americana. La colección de la Biblioteca del Congreso de EE UU ha sido catalogada por Bernard F. Reilly, Jr.

- **Axis:**

<http://www.axisartists.org.uk/axishome/default.htm>

Axis es la mayor base de datos sobre artistas internacionales británicos actualmente disponible en Internet.

- **Collage (Cooperación de bibliotecas de Londres y galerías de arte en Internet)**

<http://collage.nhil.com/collagedev/index.html>

Base de datos visual que contiene más de 20.000 trabajos de la biblioteca Guildhall y de la galería de arte del mismo nombre en Londres.

- **ExCALENDAR:**

<http://www.excalendar.net/>

Calendario oficial de los mejores museos de arte del mundo.

- **Base de datos del Museo Nacional de Kyoto:**

<http://www.kyohaku.go.jp/>

Acceso en línea al catálogo de más de 5.000 trabajos del Museo Nacional de Kyoto.

- **Museo metropolitano de arte a través de Internet:**

<http://www.metmuseum.org/>

Esta base de datos incluye una importante colección de pinturas de artistas europeos.

- **Galería Nacional de Arte de Londres:**

<http://www.nationalgallery.org.uk/>

Esta colección comprende pinturas de artistas europeos desde el año 1260 hasta el 1900. La parte correspondiente al siglo XX a la galería de arte moderno Tate en Londres. Permite la búsqueda por el nombre del artista o del cuadro.

- **Galería Nacional de Arte de EE UU**

<http://www.nga.gov/>

Permite la búsqueda por artista o nombre del trabajo del la Galería Nacional de Arte de Washington D.C.

- **Archivo de la revista Rolling Stone:**

<http://www.rollingstone.com/>

Ofrece al usuario la oportunidad de acceder al archivo de la revista de música Rolling Stone desde sus comienzos en 1967.

- **Guía de arte de Canadá:**

<http://www.chin.gc.ca/>

Acceso a través de Internet a museos y galerías de Canada.

- **Colección del Museo estatal de Hermitage:**

<http://www.hermitagemuseum.org/>

Galería virtual de imágenes del Museo estatal de Hermitage.

.

- **The Thinker ImageBase:** <http://www.thinker.org/fam/thinker.html>

Base de datos visual de las colecciones de objetos de arte del Museo de Arte de San Francisco (EE UU).

➤ **DIRECTORIO DE RECURSOS DE ARTE Y ARQUITECTURA**

- **ADAM :**

<http://adam.ac.uk/>

ADAM es un catálogo de 2546 fuentes en Internet cuidadosamente seleccionados por bibliotecarios británicos profesionales.

- **Directorio de Arte Americano**

<http://www.artstar.com/>

Este directorio, ahora disponible a través de Internet fue publicado por primera vez en 1898. Aquí el usuario podrá encontrar información sobre el contacto, dirección y actividades de más de 7.500 museos, bibliotecas, organizaciones y escuelas de todo el mundo.

- **Directorio de arte IFLA:**

<http://iberia.vassar.edu/ifla-idal/>

Este directorio facilita el acceso a más de 3.000 bibliotecas especializadas en arte, arquitectura y arqueología de todo el mundo. Los datos aquí recopilados incluyen direcciones, teléfonos, horarios, listas del personal empleado, correo electrónico y enlaces al sitio web principal de cada institución. Esta base de datos ha sido realizada por la Federación Internacional de Bibliotecas, Asociaciones e Instituciones (IFLA).

- **Artcyclopedia:**

<http://www.artcyclopedia.com/>

Índice de artistas de cientos de museos, archivos de imágenes y otras fuentes en línea.

➤ **REFERENCIAS SOBRE ARTE Y ARQUITECTURA**

- **Thesaurus de Arte y Arquitectura (AAT):**
<http://www.getty.edu/gri/vocabularies/index.htm>

El AAT es una base de datos estructurada que contiene más de 125.000 términos sobre arquitectura y artes decorativas, además de otras informaciones.

BIBLIOGRAFÍAS Y CATÁLOGOS

➤ **BIBLIOGRAFÍAS**

- **21 North Main:**
<http://www.21northmain.com/>

Es una de las bases de datos más voluminosas. Permite además la posibilidad de comprar libros usados a través de este sitio web.

- **Achron (Manuscritos Históricos):**
<http://www.hmc.gov.uk/archon/archon.htm>

ARCHON es el principal recurso de información para la búsqueda de manuscritos sobre la historia británica. Esta base de datos está mantenida por la Comisión de Manuscritos Históricos de Gran Bretaña.

- **Archisplus:**

http://europa.eu.int/comm/secretariat_general/sg1/archives/home-en.htm

La Base de datos del archivo histórico de la Comisión Europea contiene detalles de los archivos de la Comunidad Europea del Carbón y del Acero (CECA), la Comisión Económica Europea (CEE), de la Comunidad Europea de Energía Atómica (EURATOM) y de la Unión Europea en general.

- **BOPCRIS (British Official Publications Collaborative Reader Information Service):**

<http://www.bopcris.ac.uk/>

Permite el acceso a publicaciones oficiales británicas relevantes entre el periodo de 1688 –1995.

- **Alicer Biblioteca**

<http://www.alicer.es/cgi-bin/p8-2.asp>

Consulta al fondo bibliográfico del ALICER. Contiene monografías sobre cerámica, diseño, arquitectura, historia del arte, informática, gestión... Se puede buscar por título, autor, lugar de publicación, año y materia.

- **British Library Public Catalogue:**

<http://blpc.bl.uk/>

Acceso a información sobre el contenido de las mayores bibliotecas británicas.

- **COPAC (Consortio de Bibliotecas de Universidades):** <http://copac.ac.uk/>

COPA permite el acceso libre, a través de Internet, a los catálogos de 19 de las mayores bibliotecas de universidades del Reino Unido e Irlanda.

- **Catálogo en línea de la Biblioteca Folger Shakespeare (HAMNET):**
<http://www.folger.edu/>

La Biblioteca Folger Shakespeare, centro independiente, contiene una de las mayores colecciones del mundo sobre los trabajos de Shakespeare, además de una magnífica colección de libros raros del renacimiento y manuscritos de todas las disciplinas: historia, política, teología, derecho y arte. La colección contiene aproximadamente 280.000 libros y manuscritos, y 27.000 pinturas, cuadros y grabados.

- **Bibliomedmar**

<http://www.semm.org/bibliom.html>

Base de datos de bibliografía sobre medicina marítima. Elaborado por la Sociedad Española de Medicina Marítima. SEMM.

- **Psicodoc**

<http://psicodoc.idbaratz.com>

Base de datos de referencias bibliográficas de revistas, congresos y monografías

relacionados con la psicología en España y América Latina. Permite el acceso gratuito limitado a la visualización de 3 registros y un acceso por un periodo de prueba de quince días mediante registro. Elaborado por el Colegio Oficial de Psicólogos de Madrid.

➤ **CATÁLOGOS**

- **IVAM catálogo**

<http://ivam.absysnet.com/>

Consulta al catálogo del fondo bibliográfico de la biblioteca del IVAM (Instituto Valenciano de Arte Moderno) especializado en Arte Moderno. Contiene monografías sobre teoría, estética, historia del arte y de los movimientos artísticos contemporáneos, audiovisuales y publicaciones periódicas.

NEGOCIOS E INVERSIONES

➤ **INFORMACIÓN Y BÚSQUEDA DE COMPAÑÍAS:**

- **10wizard.com** (Archivos públicos de las compañías):

<http://www.10kwizard.com>

Uno de los muchas formas de acceso al material de SEC EDGAR. 10K Wizard ofrece acceso libre y en tiempo real y búsquedas "a texto completo" del sistema EDGAR y enlaces a los archivos públicos de SEC. Además ofrece la posibilidad de acceder a los últimos archivos SEC de más de 68.000 compañías y de buscar archivos históricos desde el comienzo de la existencia de cada una de estas empresas por palabras clave, frases o nombres. Tiene disponibles otras fuentes tras completar un proceso de registro gratuito.

- **Australian business register.** <http://abr.business.gov.au/>.

La base de datos del registro de negocios australiano (ABR) está abierto al público a través de Internet y contiene parte de la información que ofrecen los negocios cuando se registran en este centro (ABN).

- **Better business bureau company reports.** <http://www.bbb.com>
<http://www.bbb.org/reports/bizreports.asp>

Se trata de una base de datos centralizada ya que el Better Bureau Company Reports no existe en la Web. Sin embargo, esta página pone a disposición del internauta enlaces a diversas oficinas en Estados Unidos y Cánada que hacen informes disponibles en la Red.

- **Business credit usa.** <http://www.businesscreditusa.com/>

Directorio básico e información de crédito de 12 millones de empresas estadounidenses. Ofrece más información bajo pago de una cuota.

- **Company sleuth.** <http://www.companysleuth.com>.

Ofrece información legal gratuita de empresas públicas de comercio. Company Sleuth recorre Internet buscando información especializada sobre los negocios, sus competidores, sus clientes, sus inversiones, etc.

- **.Disqualified directory register.** (Reino Unido).
<http://ws3.companieshouse.gov.uk/free/>

La información aquí disponible es una parte del Registro de Compañías del Reino Unido y del Registro de Directores Discapacitados, que es actualizado regularmente.

- **Directorio dot com**
<http://www.dotcomdirectory.com/nsi/advanced.htm>

Este directorio utiliza la base de datos de la "Network Solutions Domain Name Registration" para ofrecer información básica sobre las empresas.

- **Base de datos ecomp de compensación para ejecutivos.**

<http://www.ecomponline.com>

Esta base de datos ofrece datos de las compensaciones para ejecutivos de las compañías públicas de Estados Unidos.

- **Base de datos de la industria de alta tecnología europea**

<http://www.tornado-insider.com/radar/>.

Permite la búsqueda de notas de prensa de compañías de alta tecnología europeas y la oportunidad de leer sus perfiles en la base de datos Radar.

- **Empresas federales de Canadá**

Página principal: <http://strategis.ic.gc.ca>

http://strategis.ic.gc.ca/cgi-bin/sc_mrksv/corpdire/dataOnline/corpns-se

Esta base de datos elaborada por el gobierno de Canadá permite la búsqueda por el nombre de la corporación, su localización, su dedicación, etc. Permite, además, establecer los resultados por el nombre de la empresa o por el número de la misma.

- **Fortune 500**

Página principal: <http://www.fortune.com>

<http://www.fortune.com/fortune/fortune500/>

Lista de las mayores compañías públicas de comercio de Estados Unidos.

- **Herrington**

Página principal: <http://www.redherring.com>

<http://www.redherring.com/herrington/home/home.jsp>

"The Red Herring" es una publicación que ofrece bases de datos sobre compañías tecnológicas.

- **Kompass**

<http://www.kompass.com>

Cada compañía del mundo que participa en el negocio del comercio B2B (Business to Business) debe de ser registrada en la base de datos Kompass. La versión gratuita en Internet tiene información limitada. El resto de datos que contiene es accesible en la Red mediante suscripción.

- **Price waterhouse coopers money tree survey**

<http://www0.mercurycenter.com/svtech/>

<http://www.mercurycenter.com/svtech/companies/moneytree/>

Base de datos de las firmas estadounidenses que reciben capital de riesgo para financiarse. Esta versión está disponible a través de la web de San Jose Mercury News.

- **Compañía Red Herring y búsqueda de personas**

Página principal: <http://www.redherring.com>

<http://www.redherring.com/companies/>

Localiza datos superficiales de compañías y biografías de ejecutivos. "Red Herring" cubre la industria tecnológica.

- **Sedar (Archivos públicos de empresas de Canada)**

Página principal: <http://www.sedar.com>

http://www.sedar.com/search/search_form_pc.htm

SEDAR es el sistema de archivo electrónico para la revelación de documentos de compañías públicas y fundaciones de Canada.

- **Bases de datos de las empresas de Silicon Valley**

Página principal: <http://www0.mercurycenter.com/svtech/>

<http://www.mercurycenter.com/svtech/companies/db/>

Base de datos de las 150 mayores empresas públicas de comercio de Silicom Valley que contiene su historial e información financiera.

- **Registro Thomas de manufacturas americanas**

Página principal: <http://www.thomasregister.com/>

Esta web permite buscar 158000 empresas, 63669 productos y 135415 marcas localizadas en Canadá y Estados Unidos. Es necesario un registro previo gratuito para acceder a esta información.

- **Base de datos mundial de marcas**

http://adage.com/international/world_brands/index.html

Busca en la Base de datos de 400 anunciantes y 22 redes de agencias que llevan la cuenta de dichas marcas.

➤ FUENTES PARA EL CONSUMIDOR

- **Consumer Product Safety Commission Product Recalls:**

<http://www.cpsc.gov/cgi-bin/recalldb/prod.asp>

Permite la búsqueda de memorias ordenadas por la Comisión de Seguridad de Productos para el Consumidor (CPSC) desde 1977.

- **Strong Numbers:**

<http://www.strongnumbers/>

Permite el cálculo de valores de una gran variedad de artículos basados en los precios de más de cinco millones de subastas en línea cada semana.

➤ **ECONOMÍA ESTADOS UNIDOS**

- **Archivo del libro Beige:**

<http://minneapolisfed.org/bb/>

El Libro Beige es publicado dos semanas antes de cada reunión FOMC que se produce ocho veces al año. Cada banco de la Reserva Federal recopila información anecdótica sobre las condiciones económicas en su distrito, además de entrevistas con hombres de negocio claves, economistas, expertos de mercado, y otras fuentes.

- **Base de datos sobre préstamos de negocios pequeños**

<http://www.sba.gov/loans/business/>

Estadística sobre préstamos recientemente aprobados hechos a pequeños negocios en Estados Unidos. Es actualizado mensualmente.

- **Base de datos de economía y datos estadísticos:**

http://fisher.lib.virginia.edu/active_data/index.html

El Centro de Datos Estadísticos de la Universidad de Virginia proporcionan el acceso a varios recursos económicos. Se requiere un registro previo.

- **Datos económicos del Gobierno de EE UU:**

<http://govinfo.kerr.orst.edu/>

Es un proyecto del Gobierno de EE UU para compartir información que es está realizando físicamente desde la Universidad de Oregón. Permite el acceso a numerosas bases de datos del Gobierno de EE UU relativas a la economía, educación y demografía.

- **Base de datos de la Reserva Federal:**

<http://www.frbsf.org/system/fedinprint/>

Índice que permite la búsqueda en la base de datos de la Reserva Federal estadounidense.

- **Free Lunch:**

<http://www.economy.com/freelunch/default.asp>

Acceso libre a más de un millón de series de datos sobre economía y finanzas. Tras un registro gratuito el usuario puede acceder también a gráficos.

- **Base de datos del Departamento de Agricultura de Estados Unidos (USDA):**

<http://www.nass.usda.gov:81/ipedb/>

Permite el acceso a la base de datos del Departamento de Agricultura estadounidense, a sus estadísticas y a otros datos de interés en esta materia.

- **Tablas de la Renta Nacional y Cuentas de Productos (NIPA) :**

<http://www.bea.doc.gov/bea/dn/nipaweb/>

Acceso a través de Internet a tablas que muestran los datos económicos estadounidenses.

- **Estadísticas pesqueras del Servicio Nacional de Marina de EE UU (NMFS)**

<http://www.st.nmfs.gov/st1/commercial/>

El Servicio Nacional de Marina de EE UU (NMFS) ha automatizado este resumen de estadísticas de la industria pesquera estadounidense al que se puede acceder libremente.

- **Base de datos de búsqueda de sistemas de pago (PSRC):**

<http://www.chicagofed.org/paymentsystems/search1.cfm>

El PSRC (Payment Systems Resource Center) incluye una base de datos que permite la búsqueda de sistemas de pago. Posee, además, enlaces a otros sitios web de interés, entre otros datos.

- **RECON (Datos Económicos Regionales de EE UU):**

<http://www2.fdic.gov/recon/>

Regional Economic Conditions (RECON) fue diseñado inicialmente para la FDIC (Federal Deposit Insurance Corporation) con el objetivo de proveer de de determinada información económica a cada Estado de EE UU. Actualmente es útil en el análisis de riesgo de las instituciones financieras.

- **Datos económicos regionales de EE UU :**

<http://www.dismal.com/regions/regions.stm>

Permite visualizar datos económicos de 50 Estados de EE UU usando hasta 130 criterios diferentes. El usuario también puede acceder a 257 diferentes áreas de metro o encontrar un área de metro específica introduciendo el código de área.

➤ **ECONOMÍA EN EL MUNDO**

- **Banco de desarrollo de Asia**

<http://www.adb.org/Statistics/country.asp>

Esta base de datos permite el acceso a estadísticas básicas de economía y comercio en formatos Adobe Acrobat o Microsoft Excel.

- **Econbase**

<http://www.elsevier.nl/homepage/sae/econbase/menu.sht>

Econbase permite el acceso a 64 periódicos económicos publicados por Elsevier North-Holland Pergamon. Este servicio ofrece al usuario la oportunidad de buscar y leer la versión completa de estas publicaciones.

- **Convertor de códigos CNAE a IAE y viceversa**

<http://www.iberaval.es/convertor/index.asp>

Convertor de la clasificación española del Código Nacional de Actividades Económicas (CNAE) al código de Impuesto de Actividades Económicas (IAE), y viceversa.

- **Estadísticas sobre extranjería en EE UU**

<http://146.142.4.24/cgi-bin/surveymost?in>

Acceso a las estadísticas de muchas naciones sobre asuntos extranjeros. La base de datos pertenece al Bureau of Labor Statistics de Estados Unidos.

- **Calculadora de inflación (Canadá)**

http://www.bankofcanada.ca/en/inflation_calc.htm

Esta calculadora de la inflación en Canadá utiliza el precio de los consumidores que se publica mensualmente desde 1914 para presentar los cambios en la cesta de la compra de los consumidores canadienses.

- **Base de datos económica y social del del Banco Interamericano de Desarrollo**

<http://database.iadb.org/esdbweb/scripts/esdbweb.exe>

Esta base de datos de datos económicos y sociales (ESDB) ha sido creada por el Banco Interamericano de Desarrollo. Los países integrantes son los de América Latina y el Caribe.

- **Laborsta (Estadísticas de Trabajo)**

<http://laborsta.ilo.org>

Esta aplicación permite el acceso a datos del Despacho

Esta aplicación dinámica permite al usuario para tener acceso a datos de la Oficina de Organización Internacional de trabajo de Estadística. Los datos disponibles aquí son un extracto de LABORSTA, la base de datos principal de la Oficina. Consiste principalmente en series anuales que se publican en el Anuario de la Organización Internacional del Trabajo.

- **Penn World Tables**

<http://datacentre.chass.utoronto.ca:5680/pwt/>

Estadísticas detalladas de la economía mundial. Estas tablas continene actualmente datos de 152 países y 29 temas diferentes.

➤ **INSTITUCIONES FINANCIERAS**

- **Bases de datos de la Reserva Federal de Estados Unidos**

<http://www.ffiec.gov/nic/>

Contiene recursos que proporcionan datos estadísticos detallados estadísticos de bancos estadounidenses.

- **Financial Institution and Branch Office Data, Federal Deposit Insurance Corporation (FDIC)**

<http://www.fdic.gov/bank/individual/index.html>

Estas bases de datos permiten a los usuarios localizar instituciones para determinar su su condición financiera y su relación con otras instituciones.

- **Monetary Financial Instititions in the European Union**

https://mfi-assets.ecb.int/query_MFI.htm

Permite el acceso a datos del Banco Central Europeo y de otras instituciones de la Unión Europea.

- **GEA**

<http://www.partnersconsult.com/impiva/impiva.asp>

Base de datos de programas e instrumentos financieros de la Unión Europea, Estado Español y Administración Autonómica para la gestión de ayudas y subvenciones a industrias de la Comunidad Valenciana. Contiene información sobre el título y tipo de la subvención, organismo otorgante, fuente legal, duración, objetivo de la ayuda, zona geográfica beneficiaria, acciones apoyables, gastos subvencionables, condiciones de pago, cuantía máxima subvencionable, porcentajes, características de los beneficiarios, plazos...

➤ **FUENTES DE NEGOCIOS GENERALES**

- **American Community Network:**

<http://www.acn.net/>

Esta fuente ofrece acceso a varias bases de datos con información sobre la mayor parte ciudades americanas. Es necesario un registro previo de libre acceso.

- **Directorio Mundial de Cámaras de Comercio**

<http://www.worldchambers.com/CF/index.htm>

Este directorio electrónico contiene la información sobre más de 10.000 Cámaras en el mundo entero, con hiperenlaces a sus sitios web.

➤ **FUENTES ESPECÍFICAS DE LA INDUSTRIA**

- **AAAAgency Search (Publicidad)**

<http://www.aaaagencysearch.com/>

El sitio web de la Asociación Americana de Agencias Publicitarias ofrece acceso a rigurosos perfiles de las mejores agencias publicitarias de todo el mundo, y permite búsquedas sencillas y complejas. Todas las agencias listadas son miembros de esta asociación.

- **Base de datos de Biotech Alliance**

<http://www.recap.com/mainweb.nsf/HTML/alliance+frame?OpenDocument>

Esta base de datos contiene sumarios de alto nivel de más de 7.900 alianzas en las ciencias de vida que han sido formadas desde 1978.

- **European Case Clearing House**

<http://www.ecch.cranfield.ac.uk/>

Creada hace 25 años, la European Case Clearing House (ECCH) es una fuente única para el estudio del material relativo a la gestión de la educación. ECCH es una organización no lucrativa, sin embargo, su base de datos tiene disponibles más de 16,000 títulos destinados a la venta.

- **Base de datos de hoteles y propiedades**

<http://www.eventsource.com/Main/ResourceCenterMain.jsp?COMMAND=HOTELPROPERTY>

Esta base de datos contiene información de lugares destinados a convenciones de todo el mundo. Incluye descripciones detalladas y fotos.

- **Instituto para Investigación de Operaciones y las Ciencias de Gestión (INFORMS) Índice Bibliográfico Anual Completo de Base de datos**

<http://www.informs.org/Biblio/ACI.html>

Esta base de datos contiene aproximadamente 33.198 fuentes bibliográficas. Las fuentes en el periodo que va desde 1976 a 1987 son procedentes de informes de periódicos, mientras que las correspondientes al período entre 1988 y 1999 provienen de otros 614 tipos de informes diferentes.

- **Bases de datos de perfiles de firmas de arquitectura de EE UU**

<http://www.cmdg.com/profile/search.html>

Permite localizar empresas dedicadas a la arquitectura en Estados Unidos

- **Base de datos del Registro de Números de Identificación**

[https://rn.ftc.gov/TextileRN/wrnquery\\$.startup](https://rn.ftc.gov/TextileRN/wrnquery$.startup)

Un número de identificación certificado o RN es un número emitido por la Comisión Federal Comercial de Estados Unidos, previa petición, a la apertura de un negocio relacionado con la fabricación, importación, distribución, o la venta de textil, la lana, o productos de piel en Estados Unidos. No es una exigencia, sin embargo las empresas pueden usar el RN en lugar de su nombre sobre la etiqueta de sus productos.

- **Base de datos de televisión (Comisión Federal de Comunicaciones de EE UU)**

<http://www.fcc.gov/mmb/vsd/tvq.html>

Esta base de datos contiene un directorio e información técnica de las licencias de televisión de las cadenas estadounidenses.

➤ **FUENTES PARA LA INVERSIÓN**

- **Búsqueda de promedios del índice Dow Jones de Wall Street**

<http://averages.dowjones.com/screen1.htm>

Obtienen los promedios altos y bajos del índice Dow Jones de Industriales de cualquier fecha desde el 26 de mayo de 1896.

- **Base de datos financiera de la empresa Financial Times**

http://www.globalarchive.ft.com/cb/cb_search.htm

Esta base de datos permite al usuario obtener datos públicos de empresas de numerosos países

- **Base de datos de las cotizaciones de las acciones históricas**

<http://www.bigcharts.com/>

Esta base de datos proporciona acceso a datos de las acciones históricas de los mercados estadounidenses. En algunos casos proporciona acceso a material

desde 1985.

- **Hoover's StockScreener**

<http://www.stockscreener.com/>

Esta herramienta permite identificar compañías públicas de Estados Unidos usando varios criterios financieros.

- **InsiderScores.Com**

<http://www.insiderscores.com/>

Rastrea la habilidad comercial y el engranaje de distribución de mercado de más de 150.000 corporaciones en los últimos 14 años. Permite a los usuarios el acceso a las últimas informaciones de la Comisión de Cambio y Seguridad en más de 10.000 empresas públicas.

- **IPO Super Search**

<http://www.edgar-online.com/ipoexpress/supersearch.asp>

Permite la búsqueda de información reciente de Ofertas Públicas de Adquisición usando más de 20 tipos de criterios diferentes.

- **JustQuotes.Com**

<http://www.justquotes.com/>

Esta herramienta permite preconfigurar el nombre de una compañía o su símbolo para obtener la cotización bursátil en una base de datos que ofrece acceso al usuario mediante enlaces a una amplia lista de recursos de investigación.

- **Base de datos de la Actividad Mensual del Mercado NASDAQ**

<http://www.nasdaqtrader.com/asp/tdMarkSpec.asp>

Permite visualizar los informes de actividad del mercado de valores tecnológicos Nasdaq. Contiene información desde enero de 1997.

- **Clasificación Nelson de los mejores inversores a nivel mundial**

<http://www.nelnet.com/wbmm/intro.htm>

Cada quincena, esta base de datos recoge información del rendimiento de más de 1.500 gerentes de la inversión sobre más de 4.500 carteras. Este servicio clasifica sus rendimientos en una serie de 200 informes. Requiere un registro previo gratuito.

- **Calculadora de los valores en Bolsa**

<http://www.dismalscientist.com/cgi/stocks.asp>

Esta herramienta permite calcular el valor de las acciones.

- **Directorio de las empresas que cotizan en la Bolsa de Tokio**

<http://www.tse.or.jp/ec/listed/eindex.html>

Permite al usuario la búsqueda de empresas que cotizan en la Bolsa de Tokio usando diferentes criterios que incluyen su clasificación industrial, su inicial y el nombre de la

compañía.

➤ **INFORMACIÓN SOBRE TRABAJOS Y CARRERAS PROFESIONALES**

- **EURES jobsearch**

<http://www.inem.es/ciudadano/empleo/busqueda/eures.html>

Ofertas de empleo en Europa a través de la web. EURES es una red europea de servicios para la búsqueda de empleo que facilita la movilidad y colocación de los trabajadores. En inglés, francés y alemán.

- **EQUAL**

http://europa.eu.int/comm/employment_social/equal/index_es.html

Base de datos de iniciativas ADAPT / EMPLEO sobre empleo, incorporación al mercado de trabajo, integración social, lucha contra la discriminación y la desigualdad e innovación en Europa. Contiene también acceso al catálogo de la biblioteca de esta iniciativa.

- **Banco de Trabajo americano**

<http://www.ajb.org/>

El Banco de Trabajo Americano contiene la mayor base de datos sobre empleo disponible en Internet. Permite la inclusión del curriculum vitae al que acceden empresas en busca del perfil que necesitan. Permite búsquedas de trabajo automáticas.

- **Flipdog**

<http://www.flipdog.com/js/loc.html>

Las fuentes aquí listadas son representativas de miles de sitios web dedicados a la búsqueda de empleo en Internet. La mayoría de la información de estas bases de datos reside en la Red profunda.

- **Calculadora Internacional de Salarios**

http://www2.homefair.com/calc/intsalcalc.html?NETSCAPE_LIVEWIRE.src=homefair

Compara el coste de la vida en ciudades de todo el mundo.

- **O*Net**

<http://online.onetcenter.org/>

La base de datos de O*NET incluye información sobre habilidades, capacidades, conocimientos y actividades de trabajo asociadas con ocupaciones. Esta información puede ser usada para facilitar la exploración de una carrera, el asesoramiento profesional, y una amplia variedad de funciones de recursos humanas.

- **Salary.Com Salary Wizard**

http://swz.salary.com/salarywizard/layoutscripts/swzl_newsearch.asp

Esta base de datos contiene información de los salarios equivalentes a miles de puestos de trabajo. Además ofrece la posibilidad al usuario de calcular el salario basado en el puesto de trabajo según la localización geográfica.

➤ DIRECTORIOS DE SERVICIOS

- **192.Com**

<http://www.192.com/>

192. COM proporciona a sus usuarios acceso rápido a la base de datos más grande de teléfono y dirigir la información sobre Internet. El contenido continuamente es puesto al día. Es necesario un registro previo aunque gratuito.

- **Anywho.Com (Directorio telefónico)**

<http://www.anywho.com/>

Base de datos que ofrece acceso a una guía telefónica. Anywho.com proporciona teléfonos de empresas y particulares, listados residenciales y de negocio.

- **Directorio telefónico Reverse**

<http://www.anywho.com/telq.html>

Permite la búsqueda en su directorio telefónico introduciendo el número de teléfono.

➤ **FUENTES DE MARKETING**

- **New York Green Book**

<http://www.greenbook.org/greenbook/search.cfm>

The New York Chapter of the American Marketing Association permite la búsqueda en su directorio de forma gratuita. Esto incluye una amplia gama de empresas de estudios del mercado y servicios por todo el mundo.

- **Zip Code Business Profiles (Censo estadounidense)**

<http://tier2.census.gov/zbp/index.html?>

El modelo de Negocio de Zona Postal proporciona datos sobre el número total de establecimientos en Estados Unidos y su tamaño en cuanto a número de empleados.

➤ **FUENTES SOBRE PENSIONES**

- **Employee Benefits INFOSOURCE™**

<http://www.ifebp.org/infosource/default.asp>

Permite encontrar información sobre pensiones y fondos de compensación de EE UU y Canadá. Contiene más de 60.000 sumarios de más de 350 diarios en lengua inglesa, artículos, reportajes y libros. Es necesario un registro previo.

- **FreeErisa.com**

<http://www.freerisa.com/>

Ofrece el acceso a una base de datos con los últimos datos disponibles en materia de pensiones. Es necesario un registro previo para acceder a estos datos.

➤ **FINANZAS PERSONALES**

- **RateNet**

<http://www.rate.net/>

Permite al usuario acceder a los tipos de interés y otra información financiera de más de 11.000 instituciones financieras en todo Estados Unidos.

➤ **FILANTROPÍA Y FUENTES NO LUCRATIVAS**

- **Foundation Finder**

<http://lnp.fdncenter.org/finder.html>

Permite el acceso a información básica sobre más de 59.000 fundaciones en Estados Unidos.

- **Guidestar.Org**

<http://www.guidestar.org/search/>

Ofrece información de más de 640.000 organizaciones sin ánimo de lucro en los Estados Unidos.

- **Idealist**

http://www.idealist.org/is/org_search.html

Base de datos de más de 20.000 organizaciones sin ánimo de lucro en 150 ciudades de Estados Unidos.

➤ **BÚSQUEDA Y DESARROLLO**

- **CORDIS (Community Research & Development Information System)**

http://dbs.cordis.lu/EN_GLOBALsearch.html

Este servicio provee de información sobre la búsqueda en general y el desarrollo. Está subvencionado por la Unión Europea.

- **TECH-net (U.S. Small Business Administration)**

<http://tech-net.sba.gov/tech-net/search.html>

Tech-Net es una base de datos que contiene información sobre tecnología y fuentes útiles para negocios pequeños especializados en alta tecnología. Esta web está provista de un motor de búsqueda especialmente indicado para documentalistas, científicos y oficiales del Gobierno norteamericano. Además, se trata de un sitio web interesante que ofrece oportunidades de inversión.

- **TechTracS/TechFinder (NASA)**

<http://technology.larc.nasa.gov/techfinder/>

NASATechTracS es una base de datos continuamente actualizada que contienen todos los programas de la NASA y tecnologías empleadas por este organismo.

➤ **COMERCIO**

- **CTM-ONLINE servicio de consulta de marcas comunitarias**

http://oami.eu.int/search/trademark/la/es_tm_search.cfm

Proporciona acceso a los datos de las marcas comunitarias. Se puede buscar por marca o número de solicitud. La información suministrada proviene de los datos contenidos en la base de datos EUROMARC.

- **Canadian Importers Database**

http://strategis.ic.gc.ca/sc_mrkti/cid/engdoc/index.html

Esta base de datos proporciona una lista de compañías que importan mercancías en Canadá. Permite la búsqueda por producto y por ciudad.

- **Trade Data Online (Canadá)**

http://strategis.ic.gc.ca/sc_mrkti/tdst/engdoc/tr_homep.html

Genera informes del comercio de Estados Unidos y Canadá con más de 200 países.

- **Foreign Agricultural Service Import/Export Data (USDA) (Estados Unidos)**

<http://ffas.usda.gov/country.html>

Estas bases de datos permite el acceso a informes sobre las importaciones y exportaciones de mercancías derivadas de la agricultura en EE UU.

- **International Trade Commission Interactive Tariff and Trade DataWeb (EE**

<http://dataweb.usitc.gov/>

Este sitio web permite el acceso a varias bases de datos que incluye información sobre tarifas y estadísticas.

- **U.S. State Exports Database (U.S. Department of Commerce)**

http://ita.mapinfo.com/SCRIPTS/hsrun.hse/single/ITA/MapXtreme.htx;start=HS_Intro

Ofrece acceso a dos bases de datos. Una de ellas permite visualizar las exportaciones de cada Estado de Estados Unidos al resto del mundo; mientras que la otra muestra las exportaciones desde cada Estado a otro país específicamente determinado. Es dependiente del Departamento de Comercio de Estados Unidos.

- **Base de datos de acuerdos comerciales de EE UU**

<http://www.mac.doc.gov/tcc/data/index.html>

Base de datos centralizada que permite el acceso integro a más de 250 tratados comerciales de Estados Unidos, y que pueden ser de gran utilidad en los negocios.

➤ **CONVENCIONES**

- **Centro de Convención y Directorio de Oficina de Visitantes**

http://info.asaenet.org/convctrs/cvbdire_SQL.cfm/

Localiza convenciones en todo el mundo. La base de datos es actualizada por la Sociedad Americana de Ejecutivos.

- **ExhibitorNet.Com: Directorios de Ferias de Muestras**

<http://www.tscentral.com/EventCenter/>

Esta base de datos ofrece un directorio de ferias de muestras de todo el mundo. Ofrece la posibilidad de hacer búsquedas específicas, mediante herramientas de restricción de búsqueda.

INFORMÁTICA E INTERNET

➤ **INFORMÁTICA**

- **Asociación de la Biblioteca Digital de Ordenadores (ACM)**

<http://www.acm.org/>

Esta biblioteca digital permite la búsqueda en su base de datos bibliográfica a cualquier usuario de forma gratuita, aunque es requerido un registro previo. El texto completo se puede adquirir *online* previo pago.

- **Bitpipe**

<http://www.bitpipe.com/>

Bitpipe reúne catálogos, y distribuye literatura corporativa a otros sitios web. Puede considerarse un servicio de telégrafo para la literatura profesional destinada a periódicos e investigadores.

- **CORA: Computer Science Research Paper Search Engine**

<http://cora.whizbang.com/>

Cora es un motor de búsqueda especializado en investigación de temas relacionados con la informática. Permite la búsqueda de palabra clave y proporciona acceso a más de 50.000 informes de investigación sobre este tema.

- **DriverGuide.Com**

<http://www.driverguide.com/>

Base de datos de conductores y recursos sobre conducción. Aunque es gratuita se requiere un registro previo.

- **Mapa Mundial de virus informáticos de McAfee**

http://mast.mcafee.com/mast/mass_map.asp?

Permite el acceso a una lista actualizada de los virus informático que están infectando los ordenadores de todo el mundo.

- **Networked Computer Science Technical Reference Library**

<http://cs-tr.cs.cornell.edu/>

NCSTRL es una colección internacional de informes de investigación de informática para uso no comercial.

- **ResearchIndex**

<http://www.researchindex.com/>

ResearchIndex es una biblioteca científica digital de la literatura que ayuda a la divulgación de la literatura científica.

- **SecuritySearch.net**

<http://www.securitysearch.net/>

Ofrece el acceso a una base de datos que contiene un directorio de sitios web relacionados con temas sobre seguridad.

- **Bibliografías de la Colección de Ciencias de los Ordenadores**

<http://iinwww.ira.uka.de/bibliography/index.html>

Colección de bibliografías sobre temas relacionados con la informática de varias fuentes. Contiene aproximadamente 1400 bibliografías puestas al día mensualmente.

➤ **FUENTES DE INTERNET**

- **Allwhois.com**

<http://www.allwhois.com/>

Permite revisar la base de datos “whois” de cualquier dominio del mundo.

- **Cybercafe Search Engine**

<http://www.cybercaptive.com/>

Esta base de dato ofrece un listado de más de 6.000 cybercafés en 168 países.

- **Internet Traffic Report**

<http://www.internettrafficreport.com/>

The Internet Traffic Report permite visualizar el tráfico de datos a través de Internet en todo el mundo. Ofrece un valor entre 0 y 100 que muestra los niveles de rapidez de conexión en los distintos lugares.

- **ISPs.com**

<http://www.isps.com/>

Permite localizar servidores de Internet.

- **Meta-List**

<http://www.meta-list.net/query?acc=110en>

Meta-List permite buscar información en más de 170.000 listas de discusión públicas.

- **NetLingo**

<http://www.netlingo.com/>

NetLingo es un diccionario sobre términos relacionados con Internet. Contiene cientos de estos términos que utiliza la gente cuando se comunica a través de Internet o realiza transacciones de comercio electrónico.

- **Network Solutions USA Internet Facts**

<http://www.dotcom.com/facts/usmap.html#>

Permite el acceso a gráficos que ilustran los registros de dominios de Estados Unidos por ciudad, estado y área metropolitana.

- **Search Engine Guide**

<http://www.searchengineguide.com/searchengines.html>

Encuentra motores de búsqueda, portales y directorios de una gran variedad de tópicos. Posee un directorio con más de 3.500 fuentes.

EDUCACIÓN

➤ **DIRECTORIOS**

- **College Opportunities Online (National Center for Education Statistics)**

<http://nces.ed.gov/ipeds/cool/Search.asp>

Ofrece enlaces a más de 9.000 universidades en los Estados Unidos.

- **Directory of Resources for Foreign Language Programs**

<http://www.cal.org/ericcll/ncbe/fldirectory/>

Acceso a un directorio de programas de lengua extranjera. Provee información sobre asociaciones y organizaciones profesionales.

- **Education Resource Organizations Directory**

<http://www.ed.gov/Programs/EROD/>

Directorio de organizaciones relacionadas con la educación

- **ERIC/AE Test Locator**

<http://ericae.net/testcol.htm>

Tres bases de datos separadas sobre temas educativos.

- **Online Distance Education Catalog (Globewide Network Academy)**

<http://www.gnacademy.org/mason/catalog/front.html>

Directorio con enlaces de oportunidades de aprendizaje a distancia en todo el mundo. Hay listados más de 23.000 cursos.

- **Shaw Guides**

<http://www.shawguides.com/>

Ofrece acceso a más de 4.300 programas sobre una gran variedad de temas de aprendizaje para vacaciones y otros cursos creativos en todo el mundo.

- **Educ@ment**

<http://teledoc.urv.es/educ@ment/>

Base de datos de referencias bibliográficas de publicaciones periódicas sobre educación y psicología.

- **Becas, lectorados e intercambios**

<http://www.aeci.es/7-Becas/index-becas.htm>

Consulta asistida de las últimas convocatorias y concesiones de becas, lectorados e intercambios, convocados por el Ministerio. Elaborada por el Ministerio de Asuntos Exteriores de España.

- **Cursos**

<http://www.aragob.es/sid/subv/subocurs.htm>

Consulta de los cursos convocadas por la Comunidad Autónoma de Aragón y la Administración General del Estado, con acceso al texto completo de la convocatoria.

➤ **PASATIEMPOS**

- **All Game Guide**

<http://allgame.com/>

Con un alcance que abarca desde Pong hasta los productos más modernos de nueva generación, All Game Guide presenta un inventario de constante crecimiento, con más de 25000 juegos de más de 85 plataformas.

- **Internet Anagram Server**

<http://www.wordsmith.org/anagram/>

Escribiendo cualquier palabra o frase genera una lista de anagramas que la contienen, incluso reordenando las letras o las palabras.

- **The Roller Coaster Database**

<http://www.rcdb.com/>

Roller Coaster es una base de datos extensa para buscar información y estadística de más de 1000 “roller coasters” en Norte América y Europa.

➤ **FUENTES DE ENTRETENIMIENTO GENERAL**

- **Searchable Television Listings**

<http://www.clicktv.com/search.asp>

Un ejemplo de guía de T.V.interactiva.

➤ **PELÍCULAS Y CINES**

- **Cinefiles (Pacific Film Archive)**

<http://www.mip.berkeley.edu/cinefiles/>

Una base de datos de revistas, artículos de periódicos, notas de programas varios y otros documentos de los archivos de la Biblioteca de Pacific Film. Los archivos contienen

documentos de un amplio elenco de fuentes que abarcan el mundo del cine, pasado y presente.

- **Cinema FreeNet**

<http://www.cinfn.com/>

Esta página permite buscar relaciones entre películas, actores, directores y productores.

- **Current Film in the Works**

<http://www.boxoffice.com/scripts/fiw.dll?DoSearch>

Boxoffice, una publicación registrada de la industria del cine, proporciona acceso a una lista de películas que se estrenarán en un futuro. Cada lista incluye un pequeño resumen del argumento y una lista de actores que toman parte en la producción.

- **Directory of International Film and Video Festivals (The)**

<http://www.britfilms.com/fv/home.lasso>

La guía lista más de 500 películas internacionales, producciones de televisión y video dando detalles de cómo y cuando se representaron. Está recopilada y publicada por el Consejo de Cine británico.

- **Internet Movie Database**

<http://us.imdb.com/Search/>

IMDB empezó como un pequeño proyecto en el Reino Unido y ahora es un servicio que forma parte de la familia Amazon.com.

- **Inter-Play B**

<http://www.portals.org/interplay/play.html>

Inter-Play proporciona fácil acceso a localizaciones de producciones que forman parte de colecciones o antologías. La mayoría de estas producciones citadas no están indexadas en listas tales como "Ottemiller's Index to Plays in Collections or H.W.

Wilson's Play Index. Para obras de artistas más conocidos como Shakespeare, se pueden consultar catálogos de librerías locales. Aproximadamente 18000 localizaciones están citadas, incluyendo además distintos lenguajes. Inter-Play es actualizada frecuentemente para incluir las publicaciones más recientes.

- **Movie Review Query Engine**

<http://www.mrqe.com>

Esta base de datos contiene información acerca de unas 21000 películas. Esta información proviene de periódicos y de fuentes de internet.

- **MovieFone**

<http://www.moviefone.com>

Una guía gratuita de localización de espectáculos, cines y teatros a través de EE.UU.

- **Video Distributors**

[http:// www.videolibrarian.com/ producers.html](http://www.videolibrarian.com/producers.html)

Guía básica de información de compañías distribuidoras de videos.

➤ **MÚSICA**

- **All Music Guide**

<http://allmusic.com/>

Esta guía contiene una amplia información de músicos, albums y canciones.

- **ASCAP Music License Database**

[http:// www.ascap.com/ace/search.cfm?mode=search](http://www.ascap.com/ace/search.cfm?mode=search)

ACE es una base de datos de canciones publicadas con la licencia de ASCAP (Sociedad Americana de compositores, artistas y publicantes) en los EE.UU. Para cada título, se pueden encontrar los nombres de las personas que escribieron las canciones, así como las personas de contacto, direcciones y teléfonos de editores para contactar con ellos, si

se quiere usar este material. Para la mayoría de los títulos encontrará algunos artistas que han hecho una grabación comercial.

- **Beethoven Bibliography Database**

<http://www.sjsu.edu/depts/beethoven/database/database.html>

La base de datos Beethoven Bibliography es un emocionante proyecto que une el continuo interés en la vida y trabajos de Ludwig van Beethoven con las ventajas de los ordenadores e Internet. Contiene unas 10000 entradas.

- **Billboard Spotlight Reviews**

<http://www.billboard.com/reviews/finder.asp>.

Contiene más de 30000 artículos sobre los álbumes originales de Billboard, desde 1970 hasta hoy.

- **Canadian Music Periodical Index**

http://www.nlc-bnc.ca/cmpi/about_e.htm

Esta base de datos incluye más de 25000 entradas indexadas a partir de 475 revistas, periodicos y publicaciones musicales canadienses, desde finales del siglo XIX hasta el presente. Está enfocada a cubrir todos los aspectos de la actividad musical en Canadá. Más de 200 títulos están siendo indexados en la actualidad.

- **Hoagy Carmichael Collection (The)**

<http://www.dlib.indiana.edu/collections/hoagy/search/index.html>

Un completo catálogo de toda la Colección Carmichael, con acceso a objetos digitales, información suplementaria para investigar, así como genealógica. La Colección Hoagy Carmichael contiene sonido de la música de Hoagy, letras , fotografías de él y su familia, composiciones musicales manuscritas y más.

- **Mudcat Café Digital Tradition Folksong Database (The)**

<http://www.mudcat.org/>

Se puede hojear esta lista que contiene la letra de unas 8000 canciones folk.

- **Musica: The International Database of Choral Repertoire**

<http://www.musicanet.org/en/cherhgb.htm>

MUSICA es actualmente una herramienta de búsqueda de música coral, así como una pedagógica herramienta para directores, escuelas, tiendas y federaciones de música.

Pero también lo es para amateurs y gente ilusionada en el repertorio de música coral.

- **Operabase**

<http://www.operabase.com/en/>

Esta base de datos ofrece numerosas opciones de búsqueda para localizar datos sobre la ópera. Notar el acuerdo especial con “The Grove Dictionary of Opera”, que proporciona muchos datos en esta dirección.

- **Pollstar Concert Database**

<http://www.pollstar.com>

Búsqueda de datos de conciertos por medio de varios criterios, incluyendo la ciudad y lugar del mismo.

- **RIAA Gold and Platinum Database**

<http://www.riaa.com/Gold-Intro-2.cfm>

Información sobre grabaciones certificadas con discos de oro o platino por la Asociación de la Industria de la Grabación Americana.

- **Themefinder**

<http://www.themefinder.com>

La base de datos Themefinder contiene el principio de unos 20000 temas de repertorios de música clásica y folk. Para buscar información sobre algún tema normalmente se confía en el título del mismo para encontrarla; Themefider se usa cuando no se dispone del mismo, pero la melodía puede ser recordada.

- **UBL Ultimate Band List**

<http://ubl.artistdirect.com/>

Una tienda para información y fuentes de más de 10000 artistas.

➤ **REPRESENTACIONES Y EVENTOS**

- **CultureFinder**

<http://www.culturefinder.com>

Localización de eventos y entradas para representaciones en la mayoría de las ciudades de EE.UU.

- **Festivals.com**

<http://www.festivals.com/>

Los festivales son la forma universal de celebración en una comunidad. Si los llamamos ferias o fiestas traen gente a un espacio público. Esta página busca más de 37000 eventos en todo el mundo.

- **London Theatre Guide**

<http://www.officiallondontheatre.co.uk/index.cfm>

Encuentra qué se está representando con esta guía oficial de los 50 teatros más importantes de Londres.

- **Playbill Theatre Listings**

<http://www.playbill.com>

Esta conocida publicación de teatro proporciona una lista donde buscar las obras representadas en Londres y Norte América.

- **Teatro Español**

http://parnaseo.uv.es:82/teatro_modif.asp

Base de datos de bibliografía sobre teatro español, crítica y teoría teatral, actores, etc.

INFORMACIÓN SOBRE GOBIERNOS

➤ **DOCUMENTOS DEL GOBIERNO**

- **Directorio de política internacional**

<http://www.diplomaticnet.com/es/>

Páginas web de gobiernos, embajadas, consulados, ministerios de asuntos exteriores, organizaciones internacionales, ONG especializadas, prensa especializada y organismos relacionados con la formación e investigación. Se puede buscar por país o por tema. Dispone también de un buscador.

- **10 Downing Street News Search**

<http://www.number-10.gov.uk/search2.asp>

Acceso a las notas de prensa del Primer Ministro del Reino Unido.

- **1040.com**

<http://www.1040.com/>

Encuentra, descarga e imprime cualquier documento relativo a impuestos. Asimismo permite la búsqueda sobre cuestiones relativas a impuestos en cualquier localidad de los Estados Unidos.

- **AGIP (Australian Government Index of Publications)**

<http://203.2.143.24/webpac-bin/wgbroker?new+-access+top>

Ofrece acceso a documentos y publicaciones del Gobierno de Australia.

- **Catalog of U.S. Government Publications**

http://www.access.gpo.gov/su_docs/locators/cgp/index.html

Este catálogo ofrece bibliografía de archivos del Gobierno de Estados Unidos, desde enero de 1994.

- **European Foreign Policy Bulletin**

<http://wwwarc1.iue.it/iue/fillf?form.html=form2efpb.html>

Esta base de datos contiene documentos elaborados por la Unión Europea relativos a la política de extranjería desde 1985.

- **RAPID (European Union News)**

<http://europa.eu.int/rapid/start/welcome.htm>

RAPID es una base de datos continuamente actualizada que permite visualizar las actividades de la Unión Europea y de las instituciones que la integran, a través de las notas de prensa que estas emiten.

- **THOMAS: Legislative Information on the Internet**

<http://thomas.loc.gov/>

Este servicio de la Biblioteca del Congreso de EE UU ofrece acceso a la legislación de EEUU

- **United Nations Voting Records**

<http://unbisnet.un.org/webpac-bin/wgbroker?new+-access+top.vote>

Base de datos que contiene los resultados de las votaciones de todas las resoluciones que han sido adoptadas por las Naciones Unidas desde 1983 y por el Consejo de Seguridad desde 1946.

➤ **POLÍTICA Y RELACIONES INTERNACIONES**

- **Country Indicators for Foreign Policy**

<http://www.carleton.ca/cifp/>

Ofrece información estadística en materia de economía, política, sociedad y cultura de países de todo el mundo. Es necesario un registro gratuito previo.

- **Database KOSIMO (Political Conflicts)**

http://www.hiik.de/en/kosimo/kosimo_query.htm

KOSIMO es el nombre de una base de datos que contiene 693 conflictos políticos desde 1945 a 1999.

- **FIRST (Facts on International Relations and Security Trends)**

<http://first.sipri.se>

Es un servicio gratuito para políticos, periodistas e investigadores realizado por el International Relations and Security Network (ISN) y el Stockholm International Peace Research Institute (SIPRI).

- **International Boundary News Database**

<http://www-ibru.dur.ac.uk/database/data.html>

Esta base de datos contiene más de 10.000 informes sobre fronteras de todo el mundo desde 1991.

- **Women in Politics Bibliographic Database**

<http://www.ipu.org/bdf-e/BDFsearch.asp>

Esta base de datos es regularmente actualizada par incluir nuevos libros y artículos producidos en todo el mundo en materia de la mujer en política. Hasta ahora cubre más de 650 títulos.

➤ **ESTADÍSTICAS**

- **FedScope**

<http://www.fedscope.opm.gov/>

Base de datos que incluye una guía con estadísticas locales y otros datos sobre el empleo en los Estados Unidos.

- **Fedstats (United States Statistics)**

<http://www.fedstats.gov/>

FedStats ofrece información pública estadística del Gobierno Federal. Ofrece acceso a la colección de estadísticas públicas de más de 70 Agencias federales sobre economía, población, costes de sanidad, seguridad aérea, comercio exterior y uso de energía, entre otros.

- **Infonation**

http://www.un.org/Pubs/CyberSchoolBus/infonation/e_infonation.htm

InfoNation es una base de datos fácil de utilizar que permite ver y comparar datos estadísticos de los Estados miembros de las Naciones Unidas.

SALUD E INFORMACIÓN MÉDICA

➤ **ENFERMEDADES**

- **Imágenes de plagas y enfermedades**

<http://www.iicasaninet.net/pub/sanveg/html/imgplaenf/index.html>

Archivo de imágenes de plagas y enfermedades ordenadas alfabéticamente de acuerdo al nombre científico del organismo.

- **AIDS Economics Bibliographic Search**

<http://www.worldbank.org/aids-econ/biblio.htm>

Esta base de datos contiene publicaciones relevantes sobre el sida.

- **Disease Surveillance Online**

<http://www.hc-sc.gc.ca/hpb/lcdc/webmap/index.html>

Tres bases de datos con información estadística sobre el cáncer y enfermedades cardiovasculares.

- **Economics of Tobacco Control Database**

<http://www1.worldbank.org/tobacco/database.asp>

Ofrece estadísticas relacionadas con el tabaco de numerosos países.

- **UNAIDS bibliographic database**

http://www.unaids.org/aidspub/publication_search.asp

Base de datos de publicaciones e investigaciones sobre el VIH y el SIDA realizadas por los organismos de las Naciones Unidas. En inglés.

- **European Database on AIDS and HIV Infection**

<http://www.edoa.org/>

Base de datos que ofrece documentación sobre el virus VIH.

- **HIV/Aids Treatment Directory (AmfAR's)**

<http://199.105.91.6/treatment/mainframe.asp>

Esta base de datos contiene noticias recientes sobre el SIDA, estudios clínicos y tratamientos, entre otras cosas.

- **PDQ (Physician Data Query) Comprehensive Cancer Database**

<http://cancernet.nci.nih.gov/pdq.html>

Base de datos sobre el cáncer. Contiene informes sobre la enfermedad, tratamientos prevención. Además, ofrece acceso libre a un registro de más de 1.800 juicios clínicos de todo el mundo y a más de 10.300 previo pago.

- **PIE Database (Mental Health Policy)**

<http://www.mimh.edu/mimhweb/pie/database/database.htm>

Base de datos sobre enfermedades mentales.

- **Rare Disease Database (National Organization for Rare Disorders)**

<http://www.stepstn.com/nord/db/dbsearch/search.htm>

Esta base de datos proporciona acceso a información de más de 1.100 enfermedades. El
Los resúmenes son gratuitos pero para acceder al texto completo hay que pagar.

- **Tobacco Industry Documents**

<http://www.cdc.gov/tobacco/industrydocs/index.htm>

Acceso a documentos relacionados con la industria del tabaco.

- **WHO (World Health Organization) Cancer Mortality Databank**

<http://www-dep.iarc.fr/dataava/globocan/who.htm>

Datos sobre mortalidad por Cáncer de varios países.

- **HealthComm KEY Database**

<http://www.cdc.gov/od/oc/hcomm/>

Esta base de datos contiene sumarios de más de 200 artículos sobre investigaciones en materia de salud.

- **International Digest of Health Legislation**

<http://www-nt.who.int/idhl/en/ConsultIDHL.cfm>

Contiene una selección sobre legislación internacional en materia de salud.

- **Centro de documentación e información del Plan Nacional sobre Drogas**

<http://www.sindrogas.es/pnd/centro/index.jsp>

Centro de documentación e información especializado en drogodependencias, con fondos bibliográficos en diferentes lenguas nacionales y extranjeras. Los fondos bibliográficos abarcan todo tipo de sustancias adictivas (legales e ilegales), así como las políticas de actuación, tanto en España como en los principales países occidentales, en relación con la prevención, tratamiento, reinserción, etc. de las drogodependencias y el control del tráfico ilícito de drogas. También cuenta con documentación sobre la actividad desarrollada por los Organismos internacionales competentes en esta materia: Naciones Unidas, Unión Europea, Consejo de Europa, etc.

- **Observatorio interamericano sobre drogas**

<http://www.cicad.oas.org/OID/defaultespa%F1ol.htm>

Base de datos de recursos web sobre drogodependencias en América. Información sobre instituciones y organismos gubernamentales y no gubernamentales, legislación, documentos sobre prevención y tratamiento, pandillas y drogas, investigaciones en uso y abuso de drogas... Consulta por países y temas.

- **Drogodependencias. Planes y programas municipales**

<http://www.femp.es/drogodep/index.htm>

Base de datos que recoge los planes de drogas y programas desarrollados en este campo en el ámbito municipal. Información sobre municipios, planes, programas y cursos de formación.

- **Medidas sanitarias y fitosanitarias**

http://www.wto.org/spanish/tratop_s/sps_s/sps_s.htm

Documentos sobre medidas relativas a la inocuidad de los alimentos, la salud de los animales y la preservación de los vegetales (medidas sanitarias y fitosanitarias o MSF).

➤ **FUENTES SOBRE SALUD PROFESIONALES**

- **International Classification of Functioning, Disability and Health**

<http://www.who.int/icidh/>

Contiene la Clasificación Internacional conocida como ICIDH-2 que sirve para unificar el lenguaje y el marco para la descripción de ciertos términos relacionados con la salud.

- **Internet Grateful Med**

<http://igm.nlm.nih.gov/>

Ofrece acceso a 13 bases de datos relacionadas con la medicina.

- **PubMed/Medline**

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>

Medline es una de las bases de datos más importantes en el mundo sobre literatura médica.

- **TRIP Database**

<http://www.tripdatabase.com/>

TRIP Database es un metabuscador que permite la búsqueda en 61 sitios web de alta calidad médica. Ofrece acceso directo a una larga colección de casos prácticos y artículos de periódicos digitales especializados, como son: BMJ, JAMA, NEJM, entre otros.

➤ **IMÁGENES**

- **Images from the History of Medicine**

<http://www.ihm.nlm.nih.gov/>

Ofrece acceso a casi 60.000 imágenes relacionadas con la historia de la medicina.

- **Photoshare**

<http://db.jhuccp.org/mmc/photoshare/search.stm>

Contiene miles de fotografías de población, salud pública y otros asuntos relacionados con la salud en los países desarrollados.

➤ **NUTRICIÓN**

- **International Bibliographic Information on Dietary Supplements (IBIDS)**

<http://ods.od.nih.gov/databases/ibids.html>

Esta base de datos ofrece información a nivel internacional de dietas complementarias, vitaminas y minerales.

➤ **FUENTES DE INFORMACIÓN SOBRE PACIENTES Y CONSUMIDORES**

- **Medicare Health Plan Compare**

<http://www.medicare.gov/MPHcompare/Home.asp>

Medicare Health Plan Compare ofrece información detallada sobre temas relacionados con la salud.

- **Transplant Patient Datasource, United Network for Organ Sharing (UNOS)**

<http://www.patients.unos.org/tpd/>

Este sitio permite al usuario acceder a información sobre transplantes.

➤ **MEDICAMENTOS**

- **Base de datos de medicamentos SEFH**

<http://www.sefh.es/buscador/index.htm>

Fichas técnicas de medicamentos. Se puede buscar por principio activo o por nombre comercial. Proporciona información sobre su nombre, composición cualitativa y cuantitativa, forma farmacéutica, datos clínicos, indicaciones terapéuticas, posología y forma de administración, contraindicaciones, advertencias y precauciones especiales de empleo, interacción con otros medicamentos y otras formas de interacción, embarazo y lactancia, efectos sobre la capacidad para conducir y utilizar maquinaria, reacciones, sobredosificación, propiedades farmacológicas, farmacodinámicas y farmacocinéticas, incompatibilidades, periodo de validez, conservación, instrucciones de uso, titular de la autorización y comercialización de cada medicamento.

- **Drug Information Database**

<http://www.familydoctor.org/cgi-bin/drugsearch.pl>

Esta base de datos contiene información sobre medicamentos, incluyendo uso adecuado y posibles efectos.

➤ **INVESTIGACIÓN**

- **Euroethics**

<http://www.spri.se/spriline/sokforie.htm>

Es una base de datos europea sobre temas de salud e investigación biomédica.

Los países participantes son: Francia, Alemania, Holanda y Suiza.

- **New Medicines in Development Database**

<http://www.phrma.org/searchcures/newmeds/webdb/>

Contiene información sobre productos farmacéuticos

➤ **SALUD Y SEGURIDAD**

- **Mining Safety and Health Research: Common Information Service System**

<http://outside.cdc.gov:8000/BASIS/ciss/pubs/pubs/SF>

Esta base de datos permite el acceso a la bibliografía de miles de publicaciones relacionadas con la seguridad y la salud. Ofrece un extracto de cada una de ellas y permite obtener el texto completo. Estas publicaciones pueden visitarse a través de Internet, pueden ser impresas o guardadas como referencia para ser vistas en un futuro.

INFORMACIÓN LEGAL Y CRIMINAL

➤ DIRECTORIOS DE ABOGADOS Y FIRMAS DE ABOGADOS

- **Kime's International Law Directory**

<http://www.smlawpub.co.uk/kimes/search.cfm>

Este directorio internacional elaborado anualmente ofrece información de calidad y actualizada de firmas de abogados y bufetes de todo el mundo. Cubre más de 250 países.

- **West Legal Directory**

<http://www.lawoffice.com/direct/direct.asp?form=name>

Ofrece perfiles de más de un millón de abogados y bufetes.

➤ **CRÍMENES Y CRIMINALES**

• **Interpol Most Wanted**

<http://www.interpol.int/Public/Wanted/Search/Form.asp>

Interpol 'Red Notices' es una base de datos que incluye un listado de las personas buscadas por la Interpol.

➤ **DECISIONES JUDICIALES**

• **Jurisprudencia del Tribunal Constitucional de España**

<http://www.boe.es/tc/>

Base de datos de sentencias y autos del Tribunal Constitucional, desde 1980.

• **Tribunal Internacional de Justicia**

<http://www.icj-cij.org/cijwww/cijhome.htm>

Base de datos de sentencias del Tribunal Internacional. En francés e inglés.
Elaborado por la Organización de las Naciones Unidas ONU.

• **United States Supreme Court Opinions**

<http://guide.lp.findlaw.com/cascode/supreme.html>

Esta base de datos incluye decisiones del Tribunal Supremo de Estados Unidos desde 1893.

➤ **FUENTES GENERALES SOBRE LEYES**

- **Sociología Jurídica**

<http://iisj.es/bdiisj/>

Base de datos de la biblioteca y centro de documentación del IISJ, especializada en sociología jurídica. Contiene referencias bibliográficas de monografías, artículos de revistas, literatura gris o materiales no publicados.

- **Child Abuse and Neglect Clearinghouse Organizations Database**

<http://www.calib.com/nccanch/scripts/SearchPg.cfm>

Ofrece un listado de 129 organizaciones estadounidenses centradas en la defensa del abuso infantil. Además ofrece una amplia información en esta materia.

- **HUDOC**

<http://hudoc.echr.coe.int/hudoc/default.asp?Language=en&Advanced=1>

Base de datos que incluye un listado de leyes de los organismos supervisores de la Convención Europea Derechos Humanos.

- **Oran's Law Dictionary**

<http://www.wld.com/conbus/orans/Welcome.asp>

Glosario completo con más de 6.000 definiciones legales escritas con lenguaje común y poco técnico para facilitar su comprensión.

➤ **PROPIEDAD INTELECTUAL**

- **ESP@CENET (European Patent Office) Patent Database**

<http://ep.espacenet.com/>

esp@cene es un servicio gratuito que ofrece la Organización Europea de patentes para sus Estados miembros: Austria, Bélgica, Chipre, Alemania, Finlandia, Francia, Alemania, Grecia, Irlanda, Italia, Liechtenstein, Luxemburgo, Mónaco, Portugal, España, Suecia, Suiza y Reino Unido.

- **U.S. Patent Databases (U.S. Patent and Trademark Office)**

<http://www.uspto.gov/>

La base de datos oficial de las patentes en EEUU ofrece numerosas opciones de búsqueda incluyendo el texto completo y bibliografía. Permite el acceso al texto completo de todas las patentes de EEUU realizadas desde el 1 de enero de 1976. Con anterioridad, desde 1790, pueden consultarse las páginas de las patentes a través de imágenes.

- **WIPO (World Intellectual Property Organization) Digital Library**

<http://ipdl.wipo.int/>

Este proyecto ha sido desarrollado para proveer información sobre la propiedad intelectual tanto a los gobiernos, como a usuarios individuales.

➤ LEYES Y TRATADOS

- **CALEX**

<http://www.senado.es/brsweb/CALEX/>

Elaborada por el Senado de España. Legislación, jurisprudencia, procesos y sentencias del Tribunal Constitucional referida a las Comunidades Autónomas. Contiene las disposiciones del Estado de rango superior (Estatutos de Autonomía, Reales Decretos de Traspasos, Leyes...), normas de rango inferior, normativa generada por las Comunidades Autónomas y por las Ciudades Autónomas de Ceuta y Melilla y una selección de normativas. También recoge la Jurisprudencia del Tribunal Constitucional sobre recursos y cuestiones de inconstitucionalidad, impugnaciones, recursos de amparo y conflictos de competencias relacionados con las Comunidades Autónomas.

- **CATA / OPAC Biblioteca del Senado**

<http://www.senado.es/cgi-bin/abweb/L1/T1/G0>

Catálogo de la Biblioteca del Senado especializada en temas parlamentarios, autonómicos, jurídicos y políticos de interés para el trabajo parlamentario. Se puede buscar mediante tres tipos de búsqueda: asistida, avanzada y experta. Cuenta con fondos históricos y modernos. Contiene libros, revistas, colecciones legislativas, incunables, libros del siglo XVI, manuscritos, obras musicales, láminas y grabados,

mapas, medallas, monografías posteriores al siglo XVI, publicaciones periódicas y prensa.

- **COPA**

<http://www.senado.es/brsweb/COPA/formulario.html>

Base de datos de referencias de la documentación del Consejo de Europa que se recibe en el Departamento de Documentación del Senado de España. Contiene convenios del Consejo de Europa, sentencias del Tribunal Europeo de Derechos Humanos, documentación de la Asamblea Parlamentaria, del Comité de Ministros y del Congreso de Poderes Locales y Regionales de Europa (CPLRE).

- **Annual Review of Population Law**

<http://cyber.law.harvard.edu/population/arpl.htm>

Esta base de datos contiene sumarios, legislación, constituciones y otros documentos oficiales de los gobiernos de cada país del mundo, en relación con asuntos populares y derechos de las mujeres.

- **CIS (Commonwealth Of Independent States) Migration Legislation Database**

<http://www.iom.int/migrationweb/documents/Legislation/default.htm>

La base de datos de la Organización Internacional para la Inmigración contiene legislación relacionada con la inmigración de los países pertenecientes a la Commonwealth.

- **Global Banking Law Database**

<http://www.gbld.org/>

Es un proyecto del Banco Mundial y la Fundación Monetaria Internacional. Esta base de datos ofrece una colección de leyes relacionadas con asuntos financieros, banca commercial, bancos centrales, depósitos, etc. Las leyes están disponibles en inglés en formato de Word y PDF.

- **Global Legal Information Network (GLIN)**

<http://lcweb2.loc.gov/law/GLINv1/GLIN.html>

Base de datos sobre leyes, regulaciones y otras fuentes legales complementarias. Los documentos aquí incluidos han sido proporcionados por los gobiernos de los países miembros. Los originales se encuentran en la Biblioteca del Congreso de los Estados Unidos.

- **National Criminal Justice Reference Service Full-Text Search**

<http://excalib1.aspensys.com/rware/login.htm>

Ofrece información sobre crímenes y justicia en Estados Unidos. Contiene además datos estadísticos.

- **NATLEX (Labor Law)**

<http://ilis.ilo.org/ilis/natlex/ilsearna.html#SearchFormE>

NATLEX es una base de datos bibliográfica que contiene información sobre leyes estadounidenses relacionadas con los derechos humanos, seguridad social y trabajo. Contiene además textos legales de más de 180 países, en más de 40 lenguas diferentes.

- **Refugee Caselaw Site**

<http://www.refugeecaselaw.org/>

Este sitio recoge textos de recientes decisiones judiciales e interpretaciones de definiciones legales de los refugiados. Contiene casos de: Australia, Austria, Canada, Alemania, Nueva Zelanda, Suiza, Reino Unido y Estados Unidos.

- **Women's Right To Maternity Protection Database**

<http://www.cdinet.com/womensrights/database.html>

Contiene textos de las leyes de protección de la maternidad de numerosos países de todo el mundo.

NOTICIAS Y EVENTOS ACTUALES

➤ **AUDIO.**

- **Búsqueda de Archivos de la Radio Nacional Pública (NPR).**

<http://www.npr.org/search/>

La radio nacional pública proporciona acceso a los archivos de audio de la mayoría de informativos de EE.UU. El interfaz sencillo de búsqueda permite buscar los resúmenes de texto de los programas y escuchar después el programa entero o un fragmento del mismo. El archivo de programas empezó en 1996.

- **Speechbot**

<http://speechbot.research.compaq.com/>

Se trata de un proyecto que crea una base de datos para buscar la programación de la radio usando la tecnología de reconocimiento de la voz.

➤ **DIRECTORIOS Y LOCALIZADORES.**

- **Directorio de Medios de Comunicación On –line.**

<http://emedia1.mediainfo.com/emedial/>

Localiza medios de comunicación on – line situados en todo el mundo realizando la búsqueda a través de diferentes opciones.

➤ **PERIÓDICOS.**

- **BizJournals.Com**

<http://bizjournals.bcentral.com/search.html>

Permite la búsqueda on – line de contenidos de más de 40 periódicos de negocios. Esta búsqueda no suministra el texto completo del periódico cada jornada.

- **Búsqueda de transcripciones de Bloomberg Televisión.**

http://www.bloomberg.tveyes.com/trans_search.asp

Bloomberg Televisión, un servicio de noticias de negocios, proporciona acceso a la transcripción de textos de sus programas. La búsqueda se realiza a través de diferentes campos.

- **Búsqueda de Noticias CNN.**

<http://207.25.71.29/SEARCH/>

Búsqueda en la gran biblioteca de material informativo de los canales CNN, CNNSI (deportes) y CNNFn (Negocios), que consta en sus páginas web.

- **Búsqueda en el “Herald Tribune Internacional”.**

<http://www.iht.com/advancedsearch.html>

Búsqueda de material publicado en este periódico. Este archivo se empezó a realizar en 1.996.

- **Herencia.com. Búsqueda de necrológicas en periódicos.**

<http://www.legacy.com/LegacySubPage1.asp?Page=ObitFinder>

Permite encontrar necrológicas recientes de más de 1.000 periódicos americanos y canadienses.

- **Archivo de Búsqueda de “The Times”.**

<http://www.thetimes.co.uk/section/0,,103,00.html>

Hay varias opciones de búsqueda para buscar publicaciones de este periódico desde enero de 1996.

➤ **VIDEO.**

- **Búsqueda de videos del canal de noticias estadounidense ABC.**

http://abcnews.go.com/sections/us/video_index/video_index.html

Este motor de búsqueda permite acceder a una base de datos de videos realizados por el canal ABC. El enlace de búsqueda está en el lado derecho de la página y se llama “Virage Video”.

- **Archivos de Vídeos de Científicos Americanos.**

<http://www.pbs.org/saf/archive.htm>

Búsqueda a través de un archivo expandido de material del programa PBS, siendo posible una visión global o parcial del resultado.

REGISTROS PÚBLICOS

➤ **REGISTROS PÚBLICOS.**

- **Registros de Tierras.**

<http://www.dnr.state.ak.us/cgi-bin/Iris/landrecords>

Este sitio ofrece acceso a DNR, que son índices históricos de terrenos propiedad del Estado y que fueron vendidos a particulares. La búsqueda se puede realizar a través de distintos campos, como el código postal.

- **Búsqueda de Licencias Ocupacionales.**

<http://www.dced.state.ak.us/occ/search3.htm>

Esta base de datos contiene más de 30.000 licencias ocupacionales emitidas por la División de la Ocupación para Licenciados de EE.UU.

- **Base de Datos de Organizaciones Caritativas.**

http://www.sosaz.com/scripts/Charity_search_engine.cgi/

La Oficina del Secretario Estatal de Organizaciones Caritativas proporciona información pública sobre estas organizaciones y su misión.

- **Verificación de Matrimonios y Divorcios.**

<http://www.quickinfo.net/madi/comadi.html>

Búsqueda de todos los matrimonios desde 1975 hasta Noviembre del 2000 y divorcios desde 1968 hasta Noviembre del 2000 en el Estado de Colorado de EE.UU.

- **Análisis en profundidad de Réditos.**

<http://www.ioc.state.il.us/iw/Expert/Rev/ERSummary.cfm>

Informes adjuntos para analizar el estado de réditos a través de datos específicos. Así mismo hay posibilidad de conocer más información a través de la página web del Interventor estatal.

- **Instalaciones autorizadas para el cuidado de niños.**

<http://www.state.in.us/fssa/database/homes.html>

Búsqueda de instalaciones autorizadas por la Administración de Servicios Sociales y Familiares de Indiana.

- **Base de Datos de Perfiles de Comunidad.**

<http://www.ecodev.state.mo.us/medms/comm.htm>

Obtiene datos demográficos básicos sobre todas las comunidades de EE.UU.

- **Base de Datos de Inspecciones de Restaurantes por el Departamento de Salud de Nueva York.**

<http://207.127.96.244/scripts/webfood.pl>

Búsqueda del nombre del restaurante o de su localización para encontrar las recientes infracciones de la ley de higiene en la ciudad de Nueva York.

INFORMACIÓN EN TIEMPO REAL

➤ **MEDIO AMBIENTE**

- **AirNow (Datos en tiempo real de la Polución del aire).**

<http://www.epa.gov/airnow/where/>

Varias fuentes están disponibles para analizar la calidad del aire en tiempo real según el sitio de residencia en EE.UU.

- **Datos en tiempo real del flujo de las corrientes de vapor de agua.**

<http://water.usgs.gov/realtime.html>

El Instituto Geológico de EE.UU. proporciona un programa que suministra datos sobre las corrientes de vapor de agua en tiempo real para predecir inundaciones, detección de cambios en la corriente o estudiar el calentamiento global del planeta.

- **Centro de datos de boyas nacionales para recoger datos marinos.**

<http://seaboard.ndbc.noaa.gov/Maps/rmd.shtml>

Datos en tiempo real del estado de la mar a través de datos recogidos en boyas localizadas en aguas de EE.UU., Canadá y Reino Unido.

- **Mareas on-line.**

<http://tidesonline.nos.noaa.gov/>

Ofrece información en tiempo real sobre mareas, nivel del agua y tormentas en alta mar.

- **Atlas of the Oceans**
<http://www.oceansatlas.com/>

El Atlas proporciona información relevante sobre desarrollo sostenido en los océanos, organizada en cuatro secciones principales: usos (gestión de los desechos procedentes de la tierra, energía, pesca, acuicultura, biotecnología marina...), temas (cambios climáticos, economía, emergencias, seguridad alimenticia...), información de base (biología, ecología, formación y evolución de los océanos..) e información geográfica (información clasificada por áreas geográficas). En inglés.

➤ **GOBIERNO**

- **Actas de Pisos de EE.UU.**

<http://clerkweb.house.gov/floor/current.htm>

La Oficina del empleado de Registros de Casas de EE.UU. proporciona esta base de datos para encontrar en tiempo real la legislación pertinente sobre las viviendas en EE.UU.

- **Llamadas en tiempo real al Servicio de Emergencia, Departamento de Bomberos de Seattle.**

<http://www2.cityofseattle.net/fire/GetDatePubTab.asp>

Acceso en tiempo real a las llamadas al Servicio de Emergencia suministradas por el Departamento de Bomberos de Seattle.

➤ **MISCELÁNEA**

- **Rastreo de Paquetes de Programa.**

Es una base de datos de programas usadas por 38 usuarios.

- **PublicRadioFan.Com**

<http://www.publicradiofan.com/>

Base de datos de listados de programas de cientos de emisoras de radio públicas en el mundo entero.

➤ **ESPACIO**

- **Vista de la Tierra.**

<http://www.fourmilab.ch/cgi-bin/uncgi/Earth/action?opt=-p>

Esta asombrosa herramienta permite jugar con vistas en tiempo real de la Tierra, el Sol y la Luna, ver el lado nocturno del planeta, cubrirlo de nubes o ampliar el zoom.

- **Localizador de satélites J-Track.**

<http://liftoff.msfc.nasa.gov/realtime/JTrack/Spacecraft.html>

El sistema de rastreo permite calcular de forma aproximada la posición de los satélites incluidos en una lista. Cada satélite tiene asignado un color según el planeta al que pertenezca.

- **Datos en tiempo real de la NASA.**

<http://spaceflight.nasa.gov/realdata/index.html>

Varias bases de datos incluyen fuentes que proporcionan información en tiempo real del Tránsito y de la Estación Internacional Espacial.

- **Atlas de Imágenes Planetarias PDS.**

<http://www-pdsimage.wr.usgs.gov/ATLAS.html>

El atlas de Imágenes Planetarias PDS usa imágenes producidas por el software MapMaker que genera imágenes de gran calidad de los distintos planetas. Se puede elegir los límites de latitud y longitud del mapa, así como su escala.

➤ **TRANSPORTE**

- **Información sobre llegadas de vuelos en el aeropuerto de Heathrow del Reino Unido.**

http://www.baa.co.uk/main/airports/heathrow/flight_arrivals_frame.html

Base de datos constantemente actualizada con información de la llegadas de los vuelos. Además existen datos sobre llegadas y salidas de otros aeropuertos.

- **Localizador de vuelos.**

<http://www.trip.com/ft/home/0,2096,1-1,00.shtml>

Se puede encontrar el estado actual de cualquier vuelo de líneas aéreas de EE.UU. o internacionales, sabiendo además la situación del avión a través de una aplicación Java.

- **Estado de los Aeropuertos de EE.UU. en tiempo real.**

<http://www.fly.faa.gov/>

Se puede ver el estado sobre los retrasos en los aeropuertos estadounidenses en tiempo real.

- **Información sobre Incidencias de Tráfico de la Patrulla de Carreteras de California.**

[http:// cad.chp.ca.gov/](http://cad.chp.ca.gov/)

Se trata de un display en tiempo real de la actividad de la Policía de Carretera de California.

- **Información sobre llegada de trenes.**

<http://reservations.amtrak.com/novus/process-form?home&arrival>

Localiza el tiempo de llegada de cualquier estación o cualquier tren en estaciones de EE.UU.

➤ **METEOROLOGÍA**

- **Condiciones del Tiempo atmosférico Internacional.**

<http://weather.noaa.gov/pd/usframes.html>

Se pueden estudiar las condiciones del tiempo atmosférico en cientos de localidades de EE.UU.

- **Escuela Mundial WeatherNet.**

<http://aws.com/globalwx.html>

Observaciones en tiempo real del estado de la atmósfera desde distintas estaciones localizadas alrededor de EE.UU. y Canadá.

- **Biblioteca Técnica de la OMM**
<http://library.wmo.ch/>

Catálogo de la Biblioteca Técnica de la OMM sobre investigaciones en ciencias meteorológicas y atmosféricas, oceanografía, medio ambiente, hidrología, geofísica y otras disciplinas conexas.

REFERENCIAS

➤ ASOCIACIONES

- **Búsqueda de Inspectores de la Sociedad Americana de Inspectores de Casas (ASHI).**

<http://www.ashi.com/fi.cfm>

Se buscan miembros dentro de un cierto radio a partir del Código Postal. ASHI es la Sociedad de Inspectores del Hogar más grande y con los profesionales más respetados en toda Norteamérica con cerca de 5.500 miembros y candidatos.

- **QueerAmerica**

<http://www.queeramerica.com/>

QueerAmerica es una base de datos publicada por OutProud, la Coalición Nacional de Gays, Lesbianas y Bisexuales. Con más de 4.000 entradas es la colección más grande de documentos de este colectivo en todo EE.UU. Incluye información de centros comunitarios, organizaciones y asociaciones relacionadas como PFLAG (Padres, Familiares y Amigos de Lesbianas y Gays) y más. Es un fantástico enlace para hacer amigos, encontrar respuestas o buscar información.

- **El Proyecto de Sociedades de Estudiantes.**

<http://www.lib.uwaterloo.ca/cgi-bin/public/opentext/pat2db?specfile=/fsys2/opentext/ssp/full/ssp.p2w>

Contiene una base de datos de búsqueda y varios directorios. Hay más de 1.600 sociedades de todo el mundo.

➤ **PREMIOS.**

- **Base de Datos de los Premios Oscar de la Academia.**

http://www.oscars.org/awards_db/index.html

La base de datos contiene información concerniente con los Premios Oscar de la Academia desde 1927 hasta ahora. Una base de datos separada está disponible para premios técnicos.

- **Base de datos de Premios de la Literatura Infantil.**

<http://www2.wcoil.com/~ellerbee/childlit.html>

Se pueden crear listas de lectura personalizadas usando como entradas numerosos ganadores de premios de la literatura infantil.

- **Base de datos de Ganadores Olímpicos.**

<http://www.britannica.com/olympics/record/>

Una base de datos de los ganadores de medallas olímpicas desde 1896 hasta el presente.

- **Base de datos del Premio Pulitzer.**

<http://www.pulitzer.org/Archive/Search/search.html>

La página web del Premio Pulitzer incluye una base de datos de los ganadores y los finalistas nominados para cada categoría. Además de los ganadores desde 1995 hasta ahora, hay una completa colección de fotografías, dibujos y artículos ganadores.

➤ **LIBROS.**

- **Directorio de Miembros de la Asociación de Libreros Americanos.**

http://www.bookweb.org/bd-bin/search_bd

Información de más de 4500 libreros independientes de EE.UU.

- **Calendario de Eventos y firmas del Autor de Libros.**

<http://www.bookbrowser.com/Storytellers/Calendar.html>

Localiza firmas de autor, apariciones en televisión, lecturas de la obra de un autor y más.

- **Crónicas de libros del New York Times.**

<http://search.nytimes.com/books/search/>

Permite encontrara información de la sección de libros del New York Times desde 1997, con un acceso a los textos completos desde 1980. Es necesario un registro gratuito para leer esta documentación.

- **Lista de Bestsellers semanal del Editor.**

<http://www.bookwire.com/bookinfo/searchformPW.html>

Se pueden buscar bestsellers recomendados por editores y encuadernados en tapa dura desde 1991 hasta la actualidad.

➤ **CÁLCULOS.**

- **Centro On-line de calculadoras Martindale's.**

<http://www-sci.lib.uci.edu/HSG/RefCalculators.html>

La Web invisible es el hogar de muchos calculadoras interactivas. Aunque este directorio no está en la Web invisible, los datos que pueden ser encontrados sí lo están. Esta página tiene una compilación de enlaces a más de 12.000 cálculos.

- **Economía del consumo de fuel del Automóvil)departamento de energía de EE.UU.).**

<http://www.fueleconomy.gov/feg/findacar.htm>

Permite encontrar las millas por galón de consumo para coches y camiones. Los datos están disponibles desde 1985 hasta el 2001.

- **Convertidor de dinero universal.**

<http://www.xe.net/ucc/full.shtml>

El Convertidor de dinero universal permite obtener la relación de cambio de divisas extranjeras a través de Internet. Hay más de 180 monedas disponibles.

➤ **FUENTES DE CONSUMO.**

• **NADAGuides.Com**

<http://www.nadaguides.com/>

Se trata de una página de suministro de precios para la adquisición o alquiler de diferentes vehículos. Así hay precios de barcos, camiones, caravanas, motocicletas, motos de nieve y especial atención a vehículos clásicos y de coleccionistas.

• **Base de datos de automóviles de EE.UU.**

http://www.nhtsa.dot.gov/cars/problems/recalls/recall_links.html

Se puede buscar información del parque automovilístico proporcionada por la Administración de Seguridad en el Tráfico de Grandes Autopistas.

• **Base de datos de suscriptores de Laboratorios de Certificación.**

<http://www.ul.com/database/>

Se tiene acceso a la información más actual de más de 110.000 productos certificados y 4.000 empresas registradas mediante el nombre de la compañía, del producto o su localización geográfica.

➤ **DICCIONARIOS, GLOSARIOS Y TRADUCTORES.**

- **Buscador de Siglas.**

<http://www.acronymfinder.com/>

La base de datos más completa en una página web de búsqueda de acrónimos, abreviaturas e iniciales, con más de 180.000 entradas.

- **Diccionario de Cambridge on-line.**

<http://dictionary.cambridge.org/>

En este diccionario la búsqueda de la palabra se realiza en cuatro publicaciones distintas: Diccionario Cambridge de Inglés Internacional, Diccionario Cambridge de Inglés Americano, Diccionario Cambridge Internacional de Idiomas y Diccionario Cambridge Internacional de Verbos con Partícula (Phrasal Verbs).

- **Glosario de Términos de Comunicación.**

<http://www.getcommstuff.com/glossary/>

Este glosario está realizado por la Asociación de Industrias de Telecomunicaciones y consta de más de 5.800 entradas.

- **Verbix**

<http://www.verbix.com/languages/index.html>

Se puede obtener la conjugación on-line de verbos en más de 50 idiomas.

➤ **ALIMENTACIÓN.**

- **Codex**

Alimentarius

http://www.codexalimentarius.net/standards_search_es.asp

Normas alimentarias, reglamentos y otros textos relacionados tales como códigos de prácticas bajo el Programa Conjunto FAO/OMS de Normas Alimentarias. Las materias principales de este Programa son la protección de la salud de los consumidores, asegurar unas prácticas de comercio claras y promocionar la coordinación de todas las normas alimentarias acordadas por las organizaciones gubernamentales y no gubernamentales.

- **Lenguaje de la Industria Alimenticia: Glosario de Términos de Supermercados.**

http://www.fmi.org/facts_figs/glossary_search.cfm

El Instituto de Marketing en Alimentación proporciona la posibilidad de búsqueda gratuita de términos en esta página que contiene más de 20.000 datos indexados.

- **Cosecha Local.**

<http://www.localharvest.org/>

Es una base de datos de granjas y otros servicios a lo largo de EE.UU. donde se pueden adquirir productos directamente de la zona de cultivo.

- **Archivo de vinos (Winefiles.Org)**

<http://www.winefiles.org/>

Información sobre vino, fabricación del vino y crecimiento de uva, incluso del negocio, tecnología e historia asociada con el vino... La cobertura es mundial con un énfasis especial sobre California.

- **Listado de Restaurantes Zagat.**

<http://www.zagat.com/index.asp>

Proporciona acceso a una lista de restaurantes de 45 ciudades y regiones.

➤ **FUENTES DE NEGOCIO GENERAL.**

- **Búsqueda de nombres del Memorial Quilt sobre SIDA.**

<http://www.aidsquilt.org/Newsite/searchquilt.htm>

Los nombres que forman esta base de datos son aquellos individuos inmortalizados en el Memorial Quilt sobre SIDA. En total hay más de 83.000 nombres.

- **Búsqueda de la Clasificación Industrial Standard (SIC).**

<http://www.osha.gov/oshstats/sicser.html>

Esta página permite al usuario buscar desde 1987 el manual SIC para obtener la información de descripción de las cuatro cifras asociadas al código SIC de la empresa.

- **Base de datos de las directrices de un Editor.**

<http://www.writersdigest.com/guidelines/index.htm>

Proporciona más de 1.500 directrices para la preparación de un libro suministradas directamente por el editor.

➤ **PUBLICACIONES Y PERIÓDICOS.**

- **Jake**

<http://jake.med.yale.edu/>

Jake es una referencia de búsqueda la cual permite encontrar enlaces on-line de artículos para estudiantes, investigadores y librerías. Se trata de una lista de enlaces y herramientas para la búsqueda de documentación.

- **Publist**

<http://www.publist.com/search.html>

Esta base de datos proporciona información de las publicaciones de más de 150.000 revistas, periódicos, cartas y otras publicaciones.

- **Noticias**

<http://www.imsersomayores.csic.es/basisbwdocs/noticias.htm><http://www.metabase.net/>

Mayores. Una iniciativa del IMSERSO y del CSIC

Actualidad en prensa sobre el campo de la vejez, mayores y tercera edad.

También se puede buscar en la base de datos histórica de noticias.

➤ **BIBLIOTECA ON-LINE.**

- **Base de datos de textos completos DataStar.**

<http://library.dialog.com/>

Se trata de una herramienta de recuperación de información de más de 30.000 fuentes.

- **Las Páginas Amarillas del Bibliotecario.**

<http://www.librariansyellowpages.com/>

Base de datos constantemente actualizada que tiene productos y servicios de más de 2000 bibliotecarios y libreros con una lista de enlaces accesibles.

- **Noodlebib (Creador de Bibliografía).**

<http://www.noodletools.com/noodlebib/index.html>

NoodleBib es una aplicación que permite crear y editar una bibliografía propia on-line.

- **Internet Archive: wayback machine**

URL:<http://www.archive.org/>

Archivo digital de sitios web que han desaparecido, remontándose al año 1996. En inglés.

- **Digital Library Documents**

<http://www.hti.umich.edu/cgi/b/bib/bib-idx?c=dlf>

Base de datos de páginas web sobre bibliotecas digitales con información sobre políticas, estrategias, informes, normas, directrices y documentación técnica, desarrollada por los miembros de la Federación. En inglés.

- **Audiovisual Library**

http://europa.eu.int/comm/mediatheque/index_en.html

Biblioteca de materiales multimedia relacionados con la creación, desarrollo, historia y

personajes destacados de la Unión Europea. Contiene sonido, fotografías, videos, mapas, discursos, audiovisuales, etc... Se puede buscar por palabra clave, materia, personalidad, lugar y fecha. En inglés y francés.

➤ **LOCALIZADORES.**

- **Guía sobre instalaciones militares de EE.UU. en el mundo.**

<http://www.militarycity.com/moves/baseguide.html>

Las publicaciones periódicas de la Armada proporcionan el directorio de las instalaciones militares de EE.UU. en su territorio y fuera de él.

- **Localizador de Museos de EE.UU.**

<http://seeing2020.com/museums/lookup.htm>

Permite encontrar la información de cerca de 7.500 museos a lo largo de EE.UU.

- **Localizador de sedes de la Cruz Roja.**

<http://www.redcross.org/where/where.html>

Localiza las sedes locales de la Cruz Roja en EE.UU. La búsqueda se realiza por Código Postal.

- **Localizador de Oficinas de Correo de EE.UU.**

<http://www.framed.usps.com/ncsc/locators/find-po.html>

En esta base de datos se encuentran las direcciones de las oficinas postales que entregan el correo en tu domicilio, así como de otras oficinas localizadas en la zona.

- **Base de datos de Códigos Postales de EE.UU.**

http://www.usps.gov/ncsc/lookups/lookup_zip+4.html

Permite encontrar el Código Postal de cualquier dirección de EE.UU.

- **Localizador de cajeros automáticos de tarjetas Visa.**

<http://www.visa.com/pd/atm/main.html>

Localiza cajeros automáticos en muchos países.

➤ **MAPAS Y GEOGRAFÍA.**

- **¿Cuánto está de lejos?**

<http://www.indo.com/distance/>

Este servicio utiliza datos del Censo de EE.UU. y una lista suplementaria de ciudades alrededor del mundo para encontrar la latitud y longitud de dos lugares, y entonces calcular la distancia entre ellos, en línea recta.

- **Servidor de trazados de mapas Maptech.**

<http://navigator2.maptech.com/homepage/index.cfm>

Es el enlace más grande para el trazado de mapas incluyendo representaciones topográficas, cartas terrestres, marinas y aéreas. Gratuitamente se pueden ver, imprimir y mandar por e-mail estos mapas.

- **Base de datos estadounidense de construcción de carreteras RandMcNally.**

<http://www.randmcnally.com/rmc/tools/roadConstructionSearch.jsp>

Se trata de una base de datos para evitar incidencias en los viajes y evitar posibles retrasos. Se encuentran los proyectos de construcción de carreteras de EE.UU. y Canadá que afecten al viaje del usuario. Se actualiza dos veces al mes.

➤ **DEPORTES.**

- **Base de datos de tensitas de la Federación Internacional de Tenis.**

<http://onlinesql.itftennis.com/select.htm>

Remontándose sus orígenes a 1977 la base de datos de la Federación Internacional de Tenis (ITF) cubre más de 25 años de datos de todos los jugadores de tenis masculinos y femeninos que han competido en un torneo del Grand Slam en ese tiempo.

- **Búsqueda de jugadores de la Liga Mayor de Béisbol.**

<http://www.totalbaseball.com/stats/mlbstat.html>

Permite encontrar información bibliográfica y estadística básica de los mejores y más actuales jugadores de las Ligas Mayores de Béisbol.

- **Estadísticas NCAA.**

<http://www.ncaa.org/stats/>

Estadística actualizada del deporte universitario de EE.UU. Dichos datos para algunos deportes como baloncesto y béisbol, están en bases de datos de la Web Invisible. Sin embargo, debido a la naturaleza de esta de información puede echarse en falta una actualización más regular.

➤ **VIAJES.**

- **Directorio de Códigos de Líneas Aéreas.**

<http://www.iata.org/codes/index.asp>

Se pueden obtener los códigos oficiales de más de 9.000 localizaciones en el mundo entero.

- **Horario Amtrak.**

<http://reservations.amtrak.com/>

Lista de horarios, tarifas, compra de tickets y reservas previas en distintos medios de viaje.

- **Hotelguide.com.**

<http://www.hotelguide.com/search.cfm>

Permite encontrar alojamiento en cualquier parte del mundo. Se trata del directorio de hoteles más grande de la red.

- **Viajes simples realizados.**

http://www.transportforlondon.gov.uk/jp_index.html

Se puede planear el itinerario a seguir en el Sistema de Transporte de Londres.

- **Buscador de eventos.**

http://gorptools.gorp.com/GORPApps/trails/search_form.asp

Un directorio interactivo que permite localizar eventos, como carreras, competiciones ciclistas, o acontecimientos invernales en Columbia, así como ofrece la posibilidad de información de cómo llegar.

- **WhatsgoingOn.Com**

<http://www.whatsgoingon.com/>

Se trata de una herramienta para viajar a los sitios más interesantes, extravagantes y rituales en todo el planeta.

➤ **METEOROLOGÍA**

- **Base de datos sobre tiempo atmosférico.**

<http://www.weatherbase.com/>

Se trata de una base de datos de búsqueda que permite encontrar datos mensuales de tiempo atmosférico así como un historial del mismo en más de 10.200 ciudades del mundo.

- **Predicción de tiempo atmosférico.**

<http://www.weatherplanner.com/wp/start.jsp>

Una herramienta para predecir el tiempo atmosférico en EE.UU. con un año de adelanto.

CIENCIA

➤ **MEDIO AMBIENTE.**

- **Base de Datos de Tratados Medioambientales ENTRI.**

<http://sedac.ciesin.org/entri/>

Un completo servicio de búsqueda on-line para encontrar información de tratados medioambientales. El sistema ENTRI es único en permitir un sistema de búsqueda a través de diferentes preguntas.

- **Envirofacts (Ministerio de Medio Ambiente)**

http://www.epa.gov/enviro/index_java.html

Un enlace para obtener datos medioambientales del Ministerio de EE.UU. Este sitio Web proporciona acceso a varias bases de datos del Ministerio para acceder a información medioambiental y actividades relacionadas con el aire, agua y tierra en cualquier lugar de EE.UU.

- **GeoMAC (Lucha contra el fuego)**

<http://wildfire.usgs.gov/website/fireinfo/viewer.htm?Title=Wildfire%20Information>

GeoMAC es una aplicación de Internet diseñada para la lucha contra el fuego y la coordinación de los servicio de extinción. Además permite un acceso on-line a estos equipos a mapas de localización de incendios.

- **Directorio Nacional de Medioambiente.**

Un directorio de más de 13.000 organizaciones estadounidenses relacionadas con cuestiones del medio ambiente y con la educación ambiental.

- **Base de datos de pesticidas (Red de acción contra pesticidas)**

Un enlace para encontrar información actual sobre la toxicidad y la regulación de pesticidas.

- **Base de datos de los lagos del mundo.**

<http://www.ilec.or.jp/database/database.html>

Una base de datos de más de 500 lagos de más de 73 países.

- **DIRMEDA Directorio de empresas de servicios ambientales**

<http://www.impiva.es/Medioambiente/impivaMA.nsf>

Directorio de empresas de servicios, entidades sin ánimo de lucro que operan en el sector medioambiental en la Comunidad Valenciana y que realizan actividades y servicios de consultoría medioambiental, ingeniería ambiental, gestión de residuos, laboratorios de análisis y Organismos de Control Autorizados OCA. También contiene servicios complementarios de instituciones de la administración autonómica, estatal y europea, ferias y congresos, publicaciones y centros de I+D y formación. Se puede buscar alfabéticamente o por servicio ofertado, municipio y texto libre.

- **Normativa Ambiental**

http://www.cma.gva.es:8081/pls/cidam/pkt_normativa.showform

Buscador legislativo de normativa ambiental. Se puede buscar por rango, publicación, ámbito (DOCE, BOE DOGV, nacional e internacional), fecha publicación o aprobación, título y texto.

➤ **AGRICULTURA.**

- **AGRICOLA**

<http://www.nal.usda.gov/ag98/ag98.html>

AGRICOLA (AGRICultural Acceso On-Line) es una base de datos bibliográfica de documentos de la literatura agrícola creada por la Biblioteca Nacional de Agricultura y sus cooperantes. La producción de este material en forma electrónica comenzó en 1970, pero la base de datos cubre material de todas las épocas incluyendo documentos del siglo XV.

- **Colección de bases de datos Estadísticas FAOSTAT.**

<http://apps.fao.org/page/collections>

FAOSTAT es una base de datos on-line y multilingüe que contiene más de un millón de documentos que abarcan estadísticas internacionales de las siguientes áreas : Producción, Comercio, Alimentación, Fertilizantes y Pesticidas, Uso de la Tierra, Maquinaria Agrícola, etcétera.

- **CIAT** **Catálogo** **General**

<http://librum.ciat.cgiar.org:4505/ALEPH/SESSION-32145/start/cia01><http://www.eblul.org/linksearch.asp>

Referencias bibliográficas de monografías, revistas y documentos especializados sobre yuca, frijol, especies forrajeras tropicales, arroz, suelos, biotecnología, recursos naturales y biodiversidad disponibles en la Biblioteca del CIAT.

- **Servicio de información y documentación forestal - SEIDAL**

<http://www.colciencias.gov.co/conif/conif1a4.htm>

Material bibliográfico sobre forestería y agroforestería, con especial énfasis en el Trópico y Colombia. Se ofrecen dos bases de datos en línea: REFO que contiene referencias de artículos de revistas forestales y DOCFOR que contiene referencias de documentos forestales.

- **Biblioteca** **David** **Lubin**
http://www.fao.org/library/_Info_ServicesES/indexES.htm

Biblioteca de la FAO especializada en temas relacionados con la alimentación, la agricultura y el desarrollo internacional.

- **IPE** **Documentación**
http://www.ipe.es/documentacion/lista_documentacion.asp
<http://www.uv.es/%7Edocmed/documed/documed.html>

Fondo documental del IPE especializado en temas forestales, celulosa, papel, cartón y transformados de papel. Contiene: artículos, trabajos publicados en revistas especializadas, monografías, anuarios, patentes, normas, disposiciones legales e informes técnicos.

➤ **BIOLOGÍA.**

- **Guía de Campo E-Nature.**

http://www.enature.com/guides/select_group.asp

Es una completa guía de más de 4.800 especies de plantas y animales de EE.UU.

- **GeneCards.**

<http://bioinformatics.weizmann.ac.il/cards/>

GeneCards es una base de datos de genes humanos, su investigación y su influencia en enfermedades. Ofrece concisa información sobre las funciones de los genes humanos ya identificados así como un listado de ellos.

- **ISIS (Sistema Internacional de Información de Especies).**

<http://www.worldzoo.org/abstract/abstract.htm>

ISIS es una organización internacional no lucrativa que trabaja con casi 550 instituciones zoológicas de 54 países en todo el mundo.

- **Directorio de Zoológicos y Acuarios. (Asociación Americana de Zoo y Acuario).**

<http://www.aza.org/members/zoo/>

Permite la búsqueda de los miembros de esta asociación por nombre o Estado. Muchas entradas contienen una breve descripción del zoo o acuario mediante un enlace a su página web.

➤ **BOTÁNICA.**

- **Base de datos de la literatura del cacahuete.**

<http://nspal.cpes.peachnet.edu/home/pnutdb/>

Una colección de artículos y referencias sobre el mundo del cacahuete.

- **Base de datos de plantas.**

<http://plants.usda.gov/>

Desarrollado por el Departamento de Agricultura de EE.UU. es una excelente fuente de información de plantas, incluyendo sus características más importantes.

- **REFORGEN (Silvicultura).**

<http://www.fao.org/montes/for/form/FOGENRES/reforgen/>

Es una página del Sistema de Información Global de la FAO sobre recursos forestales. La información disponible puede dividirse en dos grandes campos: (i) datos del estado de las especies forestales de un país dado, (ii) datos de instituciones de recursos forestales de un país dado.

➤ **QUÍMICA.**

- **Resúmenes Beilstein.**

<http://www.chemweb.com/databases/beilstein>

Los usuarios son capaces de tener acceso a títulos, resúmenes y citas de publicaciones relacionadas con la química orgánica, publicados desde 1980 hasta el presente. Actualmente hay cerca de 600.000 artículos en esta base de datos, requiriéndose un registro gratuito para acceder a la misma.

- **Chemfinder.Com**

<http://www.chemweb.com/databases/beilstein>

Localiza las propiedades de muchas sustancia químicas.

- **Base de datos de sustancias solventes.**

<http://solvdb.ncms.org/solvdb.htm>

El Centro Nacional de Ciencias presenta esta base de datos donde se pueden obtener datos de este tipo de sustancias.

- **THERMODEX (Datos Termoquímicos).**

<http://thermodex.lib.utexas.edu/>

ThermoDex contiene una base de datos termoquímicos y termo físicos de varias sustancias químicas. Se puede seleccionar uno o más tipos de componentes y unirlos según una o más propiedades térmicas, y ThermoDex proporcionará una lista de sustancias que cumplen esos datos.

➤ **CIENCIA TERRESTRES.**

- **Resúmenes de Ingeniería de Terremotos.**

<http://www.eerc.berkeley.edu/cgi-bin/texhtml?form=eea>

Esta base de datos abarca la literatura mundial sobre terremotos desde 1971. Los contenidos incluyen informes técnicos, conferencias de prensa, monográficos y artículos periodísticos.

- **Búsqueda de terremotos.**

<http://wwwneic.cr.usgs.gov/neis/epic/epic.html>

Proporcionada por el Centro Nacional de Información Sobre Terremotos. Se tiene acceso a información en este campo a lo largo de la historia a través de esta base de datos.

- **Servidor de la Tierra.**

<http://www.terraserver.com/advfind.asp>

Una fuente de imágenes sobre la Tierra, con material de más de 60 países. Este servidor surgió a través de un proyecto conjunto entre las compañías Aerial Images, Microsoft, USGS, y Compaq.

- **Base de Datos de Volcanes en el Mundo.**

http://www.ngdc.noaa.gov/seg/hazard/vol_srch.shtml

Esta base de datos contiene una lista de más de 4.300 erupciones volcánicas desde 1979 hasta 1997.

➤ **ENERGÍA.**

- **Base de Datos Mensual de Energía.**

<http://tonto.eia.doe.gov/mer/>

Se trata de una base de datos interactiva de estadísticas sobre energía que comenzó en 1973.

- **Base de Datos de Explosiones nucleares.**

http://www.ausseis.gov.au/information/structure/isd/database/nukexp_query.htm

l

AGSO (Organización Australiana de Geología) mantiene una base de datos de explosiones nucleares con el lugar, tiempo y potencia de las explosiones nucleares a lo largo del mundo desde 1945.

- **Base de Datos de Centrales Nucleares.**

<http://www.insc.anl.gov/plants/index.html>

Se trata de una base de datos de centrales nucleares repartidas en todo el mundo con información básica y datos de operación de las mismas.

- **Centro de Investigación de Recursos y Consumos Energéticos CIRCE**

<http://circe.cps.unizar.es/spanish/texcast/publi/publi.html>

Base de datos de publicaciones científicas del Centro de Investigación de Recursos y Consumos Energéticos. Se puede examinar por título o por autor. Los contenidos están almacenados en formato PDF.

➤ **INGENIERÍA.**

- **Base de Datos de Ingeniería Civil (Sociedad Americana de Ingenieros Civiles ASCE).**

<http://www.pubs.asce.org/cedbsrch.html>

La base de datos de ingenieros civiles (CEDB) está diseñada para proporcionar fácil acceso bibliográfico a todas las publicaciones de ASCE. La base de datos

abarca documentos desde 1973 y proporciona acceso a todos los artículos, libros, conferencias de prensa, revistas, cartas y revistas sobre Ingeniería Civil, con más de 80.000 entradas.

- **Base de datos de Seguridad en la Información.**

<http://msds.pdc.cornell.edu/msdssrch.asp>

Una base de datos de seguridad de cerca de 250.000 entradas recopiladas de varias fuentes incluyendo la base de datos del Ministerio de Defensa de EE.UU.

- **Base de Datos Mundial de Presas.**

<http://genepi.louis-jean.com/cgi-bin/cigb-registre.pl?language=en>

Este enlace es una base de datos de Presas realizadas en todo el mundo proporcionado por el Consejo Mundial de Grandes Presas (ICOLD).

➤ **FUENTES DE NEGOCIO GENERAL.**

- **Red PrePrint.**

<http://www.osti.gov/preprint/ppnsearch.html>

La Red PrePrint del Departamento de Energía es una base de datos de servidores de información científica incluyendo campos como física, química, biología y medicina nuclear.

- **PubScience**

<http://pubsci.osti.gov/srchfrm.html>

PubSCIENCE proporciona a los usuarios la posibilidad de buscar información a lo largo de un largo compendio de literatura centrada en Ciencias Físicas y otras ciencias relacionadas con el Departamento de Energía de EE.UU (DOE).

- **sciBase**

<http://www.thescientificworld.com/scibase/search.asp>

sciBASE da acceso libre a una base de datos de literatura científica, técnica y médica. Actualmente incluye cerca de 12 millones de artículos publicados desde 1993 en más de 20.000 publicaciones, y es actualizada diariamente con aproximadamente 7.000 artículos nuevos.

➤ **MATEMÁTICAS Y FÍSICA.**

- **Archivo arXiv.org.**

<http://xxx.lanl.gov/form>

Desde agosto de 1991, arXiv.org (antes xxx.lanl.gov) es un archivo electrónico completamente automatizado que funciona como servidor de documentos de investigación. Abarca áreas como Física, Matemáticas, Lenguaje Computacional, etcétera.

- **Enciclopedia Sloane´s On-Line de Secuencias de Números Enteros.**

Se usa para buscar secuencias de números enteros.

➤ **OCEANOGRAFÍA.**

- **Calidad del Agua de baño en la Unión Europea: Atlas Turístico.**

<http://europa.eu.int/water/cgi-bin/bw.pl>

El Atlas Turístico proporciona un acceso para encontrar un resultado de la calidad del agua en un país determinado o en el lugar elegido para las vacaciones.

- **Base de datos REF. (Datos de especies marinas).**

<http://www.reef.org/data/database.htm>

Desde la base de datos se puede generar una variedad de informes de especies marinas, así como su distribución y abundancia.

- **Predictor de Mareas y Corrientes.**

After selection of a region this database predicts local tide times for locations worldwide. Chart and graph tools are also available.

- **Base de datos de Tsunami.**

<http://www.ngdc.noaa.gov/seg/hazard/tsu.shtml>

Esta base de datos proporciona información sobre tsunamis desde el año 49 D.C. hasta el presente en el Mar Mediterráneo, Mar del Caribe y océanos Atlántico, Índico y Pacífico.

➤ **INVESTIGACIÓN Y DESARROLLO.**

• **PortalTecnociencia**

<http://www.portaltecnociencia.es><http://biblio2.colmex.mx/bibdig/>

Sistema de información que recopila, estructura y difunde recursos de calidad sobre ciencia y tecnología. Centrado principalmente en España recoge también recursos europeos e internacionales. Estructurado a partir de la clasificación ISI contiene información sobre: administraciones públicas, empresas, organizaciones internacionales, congresos, asociaciones, centros tecnológicos, documentos, instrucciones académicas, universidades, listas de distribución, museos, noticias, OTRIS, proyectos, recursos web, revistas... Se puede buscar por texto libre, canal, zona geográfica, sección y fecha.

• **Resúmenes del Proyecto Federal R&D.**

<http://fedrnd.osti.gov/>

Estos resúmenes proporcionan un portal de información sobre proyectos federales de investigación a través de una búsqueda interactiva en bases de datos de diferentes agencias.

• **Directorio Internacional de Laboratorio de Pruebas (ASTM).**

<http://astm.365media.com/astm/labs/>

El Directorio Internacional de Laboratorio de Pruebas de la ASTM [Sociedad Americana de Pruebas y Materiales] es un enlace de búsqueda on-line de los servicios y localizaciones de laboratorio de pruebas.

- **Servidor de Informes Técnicos de la NASA.**

<http://techreports.larc.nasa.gov/cgi-bin/NTRS>

El Servidor de Informes Técnicos de La NASA es un servicio experimental que permite al usuario buscar informes técnicos de la NASA de diferentes materias realizados por distintos centros y programas de la NASA.

- **Investigación Científica en el Parque Nacional de Yellowstone.**

<http://www.wsulibs.wsu.edu/ris/risweb.isa>

Una base de datos bibliográfica de cerca de 10.000 enlaces de artículo técnicos, videos, libros , informes y documentación en general que tratan del Parque Nacional Yellowstone.

➤ **ESPACIO.**

- **Datos completos del Sol y la Luna durante un día.**

http://aa.usno.navy.mil/aa/data/docs/RS_OneDay.html

Se puede obtener el horario del amanecer, atardecer, salida de la luna y el sol y datos semejantes, como las fases de la luna.

- **Tierra desde el Espacio.**

<http://earth.jsc.nasa.gov/>

La base de datos fotográfica de imágenes de la Tierra desde el trasbordador espacial de la NASA contiene más de 375.000 imágenes. Dentro de esta página web se pueden encontrar diferentes caminos de búsqueda de imágenes captadas con diferente resolución.

- **Base de datos astronómica SIMBAD.**

<http://simbad.u-strasbg.fr/Simbad>

Esta base de datos proporciona información básica de bibliografía astronómica para identificar objetos en el Sistema Solar.

- **Simulador del Sistema Solar.**

<http://simbad.u-strasbg.fr/Simbad>

Vista de los planetas del Sistema Solar así como una gran variedad de datos sobre ellos.

➤ **METEOROLOGÍA.**

- **Base de datos de Imágenes de Acontecimientos Históricos y Significativos.**

<http://www5.ncdc.noaa.gov/cgi-bin/hsei/hsei.pl?directive=welcome>

Cientos de imágenes seleccionadas capturadas por satélite algunas de ellas muy importantes sobre el tiempo atmosférico y el medio ambiente de los últimos treinta años.

- **Índice de tormentas del Centro Nacional de Datos Climáticos.**

<http://www4.ncdc.noaa.gov/cgi-win/wwcgi.dll?wwEvent~Storms>

La base de datos de este Centro permite encontrar datos de las tormentas más importantes en un país o en otro sitio dependiendo del criterio de selección. Esta base de datos empezó en 1993 y sigue hasta la actualidad.

- **Índice UV.**

<http://www.epa.gov/sunwise/uvindex.html>

Desarrollado el Servicio Nacional Meteorológico de EE.UU. (NWS), este índice predice el nivel de la radiación de rayos ultravioletas el día siguiente a la consulta, para ayudar a la gente a decidir su nivel de protección contra el Sol. La búsqueda se realiza por código postal.

BÚSQUEDA DE GENTE.

- **RECURSOS GENERALES DE NEGOCIO.**

- **Lista Cyndi's de Sitios Genealógicos de Internet.**

<http://www.cyndislist.com/>

Hay cientos de miles de fuentes genealógicas, muchas de ellas en bases de datos en la Web Invisible. Aunque esta lista no pertenece a la red profunda, es un excelente lugar para empezar la investigación de sitios genealógicos. Cuando se realizó dicha lista había más de 89.000 entradas de fuentes genealógicas muchas de ellas incluidas en la web profunda.

- **Búsqueda de familia.**

http://www.familysearch.org/Eng/Search/frameset_search.asp

Más de 600 millones de nombres forman esta base de datos. La Iglesia de Jesús ha juntado registros genealógicos a lo largo del mundo durante más de un siglo para ayudar a la investigación histórica de los miembros de la familia.

- **Índice de Mortalidad de la Seguridad Social.**

<http://ssdi.rootsweb.com/cgi-bin/ssdi.cgi>

Una de las más grandes y fácil base de datos usada para la investigación genealógica es el Índice de Mortalidad de la Seguridad Social (SSDI). Las pistas y hechos del SSDI a menudo pueden ser usados para obtener información para conseguir localizar certificados de defunción, una necrología, un registro de cementerio, etcétera.

➤ **GRP.**

- **Selección de Médico de la Asociación Americana de Medicina.**

<http://www.ama-assn.org/aps/amahg.htm>

Esta Asociación proporciona información básica profesional sobre cada licencia médica y su poseedor en los EE.UU., incluyendo más de 690.000 doctores de medicina general y especialistas de medicina osteopática.

- **Base de Datos de pilotos certificados de EE.UU.**

http://www.landings.com/_landings/pages/search_amd.html

Esta base de datos contiene cerca de 600.000 pilotos y se puede especificar la búsqueda usando diferentes opciones.

- **Directorio CPA.**

http://www.cpadirectory.com/search_advanced.cfm

Directorio de búsqueda on-line de Contables certificados y firmas de auditorías en EE.UU.

- **DocFinder (EE.UU.)**

<http://www.docboard.org/>

Esta compilación está formada por un listado de licencias médicas oficiales de varios Estados de EE.UU. Está proporcionada por la Administración Americana de Medicina (AIM).

- **Jueces de la Corte de EE.UU.**

http://air.fjc.gov/history/judges_frm.html

La base de datos bibliográfica de Jueces Federales de EE.UU. contiene información de todos los jueces que han servido en las Cortes de Distrito, Cortes Supremas y de Apelación de EE.UU. desde 1789.

- **Directorio de Fuentes de Noticias del Club de Prensa Nacional.**

<http://npc.press.org/who/sources.htm>

Una de las muchas bases de datos de búsqueda de Internet que proporciona información comunitaria de varios temas con una gran variedad de expertos.

- **Directorio de Servicios de traducción. (Asociación Americana de Traductores).**

http://www.americantranslators.org/tsd_listings/

Se puede usar esta base de datos para buscar un traductor o intérprete. Este directorio on-line suministra el perfil de más de 4.000 traductores e intérpretes.

- **Base de datos de Radioaficionados Amateurs Estadounidenses.**

<http://callsign.ualr.edu/callsign.shtml>

Se puede buscar un amplio número de radioaficionados estadounidenses en esta base de datos.

➤ **PERSONAJES FAMOSOS E HISTÓRICOS.**

- **Base de Datos Biografía.Com.**

<http://www.biography.com/>

Una base de datos de más de 20.000 pequeñas biografías de una amplia variedad de gente.

- **Lista de Forbes.**

http://www.forbes.com/peopletracker/protected/exec_tracker.jhtml

Pistas de más de 120.000 ejecutivos y miembros de las lista de ricos y celebridades de la revista Forbes. Es necesario registrarse gratuitamente para acceder a esta lista.

- **El 7º Índice Biográfico Mundial.**

http://www.biblio.tu-bs.de/acwww25u/wbi_en/wbi.html

Esta base de datos está basada en la 7ª edición del CD-Rom del Índice Bibliográfico Mundial que contiene 2,8 millones de pequeñas biografías de personajes importantes que vivieron en los cinco continentes.

➤ **SERVICIOS DE CONSULTA.**

- **192.Com**

<http://www.192.com/>

Se trata de un directorio de información del Reino Unido. 192.Com proporciona a sus usuarios un acceso rápido y gratuito a su base de datos de direcciones y teléfonos en Internet. El contenido de este listado es continuamente actualizado y para acceder a él es necesario un registro de manera gratuita.

- **Directorio Telefónico Anywho.Com.**

<http://www.anywho.com/>

Uno de las muchas bases de datos de directorios telefónicos en Internet. Anywho.com proporciona listados telefónicos de residencias particulares y empresas.

- **InfoSpace**

<http://www.infospace.com/>

Búsqueda de direcciones, números de teléfono, direcciones de e-mail y otras informaciones de gente alrededor de todo el mundo.

- **Agente de Búsqueda de E-Mails (MESA).**

<http://mesa.rrzn.uni-hannover.de/>

Esta base de datos busca direcciones de e-mail en seis bases de datos diferentes.

- **Directorio Telefónico Inverso.**

<http://www.anywho.com/telq.html>

Búsqueda de información de personas a través de su número de teléfono.

- **Directorio mundial de E-mail WED.**

<http://www.worldemail.com/>

Más de 18 millones de direcciones e-mail a lo ancho del mundo.

➤ **SERVICIO DE MILITARES VETERANOS.**

- **Sistema de Búsqueda de Marineros y Soldados de La Guerra Civil Norteamericana.**

<http://www.itd.nps.gov/cwss/index.html>

El Sistema de Búsqueda de Marineros y Soldados es una base de datos computerizada que contiene información básica sobre hombres que sirvieron en ambos lados durante la Guerra Civil. Esta basada en más de 5.4 millones de nombres obtenidos de los archivos nacionales.

- **El Muro Virtual (Directorio de los Veteranos de la Guerra de Vietnam).**

http://www.thevirtualwall.org/search/search_index.htm

Búsqueda virtual de veteranos de la Guerra del Vietnam.

- **Directorio de la Marina de EE.UU.**

<http://www.navydirectory.smartlink.navy.mil/>

Este directorio está proporcionado por personal de la Marina y en él se incluyen trabajadores civiles, militares y trabajadores de Agencias asociadas. El personal que está a bordo de algún barco o en Unidades desplegadas no está accesible desde Internet.

CIENCIAS SOCIALES

➤ **ANTROPOLOGÍA.**

- **Base de Datos de Revistas de Antropología.**

<http://wings.buffalo.edu/ARD/>

La base de datos de Revistas de Antropología es utilizada para mejorar el nivel de acceso de antropólogos a la literatura antropológica para dar mayor conciencia de los documentos publicados en este campo.

➤ **ARQUITECTURA.**

- **Great Buildings Collection**
<http://www.greatbuildings.com/>

Base de datos sobre arquitectura. Proporciona información sobre el edificio, arquitecto, localidad, fecha y sistema de construcción, clima, utilización, estilo arquitectónico, imágenes y fotografías, plano y referencias bibliográficas en Internet. Da la posibilidad mediante la adquisición libre de un programa de ver el modelado del edificio en 3D. Se puede buscar por edificio, arquitecto y localidad. En inglés.

- **Base de datos de Arqueología (EE UU)**

<http://www.cast.uark.edu/other/nps/nadb/nadb.mul.html>

La base de datos Nacional de Arqueología es un dilatado inventario bibliográfico de aproximadamente 240.000 informes sobre la investigación arqueológica. Esta información representa una parte de la información disponible sobre sitios arqueológicos estadounidenses. La información puede ser buscada por Estado, condado, material, año de publicación, título o autor.

- **CEDEX Catálogo de revistas**
<http://www.cedex.es/documentacion/catalogor.htm>

Catálogo de revistas de la red de bibliotecas del CEDEX. La temática de publicaciones comprende estudios de carreteras, hidrográficos, de puertos y costas, estructuras y materiales, geotecnia, estudios históricos de obras públicas y urbanismo.

➤ **DEMOGRAFÍA.**

- **Datos demográficos. (Proyecto de Información Compartida del Gobierno de**

EE.UU.)

<http://govinfo.kerr.orst.edu/>

El Proyecto de Información Compartida del Gobierno de EE.UU. con sede en la Universidad de Oregón proporciona acceso interactivo a numerosas bases de datos del Gobierno estadounidense en temas económicos, de educación y demográficos.

- **EASI Informes rápidos y Análisis (Datos Demográficos).**

http://www.easidemographics.com/reports/easi_free_reports.phtml

El Software Analítico Easy ofrece varias bases de datos gratuitas que contienen material demográfico básico de EE.UU. Este servicio gratuito permite al usuario elegir entre 12 variables demográficas diferentes y crear 11 tipos de informes distintos.

- **Índice de Población.**

<http://popindex.princeton.edu/search/index.html>

El Índice de Población, publicado desde 1935, es la primera herramienta de referencia en la literatura demográfica. Presenta una bibliografía de libros recientemente publicados, artículos periodísticos y otros documentos sobre la demografía.

- **Base de Datos de la Oficina Internacional del Censo de EE.UU.**

<http://www.census.gov/ipc/www/idbacc.html>

Con esta base de datos se tiene acceso a estadísticas demográficas de varios países.

- **Consulta de Censo estadounidense.**

<http://venus.census.gov/cdrom/lookup>

Se pueden crear tablas personalizadas según diferentes criterios usando datos del censo estadounidense. Este sitio web es uno de los mejores ejemplos de lo que la Red Invisible puede ofrecer.

- **Estadísticas Nacionales del Reino Unido.**

<http://www.statistics.gov.uk/statbase/datasets2.asp>

Es una base de datos de estadísticas oficiales del Reino Unido. Este sitio tiene numerosas opciones de búsqueda tanto de material gratuito como de pago. El enlace de búsqueda proporciona acceso on-line a más de 1900 datos estadísticos.

- **Hoja Mundial de Datos.**

<http://www.worldpop.org/prbdata.htm>

Esta base de datos contiene información de 85 variables demográficas de 221 países en el mundo, 28 regiones geográficas, 50 Estados de EE.UU. y del Distrito de Columbia.

➤ **RECURSOS DE DESARROLLO.**

- **Base de datos: Indicadores Mundiales de Desarrollo.**

<http://devdata.worldbank.org/data-query/>

Este conjunto de datos ofrece acceso libre a los Indicadores Mundiales de Desarrollo (WDI), la colección más extensa del mundo sobre datos de desarrollo. Este segmento incluye 54 series temporales de indicadores de 206 países y 17 grupos, en el intervalo desde 1995 hasta 1999.

- **ELDIS (Datos de Desarrollo).**

<http://nt1.ids.ac.uk/eldis/eldsea.htm>

ELDIS proporciona un número creciente de descripciones y enlaces de varias fuentes de información, incluyendo documentos on-line, organización de sitios de World Wide Web, bases de datos, catálogos bibliográficos, proyectos de investigación, mapas y colecciones de periódicos.

- **Banco de Fotos – UNESCO (Organización para la Educación, Ciencia y Cultura de la Organización de Naciones Unidas) .**

<http://www2.unesco.org/photobank>

La colección cubre un amplio rango de temas relacionados con las competencias de la Organización: educación, ciencia, cultura y comunicación. Dicha colección fue empezada in 1946, fecha en la que se fundó la UNESCO y actualmente contiene más de 10.000 imágenes digitalizadas. Más imágenes estarán disponibles cuando se digitalice el resto de la colección.

- **USAID (Agencia Estadounidense de Desarrollo Internacional).**

http://www.dec.org/partners/dexs_public/

El Sistema Experimental sobre Desarrollo (DEXS) es un conjunto de Bases de Datos Bibliográficas que contienen archivos de más de 100.000 documentos de programas técnicos de la Agencia Estadounidense de Desarrollo Internacional. El propósito del Sistema DEXS es reforzar los proyectos de desarrollo, actividades y programas de la USAID para posibilitar el acceso a estos

documentos por parte de otras agencias, instituciones gubernamentales y todo el público en general.

- **InfoNación**

http://www.un.org/Pubs/CyberSchoolBus/infonation/s_infonation.htm

InfoNación es una base de datos que permite ver y comparar los datos estadísticos de los Estados Miembros de las Naciones Unidas. Se puede buscar por país, datos económicos, demográficos e indicadores sociales.

➤ **FUENTES DE NEGOCIO GENERAL.**

- **Estadística de la Asociación de Bibliotecas de Investigación.**

<http://fisher.lib.virginia.edu/newarl/>

Se pueden encontrar los datos de los 111 miembros de la Asociación desde 1962-63 hasta 1998-99 así como estudiar cualquier estadística de las diferentes categorías de datos que ofrece la Asociación. Es posible descargar estos datos en formato .txt o .wkl.

- **DARE: Directorio en Ciencias Sociales.**

<http://www.unesco.org/general/eng/infoserv/db/dare.html>

En este directorio hay 11.000 referencias de páginas web sobre investigación en ciencias sociales, documentación y servicio de información, así como referencias a instituciones que vigilan las leyes sobre paz y derechos humanos.

- **Base de Datos de Reconstrucción Nacional.**

<http://www.lin.ca/htdocs/findrs.cfm>

Es una base de datos producida por la Red de Información de Ocio. Contiene información sobre recursos prácticos en este campo de la ciencia.

- **NISSO Base de datos de sexología. (Instituto Holandés de Investigación de Sexología Social).**

http://www.nisso.nl/cgi-bin/nph-nisso_search.pl?language=us&db=nissomain

La colección de unos 55.000 documentos contiene monográficos, informes, conferencias, artículos, etc. Los resúmenes son incluidos con multitud de entradas diferentes.

- **Base de Datos de Encuestas del Washington Post.**

<http://www.washingtonpost.com/wp-srv/politics/polls/vault/vault.htm>

Los resultados de las encuestas nacionales del Washington Post incluyen una gran variedad de temas. El material empezó a recopilarse en Enero de 1998.

➤ **ESTUDIOS Y DATOS DE GENERO.**

- **Sitio sobre Género: Base de Datos sobre estudios de Género.**

http://www.uni-koeln.de/phil-fak/englisch/datenbank/e_index.htm

Este enlace es una base de datos de búsquedas que proporciona acceso a más de 6000 registros sobre la teoría y crítica literaria feminista enfocada a la literatura Inglesa y Americana.

- **Estadísticas de Género.**

<http://genderstats.worldbank.org/>

Es una base de datos del Banco Mundial. Las estadísticas son las típicas usadas en los estudios de Género y las estadísticas de varios países están disponibles.

- **Índice Medieval Feminista.**

<http://www.haverford.edu/library/reference/mschaus/mfi/mfi.html>

El Índice Medieval Feminista incluye artículos periodísticos, revistas, ensayos y libros sobre mujeres, sexualidad y diferentes géneros durante la Edad Media. Debido al auge de la investigación sobre estudios de la mujer en las últimas dos décadas, los eruditos y estudiantes interesados en la vida de la mujer durante la Edad Media pueden encontrar un gran número de documentos sobre este tema en este enlace.

- **Base de Datos sobre Género, Mujer y Trabajo.**

<http://www.wsulibs.wsu.edu/ris/risweb.isa/>

Es una base de datos de citas usadas en trabajos de investigación de estudiantes.

- **Base de Datos de los discursos de Fidel Castro.**

<http://lanic.utexas.edu/info/search/castro.html>

Se trata de una base de datos con las traducciones completas al inglés de los discursos, entrevistas y conferencias de prensas de Fidel Castro, basadas en grabaciones del Servicio Extranjero de Difusión de Información (FBIS), una Agencia del Gobierno de EE.UU. responsable de la supervisión y difusión de información a distintos medios de comunicación a lo largo de todo el mundo. Estos documentos son de dominio público.

- **Manual de Estudios Latinoamericanos.**

<http://lcweb2.loc.gov/hlas/mdbquery.html>

Este manual es una bibliografía latinoamericana consistente en trabajos seleccionados y realizados por estudiantes. Ha sido editada por la División Hispánica de la Biblioteca del Congreso, y el manual alterna anualmente entre Ciencias Sociales y Humanidades. Cada año, más de 130 academias alrededor del mundo eligen más de 5.000 trabajos para incluirlos en este manual. Se empezó a publicar en 1935 y ofrece una guía esencial de los recursos disponibles en este campo de las Ciencias Sociales.

➤ **MILITAR.**

- **Índice de Revistas Militares de la Universidad del Aire.**

<http://www.dtic.mil/search97doc/aulimp/main.htm>

Un índice substancial de artículos significativos, noticias y editoriales en lengua inglesa de revistas aeronáuticas y militares. Este índice contiene citas desde 1990 y es actualizado cada tres meses gracias al Índice de Revistas Militares de la Universidad del Aire.

- **Biblioteca de Imágenes Militares.**

<http://www.militarycity.com/newsroom/images.html>

MilitaryCity.Com proporciona esta base de datos. Hay imágenes de aviones, misiles, barcos y muchos otros artefactos militares.

➤ **SICOLOGÍA.**

- **Base de Datos Jourlit/Bookrev (Asociación Psicoanalítica Americana).**

<http://apsa.org/lit/>

La base de datos Jurlit y Bookrev, originada conjuntamente por los doctores Stanley Goodman y Vann Spruiell, constituye una enorme bibliografía de artículos, libros y revistas sicoanalíticos. Incluye archivos suministrados por otras personas y se completa con más de 30.000 entradas.

- **Enlaces de revistas de Psicología.**

<http://www.wiso.uni-augsburg.de/sozio/hartmann/psycho/journals.html>

Un índice de más de 1.600 artículos sobre psicología y ciencias sociales. Suministra enlaces de revistas y páginas web relacionadas con este tema.

➤ **INVESTIGACION Y DESARROLLO.**

- **FundSource.**

<http://www.decadeofbehavior.org/fundsource/index.html>

Una herramienta de búsqueda en el ámbito de la Ciencias Sociales y de la Conducta Humana.

➤ **RELIGIÓN**

- **Hispania**

<http://www.moderna1.ih.csic.es/hispaniasacra/default.htm>

Revista sobre historia religiosa en España y el mundo hispánico y sobre cuestiones culturales que hayan sido elementos conformadores de las sociedades española e hispánica. Búsqueda de los índices, desde 1945, por autores, autores de obras reseñadas, descriptores temáticos, onomásticos y geográficos.

Sacra

- **Archivo de Datos sobre Religión Americana.**

<http://www.wiso.uni-augsburg.de/sozio/hartmann/psycho/journals.html>

El Archivo de Datos sobre Religión Americana reúne datos cuantitativos sobre el estudio de la religión americana.

- **Diccionario Beliefnet.**

<http://www.beliefnet.com/glossary/index.asp>

Beliefnet.Com proporciona una versión de búsqueda del diccionario HarperCollins de religión.

- **Directorio de Centros Religiosos.**

<http://www.fas.harvard.edu/~pluralsm/directory/directory.cgi>

El Proyecto Pluralismo (de la Universidad de Harvard) mantiene un extenso directorio de centros religiosos en EE.UU. Actualmente, este directorio es una base de datos de más de 3.500 centros. No incluye información sobre centros cristianos o judíos ya que pueden encontrarse fácilmente en la guía local; sin embargo incluye listas de centros Budistas, Hindúes, etc.

- **Recurso Ingleses Religiosos.**

<http://etext.lib.virginia.edu/relig.browse.html>

Incluye búsqueda y análisis de datos sobre la Biblia, el libro de los Mormones, el Corán, así como una concisa enciclopedia sobre religión.

- **SESRTCIC**

<http://www.sesrtcic.org>

SESRTCIC (Investigación Estadística, Económica y Social para países Islámicos) proporciona acceso a más de 180 indicadores estadísticos de 56 países.

- **El Navegador de la Biblia.**

http://www.stg.brown.edu/webs/bible_browser/pbeasy.shtml

Posibilita la búsqueda de varias versiones de la Biblia.

- **El Navegador del Quaran.**

http://www.stg.brown.edu/webs/quran_browser/pqeasy.shtml

Búsqueda y visualización de varios pasajes o palabras individuales del Quaran. Hay varias traducciones disponibles para buscar.

TRANSPORTE

➤ **TRANSPORTE AÉREO**

- **Información Certificada de Líneas Aéreas (Administración Federal De Avia**

ción).

<http://av-info.faa.gov/OpCert.asp>

Contiene información perteneciente a las certificaciones de las líneas aéreas, sus operaciones y flotas según la fabricación y modelos de aviones.

- **AirNav: Base de Datos de Aeropuertos.**

<http://www.airnav.com/airports/>

AirNav proporciona información aeronáutica detallada y gratuita sobre aeropuertos y ayudas a la navegación en EE.UU. Se ofrece búsquedas rápidas en la base de datos que permiten al piloto obtener información útil para su plan de vuelo.

- **Resumen de Accidentes de Aviación (Consejo de Seguridad de Transporte Nacional).**

<http://www.nts.gov/aviation/Accident.htm>

La base de datos contiene información desde 1983 en adelante de accidentes de aviación civil dentro de los EE.UU., sus posesiones y territorios, así como en aguas internacionales. Así mismo se incluyen los incidentes aéreos investigados por el Consejo.

- **Base de Datos de Aviones Registrados.**

http://www.landings.com/_landings/pages/search.html

Esta página es un resumen de enlaces de bases de datos que proporcionan datos sobre las flotas de aviones de varios países.

- **Base de Datos On-line de Líneas Aéreas de EE.UU.**

http://www.bts.gov/cgi-bin/oai/ontime_js.pl

Esta base de datos se ha formado para asistir a los usuarios de transporte aéreos para tomar decisiones en cuanto a la calidad del servicio de transporte aéreo. Actualmente, la base de datos on-line incluye las líneas aéreas Alaska, America West, American,

Continental, Delta, Northwest, Southwest, TWA, United y USAIR. Estas compañías incluyen más del 90 % de las operaciones de transporte aéreo en EE.UU.

- **Resumen de Accidentes Aéreos en el mundo. (WAAS)**

<http://www.waasinfo.net/>

Contiene información sobre los accidentes aéreos en el mundo Ha sido realizado por la Autoridad de Aviación Civil Británica y proporciona breves detalles de todos los accidentes aéreos conocidos en aviones con motores de combustión y helicópteros.

➤ **TRANSPORTE AUTOMOVILÍSTICO.**

- **Safer (Información sobre transporte motorizado en EE.UU)**

<http://www.ite.org/>

La Administración Federal de Seguridad en Transporte Motorizado (FMCSA) ha proporcionado datos sobre seguridad a la industria y al público a través de encuestas telefónicas durante varios años. El Sistema de Grabaciones Electrónicas sobre Seguridad y Salud (SAFER) permite ofrecer la información electrónicamente, así las funciones de SAFER están ahora disponibles gratuitamente en Internet.

- **Librería de Tráfico Calming**

<http://www.safersys.org/>

Esta librería contiene una base de datos con informes, artículos y otros documentos relacionados con el tráfico. En algunos casos la publicación está por completo on-line y en otros sólo un resumen.

➤ **RECURSOS GENERALES DEL NEGOCIO DEL TRANSPORTE.**

- **Base de Datos de Transporte Canadiense.**

http://www.tc.gc.ca/pol/en/t-facts_e/Statistical_Data_Menu.htm

Una colección de varias bases de datos de estadísticas de numerosos medios de transporte. Estos datos también se pueden descargar en formato de hoja de cálculo.

- **Directorio de Expertos Forenses en Ingeniería de Transporte**

<http://www.ite.org/expert.htm>

El ITE (Instituto de Ingenieros de Transporte) proporciona el Directorio de Expertos Forenses en Ingeniería de Transporte para facilitar la búsqueda de datos sobre transporte. Las búsquedas pueden ser llevadas a cabo por nombre, localización o especialidad.

- **Sistema de Informe y Análisis de Fatalidades.**

<http://www-fars.nhtsa.dot.gov/>

El Sistema de Informe y Análisis de Fatalidades (FARS) contiene datos de todos los choques de vehículos en EE.UU que ocurren en una autopista pública y que incluye una fatalidad.

- **Analizador Global de Transporte.**

<http://www.joc.com/gta/>

El Diario de Comercio proporciona acceso a varias bases de datos sobre el mundo del transporte.

- **Base de Datos PATH (Sistemas de Transporte Inteligente).**

<http://www.dccdata.com/path/path.htm>

La base de datos californiana PATH proporciona acceso a la colección más grande y completa de información bibliográfica de Sistemas de Transporte Inteligentes (ITS).

- **Datos de Carga de la Superficie del Tránsito**

<http://www.bts.gov/ntda/tbscd/search.html>

Búsqueda interactiva de datos de carga de la superficie del tránsito. Se puede seleccionar el Estado o puerto de EE.UU., el modo de transporte, el periodo de tiempo y las comodidades en las que el usuario está interesado. Además ofrece la posibilidad de generar tablas con estos datos.

- **TRIS (Transportation Research Information Service)**

<http://199.79.179.82/sundev/search.cfm>

TRIS contiene más de 400.000 archivos publicados de investigación sobre transporte. Ofrece únicamente información bibliográfica.

➤ **TRANSPORTE MARÍTIMO**

- **Equasis (Barcos Mercantes)**

<http://www.equasis.org/>

Equasis proporciona información de calidad relacionada con la seguridad de los barcos mercantes. Para acceder a esta información es necesario registrarse.

- **Información sobre Cambio de Estado de Puertos.**

<http://psix.uscg.mil/Default.asp>

El Sistema de Información sobre Cambio de Estado de Puertos (PSIX) contiene información específica sobre barcos obtenida del Sistema de Información sobre Seguridad Marina del Servicio Guardacostas de Estados Unidos (MSIS).

- **Documentación de Barcos de Guardacostas de EE.UU.**

http://www.st.nmfs.gov/st1/commercial/landings/cg_vessels.html

Esta base de datos sólo retiene información de barcos de 5 toneladas de peso neto o más y son documentados por los Guardacostas americanos. Los barcos de menos de 5 toneladas de peso neto se registran individualmente.

- **Sistema de Registro de Barcos Canadiense.**

<http://www.tc.gc.ca/ShipRegistry/menu.asp?lang=e>

El Registro de Canadá es básicamente un sistema de Título de Posesión de Barcos. Es similar al registro de títulos de propiedad en tierra. El Registro es obligatorio para barcos que exceden de 15 toneladas de peso bruto y voluntario para barcos más pequeños. Hay aproximadamente 46.500 barcos con un total de 4,7 millones de toneladas de peso bruto. Además los armadores canadienses tiene 100 barcos con un total de 2,2 millones de toneladas de peso bruto operando con bandera extranjera.

➤ **TRANSPORTE FERROVIARIO.**

- **Base de datos de la Administración Federal Sobre Seguridad Ferroviaria.**

<http://safetydata.fra.dot.gov/officeofsafety/>

...Informes controlados y detallados, y sumarios de accidentes de ferrocarril de la base de datos de la Administración Federal. Los usuarios pueden optar por criterios de búsqueda específicos y obtener informes de salida de la misma. Para estos servicios se debe usar el menú de la página.

HISTORIA DEL MUNDO

- **Base de Datos de Registros Visuales Columbia Británica.**

<http://www.bcarchives.gov.bc.ca/visual/visual.htm>

La base de datos Columbia contiene más de 110.000 textos y más de 60.000 disponibles on-line.

- **Marxists Internet Archive: sección en español**

<http://www.marxists.org/espanol/>

Base de datos sobre el marxismo: archivo de autores marxistas, enciclopedia del marxismo y archivos de referencia.

- **Base de Datos Bibliográfica del Genocidio Camboyano.**

<http://www.yale.edu/cgp/>

El Programa de Genocidio Camboyano ha creado un catálogo indexado de fuentes de documentación para demostrar la grave violación de los derechos humanos durante el régimen Khmer Rouge. Actualmente la Base de datos CGP Bibliográfica (CBIB) contiene más de 3,000 registros.

- **Directorio de Herencia de Canadá.**

http://www.chin.gc.ca/Museums/CHER/e_hp_cher.html

El Directorio de Herencia proporciona la información detallada sobre organizaciones, departamentos de gobierno y agencias, principalmente en Canadá, que es contratado en actividades de herencia.

- **Base de Datos de Mujeres Inventores canadienses.**

http://napoleon.ic.gc.ca/cipo/cdn_inv.nsf/Home+Page+View/EnglishHome

Esta base de datos de mujeres inventores de Canadá. Cada entrada incluye una biografía breve.

- **Early Canadiana Online (ECO)**

<http://www.canadiana.org/eco/index.html>

Early Canadiana Online (ECO) es una colección on-line de más de 3.000 libros y folletos que documentan la historia canadiense desde el primer contacto con los europeos hasta finales del siglo XIX. La colección está en particular muy documentada en la historia de la literatura, de la mujer, estudios natales, viajes y la exploración, y la historia de Canadá francesa.

- **Proyecto de aprendizaje On-line del Atlas Histórico de Cánada.**

<http://www.canadiana.org/eco/index.html>

Este Atlas On-line realizado para estudiantes está lleno de mapas interactivos y gráficos. El sitio web sugiere el uso de Internet Explorer para un correcto

funcionamiento.

- **Base de Datos de Imágenes de Napoleón.**

http://www.napoleon.org/us/us_cd/media/me-principal.asp

Una biblioteca de imágenes relativas a los archivos civiles y militares del Primer y Segundo Imperio. Proporcionada por la Fundación Napoleón.

- **Fotos de Australia**

<http://www.pictureaustralia.org>

Fotos de Australia consiste en enlaces de imágenes de diferentes sitios de Australia, visitables todos excepto textos completamente digitalizados que son restringidos por derechos de reproducción.

- **SHIPDES (Descripción de Barco)**

<http://www.library.tudelft.nl>

SHIPDES es una base de datos que contiene mas de 19.000 descripciones de barcos específicos. La búsqueda puede ser hecha por datos específicos como el nombre del barco, el tipo de barco, la eslora, la manga, el calado, el peso, la velocidad, el número de contenedores, etc...Está proporcionada por el Centro de Información Marítimo de la Universidad Tecnológica de Holanda.

- **Búsqueda en la colección y archivo del Museo Memorial Americano del Holocausto**

http://www.ushmm.org/uia-cgi/uia_form/db_group/collections

El catálogo on-line del Museo Memorial Americano del Holocausto proporciona el medio para la búsqueda y la recuperación de información de sus colecciones. El archivo de fotos contiene una copia de la imagen.

- **Colección de Carteles de la II Guerra Mundial.**

<http://www.library.northwestern.edu/govpub/collections/wwii->

[posters/index.html](#)

El Departamento de Publicaciones de Gobierno de la Biblioteca de la Universidad del Noroeste posee una colección de 300 carteles emitidos por Agencias Federales desde el comienzo de la guerra hasta 1945.

ANEXO 2

ENTREVISTAS REALIZADAS POR EL AUTOR

ENTREVISTAS REALIZADAS POR EL AUTOR

INTRODUCCIÓN A LAS ENTREVISTAS

A continuación les exponemos una serie de entrevistas realizadas a algunos de los máximos expertos en la materia del periodismo a través de Internet. Con ellas pretendemos dar una visión realista de lo que está sucediendo en la actualidad en este campo, además de abordar perspectivas de futuro. Los entrevistados son los siguientes:

- **Mariló Ruiz de Elvira:** Fue directora de la edición digital del diario *El País* (<http://www.elpais.es>) desde sus comienzos. En 1999 pasa a ser directora de producto de Proel (Promotora de Ediciones Electrónicas del Grupo PRISA). A partir de esta empresa se crea Inicia, (<http://www.inicia.es>) portal de Internet del Grupo PRISA en el que Ruiz de Elvira ocupa hasta la actualidad el cargo de directora de contenidos.
- **Mario Tascón:** Fue director de la edición digital del diario (<http://www.elmundo.es>) *El Mundo* desde sus comienzos. En el año 2000 pasa a ser director de contenidos de Prisacom, empresa creada para agrupar todos los contenidos digitales del Grupo PRISA.
- **Ricardo Fornás:** Creador del buscador de buscadores “Buscopio” (<http://www.buscopio.com>).
- **Ramón Salaverría:** Es director del Laboratorio de Comunicación Multimedia (MMLab) de la Universidad de Navarra (<http://www.unav.es/fcom/mmlab/>). Es además autor de la *Guía de Internet para periodistas* (<http://www.unav.es/fcom/guia/>) y editor del Boletín MMLab (<http://www.unav.es/fcom/mmlab/boletin/>), publicación difundida por correo electrónico que recopila noticias nacionales e internacionales sobre comunicación.

En estas entrevistas se ha intentado poner de manifiesto las distintas perspectivas, muchas veces coincidentes, que tienen estos profesionales de la comunicación *online* con respecto al periodismo digital en general, al papel que desempeña el periodista en este nuevo ámbito y al uso de Internet como herramienta a la hora de buscar información.

De estas entrevistas, cabe destacar las siguientes aspectos:

- Tanto Ruiz de Elvira como Tascón reconocen la enorme importancia del propio usuario como fuente de información para los periodistas que trabajan con Internet, hecho motivado por la creciente interactividad que permite la Red. Ricardo Fornás reafirma esta observación pero aconseja prudencia ya que en Internet todos pueden opinar de cualquier cosa, aunque sea de un hecho falso. Ramón Salaverría señala a su vez que la clave para encontrar lo que se busca en Internet es determinar cuál es el recurso de búsqueda pertinente en cada caso.
- Con respecto a la posible desventaja que pudiera suponer para el periodista la credibilidad de las fuentes en la Red, Tascón opina que la rapidez con la que se procesa la información en un medio digital supone un mayor riesgo a la hora de publicar datos, procedentes de Internet, que no han sido lo suficientemente contrastados. Sin embargo, tanto él como Ruiz de Elvira y Fornás, coinciden en afirmar que en las paginas web hay suficientes indicios como para saber si estos datos son verdaderos o falsos. Los tres señalan que tanto en Internet como en cualquier otro medio, lo que se debe hacer es comprobar siempre las fuentes de referencia. Asimismo, apuntan que uno de los factores que en la actualidad y en un futuro indican fiabilidad son las marcas.
- Con respecto al lenguaje de Internet, Ruiz de Elvira opina que se va a internacionalizar. Según ella, va a haber una serie de símbolos o de signos, a parte una serie de palabras inglesas comunes a toda la comunidad internauta, para comunicarse a través de la Red. Afirma que será especialmente importante que los textos sean concisos y que lleven suficientes elementos gráficos. Sin

embargo, Mario Tascón señala que aún no tenemos desarrollado el lenguaje de Internet, pero que a la hora de escribir hay que tener muy en cuenta las

características del propio medio, como puede ser el hecho de que la información la va a leer, escuchar o ver un solo individuo sentado frente a su ordenador. Por su parte Salaverría señala la importancia que tendrá en la elaboración de documentos web la inclusión de multimedia y elementos que promuevan la interactividad.

- En cuanto a la gratuidad de la información distribuida a través de los medios *online*, tanto Ruiz de Elvira como Tascón, coinciden en su opinión de que, en su mayoría, no debería ser gratuita ya que la producción de información siempre supone un coste para la empresa que esta no puede amortizar con publicidad.
- En referencia a la evolución de los medios a través de la Red, Tascón apunta que dicha evolución estará íntimamente ligada a que desarrollemos el lenguaje. Mientras, Mariló Ruiz de Elvira apuesta por el papel relevante de los periodistas a la hora de procesar, seleccionar, escribir y presentar la información de la mejor forma posible al usuario, dentro del marasmo de datos que supone actualmente Internet. Asimismo, Salaverría opina que el ámbito del periodismo digital será en un futuro el de información puntual, del dato, la última hora y la documentación.

Con respecto a la evolución de los buscadores Ricardo Fornás apunta que se insistirá en dos aspectos que están íntimamente unidos: una mayor capacidad para comprender el lenguaje natural y una mayor relevancia en las respuestas. Según él, supone un reto difícil ya que entre la comunicación de dos personas intervienen algo más que palabras, como los gestos o el contexto.

ENTREVISTA A MARIO TASCÓN- DIRECTOR DE CONTENIDOS DE PRISACOM Y EXDIRECTOR DE LA EDICIÓN DIGITAL DEL DIARIO *EL MUNDO* :

1- ¿Cuáles piensa que deben ser las aptitudes que debe de tener un periodista que trabaja en un medio digital como es Internet?

Los periodistas que trabajan en este nuevo medio deben de tener las mismas aptitudes que si trabajaran para cualquier otro medio.

2- ¿Qué ventajas e inconvenientes cree que aporta Internet al periodismo tradicional?

Ventajas muchas. Por ejemplo yo creo que aporta que tienes una relación muy directa con tu publico, con tus lectores. Otra es que es el único medio en el que tienes todos los medios a la vez. Otra es que es el único medio en el que la participación con el lector sobre cómo se hacen las informaciones o incluso en cómo se muestran, es mayor. Con lo cual es un reto. Las desventajas podían estar también sumadas a las ventajas. Depende de cómo se manejen esas ventajas puedes hacer un uso no muy bueno de ellas. Con respecto a la posible desventaja de la credibilidad de las fuentes en la Red, es cierto que la rapidez que imprimes si que te hace estar con un riesgo mayor que el que tienes en un periódico. Pero también es cierto que es muy parecido al de una radio o al de una agencia, por lo que tampoco es nuevo. Otra desventaja, que realmente no es una desventaja sino una característica, es que Internet es un medio todavía muy elitista. Y no sólo Internet, sino todo lo digital. Es decir, todavía es más fácil comprar un periódico que acceder a Internet.

3- ¿A qué recursos de búsqueda suele recurrir cuando quiere buscar algo en la Red?

A mí me gusta mucho Google. De lo que yo conozco es el mejor, y el que más uso. Me parece que tiene los mejores algoritmos de búsqueda. Por lo menos en lo que yo busco habitualmente es el que mejor me encamina hacia lo que quiero encontrar. Es el que menos me suele fallar. A veces también utilizo algún directorio como Yahoo.

4- Según su experiencia profesional ¿Cuál piensa que es el mejor recurso de búsqueda para el periodista digital?

Depende de lo que el periodista digital en particular esté buscando. Debe de utilizar todas las herramientas que precise, no solo una. Existen listas de recursos para periodistas que éstos pueden utilizar para facilitar su trabajo. La misma guía de *El País* es muy básica pero para periodistas está bastante bien.

Uno de los mejores recursos para el periodista digital son los propios usuarios. Y esto ocurre gracias a una de las características del medio que es la accesibilidad. De repente el periodista es accesible, pero no accesible en el sentido tradicional. Por supuesto que te podían mandar una carta, pero eso lleva un esfuerzo, hay que buscar un sobre y un sello, mandarlo por correo, etc. Esto es más inmediato y además sabes que le llega directamente a la persona en concreto, sea bueno o malo. Por este motivo el lector como fuente para el informador se ha multiplicado. Además en la historia de Internet, que es corta, hay muchísimos casos en los que el lector se ha convertido en una fuente. Por contarle uno de ellos, recuerdo una investigación que hizo *El Mundo* sobre el tema de el Rabal en Barcelona. Gran parte de la investigación la dirigió un lector a través del correo electrónico. Fue éste el que nos puso sobre la pista, y el que nos iba guiando a medida que salían crónicas y artículos. Obviamente conocía perfectamente el tema. Fue él el que se puso en contacto con nosotros. Recuerdo que era un correo camuflado, es decir, nosotros no podíamos devolvérselo. Él fue el que estableció los códigos de la comunicación. Lo que hizo fue avisar en un primer correo todo lo que el creía que se debía investigar y nos animaba a que si todo lo que el decía, lo comprobábamos y nos parecía interesante, si se escribía un artículo que se le diera una pista, que al final se le dio. Creo que era algo así como “acaben ustedes el

artículo con esta frase...” o incluyan ustedes tal palabra. Esa pista a él le hizo ver que sí que se había leído y que no era una investigación que se hacía por casualidad. Esto animaba al lector y seguía dándonos pistas. Supongo que se lo pasó bien porque era

parecido a una novela de espionaje. Al final, por la historia, nos llegamos a imaginar quien era, pero con certeza nunca lo supimos.

5- ¿Me podría hacer una valoración de las listas de distribución, boletines, bases de datos y otros recursos no accesibles mediante buscadores generales?

Depende, pero en general las considero bastante útiles para el periodista. Las listas de distribución suelen ser muy interesantes. Y además éstas, como suelen estar cada una muy enfocadas a cosas muy específicas, sirven de mayor ayuda. También hay que tener en cuenta que suelen tener mucho ruido.

6- Está bastante claro que las posibilidades de Internet como medio de información del profesional de la información son ilimitadas, pero a tantas ventajas cabe añadirle el inconveniente del anonimato o la incierta identidad de la mayoría de dichas fuentes. ¿Qué opina de la fiabilidad de las fuentes en Internet?

Es como en el mundo real. El problema es que como es un nuevo mundo, hay una serie de parámetros que hay en el mundo real, que el periodista tiene que aprender las claves. Con relación a la fiabilidad en general, si voy por una calle, entro a un comercio y compro, hay una serie de códigos de fiabilidad que yo asumo y que he aprendido por cultura. Yo doy por supuesto que si ese local está abierto es que es legal, que ha pasado sus inspecciones técnicas, sanitarias y demás, etc. En el mundo virtual, a parte de la dificultad de la gran cantidad de información con la que nos encontramos, tenemos que ir descubriendo cuáles son los factores que nos indican

la fiabilidad. Uno de ellos son las marcas. Por ejemplo, si yo entro al web de la Agencia EFE, tengo la fiabilidad que tiene en general esta empresa. Si yo me encuentro un mensaje en un foro no le daré el mismo crédito. Pero esto es igual que en el mundo real, si me encuentro un papel manuscrito en la calle o si me lo cuenta alguien desconocido, no le doy la misma fiabilidad que si me entero por un teletipo

de EFE. No creo, por tanto, que haya un problema de fiabilidad de fuentes salvo por la cantidad, que es diferente de la que te puedes encontrar en la vida normal.

El problema se puede encontrar, sobre todo, si se recurre a fuentes de otros países y el periodista desconoce el código de ese lugar. Por ejemplo si Juan Luis Cebrián tuviera una pagina personal, para un español la información que contendría esta pagina seria fiable pero si la busca alguien de otro país que no sabe quién es él, pues no le daría el mismo crédito. Pero esto tiene más que ver con el conocimiento.

7- Hay un archivo digital llamado “The Internet Archive” que se dedica a almacenar todas las *webs* que encuentra, incluyendo información hallada en los grupos de discusión, para preservarlas antes de su posible desaparición. Su objetivo es servir como fuente de documentación para futuros trabajos de investigación. ¿Opina que es un trabajo útil o en su mayor parte inservible debido a la poca fiabilidad de la mayoría de los documentos allí guardados?

Es útil porque en los primeros años Internet era un instrumento de publicación pero se pensaba poco en el archivo. Ahora mismo encontramos que en algunos medios de comunicación es muy difícil acceder a su información histórica, porque ha cambiado mucho todo. Encontrarte una pagina de hace tres años es bastante difícil. En concreto la idea de guardar páginas de Internet, no sé si es útil, pero desde luego el hecho de que en Internet se ha pensado poco en guardar, sí es verdad. Se ha pensado mucho en publicar y poco en guardar. A la hora de guardar estos documentos es muy importante que estén bien datados.

8- ¿Cuál piensas que debe ser el lenguaje que deben utilizar los periodistas que escriben en un soporte como Internet?

Aún no tenemos desarrollado el lenguaje de Internet. Por ahora lo que hacemos es adaptar los esquemas de los medios tradicionales, pero todavía no hemos tenido tiempo suficiente para habituarnos al medio y desarrollarlo. El ejemplo que siempre se pone es el de la televisión que se sus comienzos emitía con el lenguaje de

la radio. Pues creo que estamos un poco en ese punto. Seguro que dentro de unos años nos reiremos de lo que hacemos ahora. Por poner el ejemplo del periódico, éste tiene un sistema específico de organizar la información. Hay una forma de organizar la portada y en el interior se organiza por categorías. Sin embargo en Internet hay muchísimas formas posibles de organizar. Yo puedo ordenar un periódico alfabéticamente, por el titular, por el autor o por el comienzo del párrafo. Puedo ordenar un periódico de mayor a menor extensión o de mayor a menor importancia o por categorías mezcladas, o por localización geográfica. Se puede disponer de todas estas formas de organizar a la vez. Esto no existe en ningún otro medio.

También tenemos que tener en cuenta a la hora de escribir en Internet que la relación que tiene el usuario con el medio es muy individual. Está una persona sola delante de un ordenador. No me imagino a alguien de la anterior época del PSOE entrando en la sede de Ferraz con un diario *El Mundo* debajo del brazo, porque le daría vergüenza. En cambio lo puede leer en Internet porque nadie le ve. Creo que este factor no le estamos teniendo en cuenta. Tampoco estamos teniendo en cuenta el tiempo para trabajar. Ahora hacemos la misma información para el web a las diez de la mañana que a las siete de la tarde. Y la pregunta es ¿es lo mismo lo que necesita alguien a las diez de la mañana que a las siete de la tarde?. La respuesta es no, y no lo estamos haciendo. Seguimos con una programación fija. La radio ya ha resuelto este problema pero nosotros aún ni nos lo hemos planteado y creo que es algo que deberíamos empezar a pensar.

9- ¿ Cree que toda la información que se ofrece en la Red debe de ser gratuita?. En caso negativo ¿En qué casos piensa que debería ser gratuita y en qué casos de pago?

Hay que tener en cuenta que la información es un bien que cuesta mucho producir por lo que nunca es gratuita porque siempre hay alguien que paga eso. Puede ser gratis para el lector, pero nunca es gratis. La información gratis para el lector es un modelo que nace con Internet por dos motivos principales. Por un lado una carrera

por ganar audiencia y por otra parte debido a una corriente de pensamiento, muy ligada a la cultura *hacker*, de que la información tiene que ser gratuita y estar siempre disponible para todo el mundo. Ha habido algunos fracasos en las primeras épocas de modelos de pago, que llevaron a retroceder en esta tentativa. En cuanto a que si este modelo es sostenible en el tiempo, yo creo que en general no. Sigue siendo costoso producir. Se puede llegar a modelos de poco coste, por ejemplo el volcado del diario *El País*. Pero en un plazo corto también se necesitaron incluir informaciones y otros elementos específicos para la edición digital lo que acarrea un coste mayor.

El dilema es que nadie le pone el cascabel al gato, es decir, no se juntan varios ni será fácil que se junten. Puede ser también que el modelo gratuito funcione como ha funcionado en algunas televisiones. Yo sospecho que vamos a un modelo como el de la televisión, es decir, que aunque haya una gran parte que sea gratuita, también haya parte de pago. Lo que no se puede cobrar es por lo que tienen todos los demás.

Antes tenías limitaciones físicas que te hacían ser el dueño de un mercado pero ahora no existen. Por ejemplo en la prensa tú eres dueño del mercado y puedes cobrar porque la información que das es diferente en algún sentido a la del resto de los medios. El lector decide entonces si compra *El País*, *ABC* o cualquier otro. Sin embargo en Internet tienes disponible información muy similar pero gratuita. Con lo cual si alguien cobra por una información que el de al lado da gratis está clara cual va a ser la elección del lector. El factor diferencial que le puedas añadir tiene que ser muy

grande para que realmente el lector pague. Encontrar cuál puede ser ese factor diferencial es difícil. Es indudable que influirá el nivel de calidad de las informaciones. Pero el punto en el que pagas por cambiar de nivel de calidad ¿cuál es?. Un valor añadido que podríamos cobrar es la información del archivo. Pero en este tema podríamos encontrarnos con un problema de derechos de autor. Hay muchos sitios que están haciendo negocio porque les cuesta poco producir información; el motivo es que la roban.

10- El periodismo a través de Internet esta prácticamente en su fase inicial, ¿Cuál piensa que será su evolución?

Por una parte se facilitará el trabajo y por otra su evolución estará íntimamente ligada a que desarrollemos el lenguaje. En un futuro, los lectores tendrán la posibilidad de personalizar la información que desean leer. Y yo creo que lo harán pero con algún matiz. Sólo habrá un porcentaje muy pequeño que utilizará todas las herramientas que el medio en cuestión ponga a su disposición, para hacer un periódico muy personalizado. La mayoría de la gente hará una selección más genérica. Esto estoy casi seguro que va a suceder así. Un ejemplo que para mí ha sido muy revelador es el de la personalización actual del ordenador por parte del usuario. Hoy día cuando alguien compra un ordenador lo puede personalizar muchísimo. La pregunta es ¿la gente lo hace?. La respuesta es no. La gente va a por el modelo. El ejemplo llevado al límite es lo que ha hecho Apple con los Macintosh de colores. Aquí, directamente eliges un Macintosh azul, naranja, verde...Quiero decir con esto que normalmente el usuario busca un perfil y se acopla a él. Ésto es lo que ocurre con la gente en general, aunque haya un porcentaje pequeño que si personalice lo más posible.

ENTREVISTA A MARILÓ RUIZ DE ELVIRA - DIRECTORA DE CONTENIDOS DEL PORTAL DE PRISA INICIA Y EXDIRECTORA DE LA EDICIÓN DIGITAL DEL DIARIO *EL PAÍS*:

1- ¿Cuáles piensa que deben ser las aptitudes que debe de tener un periodista que trabaja en un medio digital como es Internet?

Lo fundamental es que no tenga miedo a la tecnología. Se debe considerar periodista ante todo, es decir, saber en qué consiste ser periodista. Un periodista digital tiene que tener la mentalidad de que las cosas se pueden contar con muchos elementos – multimedia, audio, vídeo, enlaces-, no solamente con texto. Es necesario que le guste enriquecer la información, que no le baste solamente con lo que haya cubierto y que vaya sólo su firma, sino que de alguna forma no le importe hacer enlaces a otros medios o a otros *webs*, con cosas que él no tiene por que contar porque ya están contadas en otro sitio. Ésta es una de las ventajas de Internet y hay que aprovecharla. Un periodista digital debe de pensar también mucho en el usuario y cómo éste va a navegar por esa información y también en el soporte; no es lo mismo que se esté informando a través de un *wap* que a través de un *web*. La información la debe hacer lo mas corta posible, pero eso es lo mismo que en la prensa tradicional.

La información tiene que ser concisa y responder a las preguntas que el periodista cree que el lector se puede hacer no siempre con texto, sino también con gráficos y otros elementos multimedia que pudieran suponer un valor añadido para el usuario-lector. Hay que tener en cuenta que Internet abre muchas más posibilidades a la hora de expresarse. Tampoco debe importarle al profesional de la información poner enlaces a *webs* que están en inglés o en otros idiomas porque siempre habrá alguien que lo agradezca. En definitiva, todo lo que sea enriquecer la información con otras cosas que haya en Internet, hay que hacerlo.

2- ¿Qué ventajas e inconvenientes cree que aporta Internet al periodismo tradicional?

Desde el punto de vista empresarial aporta muchísimas ventajas porque reduce costes en papel, distribución, en imprenta, etc. Desde el punto de vista del periodista la ventaja es que puede expresarse con más elementos que el puro texto; que su información está disponible instantáneamente, que está accesible 24 horas al día desde todo el mundo; que se puede almacenar en base de datos para leer más adelante, y por supuesto que puede buscar información por Internet, siempre que sepa donde la busca.

Con respecto a los inconvenientes, yo no le veo ninguno. Acaso la fiabilidad de las fuentes a las que puedas recurrir a través de Internet, pero ni siquiera. Yo creo que hay suficientes guiños en una pagina como para saber si lo que estás viendo es verdad o no lo es. Y además esto es igual que en la prensa escrita, el periodista necesitará al menos dos fuentes que digan lo mismo. Siempre se deben comprobar las fuentes, sea el medio que sea.

3- ¿A qué recursos de búsqueda suele recurrir cuando quiere buscar algo en la Red?

Uso mucho Copernic, Google, Buscopio y Altavista. Uso Google porque pienso que es muy bueno y muy rápido. Copernic porque te busca en muchos buscadores a la vez y tiene muchas utilidades, como guardar las búsquedas. Altavista porque, para cosas generales, es con el que me entiendo mejor.

4- Según su experiencia profesional ¿Cuál piensa que es el mejor recurso de búsqueda para el periodista digital?

Depende. Yo creo que hay que recurrir a buscadores especializados, para hacer informaciones especializadas. Hay que saber buscar bien. Utilizando las propias herramientas de los buscadores y otros recursos se puede acotar bastante la búsqueda.

Otro de los mejores recursos de búsqueda en Internet son los propios lectores. Eso se ve mucho en los foros; un usuario lanza una pregunta y siempre hay alguien que le contesta. Hay que tener en cuenta que a ese foro puede estar accediendo gente

del mundo entero, por lo que la posibilidad de respuestas a cualquier tipo de preguntas específicas es muy alta.

Como ejemplo de esto, yo organicé un viaje cuando estaba en *El País Digital* a las redacciones de los periódicos más importantes de EEUU, y lo hice todo directamente por Internet. No recurrí a la secretaria para nada. Además empezamos los directores a escribirnos asiduamente, aparte de sobre cosas del trabajo, también sobre cosas de nuestras familias, otras cosas personales, o simplemente tonterías o curiosidades, cosa que no hubiéramos hecho por teléfono. Cuando llegamos allí parecía como si nos conociéramos de toda la vida. Esto igual no habría sido posible si lo hubiéramos hecho por teléfono. Siempre encuentras un hueco para escribir un e-mail, al que le dedicas un tiempo específico del que la persona dispone en ese momento. Pero si se llama por teléfono, la llamada pasará por una secretaria o puede que la persona en cuestión esté en una reunión u otras mil cosas. A través del correo electrónico llegas directamente a la persona que te interesa y tienes mucha más confidencialidad.

Saliéndonos del tema profesional, pero que también ilustra la alta capacidad de contactos que se puede tener a través de este nuevo medio, yo he hecho un montón de amigos por Internet. El 60 por ciento de la gente que tengo guardada en mi libreta de direcciones, los he conocido a través de Internet o gracias a Internet.

5- ¿Me podría hacer una valoración de las listas de distribución, boletines, bases de datos y otros recursos no accesibles mediante buscadores generales?

Las bases de datos me parecen fundamentales, pero hay que saber donde encontrarlas. Las listas de distribución personalmente no me gustan nada. Pienso que están muy bien para gente especializada en un tema muy concreto.

6- Está bastante claro que las posibilidades de Internet como medio de información del profesional de la información son ilimitadas, pero a tantas ventajas cabe añadirle el inconveniente del anonimato o la incierta

identidad de la mayoría de dichas fuentes. ¿Qué opina de la fiabilidad de las fuentes en Internet?

Yo no veo tan difícil la credibilidad de las fuentes por Internet. Al final lo que van a predominar son las marcas importantes o alguien que se hace una marca a si mismo. El problema es que te hagan un web paralelo. Por ejemplo: que entres en El País.com y hayan hecho una copia de *El País digital* tergiversando todos los textos. De todas formas creo que la gente es lo suficientemente racional como para darse cuenta si una cosa es o no es falsa. En el momento en que ves que te están dando gato por liebre, simplemente no vuelves a entrar. Eso un periodista se da cuenta sin problemas. Y en el caso de que no se dé cuenta tampoco es mucho problema ya que siempre hay que comprobar las fuentes.

7- Hay un archivo digital llamado “The Internet Archive” que se dedica a almacenar todas las webs que encuentra, incluyendo información hallada en los grupos de discusión, para preservarlas antes de su posible desaparición. Su objetivo es servir como fuente de documentación para futuros trabajos de investigación. ¿Opina que es un trabajo útil o en su mayor parte inservible debido a la poca fiabilidad de la mayoría de los documentos allí guardados?

Sobre todo pienso que será útil para hacer investigaciones o tesis, o para los documentalistas. No creo que vayan a tener ninguna otra utilidad. Entre otras cosas porque las cosas evolucionan tanto que un web que ha salido hace tres años ahora a ti no te vale de nada.

8- ¿Cuál piensas que debe ser el lenguaje que deben utilizar los periodistas que escriben en un soporte como Internet?

El lenguaje de Internet se va a internacionalizar, de alguna forma. Va a haber una serie de símbolos o de signos, pero a parte una serie de palabras inglesas sobre todo, para comunicarse. En Internet cuanto más breves, más despiezados y más incisivos

sean los textos mejor. Y por supuesto aquí creo que sigue valiendo la frase de “una imagen vale más que mil palabras”. Debe de haber un texto principal corto y enlaces o links que te lleven a textos secundarios que completen la crónica. También debería haber enlaces en palabras del texto que no fueran muy conocidas en las que al pulsar se abriera una ventana flotante con la explicación de la palabra en cuestión. Lo más importante es que el periodista piense lo que al lector le interesa.

9- ¿Cree que toda la información que se ofrece en la Red debe de ser gratuita?. En caso negativo ¿En qué casos piensa que debería ser gratuita y en que casos de pago?

Creo que la información en Internet no debería ser gratuita porque siempre hay un coste a la hora de buscar información y pagar a los periodistas. Mientras tengan un soporte papel que les compense estos gastos bien, pero sino ¿cómo pagas a una redacción?. No está nada claro que se pueda financiar con publicidad y lo lógico es que si es información válida y con valor añadido, la gente que le interese estará dispuesta a pagar por ella. Pienso que a medio plazo se empezará a cobrar. Entre otras cosas porque como acaben siendo sólo redacciones digitales ¿cómo pagas un corresponsal en Moscú.

10- El periodismo a través de Internet esta prácticamente en su fase inicial, ¿Cuál piensa que será su evolución?

Pienso que los periodistas van a jugar un papel muy importante en la evolución de Internet. La gente siempre va a estar buscando información y no información en bruto. Va a ser labor de los profesionales de la información buscar con exactitud informaciones específicas, procesarla, organizarla, escribirla y presentarla. Internet ha abierto un gran campo para los periodistas. Lo importante es acabar dando, a través del medio digital, información de calidad. Es la única forma de fidelizar al

usuario en tu web, y los que pueden hacer esto son los periodistas. Internet, en definitiva, es y será comunicación. Tener fácil acceso inmediato a otras personas, a otras formas de escribir u otras opiniones.

ENTREVISTA A RICARDO FORNÁS – CREADOR DEL BUSCADOR DE BUSCADORES “BUSCOPIO” ([HTTP://WWW.BUSCOPIO.COM](http://www.buscopio.com)):

1- ¿Con qué objetivo se creó Buscopio?

Buscopio, buscador de buscadores surgió de la necesidad de dotarse de una página desde la cual se pudiera localizar toda una extensa recopilación de motores de búsqueda, de este modo, pretendía ahorrarme tiempo en todos los procesos para localizar información, existían algunas páginas en ese sentido pero todas estaban totalmente orientadas hacia el mundo anglosajón y tampoco llegaban a satisfacerme del todo porque en el fondo eran, y son, muy localistas. Como Internet ofrece la oportunidad de compartir ideas, eso mismo hice y lo que en principio era una guía de direcciones para uso privado decidí publicarla en la Red.

2- ¿Cómo es el funcionamiento interno de Buscopio?

Buscopio se configura como un director donde todas las direcciones se encuadran en una clasificación temática y también dispone de un motor de búsqueda interno con la particularidad de que cuando se interroga a la base de datos esta busca en el título, url, descripción y descriptores asociados a cada registro. Tanto la descripción como los descriptores son criterios que aplico en base a mis conocimientos y a los propios contenidos de cada buscador.

3- ¿Cómo sería, según su opinión, el buscador ideal?

El buscador ideal sería aquél que ofreciese respuestas relevantes siempre que le planteáramos cualquier cuestión. En ese caso debería ser un buscador donde prácticamente abarcase el 95% de los contenidos globales de la Red, los que existen ahora apenas abarcan un 15% o menos, fuese rápido y dispusiese de la capacidad de comprender exactamente cuál es la pregunta que le formulamos.

4- ¿A qué recursos de búsqueda suele recurrir cuando quiere buscar algo en la Red?

Las fuentes que utilizo se caracterizan por su extremada variedad, entre otras cosas porque según lo que uno busque deberá seguir una estrategia diferente. No es lo mismo buscar cuestiones legislativas, datos sobre un actor, una sonata de Beethoven o imágenes sobre murciélagos.

5- Según su experiencia profesional ¿Cuál piensa que es el mejor recurso de búsqueda para el periodista digital?

Esta respuesta viene asociada con la anterior. No hay un único recurso que sea capaz de satisfacer todas las peticiones de información. En todo caso, debería guiarse por el prudente criterio de selección de direcciones realizada por profesionales y especialistas en documentación. Si no sabemos cuáles son las mejores páginas de nanotecnología deberemos optar por un criterio más o menos lógico y no guiarnos exclusivamente por la primera respuesta que nos otorgue cualquier buscador. El factor humano aún no ha podido ser desplazado por procesos automatizados.

6-¿Qué opina del papel de los propios usuarios como herramienta de búsqueda? (Foros, listas de distribución, etc)

En principio en Internet todos pueden opinar de todo pero se impone lógicamente la prudencia en considerar a los usuarios como fuente de primera mano. En esto interviene mucho el sentido común porque las intervenciones en este tipo de áreas de participación sólo pueden considerarse como punto de referencia ya que en algunos casos es muy difícil verificar la información. No es lo mismo leer recomendaciones sobre cómo comprimir una imagen de ordenador, frente a otro internauta que comunique la inminente fusión empresarial entre dos compañías eléctricas.

7-¿Me podría hacer una valoración de la llamada Internet Invisible? Es decir, de aquella parte de Internet no accesible mediante los buscadores.

El fenómeno de la Internet Invisible simplemente quiere decir que es aquellos contenidos de la Red que quedan fuera del alcance habitual de los buscadores. Para ser más precisos habría que decir que no existiría una única Internet Invisible sino tantas como buscadores existen. Es decir, aquella información que un buscador está imposibilitado de recopilar. Pero es probable que los contenidos en chino de un buscador genérico como Google o AltaVista sean inferiores a los de un buscador originario de ese país. Ciertamente es muchísimo el volumen de información que queda fuera de las áreas de acción de un buscador pero ello no quiere decir que esa misma información sea imposible de encontrar porque esta puede localizarse en bases de datos que por sus especiales características no son accesibles directamente a un buscador pero sí puede hacerlo un internauta si sabe acceder a ella. Existen muchos ejemplos como el archivo histórico de un periódico, una base de datos de una filmoteca nacional, etcétera. En todo caso los buscadores únicamente apuntarían a la dirección del periódico o de la filmoteca nacional pero no a los contenidos de las mismas y a partir de aquí empieza a surgir el fenómeno de Internet Invisible.

8-En la actualidad para cualquier empresa que tenga pagina web hay una gran diferencia entre aparecer el primero como resultado de una búsqueda o el décimo. ¿Piensa que llegara un momento en que estas empresas paguen a los buscadores por aparecer en los primeros puestos?

Si fuese por parte de las empresas haría tiempo que este procedimiento ya se habría aplicado, pero afortunadamente los buscadores deben mantener un criterio imparcial cuando presenta los resultados de una búsqueda porque su única fuerza, lo que le mantiene en los altos puestos en el ránking de visitas es por la credibilidad y la aprobación que goza entre los internautas. Son ellos los que sostienen el valor de un buscador frente a otro. Si como usuario, cada vez que busco información, ésta me es absolutamente irrelevante no volveré a visitar ese buscador. Si busco información sobre Marte y los primeros puestos están copados por empresas de venta de productos relacionados con el planeta rojo

(camisetas, fotografías, carteles, etcétera) no me aportan valor informativo directo y lo que es una herejía en Internet, me están haciendo perder el tiempo.

9-Está bastante claro que las posibilidades de Internet como medio de información del profesional de la información son ilimitadas, pero a tantas ventajas cabe añadirle el inconveniente del anonimato o la incierta identidad de la mayoría de dichas fuentes. ¿Qué opina de la fiabilidad de las fuentes en Internet?

En este caso se deben aplicar los mismos criterios que en la vida real. Usted como periodista siempre deberá verificar sus fuentes, ¿publican los periódicos informaciones realizadas por un informante anónimo?. Otro tanto sucede con la credibilidad a aplicar ¿quién es más fiable para cuestiones de salud un curandero o un médico, una página sobre el cáncer de un paciente o de un centro de investigación oncológica?. Por tanto no deberían existir grandes problemas respecto al tema. Si un científico propone una teoría bien sea por Internet o en un congreso será cuidadosamente escrutada y comprobada. Por ello, el medio en que se divulgue es lo de menos.

10-¿Cuál piensa que será la evolución de los buscadores y de los sistemas de búsqueda en general en Internet?

La lógica evolución sería insistir en dos aspectos que están íntimamente unidos: una mayor capacidad para comprender el lenguaje natural y una mayor relevancia en las respuestas. Un reto ciertamente difícil porque entre la comunicación de dos personas intervienen algo más que palabras como los gestos o el contexto y aún así la cosa se complica. Todos los días podemos experimentarlo, por ejemplo, a una simple petición en un restaurante de un bistec el camarero probablemente le pregunte a su vez cómo lo quiere, qué guarnición, etcétera. La relevancia va en función de la primera cuestión, si un

buscador no interpreta correctamente la pregunta probablemente la respuesta será errónea y aunque progresivamente se van incorporando técnicas propias de

la documentación para evaluar cada página en su relación con las preguntas planteadas queda mucho por hacer.

Otro aspecto en el que se profundizará más si cabe es en la especialización. Si busco información de geología tendré que acudir a un directorio y buscador especializado y me olvidaré de los buscadores genéricos. Es un problema de crecimiento, Internet tiene que evolucionar y madurar para que progresivamente los contenidos sean mejor organizados en torno a temas de interés afín. Y es posible que de un buscador de geología pasemos a otro más concreto de vulcanología y así sucesivamente.

1.5 OTRAS ENTREVISTAS:

En este apartado el autor ha creído conveniente incorporar una serie de entrevistas, no realizadas expresamente para este trabajo de investigación, pero de significativa relevancia para el tema tratado.

ENTREVISTA A RAMÓN SALAVERRÍA: DIRECTOR DEL LABORATORIO DE COMUNICACIÓN MULTIMEDIA (MMLAB) DE LA UNIVERSIDAD DE NAVARRA

Entrevista realizada por el autor de este proyecto de investigación y publicada en Inicia (Portal de Internet del Grupo Prisa: <http://www.inicia.es>) el 15/10/2000.

Ramón Salaverría es director del Laboratorio de Comunicación Multimedia (MMLab) de la Universidad de Navarra. Ha publicado varios artículos sobre géneros periodísticos y periodismo digital, y ha dictado numerosos cursos y conferencias sobre el tema. Es además miembro de la Academia de los Premios IBest 2000, uno de los certámenes más conocidos de Internet en España, autor de la *Guía de Internet para periodistas* y editor del Boletín MMLab, publicación difundida por correo electrónico y que recopila noticias nacionales e internacionales sobre comunicación.

Internet ha cambiado los hábitos de muchos profesionales que han visto en este nuevo medio una útil herramienta para ampliar perspectivas de futuro. Sin embargo para el periodismo, la Red de redes ha significado, no sólo la apertura de nuevos horizontes, sino una nueva y revolucionaria forma de trabajo.

Ramón Salaverría ve en este cambio grandes oportunidades tanto para el profesional del periodismo, como para el receptor de la información: “En los medios tradicionales es el periodista el que determina lo que el lector, oyente o televidente tiene que ver, pero, gracias a la interactividad que permite Internet, el poder de decisión de cuál es el producto informativo pasa a manos del lector. Ya no es el periodista el que determina cuál es el orden en importancia del producto informativo, sino que es el usuario de la información quien decide cuáles son en cada momento los elementos informativos que más le interesan”, señala.

Prensa, radio y televisión han volcado sus contenidos en Internet, aprovechando su capacidad y facilidad de difusión. Sin embargo, no se debería olvidar la esencia de Internet como medio de comunicación en sí mismo. No se trata simplemente de un

soporte en el que adaptar los medios tradicionales. Se trata de una nueva forma de informar y de distribuir dicha información.

Multimedia e interactividad

Salaverría aporta dos características fundamentales de este nuevo medio. Una tiene que ver con las características **multimedia** del nuevo soporte y otra con la **interactividad**. “El carácter multimedia existía antes también en los medios tradicionales, pero siempre eran medios bimedia. Es decir, un periódico en realidad es multimedia, tiene imagen y texto. Una emisión de televisión también es multimedia en el sentido de que tiene sonido e imagen. La radio es el único medio antiguo que únicamente se limitaba a un único soporte. La novedad más importante que presenta Internet frente a estos medios multimedia más rudimentarios es que en lugar de dos, trae la posibilidad de tres, y además con un factor potencial añadido que es el de la interactividad”, señala.

Cuando se navega por Internet, puedes encontrar páginas web que integran texto, sonido, fotos e incluso vídeos. Pero aún nos queda mucho por ver, como la posibilidad de incluir también olores, técnica que ya ha comenzado a desarrollarse. “A principios de este año 2000 salió la noticia de que había un aparato que conseguía reproducir una serie de códigos olorosos, digitalizar esos impulsos y reproducirlos digitalizados en un receptor, que podría estar en otra parte del mundo que dispusiera de ese mismo receptor. El tacto también se está incorporando en la medida que se están desarrollando los entornos 3D. La exigencia multimedia del nuevo soporte va a permitir que los usuarios de la información tengan una riqueza comunicativa mucho mayor que la que tenían hasta ahora con los medios tradicionales”, afirma Salaverría.

Pero no sólo el carácter multimedia de Internet es importante. ¿A alguno de vosotros le han publicado alguna vez una carta al director? ¿Habéis podido hacer oír vuestra opinión en directo en algún programa de radio o televisión? En los medios tradicionales no son muchas las posibilidades de respuesta del receptor de la información. Tiene su derecho de réplica, por supuesto, pero la mayor parte de las veces no tiene la ocasión de ejercerlo. La interactividad que permite Internet es el elemento

clave en este sentido. En este nuevo medio de comunicación es el usuario el que selecciona el tipo de información que le interesa de forma completamente personalizada. Cada persona es un mundo con sus propios gustos informativos e intereses, que a través de la Red podrán verse satisfechos.

Marshall McLuhan afirmó en su libro *Galaxia Gutenberg* que todo nuevo medio trata, en una primera instancia, de integrar a los medios precedentes y referirse a ellos. Esto resulta muy claro en el caso de Internet, puesto que, como hemos podido observar, tanto la prensa como la radio y la televisión van tomando su sitio en la Red y abriendo sus respectivas sedes telemáticas, olvidando en la mayoría de los casos que Internet posee una estructura interna muy distinta a la de los medios de comunicación precedentes, derivada en parte de su origen, y en parte de su evolución.

No se trata pues de adaptar radio, televisión o prensa en Internet, sino de crear contenidos específicos para un medio completamente nuevo, aún sin nombre, pero con mayores posibilidades para realizar el trabajo periodístico que los medios tradicionales. En este sentido, Ramón Salaverría apunta que el futuro, en un plazo medio de 10-15 años, es la complementariedad de estos cuatro medios. “Un ámbito que antes estaba cubierto por 3 medios ahora va a pasar a estar cubierto por cuatro y me parece que el terreno del periódico de papel será el de la explicación, de la contextualización, del desarrollo, de la profundidad, frente al ámbito de Internet que va a estar marcado fundamentalmente por esta cuestión de interactividad. Va a ser el ámbito de la información puntual, del dato, la última hora y va a ser también Internet un ámbito en el que el usuario va a poder encontrar documentación de amplio espectro”, señala.

El auge de estos medios tradicionales en Internet ha desatado todo un cúmulo de reflexiones sobre cómo presentar de manera más efectiva e idónea la información periodística. Se trata de averiguar la forma de hacer “más atractiva” la interfaz gráfica, cómo mantener la atención del lector-internauta y otras especificaciones referidas a aspectos técnicos.

Para Salaverría, la clave del éxito del producto periodístico estará en saber dar los contenidos que realmente interesan. “Si uno es capaz de dar al lector estos contenidos es

cuando va a conseguir que el lector se vuelque en la información. El lector en aquellos asuntos que realmente le interesan no se contenta con 200-300 palabras que una noticia de prensa le ofrece. Ahí no hay posibilidad de profundización.

Además, el usuario no se quedará, como ocurre en los medios tradicionales, con un solo periódico, no existe esa especie de lector cautivo que puede tener un periódico. Uno se interesa por la información no por el medio que se la da. Esto también va a suponer una competencia mediática muy grande, porque un día será un determinado periódico el que dé mejor la información y la gente irá a ese medio porque no tiene más que hacer un clic para encontrar esa información. Además, la exigencia profesional que van a tener estos nuevos medios con los periodistas va a promover la mejora de la calidad de los contenidos”, añade.

La búsqueda de información a través de la Red

Cuántas veces se intenta buscar algo en Internet y no se consigue; simplemente se encuentran con multitud de enlaces que tienen en realidad muy poco que ver con lo que puntualmente se está buscando. Sin embargo en la actualidad prácticamente todo tipo de información está disponible en Internet. ¿Qué hacer?

La World Wide Web nos presenta una especie de exposición de fuentes universal actualizada (aunque a veces no tanto). Podemos recorrer universidades, centros de investigación, los laboratorios de la NASA, las ediciones de hoy de los diarios de todo el mundo. Y a la hora de establecer un posible contacto, sólo es cuestión de enviar un mensaje, mediante el correo electrónico.

Sin embargo son frecuentes las quejas del usuario de Internet en las que afirma que en la Red hay mucha “basura”. En este sentido Salaverría señala que “la clave para encontrar lo que buscas en Internet es determinar cuál es el recurso de búsqueda pertinente en cada caso. Lo mismo que uno no busca un teléfono de un particular en las páginas amarillas, ni busca el teléfono de una determinada empresa en las páginas blancas, hay que saber determinar muy bien cuál es el tipo de buscador o directorio

pertinente para cada caso”.

Para este propósito, existen distintos tipos de buscadores según la necesidad del usuario. Algunos de éstos nos permiten las **búsquedas de tipo temático**. Cuando uno quiere enterarse de algún tema en general puede ser lo más pertinente ir a un buscador tipo Yahoo, Galaxy, Magellan, etc. Cuando, en cambio, uno está buscando un asunto puntual, un dato concreto en cualquier parte de Internet o dentro de un determinado país, entonces es el momento de usar los **buscadores automáticos**, como Altavista, Excite, Hotbot, Infoseek. Existe un tercer tipo llamado **Metabuscaador** cuya función es la búsqueda simultánea en varios buscadores automáticos. Entre éstos, destacan: Cyber411, SavvySearch, Dogpile y MetaCrawler.

El papel del periodista digital

En la actualidad, el papel del “periodista digital” es fundamentalmente de gestión de contenidos, de coordinación con los otros medios, pero, según Salaverría, no se apuesta informativamente por Internet.

Experiencias como las recientes elecciones en EEUU son las que marcan las pautas del futuro papel de los periodistas de medios digitales. El hecho de que la CNN en Internet tuviera 10 millones de visitas por hora durante la jornada de votación de estas elecciones, significa que ahí se estaban dando contenidos. “Esa demanda que está siendo generada por la propia audiencia va a tener que ser satisfecha de algún modo, lo que va a significar elaboración de contenidos. Por lo tanto, cuando se establezcan las rutinas de trabajo de los medios digitales, la potencialidad de creación de información va a ser mucho mayor. El papel de periodista será entonces coordinar todos esos materiales, contextualizarlos y mantenerlos perfectamente actualizados”, señala Salaverría.

Entre sus funciones el periodista digital tendrá que elaborar mucha mayor cantidad y calidad de información, además de poner a disposición del usuario las fuentes de información que ha utilizado para su trabajo. Los profesionales de la información intentarán, en definitiva, buscar nuevas formas de narrar que se ajusten a las necesidades y capacidades del medio, manteniendo como punto esencial de partida la interactividad con el lector.

LEONARD KLEINROCK¹⁶⁴ • 'PADRE' DE LA RED: "INTERNET ESTÁ EN LA EDAD DE PIEDRA"¹⁶⁵

Hace 30 años un joven erudito informático pulsó una letra en el teclado de un ordenador y la L apareció en la pantalla de otro ordenador que estaba unido al primero por un conmutador del tamaño de una nevera. En ese momento nació Internet. El *padre* de la Red se llama Leonard Kleinrock, profesor de Ciencia Informática en la Universidad de Los Ángeles (UCLA) e inventor del lenguaje que permite a los ordenadores la comunicación fluida en la que se basa Internet. En una entrevista con *El País*, Kleinrock dice que Internet todavía está en una etapa primitiva de su existencia.

Hace 30 años era una idea de locos: cuatro científicos encerrados en un laboratorio de la UCLA tratando de inventar un cordón umbilical entre dos ordenadores. Su proyecto estaba financiado por el Gobierno de EE UU, lanzado al desarrollo de nuevas tecnologías una vez que había puesto su huella en la Luna. Detrás de la iniciativa estaba el Pentágono, que soñaba con la creación de un sistema de comunicaciones que le permitiese la interconexión de todos sus ordenadores. Paradójicamente, los orígenes militares de Internet contribuyeron a que la Red naciera con un carácter abierto. Sin fronteras. Para el ejército era un requisito imprescindible que la Red permitiese mantener unidos todos los ordenadores, incluidos los que estaban situados en las bases de EEUU en otros países. En esa época el objetivo era construir un mecanismo de comunicación con los aliados capaz de funcionar incluso en un escenario de guerra nuclear.

"No hay duda de que Internet ha crecido enormemente desde el día en que nació", dice Kleinrock, de 65 años, antes de agregar: "Pero a finales de este milenio mi conclusión es que todavía estamos en la Edad de Piedra. Nadie nos iba a decir en 1969 que 30 años después íbamos a acceder a aquella red primitiva desde nuestros hogares y nuestros

¹⁶⁴ Leonard Kleinrock es profesor de Ciencia Informática en la Universidad de Los Ángeles (UCLA) e inventor del lenguaje que permite a los ordenadores la comunicación fluida en la que se basa Internet.

¹⁶⁵ Entrevista realizada por Javier del Pino, corresponsal en Washington, D.C del diario *El País*, y publicada en la edición del jueves 21 de octubre de 1999.

despachos, a todas horas y desde cualquier parte del mundo. Ello nos lleva a pensar que la tecnología de Internet todavía puede avanzar enormemente en todos los terrenos: en el trabajo, en el juego, en el entretenimiento, en la vida...".

En realidad los cuatro *padres fundadores* de la Red no se ponen de acuerdo sobre cuál debe ser la fecha en que se celebre el cumpleaños de Internet, aunque sólo usan esta discusión para bromear y recordar la época. Vint Cerf y Larry Roberts mantienen que Internet nació el 2 de septiembre, cuando se unió un ordenador a un conmutador. Robert Khan y Kleinrock, más puristas, sostienen que la fecha clave fue el 20 de octubre, cuando por primera vez dos ordenadores se comunicaron entre sí.

Esto ocurrió de la siguiente manera, según lo relata Kleinrock: "Teníamos al teléfono a un tipo sentado ante un ordenador en la UCLA y a otro en la Universidad de Stanford. Cuando todo estaba preparado, el de UCLA comenzó a escribir en su ordenador la palabra *log* (la orden de entrada al sistema). Pulsó la L y preguntó al otro "¿Te llega la L?". El de Stanford respondió: "Tengo la L". Hicieron lo mismo con la O y "en ese momento, todo el sistema se vino abajo". Su primera valoración fue poco triunfalista, pero el experimento desembocó en la invención del lenguaje llamado TCP/IP que permite la existencia y funcionamiento actual de Internet.

Kleinrock trabaja 10 horas en el despacho y cuatro en casa. Dedicar dos a hacer ejercicio y tres a estar con su familia, lo que sólo le deja cinco para dormir, suficientes para alguien que ahora dedica su intelecto a crear ordenadores permanentemente conectados a Internet, a alta velocidad y sin cables ni teléfonos. A veces disfruta "del placer de leer el periódico, un libro o una carta", porque hay "algo maravilloso, humano" en el papel que no se reemplaza con nada.

Kleinrock recuerda el *boom* de la Red en la última década. "La mayoría de la gente", dice Kleinrock, "piensa que simplemente accede a Internet desde su mesa, conectando su ordenador. Pero en realidad están haciendo algo más. Cuando uno se conecta lo que hace es viajar por el mundo a través de culturas, países y personas". No hay ningún control en Internet, ni puede haberlo -"la pornografía en la Red es un ejemplo", afirma Kleinrock-, y recuerda: "Cuando lo inventamos hablábamos de conmutadores y

kilobytes, no de si el niño iba a estar haciendo los deberes o entrando en páginas pornográficas". Pero Kleinrock ve un lado humano en la tecnología y expone un hecho objetivo: Internet sólo despegó con el auge del correo electrónico, cuando la comunicación entre dos ordenadores permitió la comunicación entre dos seres humanos.

ANEXO 3

CRONOLOGÍA DE INTERNET DE HOBBS

CRONOLOGÍA DE INTERNET DE HOBBS

1957

La Unión Soviética lanza el Sputnik, el primer satélite artificial. En respuesta a este hecho, Estados Unidos crea el ARPA (Organismo de Proyectos de Investigación Avanzada) dentro del Ministerio de Defensa a fin de establecer su liderazgo en el área de la ciencia y la tecnología aplicadas a las fuerzas armadas (:amk:).

1961

Leonard Kleinrock, Massachusetts Institute of Technology MIT (Instituto tecnológico de Massachusetts): "Flujo de Información en Redes Amplias de Comunicación" ("Information Flow in Large Communication Nets") (Julio)

- Primer documento sobre la teoría de conmutación por paquetes (PS).

1962

J.C.R. Licklider y W. Clark, del MIT: "Comunicación hombre - computadora en línea" ("On-line Man Computer Communications")(Agosto)

- Concepto de *Red Galactic*(*Galactic Network*) que abarca interacciones sociales distribuidas.

1962

Paul Baran, RAND: "Redes de Comunicación Distribuida"("On Distributed Communication Networks")

- Redes conmutadas por paquetes, sin punto único de interrupción.

1965

El ARPA promueve un estudio sobre "Redes cooperativas de computadoras de tiempo compartido".

- El TX-2 en el laboratorio Lincoln del MIT y el AN/FSQ-32 de la System Development Corporation (Santa Mónica, California) quedan vinculadas directamente (sin conmutación por paquetes) por medio de una línea telefónica dedicada de 1200 bps; más tarde se agrega la computadora de la Digital Equipment Corporation (DEC) en ARPA y así conforma la red experimental ("The experimental Network")

1966

Lawrence G. Roberts, del MIT: "Hacia una Red Cooperativa de Computadoras de tiempo Compartido" ("Towards a Cooperative Network of Time-Shared Computers") (Octubre)

- Primer plan de ARPANET, Red del Organismo de Investigaciones Avanzadas (Advanced Research Projects Agency Network).

1967

Larry Roberts lleva a cabo las negociaciones sobre el diseño de ARPANET en la asamblea ARPA IPTO PI en Ann Arbor, Michigan (Abril)

ACM Simposio sobre Principios Operativos en Gatlinburg, Tennessee (Octubre)

- First design paper on ARPANET published by Larry Roberts: "Multiple Computer Networks and Intercomputer Communication"
- Se lleva a cabo la Primera Asamblea sobre los tres equipos independientes de redes por paquetes (RAND, NPL, ARPA)

El Laboratorio Nacional de Física (National Physical Laboratory)(NPL) en Middlesex, Inglaterra desarrolla la red NPL Data Network supervisada por Donald Watts Davies quien introdujo el termino "paquete". La red NPL, un experimento en conmutación por paquetes utilizaba, líneas telefónicas de 768 kbps.

1968

Se presenta la red conmutada por paquetes (PS - Network) ante el ARPA.

En Agosto se envían los pedidos de propuestas para ARPANET; las respuestas se reciben en Septiembre

En Octubre le es otorgado a la Universidad de California Los Angeles (UCLA) el contrato para el Centro de evaluación de Redes (Network Measurement Center)

A Bolt Beranek y Newman, Inc. (BBN) le es otorgado el contrato de Conmutación por Paquetes a fin de crear una Interfase Procesadora de Mensajes (Interface Message Processors)(IMPs)

El senador estadounidense Edward Kennedy envía un telegrama de felicitación a BBN por su contrato de ARPA por un millón de dólares para crear el Procesador de Mensajes "Interfaith". También les agradece por sus esfuerzos

El Network Working Group (NWG)(Grupo de trabajo de redes), liderado por Steve Crocker, se organiza a fin de desarrollar protocolos a nivel host para establecer comunicaciones en ARPANET. (:vgc:)

1969

El Ministerio de Defensa designa a ARPANET para la tarea de investigación de redes.

Se ponen en servicio los nodos a medida que BBN construye cada IMP [Honeywell DDP-516 con 12 K de memoria]; AT&T provee líneas de 50 kpbs

Nodo 1: UCLA - Universidad de Los Ángeles, California. (30 de Agosto)

- *Función:* Centro de evaluación de redes.
- *Sistema, Sistema operativo:* SDS SIGMA 7, SEX.

Nodo 2: Instituto de Investigaciones de Stanford.(SRI) (1 de Octubre)

- Centro de Información de Redes (Network Information Center)(NIC)
- SDS940/Genie
- Proyecto de Doug Engelbart sobre "Debate sobre el intelecto humano" ("Argumentation of Human Intellect")

Node 3: Universidad de California Santa Barbara (UCSB) (1 de Noviembre)

- Matemática Interactiva de Culler - Fried.
- IBM 360/75, OS/MVT

Nodo 4: Universidad de Utah. (Diciembre)

- Gráficos.
- DEC PDP-10, Tenex

Primera Solicitud de Comentarios (RFC): "Host Software" por Steve Crocker.

RFC 4: Cronología de la Red

Los primeros paquetes se enviaron por Charley Kline en la UCLA tratando de conectarse al SRI. El primer intento resultó en un colapso del sistema en el momento en el que se ingresó la letra G de la palabra LOGIN. (Octubre)

La Universidad de Michigan, el estado de Michigan y La universidad del Estado de Wayne establecen una red Merit con base X.25 para los estudiantes, el cuerpo docente y los ex alumnos (:sw1:)

1970

Aparece el primer informe sobre ARPANET en AFIPS: "Desarrollo de Redes de Computadoras para lograr recursos compartidos" ("Computer Network Development to Achieve Resource Sharing") (Marzo)

La red ALOHAnet, la primera red de radio por paquetes, desarrollada por Norman Abrahamson, de la Univ of Hawaii, comienza a operar (Julio) (:sk2:)

- Se conecta a la ARPANET en 1972.

Las hosts de ARPANET comienzan a utilizar los Protocolos de Control de Redes (NCP). Primer protocolo host-a-host

AT&T instala el primer vínculo costa a costa entre la UCLA y BBN a 56 kbps. Está línea fue reemplazada más tarde por otra entre BBN y RAND. Se agrega una segunda línea entre MIT y Utah

1971

15 nodos (23 hosts): UCLA, SRI, UCSB, Universidad de Utha, BBN, MIT, RAND, SDC, Harvard, Laboratorio Lincoln, Stanford, UIU(C), CWRU, CMU, NASA/Ames.

BBN comienza a crear IMPs utilizando los Honeywell 316 más baratos. Aunque las IMPs están limitadas a 4 conexiones a hosts y por lo tanto BBN desarrolla una terminal IMP(TIP) que soporta hasta 64 hosts (Septiembre)

Ray Tomlinson de BBN inventa un programa de correo electrónico para mandar mensajes en redes distribuidas. El programa original es producto de otros dos:

un programa interno de correo electrónico (SENDMSG) y un programa experimental de transferencia de archivos (CPYNET) (:amk:irh:)

1972

Ray Tomlinson (BBN) modifica el programa de correo electrónico para ARPANET donde se transforma en un éxito. Se elige el signo @ entre los signos de puntuación de la máquina de teletipos Tomlinson Modelo 33 para representar el "en" (Marzo)

Larry Roberts crea el primer programa de administración de correo electrónico (RD) para listar, leer selectivamente, guardar, re enviar y responder mensajes. (Julio)

Se lleva a cabo la conferencia Internacional sobre Comunicaciones por Computadora (International Conference on Computer Communications (ICCC)) en el Hilton de Washington D.C. con una demostración de ARPANET entre 40 máquinas y el Procesador Terminal Interfase (TIP) organizada por Bob Kahn. (Octubre)

Se lleva a cabo el primer chat entre computadoras durante ICCC mientras que el psicólogo PARRY (en Stanford) discutía sus problemas con el Doctor (en BBN)

En Octubre se crea el Grupo de Trabajo de Redes Internacional (International Network Working Group (INWG)) como resultado de una Asamblea en ICCC en la que surge la necesidad de realizar un esfuerzo conjunto a fin de lograr un avance en las tecnologías de redes. Vint Cerf fue el primer Presidente. En 1994, INWG se convierte en IFIPWG 6.1 (:vgc:)

Louis Pouzin lidera el proyecto francés de crear su propia ARPANET - CYCLADES

RFC 318: Especificación Telnet

1973

Se realizan las primeras conexiones internacionales a la ARPANET: University College of London (Inglaterra) y el Royal Radar Establishment (Noruega).

De la tesis PhD de Bob Metcalfe para el doctorado en Harvard surge la idea para Ethernet. Este concepto fue probado en las computadoras Xerox PARC's Alto y la primera red de Ethernet se denominó Alto Aloha System (Mayo) (:amk:)

Bob Kahn plantea el problema de Internet, comienza el programa de investigación de interacción de redes en el ARPA. En Marzo Vinton Cerf diseña la arquitectura básica de acceso sobre un sobre en el hall de un hotel en San Francisco (:vgc:)

En Septiembre, Cerf y Kahn presentan ideas basicas de Internet en INWG en la Universidad de Sussex, Brighton, Reino Unido (:vgc:)

RFC 454: Especificación de Transferencia de Archivos

Especificación (RFC 741) para Protocolo de Voz en Redes (NVP) e implementación para permitir llamadas en conferencia a través de la ARPANET. (:bb1:)

SRI (NIC) comienza a publicar ARPANET News en Marzo; el número de usuarios de ARPANET se estima en 2,000

Un estudio de ARPA muestra que el correo electrónico constituye el 75% del tráfico total de ARPANET

Christmas Day Lockup (Bloqueo de Navidad) - el problema de hardware del IMP de Harvard lleva a que este no encuentre ningún destino en ARPANET lo que hizo que todos los IMPs envíen su tráfico a Harvard (25 de Diciembre)

RFC 527: ARPAWOCKY

RFC 602: Las medias están colgadas cuidadosamente en la chimenea

Vinton Cerf y Bob Kahn publican "Protocolo para Interconexión de Redes por paquetes" ("A Protocol for Packet Network Interconnection") que especifica en detalle el diseño del Programa de Control de Transmisión (TCP). [IEEE Trans Comm] (:amk:)

BBN abre Telenet, el primer servicio público de paquetes de información (una versión comercial de ARPANET) (:sk2:)

1975

La administración operacional de Internet pasa a DCA (hoy DISA)

Steve Walker crea la primera lista de distribución de First ARPANET, MsgGroup Einar Stefferud asume rápidamente como moderador ya que la lista al principio no estaba automatizada. Una lista de ciencia ficción, SF-Lovers, se convirtió en la lista no oficial más popular por aquellos días.

John Vittal desarrolla MSG, el primer programa de correo electrónico realmente completo que incluye la posibilidad de contestar, reenviar y guardar mensajes.

Los enlaces satelitales cruzan dos océanos (a Hawaii y el Reino Unido) al tiempo que en Stanford se lleva a cabo las primeras pruebas de TCP por Stanford, BBN, y UCL

Se publica por primera vez el "Archivo de Jerga" (" Jargon File") por Raphael Finkel en SAIL. (:esr:)

Aparece Shockwave Rider por John Brunner (:pds:)

1976

Elizabeth II, Reina del Reino Unido envía un mensaje de correo electrónico en Febrero desde RSRE (Malvern)

AT&T Bell Labs desarrolla el UUCP (Unix-to-Unix Copy) (Copia de Unix a Unix) y se distribuye con UNIXun año después.

Se diseñan y distribuyen los IMPs Multiprocessing Pluribus

1977

Larry Landweber crea THEORYNET en la Universidad de Wisconsin, que provee el servicio de correo electrónico a más de 100 investigadores en ciencias de la computación (utilizando un sistema de correo electrónico desarrollado localmente sobre TELENET).

RFC 733: Especificación de correo

Tymshare lanza Tymnet

Primera demostración de la red ARPANET/SF Bay Packet Radio Net/Atlantic SATNET En Julio comienzan a operar los protocolos de Internet con los accesos que provee BBN (:vgc:)

1978

TCP se divide en TCP e IP (Marzo)

RFC 748: TELNET RANDOMLY-LOSE Option

1979

Reunión entre la Universidad de Wisconsin, DARPA, NSF, e investigadores en ciencias de la computación de diversas universidades a fin de establecer una Red de investigación del Departamento de Ciencias de la Computación (organizada por Larry Landweber).

Tom Truscott, Jim Ellis y Steve Bellovin establecen USENET utilizando UUCP entre Duke y UNC. Todos los grupos originales se encuentran en net.*hierarchy.

Aparecen los primeros MUD y MUD1 por Richard Bartle y Roy Trubshaw de la Universidad de Essex.

ARPA establece la primera Comisión de control de la Configuración de Internet (Internet Configuration Control Board) (ICCB).

Comienza el experimento de Packet Radio Network (PRNET) financiado por DARPA. La mayor parte de las comunicaciones se llevan a cabo entre furgones móviles. Aparece la conexión de ARPANET vía SRI.

El 12 de Abril, Kevin MacKenzie envía un mensaje de correo electrónico a MsgGroup en el que les sugiere agregar un poco de emoción al texto de los mensajes de correo electrónico por ejemplo el signo -) para implicar una oración irónica. A partir de este mensaje comenzaron a utilizarse los signos para demostrar emociones en los mensajes.

1980

ARPANET deja de funcionar por completo el 27 de Octubre a raíz de una advertencia de virus propagada accidentalmente

Aparece el primer IMP con base C/30 en BBN

1981

BITNET, la "Because it's time NETwork" (La red "porque es el momento")

- Comienza como una red cooperativa en la Universidad de New York, con la primera conexión con Yale (:feg:)
- La sigla original representaba 'There' (hay) en lugar de 'Time' (momento) haciendo referencia a los protocolos NJE que estaban incluidos en los sistemas de IBM.
- Provee correo electrónico y servidores listserv que distribuyen información así como también transferencia de archivos.

La CSNET (Computer Science Network) (Red de las Ciencias de la Computación) se crea gracias a la colaboración de expertos en computación de la Universidad de Delaware, la Universidad Purdue, La Universidad de Wisconsin, RAND Corporation y BBN financiado por NSF con el objeto de prestar servicios de red (especialmente de correo electrónico) a los científicos que carecían de acceso a la ARPANET. Más tarde la CSNET se conocerá como la Red de Computación y Ciencia. (:amk,lhl:)

C/30 Los IMPs predominan en la red; aparece el primer TIP C/30 en SAC

Telecom de Francia despliega Minitel (Teletel) por toda Francia.

"True Names" ("Denominaciones reales") por Vernor Vinge (:pds:)

RFC 801: Plan de transición NCTP/TCP

1982

Noruega deja la red para convertirse en una conexión a Internet via TCP/IP a través de SATNET; UCL hace lo mismo

DCA y ARPA establecen el Protocolo de Control de Transmisión (TCP) y el Protocolo de Internet (IP) como el conjunto de protocolos, conocido comúnmente como TCP/IP, para ARPANET. (:vgc:)

- Esto genera una de las primeras definiciones de Internet: "una serie de redes conectadas entre sí, específicamente aquellas que utilizan el protocolo TCP/IP" y se utiliza el término "Internet" como conectado a redes TCP/IP interconectadas.
- El Ministerio de Defensa establece que el protocolo TCP/IP será el standard para ese organismo. (:vgc:)

EUUG crea la EUnet, (European Unix Network) (Red Unix Europea) para brindar los servicios de correo electrónico y de USENET. (:glg:)

- Conexiones originales entre los Países Bajos, Dinamarca, Suecia y el Reino Unido.

Especificación para el Protocolo de Acceso Externo (RFC 827). Se utiliza el EGP para el acceso entre redes.

1983

El servidor de nombres desarrollado en la Universidad de Wisconsin ya no requiere que el usuario conozca la ruta exacta para acceder a otros sistemas.

Paso de NCP a TCP/IP (1 Enero)

Desaparecen los IMPs Honeywell o Pluribus ; los TIPs son reemplazados por TACs

Stuttgart y Korea se conectan

La red de Movimiento de Información (MINET) aparece a principio de año en Europa y se conecta a Internet en Septiembre

El acceso a CSNET/ARPANET comienza a funcionar.

ARPANET se divide en ARPANET y MILNET, esta última se integra con la Red de información de Defensa creada el año anterior. 68 de los 113 nodos existentes pasan a la red MILNET

Aparecen las estaciones de trabajo, muchas de ellas con el sistema Berkeley UNIX (4.2 BSD) que incluye software de red IP (:mpc:)

Los sistemas de redes necesitan pasar de tener grandes computadoras de tiempo compartido conectadas a la Internet en cada sitio a conectar redes locales enteras

Se establece la Comisión de Actividades de Internet Internet Activities Board (IAB) en remplazo de la ICCB.

Se establece la EARN, European Academic and Research Network (Red Académica de Investigaciones Europea). Funciona de manera similar a la BITNET con un acceso financiado por IBM

Tom Jennings desarrolla la FidoNet

1984

Se introduce el Domain Name System(DNS) (Sistema de nombre de dominio)

La cantidad de hosts supera las 1.000

Se establece la JUNET, Japan Unix Network (Red Unix de Japón) utilizando UUCP

Se establece en el Reino Unido la JANET, Joint Academic Network (Red académica conjunta) utilizando los protocolos Coloured Book, anteriormente conocida como SERCnet

Se incluyen los grupos de interés moderados en la USENET. (mod.*)

"Neuromancer" por William Gibson

Canadá comienza un proyecto de un año para instalar redes en sus universidades. La red NetNorth se conecta a la BITNET en Ithaca desde Toronto (:kf1:)

Kremvaxmensaje que anuncia la conectividad de la Unión Soviética a la USENET

1985

Se pone en funcionamiento la Whole Earth 'Lectronic Link (WELL)(Conexión electrónica de toda la Tierra).

El Instituto de Ciencias de la Información (ISI) en USC recibe la responsabilidad de administrar el arbol de DNS por medio de DCA, y SRI para los registros de DNS NIC

Symbolics.com se convierte en el primer dominio registrado el 15 de Marzo. Otros: cmu.edu, purdue.edu, rice.edu, ucla.edu (Abril); css.gov (Junio); mitre.org, .uk (Julio).

A 100 años del día en que la última espiga fuera transportada por tren a través de Canadá, se conecta la última universidad canadiense a la NetNorth en un esfuerzo por tener conexión de costa a costa que duró un año. (:kf1:)

RFC 968: Fue la noche anterior al arranque

1986

Se crea la NSFNET (Con una velocidad principal de 56Kbps).

- NSF establece 5 centros de super computadoras para proveer alto poder de proceso. (JVNC@Princeton, PSC@Pittsburgh, SDSC@UCSD, NCSA@UIUC, Theory Center@Cornell).
- Esto permite una explosión de conexiones, especialmente por parte de las universidades.

Aparecen NSF basado en SDSCNET, JVNCNET, SURANET y NYSERNET operacional. (:sw1:)

Aparecen la Internet Engineering Task Force (IETF)(Fuerza operante de Ingeniería de Internet) y la Internet Research Task Force (IRTF) (Fuerza operante para la Investigación de Internet) bajo la supervisión de la IAB. En Enero se lleva a cabo la primera reunión de la IEFT en Linkabit, San Diego.

La primera Freenet (Cleveland) se conecta en línea el 16 de Julio auspiciada por la (SoPAC) Society for Public Access Computing (Sociedad para la Computación de Acceso Público). Más tarde, la NPTN National Public Telecomputing Network (Red Nacional Pública de Telecomputación) se hace cargo de la administración del programa en 1989 (:sk2,rab:)

Se diseña el Protocolo de Transmisión de Información en Redes (NNTP) para mejorar el desempeño del Usenet en TPC/IP.

El Mail Exchanger (MX) desarrollado por Craig Partridge permite a los hosts de redes que no utilizan el IP tener direcciones de dominio.

Cambia el nombre de la USENET, los grupos de interés moderados cambian en 1987.

Se establece la BARRNET, Bay Area Regional Research Network (Red de Investigación Regional del Área de la Bahía) utilizando enlaces de alta velocidad. Se torna operacional en 1987.

New England queda afuera de la red debido a que AT&T sufre una rotura de los cables de fibra óptica que unen Newark/New Jersey y White Plains/New York. Sí, las siete líneas principales de ARPANET en New England estaban en un cable cortado. El corte se dio entre la 1:11 y las 12:11 EST el 12 de Diciembre

1987

NSF firma un contrato de colaboración para la administración de NSFNET con Merit Network, Inc., (IBM y MCI se involucraron mediante un contrato con Merit). Merit, IBM y MCI luego fundaron ANS.

UUNET se funda con fondos de Usenix para proveer UUCP comercial y acceso a la Usenet. En un principio se trataba de un experimento de Rick Adams y Mike O'Dell.

Primera conferencia de interoperabilidad TCP/IP (Marzo), el nombre se cambió en 1988 a INTEROP

Se establece una conexión de correo electrónico entre Alemania y China utilizando protocolos CSNET. El primer mensaje desde China se envió el 20 de Septiembre. (:wz1:)

1000th RFC: "Request For Comments reference guide" ("Guía de referencia para la Solicitud de Comentarios")

La cantidad de hosts supera los 10,000.

La cantidad de hosts BITNET supera los 1,000.

1988

2 de noviembre - el Internet worm aparece en la Red y afecta a 6,000 de los 60,000 hosts de Internet. (:ph1:)

DARPA crea el CERT, Computer Emergency Response Team (Equipo de Soluciones de Emergencia de Computación) en respuesta a las necesidades que se suscitaron durante el incidente del Virus Morris. El worm es la única advertencia que se publica ese año.

El Ministerio de Defensa decide adoptar OSI y utiliza el protocolo TCP/IP entretanto. El perfil OSI del gobierno de los Estados Unidos (GOSIP) define el conjunto de protocolos que estarán respaldados por productos adquiridos por el gobierno. (:gck:)

Se crea la Red Los Nettos sin financiación estatal, con el respaldo de miembros regionales (financiada por Caltech, TIS, UCLA, USC, ISI).

Se actualiza la estructura principal de la NSFNET a T1 (1.544 Mbps)

Susan Estrada funda la CERFnet, California Education and Research Federation Network (Red de la Federación de Educación e Investigación de California).

En Diciembre se crea la Autoridad de Asignación de Números de Internet (IANA)

Jarkko Oikarinen desarrolla el Internet Relay Chat (IRC) (Difusora de charlas en Internet) (:zby:)

Las primeras redes regionales de Canadá se unen a la NSFNET: ONet vía Cornell, RISQ vía Princeton, BCnet vía la Universidad de Washington. (:ec1:)

La FidoNet se conecta a la Red permitiendo el intercambio de correo e información. (:tp1:)

Canadá (CA), Dinamarca (DK), Finlandia (FI), Francia (FR), Islandia (IS), Noruega (NO) y Suecia (SE) se conectan a la NSFNET.

1989

La cantidad de hosts supera los 100,000.

RIPE(Réseaux IP Europeens) se forma (por proveedores de servicios Europeos) para asegurar la coordinación administrativa y técnica necesaria para permitir el funcionamiento de la Red Pan-European IP. (:glg:)

Se dan los primeros intercambios entre un operador comercial de correo electrónico y la Internet: MCI Mail a través de la CNRI, Corporation for the National Research Initiative (Empresa para las iniciativas de investigación nacional) y Compuserve a través de la Universidad de Ohio. (:jg1,ph1:)

Se crea la CREN, Corporation for Research and Education Networking a partir de la inclusión de la CSNET en la BITNET (Agosto)

AVCC y CSIRO crean la AARNET, Australian Academic Research Network (Red Australiana de Investigaciones Académicas) que comienza a funcionar el año siguiente. (:gmc:)

"Cuckoo's Egg" de Clifford Stoll relata la historia real de un grupo de crackers alemanes que lograron acceder a varios organismos estadounidenses.

La UCLA patrocina el Simposio Primer Acto en conmemoración del 20º aniversario de ARPANET y la creación de su comisión (Agosto)

RFC 1121: Primer Acto - Los poemas

RFC 1097: TELNET Opción de MENSAJE SUBLIMINAL

Países que se conectaron a la NSFNET: Australia (AU), Alemania (DE), Israel (IL), Italia (IT), Japón (JP), México (MX), Países Bajos (NL), Nueva Zelandia (NZ), Puerto Rico (PR), Reino Unido (UK).

1990

ARPANET deja de existir.

Mitch Kapor funda la Electronic Frontier Foundation (EFF)(Fundación Frontera Electrónica).

Peter Deutsch, Alan Emtage y Bill Heelan de McGill lanzan Archie.

Peter Scott lanza Hytelnet (Universidad de Saskatchewan).

World se pone en línea (world.std.com) y de esta manera se convierte en el primer proveedor comercial de acceso telefónico a Internet.

Se desarrolla ISODE, ISO Development Environmet, a fin de proveer un enfoque acerca de la migración de OSI del Ministerio de Defensa. El software de ISODE permite que las aplicaciones OSI funcionen en un entorno TCP/IP (:gck:)

Se crea CA*net formada por 10 redes regionales que conforman la base canadiense con conexión directa a NSFNET (:ec1:)

La primera máquina de operación remota se conecta a Internet, se trata de la Internet Toaster de John Romkey (controlada vía SNMP) y hace su debut en Interop. Imágenes: Internode, Invisible

RFC 1149: Una norma para la transmisión de Datagramas IP en Carriers Avian

RFC 1178: Elección de un nombre para su computadora

Países que se conectan a la NSFNET: Argentina (AR), Austria (AT), Bélgica (BE), Brasil (BR), Chile (CL), Grecia (GR), India (IN), Irlanda (IE), Corea (KR), España (ES), Suiza (CH).

1991

Nace la Commercial Internet eXchange (CIX) Association Inc. (Asociación de intercambios comerciales en Internet), integrada por General Atomics (CERFnet), Performance Systems International, Inc. (PSInet) y UUNET Technologies Inc. (AlterNet), después de que la NSF eliminara las restricciones comerciales que regían sobre el uso de la Red (Marzo)(:glg:)

Brewster Kahle inventa los Wide Area Information Servers (WAIS), (Servidores de Información de área amplia), que son lanzados por Thinking Machines Corporation.

Paul Linder y Mark P. McCahill de la Universidad de Minnessota lanzan Gopher

CERN lanza la World-Wide Web (WWW) creada por Tim Berners - Lee (:pb1:)

Philip Zimmerman lanza PGP (Pretty Good Privacy) (Buena privacidad) (:ad1:)

La ley estadounidense "High Performance Computing Act" (Gore 1) establece la NREN National Research and Education Network (Red Nacional de Investigación y Educación).

La base de la NSFNET se actualiza a T3 (44.736 Mbps).

El flujo de información de la NSFNET pasa de 1.000 millones de bytes/mes a 10,000 millones de paquetes/mes.

DISA concede su contrato de la Defense Data Network NIC a Government Systems Inc. quien se hace cargo desde el SRI en Mayo.

Comienza a funcionar el JANET IP Service (JIPS) hecho que marcó el paso del software Coloured Book a TCP/IP dentro de la red académica del Reino Unido. Inicialmente el protocolo IP estaba 'restringida' dentro de X.25. (:gst:)

RFC 1216: Cambios Económicos y paradigmáticos en la Red Gigabit

RFC 1217: Memo del Consorcio para la investigación de la Conmoción Lenta (CSCR)

Países que se conectan a NSFNET: Croacia (HR), República Checa (CZ), Hong Kong (HK), Hungría (HU), Polonia (PL), Portugal (PT), Singapur (SG), Sudáfrica (ZA), Taiwan (TW), Túnez (TN).

1992

Se crea la Internet Society (ISOC) (Enero)

Se restituye la IAB como la Internet Architecture Board (Comisión de arquitectura de Internet) y pasa a formar parte de la Internet Society.

La cantidad de hosts supera 1,000,000.

Aparecen el primer MBONE audio multicast (Marzo) y el video multicast (Noviembre)

En Abril se crea el Network Coordination Center RIPE(Centro de coordinación de red) a fin de brindar servicios de registro y coordinación a la comunidad de Internet Europea. (:dk1:)

La Universidad de Nevada lanza Veronica, una herramienta de búsqueda en el entorno Gopher.

El Banco Mundial World Bank se conecta en línea.

Jean Armour Polly crea la expresión "Navegar por Internet" (" Surfing the Internet") (:jap:)

Rick Gates comienza el Internet Hunt.

RFC 1300: Recuerdos de hechos del pasado

RFC 1313: Programación actual para KRFC AM 1313 - Radio en Internet

Países que se conectan con la NSFNET: Antártida (AQ), Camerún (CM), Chipre (CY), Ecuador (EC), Estonia (EE), Kuwait (KW), Letonia (LV), Luxemburgo (LU), Malasia (MY), Eslovaquia (SK), Eslovenia (SI), Tailandia (TH), Venezuela (VE)

1993

La NSF crea la InterNIC para brindar servicios específicos de Internet: (:sc1:)

- Servicios de directorio y base de datos (AT&T)
- Servicios de registro (Network Solutions Inc.)
- Servicios de información (General Atomics/CERFnet)

La Casa Blanca se conecta en línea (<http://www.whitehouse.gov/>):

- Presidente Bill Clinton: president@whitehouse.gov
- Vicepresidente Al Gore: vice-president@whitehouse.gov

Worms (gusanos) de una nueva clase aparecen en la Red - los Worms WWW (W4) a los que se les unen los Spiders (arañas) , Wanderers (vagabundos) , Crawlers (orugas) y Snakes (serpientes)...

Comienzan las transmisiones de radio por Internet (Internet Talk Radio) (:sk2:)

La Organización de las Naciones Unidas (United Nations) (UN) se conecta en línea (:vgc:)

Se sanciona la ley de infraestructura de Información Nacional en Estados Unidos (US National Information Infrastructure Act).

Las empresas y los medios comienzan a prestarle atención a la Internet

InterCon International KK (IKK) provee la primera conexión comercial para Japón. Un mes más tarde TWICS, a pesar de ser parte de IKK, comienza a ofrecer cuentas con acceso telefónico (:tb1:)

Mosaic genera un crecimiento asombroso: la WWW crece a una tasa del 341.634% anual para el flujo de servicio. Gopher crece a una tasa del 997%.

RFC 1437: La extensión de los contenidos-tipos MIME en un nuevo medio

RFC 1438: Afirmaciones del aburrimiento de IETF (SOBs)

Países que se conectan a la NSFNET: Bulgaria (BG), Costa Rica (CR), Egipto (EG), Fiji (FJ), Ghana (GH), Guam (GU), Indonesia (ID), Kazakhstan (KZ),

Kenia (KE), Liechtenstein (LI), Perú (PE), Rumania (RO), Federación Rusa (RU), Turquía (TR), Ucrania (UA), UAE (AE), Islas Vírgenes (VI).

1994

ARPANET/Internet celebra su 25° aniversario.

Las comunidades comienzan a conectarse directamente a Internet (Lexington y Cambridge, Massachusetts, Estados Unidos).

El Senado y la Casa Blanca (House) instalan servidores de información.

Los shopping malls entran en Internet.

La primera Ciberestación, RT-FM, transmite desde Interop en Las Vegas.

Vladimir Levin de San Petersburgo, Rusia, se transforma en el primer ladrón de bancos famoso por Internet al transferir electrónicamente millones de dólares del Citibank entre Junio y Agosto.

El NIST National Institute for Standards and Technology (Instituto Nacional de normas y tecnología) sugiere que GOSIP incorpore el protocolo TCP/IP y deje de utilizar los requerimientos "OSI-only". (:gck:)

La empresa de abogados Canter & Siegel de Arizona inunda Internet con e-mails no solicitados ("spams") que promocionan servicios de sorteos de Green Cards (Credenciales de Inmigración); los ciudadanos de la Red respondieron del mismo modo.

El flujo de la NSFNET supera los 10.000 millones de bytes/mes.

Si, es verdad. Ahora puede pedir su pizza en Hut on-line.

La WWW supera a telnet y se transforma en el segundo servicio más popular de la Red (después de ftp-data) basándose en el porcentaje de flujo y distribución de paquetes y bytes en la NSFNET.

El Primer Ministro Japonés entra en línea (<http://www.kantei.go.jp/>)

El Tesoro del Reino Unido se conecta en línea (<http://www.hm-treasury.gov.uk/>)

El Primer Ministro de Nueva Zelanda se conecta en línea (<http://www.govt.nz/>)

First Virtual, el primer ciberbanco, comienza a operar.

Las emisoras de radio comienzan a transmitir las 24 hs. En la red: WXYC de la Universidad de Carolina del Norte, WJHK de la Universidad de KS-Lawrence y KUGS de la Universidad del Oeste WA.

La (TERENA) Trans-European Research and Education Network Association (Asociación de Redes Transeuropeas de Educación e investigación) se forma a partir de la unión de RARE y EARN, con representantes de 38 países así como también de CERNy ECMWF. El objetivo de TERENA es "promover y participar en el desarrollo de infraestructura en información y telecomunicaciones internacionales para el beneficio de la investigación y la educación" (Octubre)

RFC 1605: SONET para la traducción de sonetos

RFC 1606: Una visión histórica sobre el uso del IP Versión 9

RFC 1607: Una visión del siglo 21

Países que se conectan a la NSFNET: Argelia (DZ), Armenia (AM), Bermuda (BM), Burkina Faso (BF), China (CH), Colombia (CO), Jamaica (JM), Jordania (JO), Líbano (LB), Lituania (LT), Macau (MO), Marruecos (MA), Nueva Caledonia, Nicaragua (NI), Nigeria (NE), Panamá (PA), Filipinas (PH), Senegal (SN), Sri Lanka (LK), Swaziland (SZ), Uruguay (UY), Uzbekistan (UZ)

1995

La NSFNET vuelve a transformarse en una red de investigación. El flujo de la estructura principal de los Estados Unidos se canaliza a través de proveedores de red.

Nace la nueva NSFNET al tiempo que NSF establece el very high speed Backbone Network Service (vBNS)(servicio de base de red de alta velocidad) que conecta centros de alto procesamiento: NCAR, NCSA, SDSC, CTC, PSC

La policía de Hong Kong desconecta a todos los proveedores de Internet menos a uno en busca de un hacker. 10,000 personas quedan sin acceso a la Red. (:api:)

Sun lanza JAVA el 23 de Mayo

RealAudio, una tecnología de audio, permite que los usuarios de la Red reciban el sonido casi en tiempo real.

Comienza a transmitir Radio HK, la primera radio comercial las 24 hs. de transmisión exclusiva por Internet.

WWW supera a ftp-data en Marzo y se transforma en el servicio de mayor flujo en la NSFNet en base al conteo de paquetes y en Abril en base al conteo de bytes.

Los sistemas tradicionales de acceso telefónico (dial-up) (Compuserve, America Online, Prodigy) comienzan a brindar el servicio de acceso a Internet.

Miles de personas en Minneapolis- St. Paul (Estados Unidos) pierden el acceso a la Red después de que alguien enciende una fogata bajo un puente en la Universidad de Minneapolis que hace que se derritan los cables de fibra óptica (30 de julio)

Una serie de empresas relacionadas a la Red se hacen famosas, Netscape lidera el grupo con el tercer valor NASDAQ IPO por acción más alto de la historia (9 de Agosto)

El registro de los nombres de los dominios deja de ser gratuito. A partir del 14 de Septiembre se impone una cuota anual de \$50.- que hasta el momento ha sido subsidiada por NSF. NSF continúa pagando la cuota correspondiente a los registros .edu y momentáneamente los .gov

El Vaticano se conecta en línea (<http://www.vatican.va/>)

El gobierno canadiense se conecta en línea (<http://canada.gc.ca/>)

La primera interceptación de líneas telefónicas oficial que se realizó por Internet contribuyó a que el Servicio Secreto y la DEA pudieran detener a tres individuos que fabricaban y vendían de manera ilegal equipos de telefonía celular y dispositivos electrónicos clonados.

Por primera vez el Operation Home Front (Frente Operación Hogar) conecta a soldados que se encontraban en el campo de batalla con sus familias vía Internet.

Richard White se transforma en la primera persona en ser declarada "munición", de acuerdo con las leyes de control de armas de los Estados Unidos, debido a una programa de codificación de seguridad RSA que se encontraba tatuado en su hombro. (:wired496:)

RFC 1882: Los 12 días de Tecnología antes de Navidad

Países que registran dominios: Etiopía (ET), Costa de Marfil (CI), Islas Cook (CK), Islas Cayman (KY), Anguillia (AI), Gibraltar (GI), vaticano (VA), Kiribati (KI), Kyrgyzstan (KG), Madagascar (MG), Mauritius (MU), Micronesia (FM), Mónaco (MC), Mongolia (MN), Nepal (NP), Nigeria (NG), Samoa Occidental (WS), San Marino (SM), Tanzania (TZ), Tonga (TO), Uganda (UG), Vanuatu (VU)

Tecnologías del año: WWW, Motores de búsqueda.

Tecnologías emergentes: Código Móvil (JAVA, JAVAscript), entornos virtuales (VRML), herramientas de colaboración.

1996

Los teléfonos de Internet llaman la atención de las empresas de telecomunicaciones que solicitan al Congreso de los Estados Unidos que prohíban esta tecnología (que ya existía desde hacía varios años).

El Primer Ministro Malayo Mahathir Mohamad, el Líder Yasser Arafat y el presidente Filipino Fidel Ramos se reúnen durante 10 minutos en una sesión de chateo interactiva el 17 de Enero.

El controvertido proyecto de ley US Communications Decency Act (CDA) (Ley de Decencia en las comunicaciones) se transforma en ley en los Estados Unidos a fin de prohibir la distribución de material obsceno a través de la Red. Algunos meses después un panel de tres jueces impone una orden en contra de su aplicación. En 1997 la Corte Suprema, por unanimidad, la decreta inconstitucional en su mayor parte

9272 organizaciones se encuentran excluidas de la Red después de que InterNIC las excluye por no haber abonado la cuota correspondiente a su nombre de dominio.

Varios ISP sufren cortes en el servicio extendido, lo que trae a colación un cuestionamiento acerca de la posibilidad de manejar la creciente cantidad de usuarios. AOL (19 horas), Netcom (13 horas), AT&T WorldNet (28 horas sólo correo electrónico)

CNET adquiere el nombre de dominio tv.com por US\$15,000

Public Access Networks Corp. de New York (PANIX) deja de funcionar debido a los repetidos ataques de un hacker que utilizó métodos publicados en una revista de hackers. (2600)

MCI actualiza la base de Internet al sumarle ~13,000 puertos que hacen que la velocidad pase de 155 Mbps a 622 Mbps.

La comisión de Internet Ad Hoc (Internet Ad Hoc Committee) anuncia un plan para agregar 7 nuevos dominios genéricos (gTLD): .firm, .store, .web, .rec, .info, .nom. El plan de la IAHC también abre la competencia para el registro de nombres de dominio en todo el mundo.

Un virus maligno (cancelbot) es liberado en la USENET y borra más de 25,000 mensajes.

La guerra del navegador WWW, principalmente entre Netscape y Microsoft, ha dado origen a una nueva era en desarrollo de programas, en la que lanzan nuevas versiones trimestralmente con la ayuda de los usuarios de Internet ávidos por probar las nuevas versiones (beta)

RFC 1925: Las Doce verdades acerca de las redes

Restricciones al uso de Internet en todo el mundo:

- *China*: Requiere que los usuarios y las ISPs estén registrados ante la policía.
- *Alemania*: Cortó el acceso a algunos grupos de interés que funcionaban bajo CompuServe.
- *Arabia Saudita*: Limitó el acceso a Internet a hospitales y Universidades.
- *Singapur*: Requiere que aquellos que publiquen contenidos políticos y religiosos estén registrados ante el Estado.
- *Nueva Zelandia*: Considera los discos de las computadoras "publicaciones" que pueden ser censuradas y secuestradas.
- *Fuente: Human Rights Watch*

Países que registraron dominios: Qatar (QA), República Centro Africana (CF), Omán (OM), Norfolk Island (NF), Tuvalu (TV), Polinesia Francesa (PF), Siria (SY), Aruba (AW), Camboya (KH), Guyana Francesa (GF), Eritrea (ER), Cabo Verde (CV), Burundi (BI), Benin (BJ), Bosnia . Hercegovina (BA), Andorra (AD), Guadalupe (GP), Guersey (GG), Isle of Man (IM), Jersey (JE), Lao (LA), Maldivas (MV), Islas Marshall (MH), Mauritania (MR), Islas Mariana del Norte (MP), Ruanda (RW), Togo (TG), Yemen (YE), Zaire (ZR)

Actos de piratería informática del año: Departamento de Justicia de los Estados Unidos (17 de Agosto), CIA (19 de Septiembre), Fuerza Aérea (29 de Diciembre), Partido Laborista del Reino Unido (6 de Diciembre)

Tecnologías del año: Motores de búsqueda, JAVA, Teléfono Internet.

Tecnologías emergentes: Entornos virtuales (VRML), Herramientas de colaboración, aplicación de Internet (Computadora de red)

1997

RFC N° 2,000: "Internet official protocol standards" ("Normas de protocolos oficiales de Internet")

Se registran 71,618 listas de correo en Liszt, un directorio de listas de distribución de correo.

Se establece el American Registry for Internet Numbers (ARIN) (Registro Estadounidense de Números de Internet) a fin de manejar la administración y registro de los números IP para las distintas áreas geográficas que anteriormente estaban manejadas por Network Solutions (InterNIC) a partir de Marzo de 1998

En Junio de lanza la CA*net II para proveer a Canada de un servicio de Internet que utilice ATM/SONET

En señal de protesta contra el monopolio de DNS , Eugene Kashpureff , dueño de AlterNIC hace que todos los usuarios que entran en www.internic.net terminen en www.alternic.net

El dominio denominado business.com se vende en US\$150,000

El 17 de julio muy temprano un error humano en Network Solutions produce la corrupción de la tabla DNS para los dominios .com y .net lo que hizo imposible comunicarse con millones de sistemas.

Se registra el nombre de dominio más largo en InterNIC :
CHALLENGER.MED.SYNAPSE.UAH.UALBERTA.CA

Se registran 101,803 Nombres de Servidores en la base de datos whois.

RFC 2100: El nombre de los Hosts

Dominios nacionales registrados: Islas Falkland (FK), Timor del Este (TP), Congo (CG), Islas Christmas (CX), Gambia (GM), Guinea - Bissau (GW), Haití (HT), Iraq (IQ), Libia (LY), Malawi (MW), Martinica (MQ), Montserrat (MS), Myanmar (MM), Isla French Reunion (RE), Seychelles (SC), Sierra Leona (SL), Somalia (SO), Sudán (SD), Tajkistan (TJ), Turkmenistán (TM), Islas Turks y Caicos (TC), Islas vírgenes Británicas (VG), Islas Heard y McDonald (HM), Territorios Franceses del Sur (TF), Territorio Británico en el Océano Indico (IO), Islas Scalbard y Jan Mayen (SJ), St Pierre y Miquelon (PM), Santa Helena (SH), Islas Georgias y Sandwich del Sur (GS), Sao Tome y Principe (ST), Isla Ascension (AC), Tajikstan (TJ), Islas Estadounidenses Menores (UM), Mayotte (YT), Islas Wallis and Futuna (WF), Islas Tokelau (TK), República de Chad (TD), Afghanistan (AF), Isla Cocos (CC), Isla Bouvet (BV), Liberia (LR), Samoa Americana (AS), Niue (NU), Nueva Guinea Ecuatorial (GQ), Bhutan (BT), Isla Pitcairn (PN), Palau (PW), DR de Congo (CD)

Actos de piratería informática del año: Gobierno de Indonesia (19 de Enero, 10 de Febrero, 24 de Abril, 30 de Junio, 22 de Noviembre), NASA (5 de Marzo), Partido Conservados del Reino Unido (27 de Abril), Spice Girls (14 de Noviembre)

Tecnologías del año: Push, Multicasting

Tecnologías emergentes: Push, Streaming Media [:twc:]

1998

Se lanza la Cronología de Internet de Hobbes como RFC 2235 & FYI 32

El Departamento de Comercio Estadounidense (DoC) lanza su Green Paper que sienta las bases para un plan para la privatización de DNS el 30 de Enero. A este informe le sigue el White Paper 5 de Junio

La Fête de l'Internet, una feria nacional de Internet se lleva a cabo en Francia del 20 al 21 de Marzo.

El tamaño de la Red se estima entre 275 (digital) y 320 (NEC) millones de páginas por 1Q

Las empresas vuelan hacia el NIC de Turmenistán a fin de registrarse bajo el dominio tm, abreviatura de Marca Registrada en Inglés.

Los usuarios de Internet pudieron participar como jurado en un campeonato mundial de patinaje sobre hielo el 27 de Marzo. Fue la primera vez que los televidentes determinaron el resultado de un concurso en un programa televisivo de deportes.

El 4 de Mayo Network Solutions registra su dominio N° 2.000.000

Las estampillas electrónicas son una realidad. El Servicio de Correos de Estados Unidos USPS permite que las estampillas se compren e impriman directamente desde la Web.

Canada lanza CA*net 3, la primera red óptica nacional

CDA II y una prohibición sobre los impuestos a la Red se convierten en ley para los Estados Unidos (21 de Octubre)

ABCNews.com accidentalmente publica los resultados de las elecciones en los Estados Unidos con un día de anticipación.(2 de Noviembre)

Departamento de Comercio Estadounidense (US DoC) suscribe un contrato con la Empresa de Internet para Números Asigandos (ICANN) a fin de establecer un proceso para el paso de DNS de manos del Estado a manos industriales (25 de Noviembre)

Los sites de San Francisco quedan desconectados de la red debido a que la ciudad sufre un apagón el 8 de Diciembre

El gobierno de china enjuicia a Lin Hai por 'incitar al derrocamiento del gobierno' por proporcionar 30.000 direcciones de correo electrónico a una revista de Internet Estadounidense. (Diciembre)

El software de fuente abierta cumple la mayoría de edad.

RFC 2321: RITA (The Reliable Internetwork Troubleshooting Agent)- El organismo confiable en la solución de problemas de Internet

RFC 2322: Gestión de números IP por peg-dhcp

RFC 2323: Lineamientos para la identificación y seguridad de IETF

RFC 2324: Protocolo de control de Hipertexto Coffe Pot (HTCPCP/1.0)

Se registran los siguiente dominio de países: Nauru (NR), Comoros (KM)

*Generadores de acho de banda:*Juegos Olímpicos de Invierno (Feb), Mundial de Fútbol(Jun-Jul), Informe de Starr (11 Sep), Lanzamiento de Glenn space

*actos de piratería del año:*Departamento de Comercio de los Estados Unidos (20 de Febrero), New York Times (13 de Septiembre), China Society para los Estudios de Derechos Humanos (26 de Octubre), UNICEF (7 de Enero)

*Tecnologías del año:*E-Commerce (comercio electrónico), E-Auctions (subastas electrónicas, Portales

*Tecnologías Emergentes:*E-Trade (mercado electrónico), XML

1999

Comienza a estar disponible el acceso a Internet para el público de Arabia Saudita (.sa) en Enero.

vBNS establece un vínculo OC48 entre CalREN Sur y Norte usando routers Juniper M40.

Primer Banco de Internet de Indiana (First Internet Bank of Indiana), el primer banco con todos los servicios disponibles únicamente en la NET se inaugura el 22 de febrero.

IBM se transforma en el primer asociado corporativo aprobado para el acceso a Internet2.

El parlamento Europeo propone la prohibición del "caching" (Reserva temporal) de paginas web por los ISPs.

La Internet Fiesta se lanza en Marzo sobre Europa como consecuencia del suceso que "La Fête de l'Internet" tuvo en 1998.

La Corte del Estado de los Estados Unidos establece que los nombres de dominio son una propiedad que debe ser guarnecida.

MCI/Worldcom, el proveedor de vBNS para NSF, comienza a actualizar el US backbone a 2.5GBps.

Una pagina web falsificada como si fuera una noticia financiera de Bloomberg aumento las acciones de una pequeña compañía de tecnología en un 31% el 7 de Abril.

El 21 de Abril el ICANN anuncia las 5 empresas registrantes de prueba para el competitivo Sistema de Registro Compartido: AOL, CORE, France Telecom/Oléane, Melbourne IT, Register.com. El 21 de abril se seleccionan 29 empresas registrantes mas, luego el 25 de Mayo se agregan 8, el 6 de Julio 15 y así sucesivamente hasta alcanzar 98 empresas a fin de año. La prueba, originalmente planificada para durar hasta el 24 de Junio, se extiende hasta el 10 de setiembre y luego hasta el 30 de noviembre. La primera empresa en estar en linea fue Register.com el 7 de Junio.

La primera Ciberguerra a gran escala tiene lugar simultaneamente a la guerra en Serbia / Kosovo.

Abilene, la cadena de Internet2, se amplía a través del Atlántico y se conecta a NORDUnet y SURFnet.

La Web se transforma en el punto central de la política inglesa cuando una lista de agentes del M16 es lanzada en un sitio web del Reino Unido. Si bien la lista fue removida del sitio, ya había sido replicada a lo largo de la red (15 de Mayo).

Activistas de toda la Net apuntan a los centros financieros del mundo el 18 de Junio, día de la cumbre del G8. Poco impacto real es reportado.

MCI/Worldcom lanza el vBNS+, una versión comercializada del vBNS dirigida a instituciones educativas pequeñas y de investigación.

Somalia obtiene su primer ISP – Computadora Olímpica (Septiembre).

ISOC aprueba la formación de la Fuerza de Trabajo Social de Internet (Internet Societal Task Force) (ISTF). Vint Cerf es elegido su presidente.

Las Computadoras gratuitas cubren todas las necesidades (Mientras se firme un contrato a largo plazo por un servicio en la Red)

.ps es registrado a Palestina (11 Octubre).

El vBNS alcanza las 101 conexiones.

business.com es vendida en U\$S 7.5 millones (fue adquirida en 1997 por U\$S 150,000 (30 Noviembre).

RFC 2549: IP sobre la empresa de transportes Avian Carriers con Calidad de Servicio

RFC 2550: Y10K más allá

RFC 2551: Proceso Standard Romano (The Roman Standards Process) - Tercera Revisión

RFC 2555: 30 años de RFCs

RFC 2626: La Internet y el problema del Milenio (Año 2000)

Mejores 10 TLDs por número de servidores: com, net, edu, jp, uk, mil, us, de, ca, au

actos de piratería del año: Star Wars (8 Enero), .tp (Enero), USIA (23 Enero), E-Bay (13 Marzo), Senado de los Estados Unidos (27 Mayo), NSI (2 Julio), Gobierno de Paraguay (20 Julio), AntiOnline (5 Agosto), Microsoft (26 Octubre), Rieles Británicos (31 Diciembre)

Tecnologías del año: Comercio electrónico (E-Trade), Bancos online (Online Banking), MP3

Tecnologías Emergentes: Teléfonos celulares de Red (Net-Cell Phones), Thin Computing (Computadora dependiente del servidor), Embedded Computing (Computadora incorporada a electrodomésticos).

Virus del año: Melissa (Marzo), ExploreZip (Junio)

2000

El Controlador de tiempo de los EE. UU. (USNO) y otros pocos servicios de tiempo de todo el mundo reportan el nuevo año como 19100 el primero de Enero.

Un ataque de rechazo de servicio masivo es lanzado contra importantes sitios web, incluyendo a Yahoo, Amazon, y eBay a comienzos de Febrero.

El tamaño de la Web estimado por NEC-RI e Inktomi sobrepasa los mil millones de paginas susceptibles de ser catalogadas.

ICANN redelega el dominio .pn, devolviéndoselo a la comunidad de la Isla de Pitcairn (Febrero)

Varios secuestros de nombres de dominio tienen lugar a finales de Marzo y comienzos de Junio, incluyendo a internet.com, bali.com, y web.net

RFC 2795: The Infinite Monkey Protocol Suite

CRECIMIENTO

Internet | Redes | WWW | USENET | Seguridad

Crecimiento de Internet:

Fecha	Servidores		Fecha	Servidores	Redes	Dominios
-----	-----	+	-----	-----	-----	-----
12/69	4		07/89	130,000	650	3,900
06/70	9		10/89	159,000	837	
10/70	11		10/90	313,000	2,063	9,300
12/70	13		01/91	376,000	2,338	
04/71	23		07/91	535,000	3,086	16,000
10/72	31		10/91	617,000	3,556	18,000
01/73	35		01/92	727,000	4,526	
06/74	62		04/92	890,000	5,291	20,000
03/77	111		07/92	992,000	6,569	16,300
12/79	188		10/92	1,136,000	7,505	18,100
08/81	213		01/93	1,313,000	8,258	21,000
05/82	235		04/93	1,486,000	9,722	22,000
08/83	562		07/93	1,776,000	13,767	26,000
10/84	1,024		10/93	2,056,000	16,533	28,000
10/85	1,961		01/94	2,217,000	20,539	30,000
02/86	2,308		07/94	3,212,000	25,210	46,000

11/86	5,089		10/94	3,864,000	37,022	56,000
12/87	28,174		01/95	4,852,000	39,410	71,000
07/88	33,000		07/95	6,642,000	61,538	120,000
10/88	56,000		01/96	9,472,000	93,671	240,000
01/89	80,000		07/96	12,881,000	134,365	488,000
			01/97	16,146,000	828,000	
			07/97	19,540,000	1,301,000	

*** Ver nota ***

Nota: En Enero de 1998 de desarrolló un mecanismo de investigación más preciso; las cifras correctas están expresadas en el siguiente cuadro y datan de Enero de 1995.

Fecha	Servidores		Fecha	Servidores		Fecha	Servidores
----	-----	+	----	-----	+	----	-----
01/95	5,846,000		01/97	21,819,000		01/99	43,230,000
07/95	8,200,000		07/97	26,053,000		07/99	56,218,000
01/96	14,352,000		01/98	29,670,000		01/00	72,398,092
07/96	16,729,000		07/98	36,739,000		07/00	93,047,785

Figura: Hosts

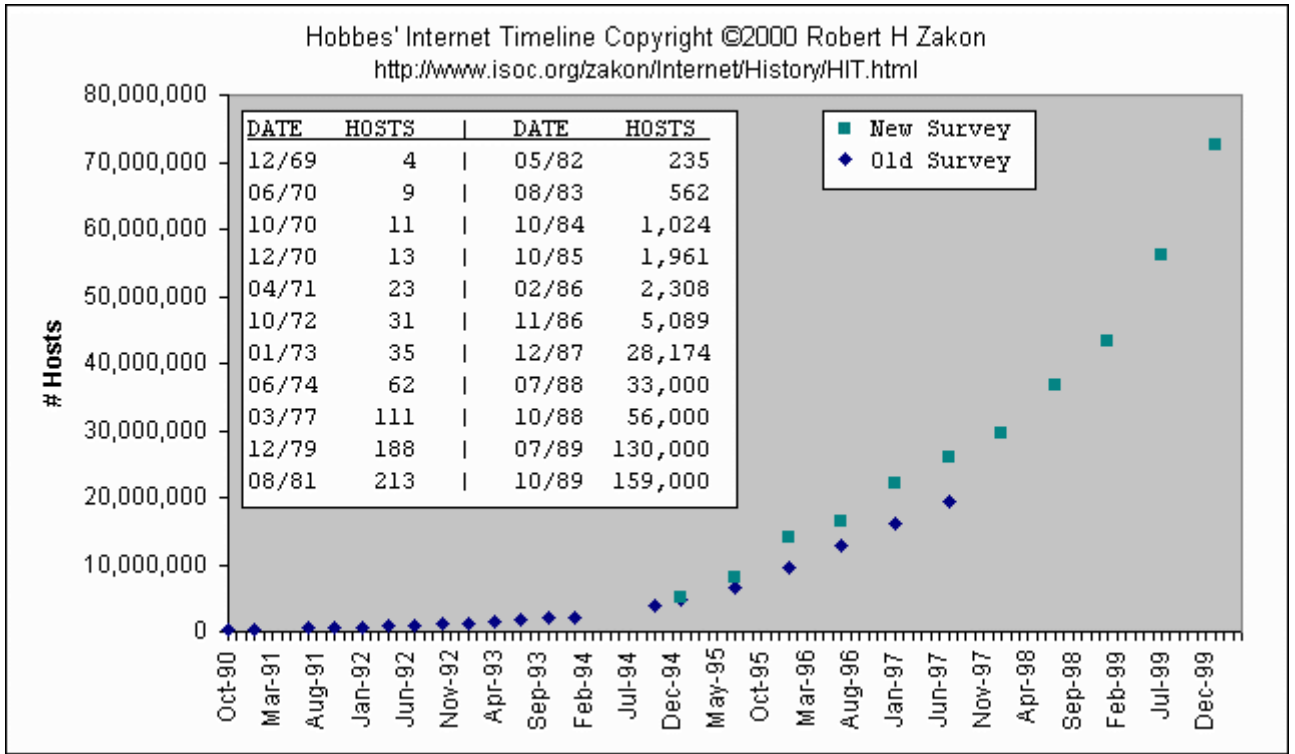


Figura: Dominios

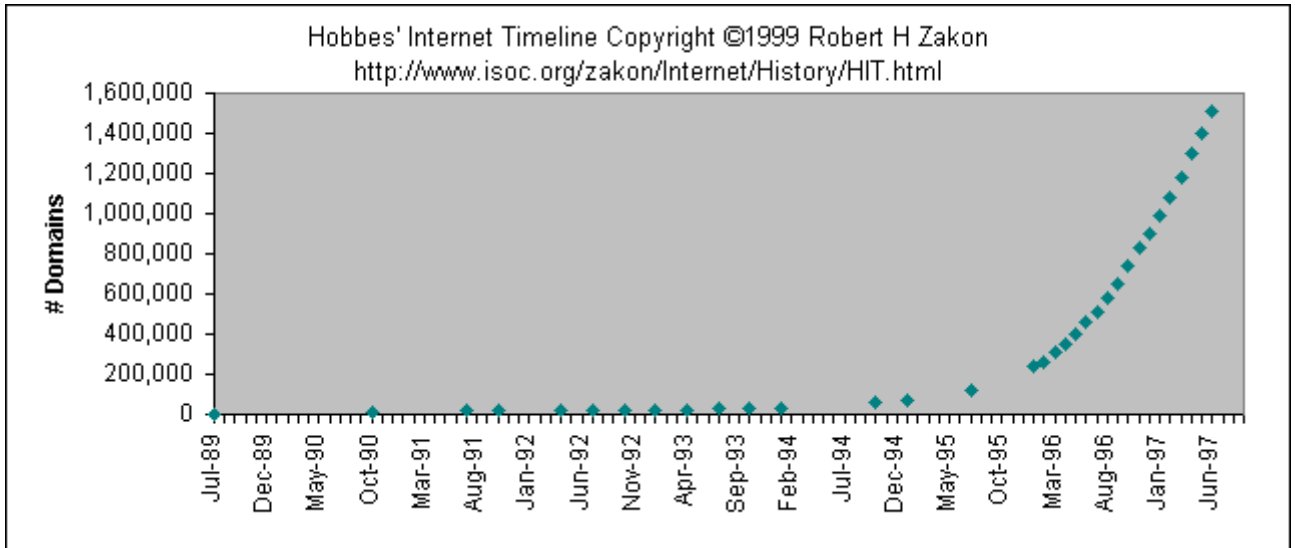
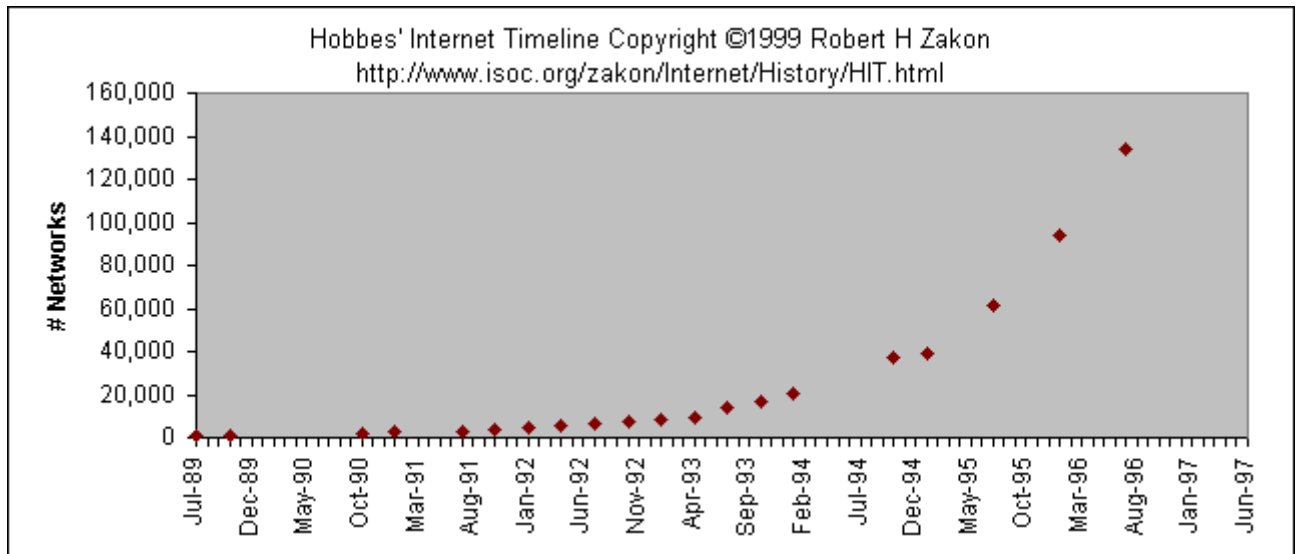


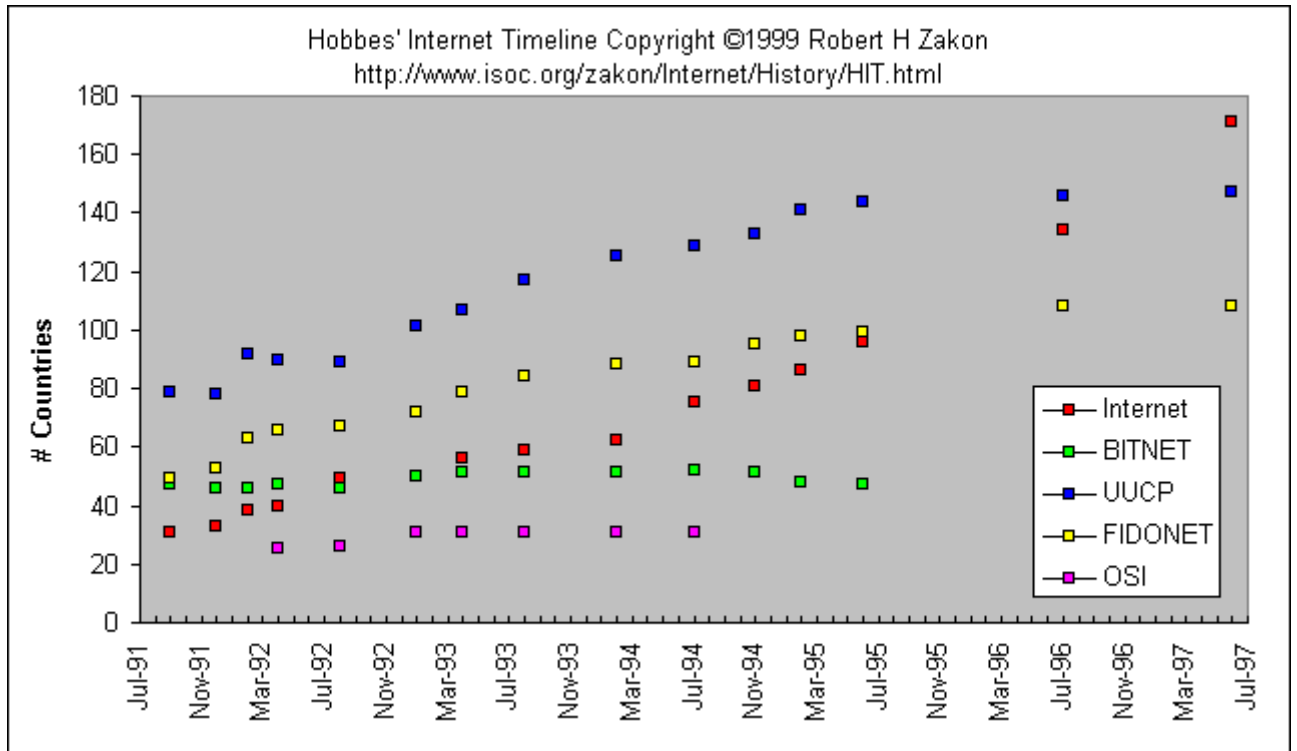
Figura: Redes



Crecimiento de las redes en el mundo:(I)nternet (B)ITNET (U)UCP (F)IDONET

(O)SI

_____# Paisés_____						_____# Paisés_____					
Date	I	B	U	F	O	Date	I	B	U	F	O
09/91	31	47	79	49		02/94	62	51	125	88	31
12/91	33	46	78	53		07/94	75	52	129	89	31
02/92	38	46	92	63		11/94	81	51	133	95	--
04/92	40	47	90	66	25	02/95	86	48	141	98	--
08/92	49	46	89	67	26	06/95	96	47	144	99	--
01/93	50	50	101	72	31	06/96	134	--	146	108	--
04/93	56	51	107	79	31	07/97	171	--	147	108	--
08/93	59	51	117	84	31						

Figura: Crecimiento de las redes en el mundo**Crecimiento de la WWW:**

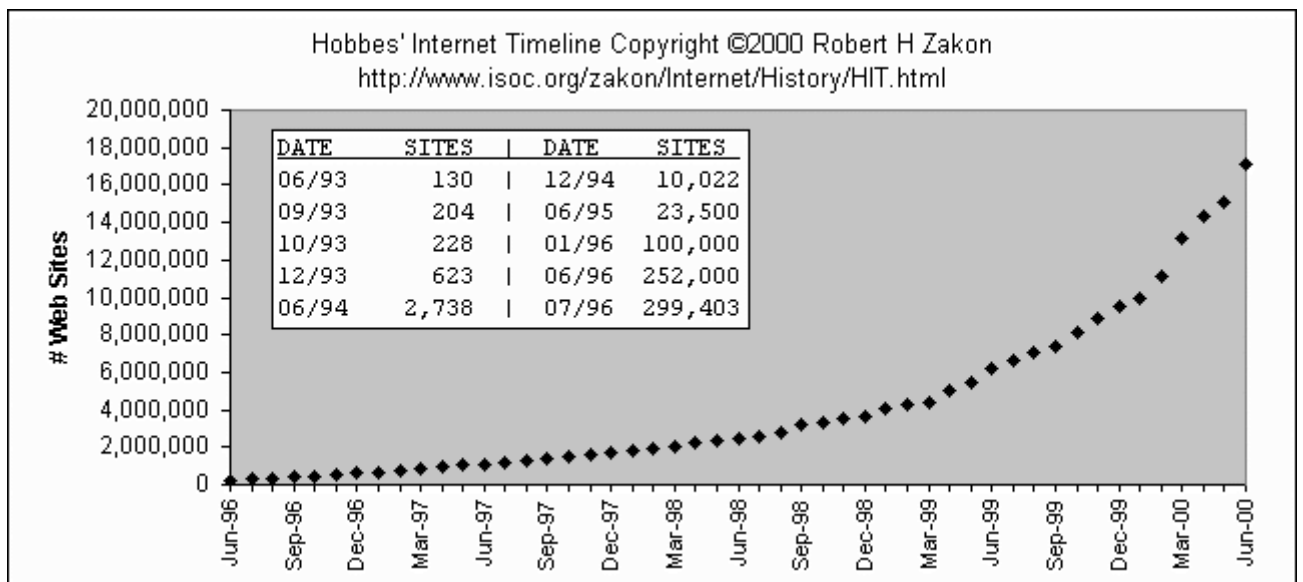
Fecha Sitios | Fecha sitios | Fecha Sitios

----- + ----- + -----

06/93	130		04/97	1,002,512		10/98	3,358,969
09/93	204		05/97	1,044,163		11/98	3,518,158
10/93	228		06/97	1,117,255		12/98	3,689,227
12/93	623		07/97	1,203,096		01/99	4,062,280
06/94	2,738		08/97	1,269,800		02/99	4,301,512
12/94	10,022		09/97	1,364,714		03/99	4,389,131
06/95	23,500		10/97	1,466,906		04/99	5,040,663
01/96	100,000		11/97	1,553,998		05/99	5,414,325
06/96	252,000		12/97	1,681,868		06/99	6,177,453

07/96	299,403		01/98	1,834,710		07/99	6,598,697
08/96	342,081		02/98	1,920,933		08/99	7,078,194
09/96	397,281		03/98	2,084,473		09/99	7,370,929
10/96	462,047		04/98	2,215,195		10/99	8,115,828
11/96	525,906		05/98	2,308,502		11/99	8,844,573
12/96	603,367		06/98	2,410,067		12/99	9,560,866
01/97	646,162		07/98	2,594,622		01/00	9,950,491
02/97	739,688		08/98	2,807,588		02/00	11,161,811
03/97	883,149		09/98	3,156,324		03/00	13,106,190
			04/00	14,322,950			
			05/00	15,049,382			
			06/00	17,119,262			
			07/00	18,169,498			
			08/00	19,823,296			
			09/00	21,166,912			

Figura: Crecimiento de Servidores WWW



Crecimiento de USENET:

Fecha	Sitios	~MB	~Mensajes	Grupos	Fecha	Sitios	~MB	~Mensajes	Grupos
1979	3	2	3		1987	5,200	2	957	259
1980	15	10			1988	7,800	4	1933	381
1981	150	0.05	20		1990	33,000	10	4,500	1,300
1982	400	35			1991	40,000	25	10,000	1,851
1983	600	120			1992	63,000	42	17,556	4,302
1984	900	225			1993	110,000	70	32,325	8,279
1985	1,300	1.0	375		1994	180,000	157	72,755	10,696
1986	2,200	2.0	946	241	1995	330,000	586	131,614	

~ aproximadamente: MB - megabytes por día, Posts - artículos por día

Incidentes de Seguridad (CERT):

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Inc	6	132	252	406	773	1334	2340	2412	2573	2134	3734	9859
Adv	1	7	12	23	21	19	15	18	27	28	13	17
Vul						171	345	311	262	417		

(Inc)idents, (Adv)isories, (Vul)nerabilities

FUENTES UTILIZADAS

La cronología de Internet de Hobbes ha sido compilada a partir de una serie de fuentes.

Las más importantes son:

Cerf, Vinton (conversaciones con Bernard Aboba). "How the Internet Came to Be."

Este artículo aparece en "The Online User's Encyclopedia," por Bernard Aboba.

Addison-Wesley, 1993.

Hardy, Henry. "The History of the Net." Master's Thesis, School of Communications, Grand Valley State University.

<http://www.ocean.ic.net/ftp/doc/nethist.html>

Hardy, Ian. "The Evolution of ARPANET email." History Thesis, UC Berkeley.

<http://www.ifla.org/documents/internet/hari1.txt>

Hauben, Ronda and Michael. "The Netizens and the Wonderful World of the Net."

<http://www.columbia.edu/~hauben/netbook/>

Kulikowski, Stan II. "A Timeline of Network History." (el email del autor abajo)

Quarterman, John. "The Matrix: Computer Networks and Conferencing Systems Worldwide." Bedford, MA: Digital Press. 1990

"ARPANET, the Defense Data Network, and Internet". Encyclopedia of Communications, Volume 1. Editors: Fritz Froehlich, Allen Kent.

New York: Marcel Dekker, Inc. 1991

El resumen del crecimiento de Internet se compiló en base a:

- Reportes de programas por zonas mantenido por Mark Lottor:

<ftp://ftp.nw.com/pub/zone/>

- Nota: En Enero de 1998 comenzó el uso de un sistema más preciso de contabilización de hosts.

- Tabla de conectabilidad mantenida por Larry Landweber:

ftp://ftp.cs.wisc.edu/connectivity_table/

- Mapas de ARPAnet publicados en varios recursos

El resumen del crecimiento de la WWW se compiló a partir de:

- La página sobre el crecimiento de la Web de Matthew Gray de MIT:

<http://www.mit.edu/people/mkgray/net/web-growth-summary.html>

- Netcraft en <http://www.netcraft.com/survey/>

El resumen del crecimiento de USENET se compiló a partir de las fuentes ya mencionadas y de los mensajes news.lists. Tom Fitzgerald colaboró con muchos mensajes históricos de USENET (fitz@wang.com).

El resumen del crecimiento de CERT se compiló a partir de los reportes de Cert en

<ftp://ftp.cert.org/>

También disponible en http://www.cert.org/stats/cert_stats.html

Arnaud Dufour colaboró con gran parte de los URLs (arnaud.dufour@hec.unil.ch)

Historias de internet de países específicos:

- Australia - "A Brief History of the Internet in Australia" por Roger Clarke

<http://www.anu.edu.au/people/Roger.Clarke/II/OzIHist.html>

- Australia - "It Started with a Ping" por Jennie Sinclair

<http://www.aarnet.edu.au/corporate/history/sinclair.html>

- Brasil - "Linha to Tempo da Internet no Brasil" por Érico Guizzo

<http://www.lsi.usp.br/~emguizzo/inetbr/>

- UK - "Early Experiences with the ARPANET and INTERNET in the UK" por Peter Kirstein

<http://www.cs.ucl.ac.uk/staff/jon/arpa/internet-history.html>

Libros adicionales sobre el tema:

- "Weaving the Web : The Original Design and Ultimate Destiny of the World Wide

Web by its Inventor"

por Tim Berners-Lee

- "Where Wizards Stay Up Late: The Origins of the Internet"

por Katie Hafner & Matthew Lyon

- "Nerds 2.0.1: A Brief History of the Internet"

por Stephen Segaller

- "Architects of the Web: 1,000 Days That Built the Future of Business"

por Robert H. Reid

- "Netizens: On the History and Impact of Usenet and the Internet"

por Michael Hauben et al

- "Exploring the Internet: A Technical Travelogue"

por Carl Malamud

Obras tempranas de interes:

- "As We May Think" por Vannevar Bush, 1945

<http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>

- "Man-Computer Symbiosis" por J.C.R. Licklider, 1960

<http://gatekeeper.dec.com/pub/DEC/SRC/research-reports/abstracts/src-rr-061.html>

CONCLUSIONES

CONCLUSIONES:**I. El usuario receptor se ha convertido en una fuente de información clave para el periodista actual.**

Gracias a la creciente interactividad que permite la Red el usuario puede participar activamente, de forma directa e inmediata, en los medios de comunicación emitiendo sus opiniones respecto a temas de actualidad o funcionamiento del medio y actuando como fuente directa para el periodista.

II. El periodista debe conocer lo mejor posible la totalidad de los recursos de búsqueda en Internet en aras de realizar de forma eficaz su trabajo.

El periodista tiene a su disposición en Internet gran cantidad de recursos de búsqueda de información, entre ellos los motores de búsqueda y los directorios temáticos. Sin embargo no se debe olvidar de aquella información no accesible mediante los buscadores convencionales, es decir, de la llamada Web profunda que posee 500 veces mas información y una mayor calidad. Asimismo debe tener en cuenta la posibilidad que ofrece Internet a los periodistas de ponerse en contacto directo a través del correo electrónico con expertos en diversas materias y fuentes de primera mano a nivel mundial.

III. La credibilidad de las fuentes y de la información en Internet no es menor que en cualquier otro medio.

El deber del periodista es comprobar las fuentes y la información se trate del medio del que se trate. Por lo tanto, las posibles informaciones falsas que se hallaran en la Red no deberían suponer un problema para este profesional.

Las marcas de los productos que se difunden a través de Internet, incluidos los medios de comunicación serán y son en la actualidad definitivos, de cara al usuario receptor, a la hora de evaluar la credibilidad del medio.

IV. La información que difunden los medios a través de Internet no debería ser gratuita.

Aunque en la actualidad la mayor parte de la información de los medios de comunicación en Internet destaca por su gratuidad, ésta no será en breve la forma de distribuir dicha información ya que el sistema de financiación por publicidad de estas *webs* no es suficiente para afrontar los gastos de producción. Por este motivo principalmente, se tenderá a un método mixto en el que parte de la información será gratuita y parte de pago. La calidad y la originalidad son los elementos clave por los que el usuario receptor pagará por la información.

V. Los elementos multimedia y la interactividad son claves a la hora de elaborar una información en la Red .

Internet ofrece la posibilidad de incorporar a las informaciones tradicionales constituidas por elementos gráficos y texto, la multimedia, es decir, audio y vídeo y la interactividad la cual permite al periodista tener una relación más directa con el usuario receptor, además de hacerle partícipe del propio medio.

VI. La posibilidad de personalizar la información constituye una de las grandes ventajas de Internet pero no será un factor importante en un futuro próximo.

La Red permite en la actualidad al usuario personalizar los contenidos que quiere visualizar en su diario personal. Sin embargo, por ahora, la mayor parte de los usuarios-receptores de un medio dejan de lado esta personalización y se atienen al criterio de selección temática tradicional de los periodistas.

VII. Uno de los papeles fundamentales del periodista que trabaja con Internet será el de servir de guía virtual al usuario-receptor.

El periodista “digital” tendrá como función principal, además de la producción de información, la contextualización añadiendo enlaces comprobados y precisos de otras *webs* que pudieran servir de referencia al usuario.

VIII. La Red profunda es una fuente de información esencial para el periodista.

Los motores de búsqueda generales suelen tener la política de que “cualquier respuesta es mejor que ninguna”, en aras de llegar a un mayor número de usuarios que puede que no utilizaran dicho motor, sino les proporcionara siempre algún tipo de resultado. Por el contrario, las bases de datos disponibles en la Red profunda intentan ofrecer información más específica y limitada dirigida a un menor número de usuarios, teniendo en cuenta que “es preferible que no haya ninguna respuesta antes de que haya muchas incompletas o imprecisas.

IX. Es muy difícil estimar el tamaño real de la Red profunda, debido a su continuo crecimiento. Sin embargo, según los estudios realizados al respecto, se estima que posee aproximadamente 500 veces más información que la parte de Internet accesible mediante buscadores generales y, sobre todo, de una mayor calidad. La Red profunda la compone fundamentalmente información disponible en bases de datos temáticas especializadas, de la cual la parte más voluminosa son imágenes.

X. La Red profunda no es la solución para todas las necesidades informativas Para obtener los resultados óptimos deben usarse conjuntamente con otras fuentes de información, incluyendo motores de búsqueda, directorios e incluso fuentes bibliográficas tradicionales (hay muchos documentos, sobre todo textos antiguos que no han sido digitalizados y por tanto no están disponibles en Internet).

BIBLIOGRAFÍA

BIBLIOGRAFÍA (FUENTES IMPRESAS):**BIBLIOGRAFÍA (Fuentes impresas):**

- Adam, N.R; Bhargava, B.K; Halem, M. y Yesha, Y. 1996. *Digital libraries*. Springer.
- Alexander, Janet E. and Marsha Ann Tate (1999): *Web wisdom: how to evaluate and Create Information Quality on the Web*. Ed: Lawrence Erlbaum Associates, Inc.
- Amat, Nuria (1990): *De la información al saber*. Ed: Fundesco. Madrid.
- Armañanzas, Emy; Díaz Noci, Javier y Meso, Koldo (1996): *El periodismo electrónico. Información y servicios multimedia en la Era del ciberespacio*. Barcelona: Ariel.
- Armentia Vizueté, José I.; Elexgaray, Jon; Pérez, Juan Carlos (1999): *Diseño periodismo electrónico*. Bilbao: Universidad del País Vasco.
- Azcona, Sergio (2001): *Internet de la @ a la Z*. Alianza Editorial. Madrid.
- Barnes, Sue (1994): "Hypertext Literacy", IPCT. *Interpersonal Computing and Technology*, Volume 2, Number 4, October 1994. Washington D.C: Center for Teaching and Technology, Academic Computer Center, Georgetown University, pp. 24-36.
- Barret, E.. 1988. *Text, ConText, and Hypertext: Writing with and for the Computer*. Cambridge, MIT Press.
- Bates, Mary Ellen (2001): *Super Searchers cover the world: The online secrets of international business researchers*. Editado por Reva Basch. CyberAge Books. Information Today, Inc.

- Bayón, Javier; Egía, Carlos (1997): *Contrainformación. Alternativas de comunicación escrita en Euskal Herria*. Bilbao: Likiniano Elkarte.
- Beaumont, José F (1994): *La prensa cambia de papel*. En *Apuntes de la sociedad interactiva*. Madrid: UIMP/Fundesco.
- Bell, H.K. (1996): "The Learned Societ, Journals and the internet", *Learned Publishing*, July 1996.
- Berganza Conde, María Rosa y Rodríguez Pardo, Julián (editores) (1997): *El comunicador ante el reto de las nuevas tecnologías*", en Actas del III Congreso sobre nuevas tecnologías. Facultad de Comunicación. Universidad de Navarra.
- Berge, Zane L.; Collins, Mauri P. (1995): *Computer Mediated Communication and the Online Classroom*. Cresskil NJ: Hampton Press.
- Bettetini, Gianfranco; Colombo, Fausto (1995): *Las nuevas tecnologías de la información*. Barcelona. Paidós.
- Bolter, J. D. (1991): *Writing Space: The Computer, Hypertext and the History of Writing*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Botas, Antonio: *La efectividad de la publicidad en Internet*. (Ponencia en el 3º Congreso Nacional de Usuarios de Internet: "Mundo Internet 98"). Madrid 1998.
- Caldwell, Barrett S.; Taha, Lias H. (1993): "Starving at he banquet: Social Isolation in Electronic Communication Media", *IPCT. Interpersonal Computing*

and Technology, Volme 1, Number 1, January 1993. Washington DC: Center for Teaching and Tecnology, Academic Computer Center, Georgetown University.

- Caridad, Mercedes, y Moscoso, Purificación (1991): *Los sistemas de hipertexto e hipermedios. Una nueva aplicación en informática documental*. Madrid: Fundación Germán Sánchez Ruipérez.
- Caruso, Denise (1997): *Show me the Money! How the FUD factor has online news in its thrall*. En *Columbia Journalism Review*, July/August 1997. New York: Columbia University.
- Carvazos, Edward A.; Gavino Morin (1994): *Cyberspace and the Law. Your Rights and Duties in the On-Line World*. Cambridge (Massachusetts): The MIT Press.
- Castells, Manuel (1996): *La Era de la Información (Economía, sociedad y cultura)*. Vol.1. La sociedad Red. Madrid. Alianza Editorial.
- Cebrián, Juan Luis (1998): *La red*. Madrid: Taurus.
- Cerdan, Michel (1990): "The role of 'Libération' in Minitel, the first interactive mass medium", IFRA. *Newspaper Techniques*, June 1990, Munich.
- Cohen, William W., Andrew McCallum, and Dallon Quass (2000): Learning to Understand the Web. *IEEE Data Engineering Bulletin* 23.3.
- Collins, Mauri P.; Berge, Zane L. (1994): "IPCT Journal: A Case Study of an Electronic Journal on the Internet", *Journal of the American Society for Information Science*, December 1994, 45 (10). Calder/Washington D.C.: Penn State University/Georgetown University.

- Cooke, Alison. Neal-Schuman (1999): *Authoritative Guide to Evaluation Information on the Internet*. Ed: Neal Schuman.
- Covi, Lisa M. (1997): *The future of Electronic Journals: Umpuzzling Researchers´ Attitudes about Electronic Journals*. En *Revista Española de Bibliología*, Vol. 1. Valencia.
- Crandall, R. y Levich, M. 1997. *The Virtual Goliath*. Springer.
- Detwiler, Susan M (2000): *Super searchers on Health & Medicine*. CyberAge Books. Information Today, Inc.
- Díaz Noci, Javier y Meso Ayerdi, Koldo (1999): *Periodismo en Internet. Modelos de la prensa digital*. Bilbao: Universidad del País Vasco.
- Díaz Noci, Javier, Koldo Meso (1997): *Medios de comunicación en Internet* Madrid: Anaya.
- Díaz Nosty, Bernardo (1998): *Las huellas digitales de Gutemberg. Informa Anual de la Comunicación. Estado y tendencias de los medios en España*. Barcelona: Zeta.
- Diffley, Mary (2001): *Internet Prophets*. CyberAge Books. Information Today, Inc.
- Doménech, Eudald y Almiron, Nuria (2001): *Negocios 3.0. Mitos y realidades de Internet y la Nueva Economía*. Ediciones B S.A. Barcelona.
- Donovan, Deirdre C. (1995): "Computer Mediated Communication and the Basic Speech Course, *IPCT. Interpersonal Computing and Tecnology*, Volume

3, Number 3, October, 1995. Washington D.C: Center for Teaching and Technology, Academic Computer Center, Georgetown University.

- Dubie, W. (1994): “Networds: The Impact of Electronic text Processing Utilities on Writing”, *Journal of Social and Evolutionary Systems*, nº17 (2). Littleton: Northhearstern University.
- Dubois, J.E. y Gershon, N. 1996. *The Information Revolution: Impact on Science and Technology*. Springer.
- Echevarría, Javier (1994): *Telépolis*. Barcelona: Destino.
- Echeverría, Javier (1999): *Los señores del aire: Telépolis y el tercer entorno*. Barcelona: Destino.
- Echeverría, Javier (1996): *Internet y el periodismo electrónico*, en *Colegio de Periodistas*, Barcelona.
- Echevarría, Javier (1995): *Cosmopolitas domésticos*. Barcelona. Barcelona: Anagrama.
- Edo, Concha (2000): *Los periódicos se instalan en definitivamente en la red*, en *Derecho y Opinión*, Universidad de Córdoba.
- Flores, José Miguel (1999): *Incidencias del sistema global de la información electrónica en el periodismo contemporáneo*. Tesis doctoral. Facultad de Ciencias de la Información, Universidad Complutense de Madrid
- Foucault, Michel. 1976. *The Archeology of Knowledge*. (trad. A.M. Sheridan Smith, Nueva York, Harper Colophon).

- Fuentes I Pujol, María Eulalia (1997): *La información en Internet*. Barcelona: Cims 97.
- Gill, K. S. 1996. *Information Society: New Media, Ethics and Postmodernism*. Springer.
- Hane, Paula J. (2000): *Super Searchers in the News: The online secrets of journalist and News Researchers*. Editado por Reva Basch. CyberAge Books. Information Today, Inc.
- Hass, S. W.. 1995. Quotations in Scholarly Text: Converting Existing Documents to Hypertext. *Computers and the Humanities*, 28: 165-176.
- Harrison, Teresa M.; Stephen, Timothy (1997): "On the future of Electronic Academic Journal Publication: Tecnology, Economics, and Sociology. En *Revista Española de Bibliología*, Vol 1. Valencia.
- Harter, S.P; "Accesing Electronic Journals and Other E Publications: An Empírica Study" *College and Research Libraries*, September 1996. Bloomington: Indiana University.
- Head, Alison J (1999): *Design Wise. A guide for evaluating the interface design of information sources*. CyberAge Books. Information Today, Inc.
-
- Hock, Randolf (2001): *The extreme searcher's guide to Web Search Engines*. CyberAge Books, Information Today, Inc.
- Hunter, Karen (1994): *Issues and experiments in Electronic publishing and dissemination*. En *Information Technology and Libraries*.

- Insa Ghisaura, Daniel y Morata, Rosario (1998): *Multimedia en Internet*. Paraninfo: Madrid.

- Kassel, Amelia (2000): *Super searchers on Wall Street. Top investment professionals share their online research secrets*. Editado por Reva Basch. CyberAge Books. Information Today, Inc.

-

- Kommers, P. A. M; Ferreira, A. y Kwak, A. K; 1996. *Document Management for Hypermedia Design*. Springer.

- Krug, Steve (2001): *No me hagas pensar. Una aproximación a la usabilidad en la Web*. Prentice Hall. Madrid.

- Landow, George P. (1995): *Hipertexto. La convergencia de la teoría crítica contemporánea y la tecnología*. Barcelona: Piados.

- Lane, Carole A (1997): *Naked in Cyberspace. How to find personal information online*. CyberAge Books. Information Today, Inc.

- Lanza, Sheri R. (2001): *International Business Information on the Web*. CyberAge Books. Information Today, Inc.

- Lévy, Pierre (1998): *La cibercultura, el segundo diluvio*. Barcelona: Universidad Oberta.

- Maher, Mike (25 de junio de 1994): *Electronic Gatekeeper for News*. En Editor & Publisher.

- Martínez Albertos, José Luis (1988): *Efectos de la tecnología electrónica sobre la comunicación periodística*. En Revista de Ciencias de la Información, nº 5. Madrid. Universidad Complutense.
- Martínez Albertos, José Luis (1999): *El periodismo en el sigloXXI: más allá del rumor y por encima del caos*, en *Estudios sobre el mensaje periodístico* nº 5, Universidad Complutense, Madrid.
- Martínez Albertos, José Luis (1998): *El ocaso del periodismo*. Barcelona: Cims 97.
- McAdams, Melinda (1995): *Inventing an Online Newspaper, IPCT. Interpersonal Computing and Tecnology*, Volume 3, Number 3, July 1995. Washington D.C: Center for Teaching and Tecnology, Academic Computer Center, Georgetown University.
- McKenna, Kate (1993): *The future is now. Newspapers are overcoming their fears of tecnology and launching a wide array of electronic products*, *American Journalism Review*, October.
- McGuire, Mary; Stilborne, Linda; McAdams, Melinda; Hyatt, Laurel: *The Internet Handbook for Writers, Researchers and journalists*. Ed: Trifolium Books. Canada 1997.
- Miller, Jerry P.(2000): *Millennium Intelligence*. CyberAge Books. Information Today, Inc.
- Nelson, Theodor H. 1981. *Literary Machines*. (Swarthmore, Pa., publicación propia.)
- Nielsen, Jakob (1995): *Multimedia and Hypertext*, AP professional.

- Negroponte, Nicholas (1995): *El Mundo Digital*, Barcelona. Ediciones B, S.A.
- O'Callaghan, Xabier (1996): *La autolimitación de la libertad de expresión*, *Revista Jurídica Española* La Ley, 1996, nº 1.
- Olaechea Labayen, Juan (1985): *Las publicaciones periódicas científicas la difusión electrónica*, en *Boletín de la Anabad*, XXXV, número 4.
- Paul, Nora and Williams, Margot: *Great Scouts! Cyberguides for subject searching on the web*. Editado por Paula Hane. CyberAge Books. Information Today, Inc.
- Pavlink, John V. (1997): *The future of Online Journalism. Bonanza or Black Hole?*. En: *Columbia Journalism Review*, July/August 1997. New York: Columbia University.
- Pérez Luque, María José y Perea Foronda, Maider (1998): *El reto de crear noticias online. Análisis de la comunicación online actual y perspectivas de futuro*, en *Cuadernos de Documentación Multimedia* 6-7, Universidad Complutense, Madrid.
- Pérez Manzanera, Leandro (2001): *Las tecnologías de la información en la nueva economía*. Ed: Díaz de Santos, S.A. Madrid
- Piedrahita, Manuel (1998): *Periodismo impreso, audiovisual y electrónico del siglo XXI*. Madrid: Madrid Universitas.
- Ramonet, Ignacio (1998): *Internet, el mundo que llega. Los nuevos caminos de la comunicación*. Madrid: Alianza.

- Reddick, Randy; Elliot King (1995): *The Online Journalist: Using the Internet and Other Electronic Resources*, London: Harcourt Brace College Publishers.
- Rodríguez de las Heras, Antonio (1991): *Navegar por la información*, Madrid: Fundesco.
- Sacks, Risa (2001): *Super Searchers go to the source*. CyberAge Books. Information Today, Inc.
- Sartori, Giovanni (1998): *Homo videns. La sociedad teledirigida*. Madrid: Taurus.
- Sharples, M. y van der Geest, T. 1996. *The New Writing Environment*. Springer.
- Sherman, Chris and Price, Gary (2001): *The Invisible Web. Uncovering information sources search engines can't see*. Ed. CyberAge Books. Information Today Inc. Medford, New Jersey (EE UU).
- Smith, Anthony (1985): *Goodbye Gutenberg. The Newspaper Revolution of the 1980's*, Oxford University Press (Traducción castellana (1983): *Goodbye Gutenberg. La revolución del periodismo electrónico*. Barcelona: Gustavo Gili).
- Smith, Anthony (1991): *Informática y comunicación*. Valencia: Generalitat de Valencia.
- Smith, Anthony (1991): *Predicciones para una década*, pp 155-162 en *Comunicación Social 1991. Tendencias. Informes anuales de Fundesco*, Madrid: Fundesco.

- Steinert, Eric C. and Bleier, Thomas E. (2000): *Net.people. The personalities and passions behing the web sites*. CyberAge Books. Information Today, Inc.
-
- Stodolsky, David S. (1994): *Telematic Journals and Organizational Control: Integrity, Authority, and Self-Regulation*, IPCT. Interpersonal Computing and Technology, Volume 2, Number 1, January 1994. Washington D.C: Center for Teaching and Tecnology, Academic Computer Center, Georgetown University.
- Terceiro, José B (1996): *Socied@d digit@l (Del homo sapiens al homo digitalis)*, Madrid. Alianza Editorial.
- Trejo Delarbre, Raúl (1996): *La nueva alfombra mágica: usos y mitos de Internet, la red de redes*. Madrid: Fundesco.
- Tudor, Jan Davis (2001): *Super Searchers on Mergers & Acquisitions*. Editado por Reva Basch. CyberAge Books. Information Today, Inc.
- Vine, David (2000): *Internet Business Intelligence. How to build a big company system on a small company budget*. CyberAge Books. Information Today, Inc.
- VV.AA (1994): *Apuntes de la sociedad interactiva. Autopistas inteligentes y negocios multimedia*. Madrid: Fundesco, Colección Encuentros, 2.
- VV.AA. (1994): *Telecommunications and Interactive newspapers*, Editor & Publisher, nº especial, February 12, 1994.
- VV.AA (1996): *El derecho de la propiedad intelectual y las nuevas tecnologías*. Madrid: Ministerio de Cultura.

- Weber, Nicolás (1997): *Cómo encontrar lo que buscas en Internet: 1.001 secretos*. Barcelona: Sirpus.
- Williams, F.; R.Rice.; E.Rogers (1988): *Research Methods and the New Media*, New York: Free Press.
- Wilson, Stephen (1995): *World Wide Web Design Guide*. Indianapolis: Hayden Books.

BIBLIOGRAFÍA (FUENTES ELECTRÓNICAS):

- Ad Resource: Información sobre publicidad en Internet.
(http://adres.internet.com/feature/article/0,1401,8961_341501,00.html)
- Armañanzas, Emy (1996): “El periodismo electrónico, resultado de una globalidad económica y social”, conferencia en el curso de verano sobre *El periodismo electrónico en la era del ciberespacio*, Universidad del País Vasco, San Sebastián, 15 de julio de 1996 (accesible en <http://www.pd.lp.ehu.es/Website/Memoria/Docentes/Armañanzas/Curso96.html>).
- *ASJA Contracts Watch*, disponible bajo solicitud mediante correo electrónico a asja-manager@silverquick.com, y en <http://www.asja.org/cwpage.htm>
- Auer, N. (1998). Bibliography on evaluating Internet resources. (Accesible en: <http://refserver.lib.vt.edu/libinst/critTHINK.HTM>)

- Berners-Lee, Tim (2000): *Subject Listing* (accessible en: <http://www.w3.org/History/19921103-hypertext/hypertext/DataSources/bySubject/Overview.html>).
- Berners-Lee, Tim (1989): *Information Management: A Proposal*, <http://www.w3.org/pub/WWW/History/1989/proposal.html>.
- Bilodeau, Anne (1994): *An Overview of the Society of Electronic News Delivery*, CMC Computer-Mediated Communication Magazine, volume 1, number 3, July 1994 (accesible en: <http://www.december.com/cmc/mag/archive/title.html>).
- Borrell, Jerry (1995): *The Emerging Role of Journalist in Multimedia*, the Quill Magazine, May 1995 (Accesible en: <http://town.hall.org/places/spj/quill.html>).
- Bray, T. (1996). *Measuring the Web. Fifth International World Wide Web Conference*. May 6-10, 1996, Paris, Francia. (Accesible en: http://www5conf.inria.fr/fich_html/papers/P9/Overview.html)
- Bright Planet (2000): *The Deep Web: Surfacing Hidden Value* (Accesible en: <http://128.121.227.57/download/deepwebwhitepaper.pdf>).
- Broder, Andrei, et al (2000): *Graph Structure in the Web*. Paper presented at Ninth International World Wide Web Conference. Amsterdam. (Accesible en: <http://www.almaden.ibm.com/cs/k53/www9.final/>).
- Bullón, p. *Algo está cambiando. La Prensa ante las nuevas tecnologías*. (Accesible en: <http://www.saladeprensa.org/art35.htm>).
- Bus, Vannevar (1945): *As We May Think* (accessible en: <http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>).

- Cabrera González, María Angeles (2000): *Retos en el diseño de periódicos en Internet*, en *Revista Latina de Comunicación Social* (Accesible en: <http://www.ull.es/publicaciones/latina/aa2000yene/149malaga.html>)
- Canga, Jesús, *et al* (1999): *¿Quién le teme al periódico digital?*, en *Revista Latina de Comunicación Social* (Accesible en: <http://www.ull.es/publicaciones/latina/a1999hmy/89pvasco.htm>).
- Carrière, J. y Kazman, R. (Octubre 1995): *Interacting with Huge Hierarchies: Beyond Cone Trees*, en *Proceedings of IEEE Information Visualization '95*, (Atlanta, GA). (Accesible en: <ftp://ftp.cgl.uwaterloo.ca/pub/users/rnkazman/fsviz.ps.Z>).
- Cerf, Vinton (2000): *A Brief History of the Internet and Related Networks*. (Accesible en: <http://www.isoc.org/internet/history/cerf.html>).
- Chesnais, Pascal R.; Mucklo, Matthew J.; Sheena, Jonathan A. (1995): "The Fishwrap Personalized news System", artículo accesible en el Massachusetts Institute of Technology (<http://nif.www.media.mit.edu>). Puede verse un ejemplo de este periódico personalizado en <http://fishwrap.mit.edu>.
- Connolly, Dan (2000): *A little History of the World Wide*. (Accesible: <http://www.w3.org/History.html>).
- Cornella, A. "Eureka: es la infonomía, y somos infonomistas". Artículo de Extra!Net (1997)<http://www.extra-net.net/articulos/en971108.htm>
- Cornellá, Alfons (1999): "Nuevas profesiones en la información", en *Revista Extra!-Net*, 26 de febrero de 1999. (Accesible en: <http://www.extra-net.net/>)

- Deutsch, Peter (2000): Archie-A Darwinian Development Process. IEEE Internet Computing 4.1. (Accesible en: <http://www.computer.org/internet/v4nl/deutsch.htm>).
- Díaz Noci, Javier (1994): “El nuevo periódico electrónico: redefinición del mensaje periodístico tradicional como producto interactivo y multimedia”, comunicación presentada a las *IX Jornadas Internacionales de Ciencias de la Información*, Universidad de Navarra, Pamplona (accesible en <http://www.pd.lp.ehu.es/Website/Memoria/Docentes/Díaz%20Doci/Textos/Conferencias/TRAINTEC.html>)
- Díaz Noci, Javier (1996): “Conceptos en torno al periodismo electrónico multimedia interactivo2, conferencia en el curso de verano sobre el periodismo electrónico en la era del ciberespacio, Universidad del País Vasco, San Sebastián, 15 de julio de 1996 (accesible en <http://www.pd.lp.ehu.es/Website/Memoria/Docentes/Díaz%20Noci/Textos/Conferencias/donosti.html>).
- Douglis, F., Feldmann, A., Krishnamurthy, B. & Mogul, J. (1997). “Rate of change and other metrics: A live study of the World Wide Web”. *USENIX Symposium on Internet Technologies and Systems*, December 1997: 147--158. (Accesible en: <http://www.research.att.com/~douglis/papers/roc/>)
- Echevarría, J. "Periodismo electrónico en la sociedad del futuro". (Accesible en: <http://www.saladeprensa.org/art08.htm>).
- ERIC. Educational Resources Information Center. 2000. (Accesible en: <http://www.accesseric.org>)

- Fernández Hermana, Luis Ángel (1996): *En busca del periodista digital*. Revista *Enredando* (Accesible en: <http://enredando.com>)
- Fernández Hermana, L.A. "De exploradores y cartógrafos" Editorial de *En.red.ando* (1997) (accesible en: <http://enredando.com/cas/cgi-bin/editorial/plantilla.pl?ident=252>)
- Fetuccino. IBM Research.2000. (Accesible en: <http://www.ibm.com/java/fetuccino/fetuccino-abstract.html>).
- Fielding, R (Mayo 1994): *Maintaining Distributed Hypertext Infostructures: Welcome to MOMspider's Web*, in Proceedings of the First World Wide Web Conference (Génova). (Accesible en: <http://www.cern.ch/PapersWWW94/fielding.ps>).
- Gach, Gary (1994): "The Internet for Journalists: The Fourth Estate in Cyberspace", *CMC Computer-Mediated Communication Magazine*, volume 1, number 6, October 1994 (Accesible en: <http://www.december.com/cmc/mag/archive/title.html>).
- Hoover's Online (2000): *The Business Network*. (Accesible en: <http://www.hoovers.com>).
- Hsieh-Yee, I. ([1996]). "Modifying catalogueing practice and OCLC infrastructure for effective organization of Internet resources". OCLC Internet catalogueing Project Colloquium Position Paper. (Accesible en: <http://www.oclc.org/oclc/man/colloq/hsieh.htm>).
- Inktomi Slurp (2000). (Accesible en: <http://www.inktomi.com/slurp.html>).
- Ipmark: Semanario de Publicidad y Marketing. (<http://www.ipmark.com/>)

- J. Carrière, R. Kazman (Octubre 1995): "Interacting with Huge Hierarchies: Beyond Cone Trees", en la revista *Proceedings of IEEE Information Visualization*. (Accesible en: <ftp://ftp.cgl.uwaterloo.ca/pub/users/rnkazman/fsviz.ps.Z>).
- Kahn, Robert E, and Vinton G Cerf (2000): *What is the Internet (And what makes it works)*. Internet Policy Institute. (Accesible en: http://www.internetpolicy.org/briefing/12_99_story.html)
- Katz, Jon (1995): "Online or Not, Newspapers Suck. How can any industry which regularly pulls Doonesbury strips for being too controversial possibly hope to survive online?", en *HotWired* (<http://www.hotwired.com>).
- Koehler, Wallace (2000): *Digital Libraries and World Wide Web Sites and Web Page Persistence*. Information Research 4.4. (Accesible en: <http://www.shef.ac.uk/~is/publications/infres/paper60.html>).
- Koll, Matthew (1998): Major trends and Issues in the Information Industry. ASIDIC Newsletter. (<http://www.asidic.org/techsumf99.html>).
- Labour Market Information (2000): (Accesible en: http://Imi-imt.hrdc-drhc.gc.ca/owa_Imi/owa/sp_show_Imi?1=e&I=1).
- [8] S. Lamm, D. Reed, W. Scullin (Mayo 1996): "Real-Time Geographic Visualization of World Wide Web Traffic", in *Proceedings of the Fifth World Wide Web Conference (Paris)*. (Accesible en: http://www5conf.inria.fr/fich_html/papers/P49/Overview.html).

- Leiner, Barry M, et al (2000): *A Brief History of the Internet*. (Accesible en: <http://www.isoc.org/internet/history/brief.html>).
- Mukherjea, S. y Foley, J. (Abril 1995): "Visualizing the World-Wide Web with the Navigational View Builder", in Proceedings of the Third World Wide Web Conference (Darmstadt, Germany). (Accesible: <http://www.igd.fhg.de/www/www95/proceedings/papers/44/mukh/mukh.html>).
- Nelson, Wade H. (1996): "Writing for Internet. An E-Style Guide?". Accesible mediante solicitud del propio autor, en la dirección de correo electrónico wadenelson@frontier.net.
- Notess, Greg (2000): Search Engine Statistics: Database Overlap. (Accesible en: <http://searchengineshowdown.com/stats/overlap.shtml>).
- Paul, Nora (1995): "A Re-Visioning Production in the Electronic Products Newsroom", discurso pronunciado en Interactive Newspapers '95 el 6 de febrero de 1995. Publicado por el Poynter Institute (<http://www.reporter.org/poynter/>).
- Pinkerton, B. (Octubre 1994), *Finding What People Want: Experiences with the WebCrawler*, in Proceedings of the Second World Wide Web Conference (Chicago). (Accesible en: <http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/Searching/pinkerton/WebCrawler.html>).
- Pirolli, P; Pitkow, J. y Rao, R. (Abril 1996): *Silk from a Sow's Ear: Extracting Usable Structures from the Web*, en Proceedings of CHI'96 (Vancouver). (Accesible en: http://www.acm.org/sigchi/chi96/proceedings/papers/Pirolli_2/pp2.html).

- Project Xanadu (2000) (Accesible en: <http://www.xanadu.net>).
- Ribas, Cristina (1999): "Las tecnologías de la Información y la Comunicación como oportunidad", *Jornadas de Periodismo Digital*, mayo de 1999, Centro Universitario de Ciencias de la Información, CEU San Pablo, Valencia. (URL: <http://members.es.tripod.de/avicpd/jornadas.html>)
- ResearchIndex. NEC Research Institute (2000). (Accesible en: <http://www.researchIndex.com>).
- Ruiz de Elvira, Mariló (1997): "The New York Times, elegido mejor diario en Internet", *El País Digital*, 17 de febrero de 1997, nº 290 (accesible en <http://www.elpais.es>).
- S. Card (Marzo 1996): *Visualizing Retrieved Information: A Survey*. Reportaje especial: *Computer Graphics and Visualization in the Global Information Infrastructure, Computer Graphics and Applications*, Vol. 16, No. 2. (Accesible en: <http://www.computer.org/pubs/cg&a/report/g20063.htm>)
- Selberg, E. y Etzioni, O. (Diciembre 1995): Multi-Service Search and Comparison Using the MetaCrawler, in Proceedings of the Fourth World Wide Web Conference (Boston). (Accesible en: <http://www.w3.org/pub/Conferences/WWW4/Papers/169/>)
- Serra, A. "La comunicación digital como nuevo campo de investigación". Congreso Internacional de la Publicación Electrónica - Maig'98. (accesible en: <http://www.gpd.org/maig98/es/comunicaciocon.html>).

- Speechbot. Compac Corporate Research (2000). (Accesible en: <http://www.speechbot.com>).
- Strom, David (1994): “The Computer Trade Press, the Web, and Our Collective Future”, CMC Computer-Mediated Communication Magazine, volume 1, number 8, December 1994 (accesible en <http://www.december.com/cmc/mag/archive/title.html>).
- T. Bray. (Mayo 1996): *Measuring the Web*, en la revista *Proceedings* de la 5ª Conferencia Mundial de la World Wide Web en Paris. (Accesible en: http://www5conf.inria.fr/fich_html/papers/P9/Overview.html).
- *The Web Robots Page*. (2000). (Accesible en: <http://info.webcrawler.com/mark/projects/robots/robots.html>).
- Tilman, Hope. (1998). “Evaluating quality on the Net”. (Accesible en: <http://www.tiac.net/users/hope/findqual.html>).
- *Understanding the Internet: Transcript*. (2000). PBS. (Accesible en: <http://www.pbs.org/uti/utitranscript.html>).