

# Aproximaciones para una historia de Internet

Lic. Rubén Cañedo Andalia<sup>1</sup>

Internet constituye uno de los fenómenos que más asombra al hombre promedio hoy día. Su desarrollo en progresión geométrica, como es lógico, impresiona y desconcierta a muchos. Y no es menos cierto, que la industria de la computación y las telecomunicaciones experimentan un avance tan acelerado que pocas disciplinas pudieran igualárseles, pero también no es menos cierto, que fueron muchas las décadas en las que el hombre puso todo su empeño por obtener y perfeccionar los primeros componentes y dispositivos. Antes de los primeros logros, tuvieron que recorrer un largo camino las matemáticas y la física; después concurren la electrónica, la computación y muchas otras disciplinas.

Sin embargo, no es precisamente el desarrollo acelerado de esta gran red la causa del asombro principal, sino su poder de penetración en la vida humana, su capacidad de revolucionar desde el puesto de trabajo hasta las relaciones sociales: su capacidad de mediar en casi todos los actos de la vida moderna.

Por otra parte, como es lógico en un fenómeno relativamente nuevo y explosivo, existe una abundante, dispersa, e inmadura bibliografía sobre el tema, por demás rápidamente desactualizable, que impide obtener una visión completa de la evolución de uno de los fenómenos más espectaculares del siglo XX.

Ello es muy necesario para todos. Internet es una revolución que arrastra tras de sí a millones de individuos, ante la cual es preciso, primero: comprender correctamente sus antecedentes, origen, evolución, estado actual y tendencias; y segundo: adoptar una posición reflexiva, que posibilite comprender el lugar del individuo y el sentido que para él tiene esta nueva tecnología.

Ante estas demandas, y a la espera de la publicación de obras completas sobre el tema, surgió la idea de elaborar una especie de cronología de los principales momentos por los que ha transcurrido el desarrollo de esta gran red, convertida en el gran nuevo espacio del progreso humano.

Fundamentalmente dirigida a estudiantes de las carreras de información, pretende, a partir de la síntesis de la bibliografía disponible, enriquecida con elementos conceptuales, facilitar la comprensión de los principales avances que ha experimentado esta red desde su nacimiento hasta el momento actual.

## Antecedentes y orígenes de Internet

Aunque, la primera vez que se emplearon las terminales (en aquel tiempo teletipos) para comunicar computadoras por medio de líneas telefónicas fue en los años 40, no es hasta la década de los 60, que se amplió considerablemente el acceso directo a las computadoras e indirecto a las bases de datos mediante

líneas telefónicas.

El lanzamiento al espacio del primer satélite soviético, el Sputnik, en octubre de 1957 puede tomarse como un punto de referencia seguro para el comienzo de la historia de Internet. La puesta en órbita terrestre de aquella pequeña bola de metal galvanizó a los Estados Unidos. La precisión que requería el envío al espacio de un cohete y su transformación en un satélite artificial de la Tierra era superior a la que exigía el lanzamiento de una bomba nuclear a un objetivo. Aquella conmoción produjo diferentes resultados, que fueron desde los subsidios para desarrollar formas efectivas de entrenamiento de los científicos hasta la creación de una nueva agencia de defensa, la Advanced Research Project Agency (ARPA - siglas en inglés).

En 1961, el sabotaje a tres torres de microondas en Utah renovó, aunque a menor escala, la sensación de pánico que produjo el Sputnik unos años antes. La preocupación sobre la confiabilidad de los mecanismos militares de "comando y control", así como el temor a nuevos ataques terroristas que pudieran dejar al país sin medios de comunicación estimularon la búsqueda de infraestructuras más resistentes. El último fin de aquella búsqueda fue el diseño de un sistema de comunicación capaz de sobrevivir a una guerra nuclear.

Es precisamente en el año 1962 que surge el proyecto de Internet, producto del interés de los Estados Unidos por crear una red de militar capaz de soportar las comunicaciones de esta esfera bajo las condiciones de un ataque nuclear procedente de la entonces Unión Soviética y otros países del campo socialista.

Entre 1962 y 1964, *Paul Baran* de la Rand Corporation, elaboró una serie de informes en los que sugirió la creación de una estructura de comunicación en forma de tela de araña con el fin de impedir la selección de objetivos para cualquier forma de ataque. Para este proyecto, se empleó el protocolo de Internet. Probablemente era la única tecnología capaz de sobrevivir a un ataque nuclear.

En 1962, ARPA creó la Information Processing Techniques Office (IPTO - siglas en inglés). La fundación de dicha institución posibilitó que la agencia de defensa se vinculara con el naciente campo de la computación. Esta tuvo una participación importante en el desarrollo del trabajo a tiempo compartido, la computación interactiva y las redes. En efecto, si es posible que varias personas se conecten simultáneamente a una misma computadora mediante diferentes terminales no es difícil pensar en la existencia de una segunda computadora que actúe como si fuera una terminal temporal de la primera y viceversa.

Desde una perspectiva económica, la creación de redes generaría una mayor explotación de los recursos con un costo menor.

El tiempo compartido permitió el uso más efectivo de las máquinas en relación al costo de su capacidad computacional.

Extender una red cuanto sea posible sobre un territorio parece una idea aún mejor. Los intereses económicos se encontraban en la base de la agenda de trabajo de la IPTO.

En 1966, esta entidad financiaba 17 sitios que mostraban un desconcertante rango de hardware y software. La incompatibilidad era la regla; continuar por ese camino crearía gastos innecesarios. Estas condiciones facilitaron la identificación de la red como una posible solución al problema creciente de los costos.

Una red puede presentar diferentes formas. Las redes se conciben generalmente como un sistema estelar con un núcleo o como una telaraña. La última topología es mucho más difícil de destruir y flexible aunque implica que los mensajes deben dirigirse de forma dinámica para poder considerar las posibles variaciones de la integridad de la red. Sus objetivos principales son comunicar fácilmente resultados e información, compartir software, facilitar la comunicación entre los individuos. El objetivo original de ARPANET era ofrecer acceso remoto a computadoras distantes para compartir tanto ficheros remotos como las experiencias de investigación en materia de computación.

A comienzo de la segunda mitad de los 60, ARPA, del Departamento de Defensa realizó un número considerable de investigaciones relacionadas con la confiabilidad y las posibilidades de supervivencia de las redes de conmutación de paquetes para circuitos de telecomunicaciones redundantes con el objetivo de planificar el desarrollo de una red experimental de conmutación de paquetes durante 1967-1968. Para ello, se concedió un contrato a *Bolt, Beranek y Newman Inc.* (BBN) en Cambridge, Massachusets.

En 1967, durante la reunión de la Association for Computing Machinery (ACM) en Tennessee se esbozaron muchas de las ideas a partir de las cuales se conformó el núcleo de la primera red de ARPA. En ella, los oradores de Gran Bretaña, *R. Scantlebury* y *D.W. Davies*, expusieron sus concepciones sobre la conmutación de paquetes - un concepto nuevo en aquella época. De acuerdo con *Peter Salus*, *Davis* utilizó el término "paquete", por primera vez, en 1966.

En aquella reunión, *Davies* definió la conmutación de paquetes como "un sistema de almacenamiento y envío de mensajes cortos". De igual forma, se introdujo la idea de utilizar computadoras que actuaran como intermediarias en el enlace de las redes locales para obtener un nivel mayor de conexión, un área más amplia y una red más rápida - idea que se adoptaría en la primera etapa del desarrollo de ARPANET.

En noviembre de 1967, *Larry Roberts* formó el Network Working Group (NWG -siglas en inglés) cuyos integrantes se reunieron el mismo mes. Dicho grupo estaba compuesto por representantes de los sitios que tomarían parte en las primeras conexiones, como eran la University of California, Stanford, la University of Utah, la UCLA, así como de la Rand Corporation.

El NWG trabajó durante el invierno que siguió a noviembre de 1967. El 1 de marzo de 1968 se encontraban listas las especificaciones para la construcción de la red.

En junio de 1968, *Larry Roberts* presentó un proyecto que contenía el programa de construcción de la futura red al director de ARPA. A este proyecto, se destinaron 2,2 millones de dólares. La firma consultora BBN (*Bolt, Beranek y Newman*) localizada cerca de Boston obtuvo el contrato para la realización del proyecto.

El proyecto era preciso. La BBN debía diseñar un equipo que posibilitara a varios hosts interconectarse mediante una red de computadores intermediarios denominadas "Procesador de Mensajes de Interfaz (IMP- Interface Message Processor, en inglés).

Paralelamente al trabajo en el IMP, se publicó el famoso informe 1822 de la BBN, titulado "Procesador de Mensajes de Interfaz: especificaciones para la interconexión de un host y un IMP (siglas en inglés)", en mayo de 1969 que de acuerdo con *Peter Salus* era la Biblia de las redes. Dicho informe se sometió a varias revisiones hasta 1978. En octubre de 1980, se le añadió un apéndice que describía la norma X.25 del Comité Consultatif Internationale de Telegraphie et Telephonie (CCITT).

Durante el transcurso del año 1969, el diseño de los programas acompañó el trabajo en el hardware. El primer modelo que apareció fue el denominado cliente-servidor, un modelo familiar por la omnipresencia de terminales que empleaban las computadoras a tiempo compartido.

Esto condujo a la creación del Network Control Program (NCP), el primer programa de control de redes que después se denominó Network Control Protocol. La expresión "protocolo" se deriva de la tradición diplomática; aparentemente surgió de una metáfora que se utilizó en una conversación entre *Vint Cerf*, *Jon Postel* y *Stephen Crocker* en fecha no determinada. En el RFC (Requests for Comments) número 33 de *Crocker* y *Cerf* con fecha 12 de febrero de 1970, se empleó el término en su título ("New Host-to-Host Protocol).

Los protocolos de comunicación son conjuntos de reglas que permiten que diferentes computadoras con distintos sistemas operativos se comuniquen y compartan recursos. Los más importantes son: IP (Internet Protocol), TCP (Transmission Control Protocol), FTP (File Transfer Protocol), TelNet (Network Terminal Protocol), y SMTP (Simple Mail Transmission Protocol).

En septiembre de 1969, se instaló el primer procesador de mensajes de interfaz, denominado ARPANET, en la Universidad de California (UCLA), en Los Angeles. Contaba con cuatro nodos enlazados mediante líneas telefónicas. La primera conexión entre un host y un IMP se produjo el dos de septiembre de 1969 en la UCLA.

El 21 de noviembre de 1969, el grupo que trabajaba en la conexión de *Stanford*, mostró un enlace estable entre esta universidad y la UCLA con el auxilio de sus respectivos IMPs ante la presencia de los expertos de ARPA. A finales de noviembre, la University of California, Santa Bárbara (UCSB) obtuvo su IMP y la de Utah en diciembre de 1969. Antes de que concluyera el año 1969, se habían instalado y se encontraban funcionando los primeros cuatro nodos de ARPANET. A partir de entonces, ARPANET creció de forma geométrica.

Como es de suponer, el programa de investigación en conmutación de paquetes de ARPANET requirió de una base económica fuerte, porque las computadoras de finales de los 60 y principios de los 70, particularmente las máquinas más avanzadas resultaban costosas. En lugar de adquirir para cada una de las instituciones que realizaban investigaciones para ARPA, era posible instalar una o dos de estas máquinas a nivel nacional y compartirlas, por ejemplo en el uso de software, que corrían solo en potentes máquinas, bases de datos e información en general, mediante una red.

En 1970, *L. G. Roberts* y *B. D. Wessler* observaron que las bases de datos podían almacenarse y mantenerse en puntos remotos en lugar de duplicar los datos necesarios en cada sitio. De esta forma, no solo pueden accederse los datos como si los usuarios fueran locales sino que un usuario desde cualquier punto de la red puede, mediante un programa, recolectar datos de distintos sitios.

En abril de 1971, el ARPANET estaba compuesta por 15 nodos. En aquel momento, ARPANET era una verdadera red de conmutación de paquetes. Sin embargo, sus diseñadores percibían que los especialistas que laboraban en el sector de las comunicaciones por medio de computadoras no habían tomado conciencia de la significación de este hecho.

*Bob Kahn*, quien se había trasladado a DARPA en 1972, decidió utilizar la International Computer Communication Conference que se celebraría en Washington a finales de octubre del mismo año para mostrar las posibilidades de la nueva tecnología. La demostración fue un éxito. *Kahn* estima que este fue el comienzo de un proceso que ha permitido convencer a la comunidad de investigadores sobre las ventajas de conectar sus máquinas a la red.

La idea de la red como un medio nunca se comprendió completamente en los inicios de ARPANET. En 1972, *S. D. Crocker* y otros publicaron una de las primeras descripciones acerca de un protocolo de aplicación relacionado con el trabajo remoto y la transferencia de ficheros donde se subrayó el punto de vista del usuario en lugar de la tecnología para la conmutación de paquetes. La mayoría de sus trabajos estaban dedicados al protocolo Telnet, para el uso de sistemas interactivos remotos.

Hacia 1973, *L. G. Roberts* estimó que ARPANET ahorró a la comunidad de investigadores norteamericanos cuatro millones de dólares, al evitar la duplicación local de recursos en el tiempo en que los costos de operación de la red eran de tres y medio millones de dólares, incluida la amortización del hardware.

Se iniciaba entonces una década difícil. El período que va desde 1972 hasta, al menos, 1983 y, de hecho, un poco más allá, corresponde a la fase de mayor complejidad en la historia de Internet. A su vez, es la parte menos documentada si se consideran los relatos e historias publicadas hasta el momento. Dicho período se caracterizó por el conflicto entre un gran número de individuos, instituciones, comunidades, sectores y organizaciones.

Establecer una fecha exacta para el nacimiento de Internet puede resultar algo complejo, porque ocurrió por partes. Sin embargo, la publicación de un trabajo conjunto elaborado por *Vinton Cerf* y *Robert Kahn*

en 1974, luego de la demostración exitosa que se realizara en Washington en 1972, puede tomarse como un punto de comienzo confiable. En este, se sugirió utilizar una forma de protocolo conjunto TCP (Transmission Control Protocol) y del IP (Internet Protocol).

La introducción del TCP/IP posibilitó la interoperabilidad entre sistemas heterogéneos, la comunicación entre terminales mediante una multitud de redes diversas y el manejo automático de los fallos en la transmisión de datos.

Dicha combinación experimentó cambios sistemáticos con vistas a su perfeccionamiento en los años siguientes. El trabajo de *Cerf Y Kahn* contenía algunas de las ideas básicas de la comunicación punta a punta y sobre la administración de los paquetes mediante protocolos en lugar de nodos en circuitos. Como tal, abrió un camino para la exploración y la experimentación en ARPANET.

Internet es un conjunto de redes interconectadas que cubren la Tierra. Su crecimiento es superior al de todas las otras redes electrónicas existentes.

Por otra parte, a inicios de los años 70, los franceses y los ingleses comenzaron sus propios proyectos.

En Francia, el Cyclades y su subred el CIGALE, fueron el resultado de estudios que comenzaron en 1970. El proyecto Cyclades experimentó con rapidez recortes en su presupuesto un año después debido posiblemente a que en 1974, el sistema de correo francés comenzó a trabajar en una red propia de conmutación de paquetes que finalmente se convirtió en Transpac, el servicio francés que se basó en el protocolo X.25.

En 1977, Gran Bretaña adoptó el protocolo X.25.

La creación del X.25 proporcionó una alternativa a la utilización del TCP/IP como una solución para la interconexión de redes. Mientras que TCP/IP era un protocolo de uso probado, disponible libremente y sin propietario, los portadores públicos deseaban un protocolo que estuviera diseñado para satisfacer sus necesidades e intereses específicos que no eran necesariamente los mismos que los de los usuarios académicos y militares, quienes apoyaban los protocolos de ARPA.

El X.25 es una norma que desarrolló CCITT, (una rama de la ITU - la International Telecommunications Union). Las diferencias entre el X.25 y el TCP/IP se resumen en dos puntos:

- En el X.25, los circuitos virtuales los establecen los nodos de la red y el control se realiza mediante los mismos nodos. En el modelo de Internet es el protocolo de TCP el que establece la vía que seguirán los datagramas - unidad de los mensajes que contiene las direcciones del origen, el destino y sus datos. La red conducirá los datagramas por sí sola sin la necesidad de proveer los circuitos virtuales requeridos para su transportación.
- En el modelo de Internet, las redes son diferentes mientras que en una red X.25 cada una de ellas trabajará como si fuera una sola.

El X.25, se basa claramente en el modelo telefónico. Descansa en la suposición de que siempre existirá alguna forma de monopolio en este servicio público. En contraste, el pensamiento que sustenta el TCP/IP surge a partir de la suposición de que las redes proliferarán más o menos de forma anárquica.

"Si surge un orden de este caos, sólo puede soportarse sobre cimientos anárquicos". El crecimiento de las redes locales demostró rápidamente la validez del principio anterior. Por razones económicas, las redes locales no emplean nodos de conmutación y el protocolo X.25, dependientes como son del poder de cálculo de los nodos, no podrían extenderse directamente a las redes locales.

Estos hechos constituyeron un argumento importante a favor del TCP/IP. ARPA comenzó a adoptar gradualmente el TCP después de 1974.

Ante la creciente discordia entre la comunidad de especialistas en computación y el Postal, Telegraph and Telephone (PTT) - un nombre europeo genérico que se utiliza para referirse a las compañías telefónicas propiedad del estado. la International Standard Organization comenzó a trabajar alrededor de un proyecto que denominado OSI - Open System Interconnection.

El esfuerzo de la ISO no se dirigió a la creación de nuevos protocolos sino a la creación de un modelo descriptivo, ordenado y abstracto que facilitara la racionalización del diseño de las redes y comparar las soluciones. Dicho modelo, con sus siete capas o niveles, se ha convertido en un patrón para los estudios de redes. Sus capas son las siguientes: la física, de enlace, de red, de transporte, de sesión, de presentación y de aplicación.

El impacto del OSI sobre el desarrollo de las redes es difícil de valorar porque los especialistas en redes relacionados con el TCP/IP confiaban en la fortaleza una base bien establecida sobre implementaciones probadas.

El eventual éxito de TCP/IP estuvo vinculado con otro factor que los europeos no pudieron percibir. En 1969, se creó un nuevo sistema operativo para las computadoras en los laboratorios Bell de la AT&T, denominado Unix. Este sistema se hizo público en otoño de 1973, cuando *Ken Thompson* y *Dennis Ritchie* expusieron un trabajo en el que presentaron sus características esenciales. Su éxito fue rápido aunque hasta 1979, sólo lo empleaba la PDP-11 de la DEC (13).

A principio de 1975, Unix se había identificado como el sistema operativo de elección para las redes del estilo ARPA. Como señalaron *Peter Salus* y otros, se pensó que se colocaría al Unix en la red cuando en realidad se colocaba la red en Unix. Este hecho constituyó otro elemento clave en el éxito para los protocolos de Internet.

En 1976, la invención del UUCP o el Unix-to-Unix Copy por *Mike Lesk*, comenzó a solucionar los problemas relacionados con la transferencia de ficheros, el envío de correos y la ejecución de comandos remotos. El UUCP se utilizó para el envío de noticias entre la University of North Carolina y de *Duke*,

idea que surgió de dos estudiantes, *Tom Truscott* y *Jim Ellis*. El UUCP creció rápidamente debido al uso que recibió en la diseminación ágil de noticias. Esto favoreció la popularidad de Unix.

En la medida que Unix envolvía al TCP/IP, se beneficiaba Internet.

La evolución de Internet en este período de tiempo no puede comprenderse sin analizar cuidadosamente la forma en la que interactuaron e interfirieron movimientos esenciales como el de Usenet con proyectos mayores concebidos dentro de grandes instituciones. Ellos son un ejemplo típico de los resultados que pueden generar la complejidad caótica y la autorganización.

La presentación del Unix en 1979 para una variedad de máquinas y en particular para las VAX, que compraban las universidades con frecuencia fue un factor importante. Mientras que los europeos estaban interesados en los productos IBM, IBM no se interesó en Unix. Ni Francia ni IBM fueron capaces de percibir el perfeccionamiento que experimentaba Unix en manos del Berkeley's Computer Systems Research Group. La versión mejorada de Unix incluyó el TCP cuyas especificaciones DARPA les había entregado gratuitamente.

Su lanzamiento, en septiembre de 1983, tuvo un gran éxito, hecho el cual puede apreciarse en que más del 90% de los departamentos de ciencias de la computación adquirieron dicha versión.

En 1979, los científicos de la University of Wisconsin decidieron organizar una red diferente que se basara en el UUCP para compensar la carencia de conexión con ARPANET. De esta forma, nació CSNET.

En 1980, *Vinton Cerf* sugirió enlazar CSNET y ARPANET por medio - una vez más - del TCP/IP por cuanto el mismo se había diseñado precisamente para solucionar este tipo de problemas. Otra vez, DARPA facilitó gratuitamente las especificaciones del TCP/IP. Entre 1980 y 1983, se realizó el enlace de DARPANET con CSNET. Dicho enlace puede considerarse como la materialización en la práctica de las ideas expuestas en el trabajo de *Cerf-Kahn* en 1974.

En enero de 1983, se decidió que todos los sitios de DARPA adoptaran el TCP/IP. El éxito de esta transición sumamente delicada favoreció al TCP/IP. El año 1983, se produjo un viraje importante en la historia de Internet, sin embargo habría que superar muchos obstáculos difíciles y los desafíos que impondría su desarrollo, para que Internet disfrutara del éxito mundial del que goza actualmente.

El concepto de una Internet con múltiples redes amplias que conectan incontables redes locales se institucionalizó alrededor de 1983, cuando el protocolo que utiliza hoy Internet, el TCP/IP, sustituyó al protocolo Network Control Program (NCP). Este protocolo se diseñó no para soportar las comunicaciones entre máquinas en una red sino entre redes que se interconectarán mediante gateways (pasarelas, porterías)

En este mismo año, la Agencia de Comunicaciones para la Defensa dividió ARPANET en dos redes: MILNET para sustentar el tráfico de información de las operaciones militares y ARPANET que

continuaría soportando la comunidad académica y científica. Ambas redes formaban el backbone nacional de Internet.

A inicio de los 80, en lugar de buscar computadoras host, costosas y grandes conectadas con los procesadores de mensajes de interfaz, empezaron a instalarse gateways que enlazaran los procesadores con redes locales completas de distintos tamaños, que pudieran interactuar con computadoras ubicadas en otras redes locales en cualquier lugar del país mediante el backbone o columna vertebral de ARPANET.

Ahora bien, la capacidad de intercambiar mensajería electrónica se identificó como una de las posibilidades más beneficiosas de ARPANET, aunque es oportuno señalar que la mayoría de las redes de computadoras se establecieron sin ninguna intención de implementación de sistemas de correo electrónico.

El problema en ARPANET era que, en la medida, en que crecía la complejidad de la red, se hacía cada vez más necesario actualizar las tablas con las direcciones de los host para conocer donde se encontraba cada uno de ellos.

Ante esta situación, en 1983, *Jon Postel* hizo público el RFC 821 en el que se especificaba el famoso Simple Mail Transfer Protocol (SMTP). Después de 1986, estas especificaciones se actualizaron particularmente con el MIME, Multipurpose Internet Mail Extensions, el cual permite enviar mediante el SMTP mensajes que en lenguas que emplean acentos como el español o el francés, imágenes, sonidos u otro tipo de información que pueda codificarse en ASCII puro.

En 1983, *Paul Mockapetris* publicó los RFC 882 y 883 que sirvieron de base al Domain Name System. Tras una serie de mejoras y extensiones que *Craig Partridge* resumió en su RFC 974 de - Mail Routing and the Domain System -, surgió un sistema que facilitó significativamente la localización y el envío de correos dentro de la red. Como afirma *Peter Salus*, estos avances comenzaron a transformar a Internet en un servicio como lo son la electricidad y el teléfono.

Hacia 1985, la National Science Foundation (NSF) financió cinco enormes centros para que sirvieran como recursos nacionales de supercomputación.

Estos sitios se enlazaron con ARPANet, pero, a mediados de 1988, la NSF había instalado una nueva red amplia llamada NSFNet para servir al sector académico que utiliza el protocolo de comunicaciones de Internet - el TCP/IP- y que como resultado de su alta velocidad de transmisión (1.54Mbits por segundo contra los 56 Kbits por segundo de ARPANET) desplazó a la vieja espina dorsal de Internet.

En el ámbito de los recursos de información, hacía la década de los 70, las redes privadas de comunicación de los Estados Unidos ofertaban un gran número de ellos con carácter comercial. Sin embargo, su falta de integración y altos costos limitaron su uso por parte de la comunidad académica.

Aun cuando los antecedentes del acceso a los recursos de información mediante las telecomunicaciones pueden ubicarse en los trabajos que desarrollaron *P. Otlet* y *W. Schuermeyer* en los años 30, en la práctica las tecnologías empleadas para ello constituían "la antítesis del acceso en red a los recursos de información", al conformar estaciones únicas de trabajo, aisladas e independientes.

Durante las décadas de los 70 y los 80, hubo un crecimiento asombroso de los recursos de información que soportaban todos los campos de la investigación y la docencia como resultado del perfeccionamiento de los sistemas de gestión de bases de datos, de la adquisición de la experiencia necesaria para estructurar, crear y manejar los recursos de información, así como del aumento del valor con que se percibían dichos recursos. En la década de los 80, con la aparición de nuevos soportes de almacenamiento, ocurrió un cambio de concepción: las bases de datos se colocaron en manos del usuario, por aquel entonces, instituciones medianas y pequeñas, aún. La información cambió también de estructurada a relacionada.

Sin embargo, el proceso de utilización de un sinnúmero de bases de datos de forma integral se hizo cada vez más difícil debido a la necesidad de dominar múltiples interfaz de búsqueda y de los problemas para combinar los resultados almacenados en diferentes sistemas de tiempo compartido (en diferentes redes) y en estaciones de trabajo aisladas, con sistemas de recuperación basados en CD-ROM, en un solo resultado, en la medida en que el número de recursos se multiplicaba.

Al parecer, cada cierto tiempo, se produce una crisis cíclica, debido al crecimiento geométrico de los recursos de información: libros, revistas, índices, bases de datos, etcétera, que obliga a la creación de nuevos tipos de recursos, cada vez más abarcadores.

Tal estado de cosas estimuló el surgimiento de una nueva visión en la que se concebía el acceso a la información para soportar la investigación y la docencia en forma universal, simple e integral.

Esta visión, reconocida generalmente como propia de los miembros de la comunidad de la educación superior, no la compartieron los especialistas en redes de computación, salvo raras excepciones, hasta recientemente. Históricamente ha existido una laguna considerable entre aquellos que planificaban, financiaban y construían redes de computación y aquellos que se preocupaban por el acceso a los recursos de información.

Ciertamente, los miembros de la comunidad de redes desde los primeros días destacaron los beneficios de la conectividad a numerosos recursos de información, pero trataron poco los problemas relacionados con su uso efectivo.

Muchos bibliotecarios concibieron el uso de las redes y el compartir recursos más en las operaciones bibliotecarias (catalogación, préstamo interbibliotecario, etc.) que sobre aquellas que posibilitaran un acceso público a los recursos de información a escala nacional. Los esfuerzos para crear lo que son, en efecto, las bases de la información bibliográfica nacional distribuidas masivamente excluyeron en gran medida el acceso directo de los usuarios de la biblioteca hasta recientemente.

Los trabajos sobre los protocolos de recuperación de información, por su parte, comenzaron a mediados de los años 70.

Estos tienen sus raíces en los esfuerzos de la Biblioteca del Congreso de los Estados Unidos, con el apoyo del Council on Library Resources. A comienzos de los 80, se iniciaron los esfuerzos para crear una base de datos distribuida en la nación que soportara el procesamiento bibliotecario y compartir recursos, los que con la ayuda de otras organizaciones. Así, se desarrolló el protocolo de aplicación Z39.50, con el objetivo de facilitar la búsqueda en bases de datos bibliográficas ubicadas en diferentes sitios, manipular el proceso de la búsqueda y transferir los registros de un lugar a otro. Dicho sistema se normalizó definitivamente en 1988.

Este protocolo se consideró como una tecnología básica para el acceso a los recursos de información de la nación norteamericana. Permitía la realización de búsquedas con operadores booleanos. Aun cuando su uso ha resultado común en el campo de las bibliotecas, como su desarrollo ocurrió al margen del crecimiento de Internet, este no se construyó sobre la base del desarrollo de la red nacional y se perdió la oportunidad de tener acceso a uno de los principales recursos surgidos en los primeros momentos.

En 1989, miembros de la Corporation for National Research Initiatives concibieron un sistema bibliotecario totalmente distribuido en el cual existieran programas activos para la búsqueda de información que interactuaran con los servers para recolectar, entregar y mantener recursos de información en interés de los usuarios o de los servers.

Entre finales de los años 80 y principios de los 90, ocurrió una explosión de nuevas aplicaciones que transformaron completamente la visión de Internet. En 1989, tuvo lugar un suceso imperceptible en el CERN(Conseil Européen pour la Recherche Nucléaire) (Organización Europea para la Investigación Nuclear) , un centro europeo de investigación en la rama de la física situado cerca de la frontera franco - suiza. *Tim Berners-Lee* y *Robert Cailiaux* comenzaron a concebir un sistema distribuido de documentos que se estructuraría en forma de hipertexto. En aquel entonces, la idea tuvo solo un alcance local.

Más espectacular fue la invención de Archie en la Universidad McGill de Montreal. *Peter Deutsch* y *Michael Emtage* descubrieron una forma de crear una base de datos que registrara los sitios anónimos FTP ubicados en la red y una forma de actualizarla mensualmente. Archie no sólo permitió el hallazgo de documentos que sin la existencia de una autoridad central - algo claramente imposible dentro de Internet - sino que mostró que Internet se había convertido en una gran biblioteca distribuida. La experiencia de Archie no se perdió. Poco tiempo después, apareció Gopher una nueva aplicación que ofrecía una extensión del protocolo FTP.

Los finales de la década de los 80 marcaron el punto de giro. Las diferencias filosóficas existentes en la concepción y en los objetivos de desarrollo de las redes entre los especialistas en computación y la comunidad académica, científica y bibliotecaria preocupada más con el uso de los recursos de información comenzaron a eliminarse.

En la medida en que se desarrollaron redes nacionales de computadoras, cada vez más amplias, robustas y capaces de servir a comunidades mayores, sus diseñadores se percataron de que si las redes eran, en realidad, un sistema de carreteras, en algún sentido, ellas debían existir mayormente para permitir que sus usuarios alcanzaran destinos de interés y valor propio. Se produce, en este momento, un cambio de concepción de la red como un fin a la red como un medio.

Ya no resultaba suficiente pensar simplemente en la red como un medio para mover información digital de un lugar a otro como un fin en sí mismo. Los defensores de las ideas sobre el acceso universal a recursos de información, que habían soñado largamente con lo que se podría lograr si estuviera disponible una conectividad de alta velocidad, costeable y de amplio alcance - largamente frustrada por la aparente imposibilidad de obtener y poder pagar por tal conectividad-, comenzaron a percatarse que se estaban logrando dichas redes de comunicación y, más aun, que "el apoyo a" y "la explotación de" podrían acelerar su crecimiento.

A finales de los años 80, cierto número de catálogos en línea extendieron su cubrimiento más allá de las colecciones monográficas mediante el montaje de bases de datos de artículos de revistas.

Hacia 1989, existían 15 ó 20 catálogos en línea en Internet, el primero y mejor conocidos de todos fue el Melvyl de la Universidad de California, que estuvo ampliamente disponible a la comunidad de Internet.

Entre 1995 y 1996, la OCLC (Online Computer Library Center), realizó un nuevo aporte en esta línea. Su objetivo fue definir un formato amigable para la catalogación distribuida de los recursos en la red, más rico que un índice creado automáticamente y menos complejo que el tradicional registro bibliográfico, y así estimular la cooperación para promover la interoperatividad semántica a través de lenguajes y disciplinas, con el fin de lograr un efectivo control del vocabulario de indización. La pregunta sería si no era demasiado tarde para ello.

Durante los primeros años de la década de 1980, algunas de las universidades principales habían desarrollado catálogos de acceso público en línea para ofrecer a sus usuarios sistemas de recuperación automatizados para la búsqueda en sus fondos.

Entre finales de los 80 y comienzos de los 90, el número de recursos en Internet creció, hecho el cual generó diversos problemas. Se dificultó a los usuarios la identificación y la localización de recursos potencialmente relevantes para satisfacer las necesidades de información en áreas específicas, particularmente si el usuario no era un especialista. Cada recurso poseía un sistema de recuperación propio e independiente; mientras que la disponibilidad de recursos sobre Internet permitía al usuario alcanzar los recursos de manera fácil y económica (al menos en términos de costos de telecomunicaciones) esto no eliminaba la necesidad de aprender a manipular una interfaz nueva y única para poder utilizar cada recurso. No existía, a su vez, una forma simple de mover los resultados de un sistema a otro para su consolidación, análisis y almacenamiento porque el acceso a cada sistema se realizaba mediante acceso remoto.

Con vistas a solucionar al menos parcialmente el primer problema se desarrolló un directorio general de recursos en Internet publicado, bajo los auspicios de la NSF, por la BBN Network Information Center en 1989. Este directorio se convirtió en una base de datos de recursos que incluía indizaciones profundas y complejas y facilidades para la búsqueda.

Existió, a su vez, una gran presión por parte de las facultades y de los estudiantes para que las instituciones pusieran en servicio bases de datos comerciales en la red - como sucedió con Medline, la principal base de datos sobre medicina que produce la Biblioteca Nacional de Medicina de los Estados Unidos - de forma tal que permitiera que estos recursos vitales estuvieran disponibles rápida y masivamente. Sin embargo, dichas instituciones enfrentaron costos significativos para la adquisición de los recursos de computación, el desarrollo, el soporte necesario para crear y operar los recursos de información para el servicio a los usuarios, así como el costo de las licencias para operar las bases de datos de distintos propietarios. Uno de los beneficios técnicos de montar localmente dichos recursos incluía la facilidad de integrar estos con los ficheros de los catálogos institucionales.

Los intentos para resolver el segundo y tercer problemas se dividen en dos categorías principales: los que tratan de orientar el servicio con los recursos existentes a las terminales remotas mediante el desarrollo de softwares para los usuarios finales y los que emplean protocolos de red para convertir los recursos de información en servers de información basados en red y ofrecer al usuario acceso a estos nuevos servers de información mediante software clientes e interfaz de protocolo.

Entre los primeros existe un gran número de sistemas desarrollados para facilitar el uso de los recursos existentes. Sistemas como Gratefull Med y Procite ofrecen interfaz más amistosas - más poderosas y sencillas- con el usuario o que facilitan el mantenimiento de bases de datos personales a partir de los registros que se extraen de sistemas remotos.

Los sistemas mas sofisticados intentaron, mediante la traducción de las solicitudes al lenguaje particular de los distintos sistemas, que el usuario final utilizara solo una interfaz, común para un grupo de recursos diferentes. Esta vertiente enfrentó la limitación de que los sistemas remotos cambian a menudo sus diálogos con las terminales, los software de este tipo deberían entonces alterarse constantemente para reconocer tales cambios. Estos sistemas no se prestan para la formulación interactiva de búsquedas donde el usuario este inseguro de la terminología de la indización y el tamaño de los resultados que pueden esperarse. Estos tienen un objetivo limitado, permitir al usuario final buscar en un recurso de información remoto en lugar de proveer una interfaz normalizada a los recursos de información.

La mayoría de los trabajos realizados con bases de datos no pueden aplicarse directamente a la solución del problema del acceso en red a los recursos de información. Un sistema cliente tendría que conocer mucho sobre la estructura de la base de información o colección de información que, en cierto sentido se organiza, articula y accede de acuerdo con la semántica de la información, montada en cada servidor. La base de información de un catálogo en línea es un conjunto complejo de bases de datos (de autores, de materias, de registros bibliográficos, de información sobre los fondos, etc.) en el que cuando se realiza una búsqueda sobre los libros de un autor determinado, es probable que ocurra una serie de búsquedas

interrelacionadas y complejas en diferentes bases de datos.

Aun cuando los años 80 fueron testigos del surgimiento de nuevas redes, estos se caracterizan por el crecimiento estable del TCP/IP y el surgimiento gradual de aplicaciones, es en los años 90 donde se produjeron adelantos que transformaron totalmente la historia de Internet.

En los Estados Unidos, el crecimiento de la red nacional de computadoras se produjo, a partir primero de las contribuciones de la Agencia para los Proyectos Avanzados de la Defensa (Defense Advanced Research Projects Agency) y más tarde de las de la Fundación Nacional para la Ciencia (National Science Foundation), la Agencia Nacional para el Espacio Aéreo (National Aerospace Agency) y el Departamento de Energía.

## **Internet como un fenómeno mundial**

En mayo de 1990, se celebró la primera conferencia sobre ciberespacio en la Universidad de Texas en Austin.

A principios de los 90, el tema del acceso a los recursos de información evolucionó de una idea a un aspecto central en el desarrollo de las redes nacionales de computadoras en la comunidad de académica y científica de los Estados Unidos. Los recursos de información crecen rápidamente, así como la base tecnológica para su integración.

La primera preocupación fue, como es lógico, crear una infraestructura tecnológica adecuada que permitiera soportar los recursos, así como la transferencia de información, después, sobre todo, con el crecimiento acelerado de los recursos, el centro de la atención se trasladó al acceso.

En 1990 surgieron algunas corporaciones comerciales con proyectos para participar en Internet. Ello, se aceleró hacia 1993. En estos años, existió también una gran demanda por la sustitución de los intermediarios tradicionales de información por software mediadores y bases de datos inteligentes para el usuario final.

El desarrollo de los sistemas tecnológicos, la facilidad de operación de los equipos y programas modernos, así como la disminución de sus costos ha hecho pensar a muchos que las funciones de los intermediarios desaparecerán.

Sin embargo, la realidad muestra otro panorama bien distinto. Por ejemplo, una búsqueda en Internet, el símbolo del progreso en materia de información, exige actualmente del auxilio de un intermediario capaz de identificar, evaluar, seleccionar y adquirir fuentes de información útiles, válidas y significativas, tarea la cual resulta, en grado sumo difícil, si se considera la multiplicidad de recursos cualitativamente diferentes disponibles, la carencia de precisión de los instrumentos de búsqueda, el volumen de tiempo necesario para realizar esta si se pretenden obtener resultados exhaustivos y precisos, entre otras limitaciones.

Y es que en la década de los 60 y los 70, se produjeron un gran número de cambios. El trabajo del profesional de la información comenzó a verse como parte integrante de la gran cadena de la comunicación social, donde interactúa de forma creciente con otros eslabones de la cadena. Esta concepción presenta actualmente una vigencia renovada.

En 1991, se aprobó la Ley Pública 102-194 que establecía el US High Performance Computing and Communication Program (HPCC).

En este año, Internet poseía un enorme potencial para el suministro de recursos de información a los profesionales del mundo; sin embargo, en aquel entonces, su acceso era posible solo mediante los protocolos de transferencia de ficheros, las listas de correo electrónico o de distribución, útiles para la discusión y la transferencia de ficheros entre direcciones conocidas, pero poco apropiadas para la búsqueda de nuevos recursos en la red para lo que resultaba necesario transferir los ficheros hasta la máquina del usuario y leer sus introducciones o guías locales de recursos disponibles.

No existía, por lo tanto, un sistema para el manejo de los recursos de información que facilitara su uso a escala internacional.

En 1991, *Linder* y *Cahil* de la University of Minnesota desarrollaron un sistema, Gopher, que permitía crear una estructura en forma de árbol que se manipulaba desde un menú. En él, los documentos, que estaban distribuidos en Internet, parecían provenir de un mismo sitio. A partir de una serie de menús jerárquicos, el usuario selecciona las entradas que desea. Un menú de Gopher es una lista de enlaces.

El cliente Gopher se presenta al usuario con una estructura de ficheros similar a la del DOS que resulta familiar para muchos. Dicha estructura contiene documentos, directorios (de distintas máquinas) y entradas para localizar información. Los documentos pueden ser ficheros texto, imágenes, servicios como Telnet. Gopher englobaba a otros protocolos como Telnet o FTP, así liberaba al usuario del aprendizaje de las aplicaciones propias de estos protocolos. Su principal inconveniente radicaba en su extrema rigidez, determinada por la estructura jerárquica en la que se basa.

A pesar de ello, Gopher, fue el programa más popular de dominio público para el almacenamiento y recuperación de la información accesible en Internet, que producían y mantenían los distintos departamentos de la universidad.

Con este sistema, cada departamento podía crear y mantener servers Gophers en sus computadoras. Toda la información se presenta en forma de directorios. El archivo Gopher se presenta como una lista de directorios que contienen ficheros u otros directorios. Los directorios, a menudo, señalan otros sitios Gopher. Uno puede moverse a través de estos directorios y de Internet seleccionando directorios y redireccionando las selecciones.

Cualquier directorio visto puede marcarse, el programa almacena su ruta de acceso y posibilita el acceso directo posterior mediante el uso automático del FTP, los ficheros pueden verse dentro de esta aplicación y salvarse. Gopher ofrece también índices. Estos pueden ser directorios y listas de correo, o

más a menudo, índices de documentos de texto completo que soportan búsquedas con operadores lógicos y el truncado.

De esta forma, la mayoría de los directorios y ficheros en el sistema mundial Gopher pueden buscarse mediante un índice único, Verónica, el equivalente Goher a Archie. El software del server Gopher ha mostrado su utilidad tanto para pequeños laboratorios que diseminan información sobre temas específicos como para aquellos que tienen grandes recursos de información y desean integrar la información de una universidad o de la nación. Los grupos pequeños tienden a crear estructuras y directorios que se relaciona estrechamente con sus propias operaciones mientras que los de grandes recursos toman la formidable tarea de dar acceso a múltiples recursos de información y mantener la actualidad y la relevancia de estos recursos.

Los numerosos puntos de acceso y facilidades Gopher Archie son una puerta en forma de directorios, a partir de la cual se obtienen distintos programas para el descubrimiento de recursos.

Pero los Gopher se multiplicaron y plantearon serios inconvenientes para el acceso.

Así surgió WAIS (Servidores de Información de Área Amplia) un sistema de recuperación de información distribuida que ayuda a los usuarios a buscar en distintas bases de datos de la red mediante una interfaz sencilla. Conecta documentos o bases de datos. Las bases de datos, denominadas fuentes, son fundamentalmente colecciones de documentos textuales pero pueden incluir también imágenes y sonidos. Dichas bases de datos se ajustan comúnmente a la norma Z39.50, que describe la forma en la que una computadora solicita una búsqueda a otra ubicada en la red.

Los documentos pueden abrirse y salvarse directamente y las solicitudes almacenadas pueden correrse periódicamente como perfiles. Utiliza el lenguaje natural para localizar documentos relevantes sobre el tema en el que se realiza la búsqueda y genera como resultado una lista de referencias procedentes de todas las bases de datos, ponderadas según el número de veces que aparece la palabra y su posición en el texto. El resultado de la búsqueda es una relación de referencias cuya relevancia se pondera entre uno y 1000. Los documentos de mayor relevancia recibirán una puntuación de 1000. Realiza búsquedas a texto completo.

En WAIS, en lugar de navegar a través de una serie de menús, una vez que el usuario se conecta a un servidor de este tipo, recibe una lista de términos que representan las temáticas sobre las cuales puede realizar su búsqueda.

WAIS ofrece una interfaz única para la búsqueda y visualización de los resultados de una variedad de bases de datos comerciales y gratuitas geográficamente distribuidas. El cliente WAIS permite al usuario construir prescripciones de búsqueda en un lenguaje natural y correrlas en un conjunto de recursos seleccionados simultáneamente. Su concepción, por tanto, es diferente a la de Gopher.

Sin embargo, las bases de datos fuentes de WAIS crecieron y se hicieron muy numerosas.

El Gopher, un sistema textual, y WAIS fueron desplazados entonces por el WWW. El WWW fue el invento más revolucionario después de la computadora personal. El WWW es un sistema hipermedia que integra en una interfase común a todos los recursos existentes en la red para su acceso en forma organizada y normalizada, cubre todos los recursos del mundo en forma hipermedia o hipertexto.

Un hipertexto es un sistema que permite la realización de conexiones lógicas entre las distintas partes de un documento o entre documentos diferentes. De hecho, funciona de forma similar a como lo hacen las referencias cruzadas: "ver..." y "ver también..." en un tesoro, con la ventaja de que en el web los documentos pueden accederse mediante un simple "click" en el mouse. En el web, se utiliza el hipertexto para navegar entre documentos relacionados mediante enlaces que establecen sus autores y editores. Los browsers sólo pueden identificar los enlaces y moverse entre ellos. A pesar de las facilidades que ofrece, la navegación hipertextual presenta el riesgo de "perder" al usuario en una sucesión de conexiones que pueden apartarlo significativamente de su fin inicial.

Hipertexto es la denominación general que se aplica a aquellos documentos o conjuntos de documentos electrónicos en los que es posible navegar entre las diferentes partes del texto o entre documentos a partir de las relaciones semánticas que entre ellos se establecen. Navegar, en este contexto, es una noción que indica que la consulta de uno o varios documentos puede realizarse "a saltos", es decir, sin seguir el texto en forma lineal, secuencial o jerárquica como es tradicional en la lectura de los textos impresos; sino en forma lógica mediante los enlaces que existen entre sus distintas partes o entre diferentes documentos. Por ejemplo, usted puede encontrarse leyendo un trabajo en el que se menciona un concepto determinado y que usted desconoce, dicho concepto aparece destacado de alguna forma - con un tipo diferente de letra, subrayado, etc.-, pues bien, con tan solo oprimir el ratón sobre ese concepto marcado, usted puede obtener su definición dentro del mismo documento o recuperar otros recursos que tratan sobre el tema.

Hipermedia, por su parte, es una extensión del hipertexto, una generalización de este concepto. Mientras que en el hipertexto es posible solo moverse de un texto a otro, en hipermedia, el movimiento puede realizarse entre piezas de información o recursos de cualquier clase, sea texto, imagen, vídeo, sonido, base de datos o programas.

Un hipermedio es un documento que se basa en los mismos principios que un hipertexto. Su diferencia consiste en que el hipermedio puede contener, además de textos, imágenes - animadas o no- y sonidos. Por ejemplo, en un documento en el que se escriba sobre China, usted, con tan solo marcar este concepto, puede obtener un mapa sobre este país y tal vez hasta escuchar su himno nacional o canciones más populares, si así su creador lo preparó.

El declive de Gopher, en especial, contribuyó a llamar la atención de algunos usuarios sobre la interesante innovación del CERN que durante años había permanecido en el anonimato. Para aquel entonces, había tomado su nombre actual, WWW (W3). El WWW, desarrollado por el CERN, el principal laboratorio de investigación en física de las partículas, presentaba un diseño original. Es un sistema de información basado en hipertexto que permite navegar a través de los documentos situados en

Internet. Los documentos preparados en forma de hipertexto se entrelazan mediante un grupo seleccionado de palabras o frases. Puede manejar distintos formatos de texto y otros objetos, imágenes, bases de datos, etcétera.

Un hipertexto para ambiente de red, permitía enlazar automáticamente documentos entre si, ubicados en diferentes computadoras. El hipertexto cambia significativamente el paradigma tradicional en el que se utiliza un documento cada vez y cada documento es completamente independiente. El resultado es una red que interconecta recursos de información a nivel mundial.

Web opera sobre una arquitectura cliente- servidor. El software server almacena y el software cliente busca y recupera. El cliente también ofrece una interfaz para el usuario final. Cada uno de estos software pueden perfeccionarse de forma independiente.

Un servidor web es un programa receptor de solicitudes de documentos hipertextos residentes en su computadora y el encargado de enviarlas a la máquina que los solicita; el cliente web, es el encargado de enviar las solicitudes hechas por el usuario a la computadora que opera como servidor y visualizar las recibidas en pantalla.

El WWW toma un enfoque algo diferente a Gopher y WAIS. El usuario comienza su navegación dentro de un documento (que puede ser multimedia). Las frases subrayadas o destacadas dentro del documento indican enlaces con otros documentos o recursos, con información relacionada, que pueden estar almacenados en cualquier punto de Internet y accederse con una simple pulsación del ratón.

El HTTP o protocolo de transferencia de hipertexto permite llamar a cualquier tipo de página web, se utiliza para transferir ficheros escritos en un idioma normalizado.

El HTML es un lenguaje desarrollado para el diseño de páginas y documentos cuya consulta se realizará en el web. Se emplea para describir y reconocer los documentos hipermedia en el web, se basa en un sistema de etiquetado. Este lenguaje es una aplicación de SGML (Standard Generalized Markup Language) o norma ISO 8879, un lenguaje mucho más amplio para la definición de lenguajes de formateo de documentos, establecida en 1986.

Utiliza una serie de marcas predefinidas que indican, tanto el comienzo y el final de cada una de las partes que distinguen un documento (título, subtítulo, notas al pie de páginas y otros) como las características tipográficas (tipo de letra, tamaño, etc.).

En este sentido, un documento en HTML puede enviarse desde una máquina a otra en forma de una cadena simple caracteres ASCII, reconstruirse y aparecer ante su solicitante con toda su riqueza original.

Los documentos HTML utilizan etiquetas para indicar el formato y la estructura de la información del documento que se presenta al usuario. Existen dos tipos de etiquetas básicas, asociadas a cada una de las instrucciones: las de comienzo y las de fin.

El lenguaje HTML consta de dos tipos de instrucciones: para presentar y estructurar el texto que se mostrará al usuario y para establecer los enlaces hipertextuales a otros documentos en Internet, utilizando los URL.

La propiedad más importante del HTML es su posibilidad de enlazar partes de un texto con otros documentos o con otras secciones del propio documento. El programa visualizador resalta de alguna manera (usualmente con color diferente o subrayado) los enlaces para indicar su condición de enlaces hipertextuales. Estas secciones o documentos pueden ser textos, imágenes animadas o no, en blanco y negro o a color. Posibilita, en general, preparar salidas, formatear, estructurar y relacionar información y recursos.

Los software clientes comprenden HTML, y en correspondencia, presentan la información siguiendo el formato descrito en las instrucciones HTML incluidas en el texto y recuperan y muestran la información demandada por el usuario cuando actúa sobre los enlaces hipertextuales a otros ficheros del hiperespacio.

Sin embargo, la creación de un sitio W3 requería de una gran cantidad de trabajo porque los documentos debían etiquetarse de acuerdo con las reglas de Hypertext Markup Language (HTML). Debido a esta razón, W3 permaneció ocioso hasta que un equipo del National Center for Supercomputing Applications (NCSA) desarrolló un browser, el Mosaic, capaz de reconocer el formato de los textos y las imágenes. A partir de aquel momento, comenzó el crecimiento explosivo del uso del WWW para la implementación de los sitios.

Aunque el WWW estuvo disponible desde 1991, el amplio uso que en la actualidad recibe el WWW es el producto de la aparición de Mosaic, el primer programa browser, en la primavera de 1993. A este le siguió el Netscape. El Mosaic y el Netscape se desarrollaron originalmente para trabajar en servidores UNIX; un sistema operativo dominante en el mundo de Internet. Sus versiones para clientes aparecieron rápidamente para el DOS/WINDOWS (primero para 16 y después para 32 bits) y la Apple Macintosh. Desde el lanzamiento del sistema operativo Windows 95 por la Microsoft, la estrategia de esta compañía, dominante en la industria mundial de microcomputadoras, fue la de promover su propio browser, el Internet Explorer.

Los clientes web se diseñan generalmente para plataformas software - hardware específicas. El término browser se usa con frecuencia como sinónimo de cliente web. Por ejemplo Mosaic para ambiente Windows, Macintosh y Unix.

Los browsers, desarrollados para las plataformas más populares (las de PC, de Macintosh, etc.), actúan como intérpretes en el web. Permiten la obtención y lectura de los documentos ubicados en los servidores hipermedia de Internet. Los documentos hipermedia, que visualizan los programas browsers, contienen enlaces que actúan como pequeñas puertas que ofrecen al usuario la posibilidad de consultar otros documentos aunque éstos se encuentren almacenados en otras máquinas, ciudades o países. Los browsers, responsables, en gran medida, por el desarrollo de Internet, son las principales herramientas

para el acceso a esta red.

Web es un concepto, Mosaic es un browser que opera sobre web. La habilidad de un browser para seguir los enlaces, visualizar gráficos multicolores, generar sonidos, procesar textos, con una interfaz coordinada requiere gastos significativos en software. Los servidores web se identifican por medio de una pantalla introductoria denominada Home Page.

El éxito del Mosaic alimentó la ambición y el deseo de independizarse del NCSA entre sus creadores. El nuevo grupo creó una versión perfeccionada - el Netscape - que disfruta de gran aceptación en la comunidad de Internet. De repente, el W3 se convirtió en el instrumento de publicación de elección. El escenario actual de Internet es el producto de estos avances recientes.

En la primavera de 1993, *M. Gray*, entonces un estudiante del Massachusetts Institute of Technology (MIT) escribió un programa llamado Wanderer para recorrer el web y contar sitios.

En junio de ese año, Wanderer había identificado 100 servidores WWW que contenían más de 200 000 documentos hipertextos. La motivación inicial de *M. Gray* fue descubrir nuevos sitios, en los tiempos cuando web era un pequeño espacio. Uno de sus resultados fue Comprehensive List of Web Sites (Lista Exhaustiva de Sitios Web).

Wanderer recorrió sistemáticamente el web desde junio de 1993 a junio de 1995 y fue el primer agente, araña o autómatas, y uno de los predecesores de los software robots que más tarde se diseñaron para construir los índices actuales al WWW como Lycos, Altavista, Hotbot, etc.

Le siguieron las denominadas guías. Las Guías Internet son un producto de Internet Resource Discovery Project (IRD) patrocinado por la Universidad de Michigan. En noviembre de 1994, existían más de 120 guías temáticas disponibles.

Estas guías pueden verse como catálogos de tópicos específicos. Cada documento contiene una lista de archivos, host Gopher, listas de correo electrónico, grupos de interés especial, grupos de discusión, grupos de noticias, servers Web y WAIS, murales, etc. todos los cuales son relevantes a áreas temáticas particulares o a disciplinas académicas. Por el contrario, las listas de Internet se organizan alfabéticamente o por el modo de acceso y no generalmente por materias.

Las guías sirven para la referencia. Ofrecen comúnmente dirección postal, número de teléfono y de fax de las personas a contactar, así como comentarios explicativos asociados con cada recurso. Estos recursos pueden ser de valor porque nadie tiene acceso a cada espacio de Internet. Es posible consultar el índice alfabético de las guías, realizar búsquedas a texto completo, seleccionar documentos por categorías en múltiples temáticas y utilizar una serie de documentos auxiliares tales como guías de guías, metaguías y otros. Un individuo puede también descargar estas guías en su máquina y realizar búsquedas por palabras claves en ellas.

Sin embargo, estas guías o facilidades- como las del FTP- no solucionaron completamente los problemas de accesibilidad.

Uno de los primeros intentos por facilitar un acceso más universal a los recursos disponibles en Internet, fue Archie. Archie, un sistema diseñado por la Universidad de Mc Gill en Toronto, Montreal, permitió crear y mantener un índice de cada uno de los directorios y ficheros - de datos, de programas u otros disponibles en los servidores de FTP anónimos accesibles en todo el mundo. Archie mantenía una base de datos con las referencias de los ficheros y las rutas para acceder a ellos, así como posibilita que sus usuarios realicen búsquedas en dicha base de datos. Archie permitía realizar búsquedas de las cadenas de texto que aparecen en sus índices y proveer al usuario con la dirección del recurso donde esta se hallaba y ruta del directorio de cada recurso recuperado.

Aparecieron, a su vez, otros motores de búsqueda como Verónica y Jughead. Los motores de búsqueda experimentaron un rápido crecimiento en número y potencia. Actualmente se dispone de motores de metabúsqueda, es decir, de una plataforma de software que opera por encima de los motores de búsqueda existentes.

Existen dos tipos básicos de motores de búsqueda. La diferencia entre estos directorios radica en su forma de indizar las páginas. Los primeros organizan la información en forma jerárquica y los segundos, a la manera de conceptos. Uno de los primeros directorios jerárquicos en crearse y de los más importantes en desarrollo es Yahoo! (<http://www.yahoo.com>). En este, se almacena la información en categorías que forman una jerarquía. Entre los directorios que ordenan la información en conceptos, uno de los más importantes y con un crecimiento espectacular es el AltaVista de Digital (<http://www.altavista.digital.com>) donde se indizan las páginas mediante palabras claves o frases.

Para realizar una búsqueda por medio de un motor jerárquico, es necesario escoger las categorías temáticas generales que se correspondan, en un primer nivel, con las materias objeto de exploración en la página principal (HomePage). A continuación de esta primera selección, aparecerá una lista de las subcategorías relacionadas con las temáticas escogidas anteriormente y así sucesivamente hasta obtener el nivel de especificidad deseado.

Para ejecutar una búsqueda con un motor del segundo tipo, se tecleará la palabra clave o frase que representa la temática en la que se realizará la exploración en el espacio que para esto se destina en la página principal. El sistema recuperará mediante su propio índice las páginas que cumplen con la prescripción de búsqueda formulada. Una vez concluida la exploración presentará sus resultados en una página que contiene la identificación y la dirección de los documentos relevantes. Las búsquedas en motores en este tipo de motor aceptan el empleo de operadores lógicos, asteriscos y otros recursos similares con el fin de precisar cuanto sea necesario sus resultados.

La creación del WWW, abrió un gran potencial para la distribución de los documentos electrónicos. Se desarrolló el protocolo HTTP, el TCP/IP y el HTML. Se creó el sistema universal de identificación de los recursos- el URI. Aparecieron los robots o agentes autómatas que indizan los recursos, mediante la

extracción de palabras claves. Ellos permiten la recuperación de recursos, a partir de estrategias prelaboradas de búsqueda . Se estableció un modelo normalizado para la representación y recuperación de información en Internet, así como para crear Intranets compatibles.

El URI (Universal Resources Identificator) tiene funciones semejantes a las del ISBN y el ISSN. El creador de un documento o de un servicio que desee que este aparezca en la base de datos de un localizador de recursos, deberá solicitar que se le asigne un identificador que distinga al documento universalmente y un URL (Universal Resources Locator) que posea los datos de su ubicación. Una vez relacionados ambos en un base de datos mediante enlaces lógicos, es posible buscar, como si tratara de un Archie por el URI y obtener como respuesta su URL.

El URL es un localizador o localización. Esencialmente es una dirección en el ciberespacio. Especifica la ruta de acceso al recurso, el sitio y su ubicación en este. La identificación de los recursos electrónicos requiere de la existencia de normas para la identificación unívoca de los recursos, es como una especie de ISBN. Esta tarea de normalización la desarrolla IEFT (Internet Engeenering Task Force), que se encarga de la creación de nuevos protocolos en Internet mediante un conjunto de normas denominadas URIs (Uniform Resources Identifiers). Son particularmente interesantes el URN (Uniform Resources Names) que es un código alfanumérico único y permanente que identifica de forma unívoca a un recurso dentro de Internet.

La combinación del URL con el HTML permite el desarrollo de una capacidad de hipermedia potente y distribuida.

En el web, se combinan los principios del hipertexto con la capacidad de localizar información (hipermedia) en cualquier parte del mundo. Es una iniciativa para la recuperación de información hipermedia en un área amplia para proveer acceso universal a un amplio universo de documentos. Un nuevo protocolo de comunicación, el HTTP que solicita y recibe documentos de host remotos en la red y posibilita la realización de saltos hipertextuales que los anteriores protocolos no permitían. Enlaza ficheros de textos y multimedia, documentos o programas de recuperación de información almacenados en diferentes computadoras conectadas a Internet en todo el mundo.

Es el intento más avanzado de crear una herramienta acceder toda la información disponible en Internet como si fuera un único documento en el que se enlazan sus diferentes partes mediante las relaciones hipertextuales.

Las principales ventajas de Web respecto a Gopher son que su protocolo, el HTTP, integra el resto de los protocolos que configuran Internet, incluido Gopher; además rompe la rigidez de Gopher al basarse en un sistema hipertexto en el que el usuario salta de un documento a otro relacionados por sus temas. Permite, a su vez, el trabajo con documentos multimedia en los que se integran texto, sonido e imágenes. Permite al usuario cierto grado de interactividad o diálogo con el servidor mediante formularios. El sistema no se basa en una estructura de menús, sino en documentos que se enlazan entre ellos. Para la realización de estos enlaces se utiliza el HTML. La inmensa parte de las revistas se distribuyen hoy

mediante web.

Hoy, se normalizan los procesos de edición con el fin de lograr uniformidad en los criterios de registro, transferencia y recuperación de la información. Un ejemplo de ello lo constituye el desarrollo de herramientas de edición electrónica y de lenguajes como el HTML y el SGML con el propósito de crear documentos normalizados que faciliten su consulta, lectura, exportación y exploración.

Para las bibliotecas y los profesionales de la información esto significa que los servicios anteriormente independientes, separados entre si como Telnet, FTP, Gopher, WAIS pueden presentarse ahora como un modelo simple de datos uniformes. El espacio web incorpora el espacio Gopher, el espacio Telnet, el espacio FTP y el espacio WAIS a la facilidad añadida de la capacidad multimedia. Es una increíble oportunidad y cambio para la gestión de información.

El World Wide Web es la responsable del crecimiento espectacular y popularidad actual de Internet.

Una publicación en mayo de 1994, introdujo a los lectores en el nuevo concepto de WWW.

A partir de 1993, con WWW, el éxito definitivo, se percibe la potencialidad de Internet, comienzan grandes inversiones del sector privado en la red. Con el advenimiento de WWW, se hizo evidente el potencial de la red para la comercialización electrónica. La columna vertebral principal de Internet pasó a principios de los 90 de solo entidades gubernamentales como la NSF a manos privadas como las de AT&T y otras. Se hizo evidente su potencial para utilizarla en funciones mas serias de los negocios como para solicitar productos, ordenar pagos, almacenar solicitudes y otras.

Para aquel entonces, hacía más de cincuenta años que se había comenzado a pensar en una base de datos de información universal, con acceso a todos y en la que mediante enlaces semánticos entre los documentos los usuarios pudieran hallar con rapidez la información buscada. En 1938, *H. Wells* había propuesto la idea de la creación de un cerebro mundial, donde se registrara todo el conocimiento acumulado hasta el momento, de una gran bibliografía mundial. Esta era una vieja idea de los bibliógrafos. Sin embargo, fue necesario que transcurrieran varias décadas para que surgieran los medio tecnológicos capaces de soportar tal volumen de información.

En este sentido, el desarrollo de las bases de datos bibliográficas en línea constituyó un paso importante en la actividad del sector de la información. Estos recursos, localizados centralmente, facilitaban la recuperación de información actualizada mediante el uso de las tecnologías de computación y las telecomunicaciones. Posteriormente, el crecimiento de las capacidades de almacenamiento en los medios individuales de trabajo y el desarrollo de los sistemas de recuperación permitió el acceso descentralizado con facilidad. Llegar a la red significó desarrollar primero las computadoras, después las capacidades de almacenamiento y finalmente las telecomunicaciones. Con la institucionalización de la red, se ha producido un cambio en la concepción del desarrollo de colecciones, la labor de identificación, la búsqueda de los recursos es cada vez más importante y la adquisición significa posibilidad de acceso.

Se reconocen tres etapas en el desarrollo tecnológico de los sistemas de información:

- Antes de 1971. El incremento de las capacidades computacionales de los sistemas de información se ajusta a la ley de Moore sobre el crecimiento exponencial que plantea períodos de duplicación de cada uno o dos años.
- De 1971 a 1980. En esta etapa, se adecuan a esta ley, las capacidades de almacenamiento.
- Desde finales de los 80 y principios de los 90. Además de los dos parámetros anteriores, se suman las comunicaciones.

En las dos primeras etapas, las tecnologías cambiaron las operaciones en el contexto de la compañía. En la tercera, las operaciones entre las compañías. En el ambiente bibliotecario, se ha producido un cambio en las concepciones de desarrollo de las colecciones, la labor de identificación y de búsqueda de los recursos se ha convertido en la más importante. Adquirir significa la posibilidad de acceder a un recurso determinado.

De la misma forma que 1995 fue el año de Internet, 1996 fue el año de las Intranets. Estas redes son una especie de mininternets que se instalan dentro de una organización o compañía que proporciona enlaces internos y a Internet en general.

Tecnológicamente, constituyen una evolución lógica de las redes locales y amplias que eliminan las limitaciones propias de sus precursoras. Presentan, como valor añadido, la posibilidad de conectarse a los enormes recursos de información y de comunicación que ofrece Internet.

El concepto de intranet se refiere a la aplicación de las herramientas utilizadas en WWW a la configuración de la red local para posibilitar la gestión de la información interna. Gestionar la intranet de una empresa le permite al centro de documentación convertirse en el articulador de la información de los distintos departamentos de la organización y, a la vez, en su nexos con el exterior.

De esta manera, las ventajas propias de Internet se trasladan al entorno inmediato de la organización: navegación asociativa e intuitiva, amigabilidad de la interfaz, multimedia, integración con los recursos externos de información presentes en Internet y otras.

Una de las ventajas de la gestión de información en una plataforma web es la posibilidad de integrar los recursos de información internos con los externos, presentes en Internet, en un mismo entorno. La selección y gestión de los recursos de información que ofrece Internet, relacionados directamente con nuestro segmento de actividad, puede ser un buen servicio complementario, sobre todo para los centros de documentación o bibliotecas especializadas que actúan como un "filtro" necesario ante la acumulación de información irrelevante para nuestros usuarios.

La creación de una red local o intranet permite integrar y estructurar los recursos de información internos y externos de la colección organizacional bajo una misma interfaz, fácil e intuitiva. Debemos apostar por Internet y colaborar activamente en la resolución de sus problemas mediante el aporte de

años de experiencia de la profesión bibliotecaria en la gestión y recuperación de información.

Mientras que los primeros públicos para la mayoría de los sitios web fueron los clientes externos, el potencial para servir a los clientes internos podría igualmente ser tan grande. La creación de las redes internas para facilitar la comunicación y las transacciones entre los empleados, los suministradores y los contratistas independientes y los distribuidores podría ser el principal valor de Internet para las compañías.

Un reciente artículo sobre las intranets reveló que las ventas de software para los servidores de las intranets habían sobrepasado las ventas para los servidores de Internet a finales de 1995 según las compañías reconocían las amplias posibilidades de uso de las aplicaciones de la intranet.

Una intranet posee un gran valor potencial para la comunicación interna en la compañía.

Hacia 1996, la Federal Communications Commission (FCC) adquiere como su responsabilidad primaria la elaboración de la política de telecomunicaciones en los Estados Unidos. En 1996, durante la Internet Society International Conference en Montreal, la FCC expresó, como sus principales preocupaciones, los problemas relacionados con el acceso a Internet y el ancho de banda de la red.

Desde el surgimiento del embrión de la red que posteriormente se convertiría en lo que hoy se denomina Internet, en 1969, el crecimiento del número usuarios y redes interconectadas se ha acelerado progresivamente.

En el año 1972, existían 10 procesadores de mensajes de interfaz, a principios de los 80, cerca de 100.

El número de redes conectadas con Internet creció rápidamente, de unas 60 en 1984, a 500 a mediados de 1988

Entre agosto de 1981 y octubre de 1991, el número de host creció de 213 000 a 617 000. De enero de 1992 a enero 1996, de 727 000 a 9 472 000. En este período, creció igualmente el número de sitios comerciales, superaron a los educacionales, así como el de usuarios de Internet no estadounidenses.

Los dominios net ,organizaciones administrativas , crecieron para las redes en el segundo semestre de 1995 a más del doble. Cerca de 76 000 sistemas ahora se nombran??? WWW (enero 1996) de 17 000 hacía 6 meses anteriores.

En mayo de 1996, *M. Gray*, informó sobre la existencia de 1 93 150 servidores web. En junio de 1993, uno de cada 13 000 host en Internet era un servidor web. En enero de 1996, eran uno de cada 94. Igualmente, sucedió con el crecimiento de los web comerciales. En Estados Unidos, en junio de 1993, sólo 1,5% eran dominios .com, en enero de 1996, el 50%.

Entre los años 1993 y 1994, el tráfico de información en Internet se duplicó de año en año.

En agosto de 1993, existían más de 14 000 redes conectadas con Internet y según las estadísticas de aquel momento, se incorporaban 1000 nuevas redes como promedio mensualmente.

En enero de 1995, la cifra total había aumentado a 26 274 redes. Habían conectados 84 países con acceso directo a esta red y 48 millones de usuarios.

En 1995, Internet alcanzaba a 148 de los 185 países miembros de las Naciones Unidas (86%) en comparación con los 73 de los 159 países miembros (46%) en 1991. Las aplicaciones disponibles cubren todos los sectores de la vida de la sociedad.

El acceso a Internet en los mercados foráneos promete crecer rápidamente según se expandan los servicios en línea y según se interesen más y más los gobiernos regionales y nacionales y las compañías de telecomunicaciones.

El crecimiento internacional a largo plazo de Internet aumenta las oportunidades del flujo y la transacción de la información a través de las fronteras. En 1995, el volumen de transacción en web se estimó en más de \$400 millones, en comparación con los menos de \$20 millones en 1994 - más del 80% de los cuales fueron a compañías de los Estados Unidos.

El volumen de transacción global alcanzó más de mil millones de dólares en 1996.

En abril de 1997, el 73% de los usuarios de Internet utilizaban WWW para realizar sus compras. No existen dudas acerca del nacimiento de un mercado electrónico, pero el trabajo de consolidarlo y establecer un conjunto mínimo de requisitos y normas aún está pendiente.

En el crecimiento acelerado de Internet han influido factores, tanto de carácter técnico como político. Entre los primeros, se destacan el éxito de los protocolos de la red, el progresivo abaratamiento de los microprocesadores y el triunfo de las nuevas concepciones de organización de redes en forma descentralizada y distribuida. Entre los factores de tipo político, aparece la acción del gobierno estadounidense y de muchas otras naciones que han apoyado incondicionalmente el desarrollo de una infraestructura de comunicaciones adecuada para soportar el uso de los recursos disponibles en Internet de forma amplia.

El rápido crecimiento de Internet se explica, en gran parte, en el éxito del protocolo TCP/IP para conectar redes y equipos de diferentes tecnologías. La incorporación de este protocolo en las configuraciones básicas de los principales equipos por parte de los proveedores comerciales impulsó decisivamente su uso.

El decrecimiento dramático de los precios del hardware ha permitido a las compañías invertir un por ciento menor de su presupuesto para sistemas de información en hardware. Sin embargo, los costos del software y de su empleo constituyen un renglón creciente de los costos totales de los sistemas de

información.

En el mundo de los negocios, el web reducirá la ventaja competitiva de las economías de escala en muchas industrias y facilitará el trabajo de los pequeños comerciantes para que compitan a nivel mundial. Los anuncios de publicidad como una barrera para entrar, se reducirán en la medida que el web posibilite alcanzar una audiencia global de forma más económica.

Una pequeña compañía puede convertirse rápidamente en un gran jugador internacional al aprovecharse de las tecnologías que permiten determinar las formas en que reaccionan los clientes o sus necesidades.

Las ventajas de tamaño de las grandes compañías desaparecerán.

El apoyo y la cooperación gubernamental foráneas determinarán cómo evolucionarán los negocios internacionales por medio de Internet. Las iniciativas tomadas por algunos gobiernos, asociaciones del comercio y compañías de telecomunicaciones son un buen presagio para la futura expansión.

Es incorrecto considerar a Internet simplemente una opción adicional para la estrategia de los negocios, aunque este pueda ser la práctica actual de muchos. Pocos negocios han reconocido el potencial de Internet y sus oportunidades para la comunicación bidireccional -entre la compañía y sus usuarios- como una ventaja seria.

Las cuestiones de seguridad ocuparon el primer lugar en las agendas de trabajo de las compañías que buscan obtener una presencia en Internet. La renuencia de muchas compañías a utilizar Internet tiene su fundamento en el temor a comprometer información interna de negocios.

En materia de publicaciones electrónicas, las redes teleinformáticas de alcance mundial se constituyeron un nuevo canal de comunicación útil para la distribución de información a millones de personas. Sin embargo, dicha distribución, no implica la realización de una publicación. Para que un documento se considere una publicación se requiere de que este transcurra por un proceso de arbitraje y edición que valide su contenido y metodología y adecue su presentación a las normas que para ello la ciencia exige.

El Directory of Electronic Journals, Newsletters and Academic Discussion List, el primero y más completo de los directorios de revistas electrónicas en el mundo que comenzó a publicarse en julio de 1991 por la Association of Research Libraries, recogió en su quinta edición de mayo de 1995 más de 675 revistas científicas y boletines electrónicos, lo que supone un incremento del 40 % con respecto a la edición anterior de 1994.

En un primer momento, la forma de distribución de estas publicaciones fue mediante el correo electrónico, los servidores FTP y sobre todo de los Gopher. Sus limitaciones eran importantes, por ejemplo, el correo electrónico permitía solo la distribución de documentos en formato ASCII.

EL Gopher se extendió rápidamente. Más adelante, con la aparición del WWW y sus ventajas con

respecto a Gopher, las revistas existentes comenzaron a migrar a servidores Web en un proceso en el que al principio convivieron los dos sistemas y poco a poco se fue imponiendo el segundo.

Por otra parte, hacia 1996, se produjeron una serie de avances tecnológicos que permitieron pronosticar que la infraestructura de televisión por cable constituirá el soporte fundamental para la implementación de servicios multimedia a nivel mundial. En particular, dichos sistemas forman una plataforma de uso general para los servicios bidireccionales como el acceso en banda ancha a Internet, y la transmisión de datos y voz. Ellos presagiaron el nacimiento de poderosísimas tecnologías de comunicación, propias de un nuevo escalón de desarrollo, Internet 2, que servirían para enfrentar los nuevos retos que plantean el acceso de millones de personas a la red.

El web ha transformado radicalmente la forma en la cual se comunican las personas. En el futuro, probablemente se considerará un adelanto comparable a los de la radio y la televisión en su tiempo.

Internet constituye un fenómeno de enorme importancia actualmente, cuyo impacto es comparable al que tuvo para el sistema feudal el descubrimiento y uso de la pólvora. Sus efectos han generado cambios trascendentales en todas las esferas de la vida de la sociedad moderna.

Conceptualmente, Internet es una red de redes de comunicación con alcance mundial que posibilita a sus usuarios el intercambio y el acceso universal a los recursos de información de manera sencilla e interactiva.

Una de las mayores críticas que enfrenta Internet es la carencia de una organización y una estructura. Los usuarios encuestados perciben una necesidad real de mejores directorios en Internet, creados por personas con algún conocimiento sobre indización y catalogación.

Muchos de los directorios en Internet son generados por individuos con un gran conocimiento en computación pero desconocedores de los problemas relacionados con la organización de la información. Internet podría ser mucho mas amistosa si las pantallas para los usuarios finales las diseñaran los profesionales de la información. La definición de mecanismos precisos de búsqueda y recuperación es un aspecto esencial para la evolución satisfactoria de Internet. Los métodos empleados actualmente, donde no existe control del vocabulario y la información no se halla estructurada, no ofrecen un resultado adecuado.

Directorios como Yahoo y Magellan, basados en la clasificación humana de los recursos informacionales de la red, no son capaces de seguir su crecimiento vertiginoso. Estos presentan problemas que se derivan de la acción de la subjetividad humana, así como del empleo de interminables categorías y subcategorías. Su mayor virtud es la precisión en los resultados de las búsquedas, aunque la posibilidad de "silencio" documental es alta.

Otra alternativa es la utilización de directorios elaborados por robots que realizan una indización automática a texto completo, en la que los resultados de la búsqueda se ordenan con frecuencia según su

nivel de relevancia a partir de algoritmos matemáticos que normalmente emplean la repetición del término buscado en el documento completo o en determinadas zonas del mismo para su clasificación. El problema es el inverso: gran exhaustividad (resultados de miles de registros), pero el "ruido" que se produce es excesivo.

La única posibilidad para solucionar este problema al menos parcialmente es el uso de buscadores que permitan acotar los resultados de nuestras estrategias de búsquedas mediante operadores booleanos y de proximidad. Altavista, Excite, Open-Text y Lycos pertenecen a este tipo de motores de búsqueda.

Altavista es la herramienta más completa, potente y precisa para la recuperación de información en el Web. Incluye la mayor parte de los recursos empleados tradicionalmente en la recuperación de información: operadores booleanos, de proximidad o adyacencia; así como la posibilidad de truncar a la derecha y por caracteres, limitar por fechas, palabras en el título y otras.

Pero, el número de motores de búsqueda crece aceleradamente su número, como ocurrió con las bases de datos en el pasado, que se multiplicaron tanto que hicieron prácticamente imposible la búsqueda, sin contar los altos índices solapamiento entre ellas.

La última opción son los metaíndices o motores de búsqueda múltiples que interrogan a otros buscadores y eliminan posteriormente los documentos duplicados. Aunque proporción una mayor exhaustividad, no responden al problema de fondo: la precisión de los resultados.

La solución al problema es difícil, si se considera que el crecimiento de la información ocurre de forma exponencial y, gran parte de ella posee un carácter eminentemente comercial. Una de las propuestas para resolver dicho problema es la creación de sistemas distribuidos de indización en los que los administradores de cada web realizarían la indización de sus páginas. La indización se aparecería en un fichero normalizado que recuperaría un robot para alimentar una base de datos global. Este fichero contendría una serie de campos semejantes a los de una ficha catalográfica.

Para esto, es necesario una normalización, establecer como describir un recurso, como clasificarlos, como indizarlos correctamente. Ello es vital para una buena recuperación, los lenguajes permutados serán esenciales para aumentar los índices de recuperación. Si todo esto no funciona, la recuperación será cada vez más deficiente.

Es indispensable igualmente que cada una de las profesiones y las bibliotecas establezcan sólidos criterios para la validación y la selección de los recursos de información, porque el desperdicio informacional crece aceleradamente. En el sector de la medicina clínica, existen notables avances en este sentido, con los trabajos realizados por el llamado club de Vancouver y las organizaciones vinculadas con el ejercicio de la medicina basada en evidencias. Hoy la mayoría de los sitios colocan sus recursos en la red sin que para ello medie un proceso de arbitraje previo.

En Internet falta más la presencia de bibliotecarios, generalmente olvidados de muchas actividades de la sociedad y que poseen la mayor experiencia en la organización del conocimiento.

Después de una pobre organización, la carencia de datos autorizados, clasifica entre las preocupaciones principales de los bibliotecarios. A menudo, no aparece cómo se recolectó la información o sus fechas, propiedades cruciales para los bibliotecarios en los negocios. Otra limitación es la existencia de demasiada información; Internet es para personas con tiempo suficiente para explorar a través de información irrelevante hasta hallar la deseada. Cada vez resulta más difícil hallar la información que se requiere en Internet. En este sentido, los bibliotecarios tienen mucho que aportar en la Era de la Información, porque ellos han manipulado información durante siglos.

Las bibliotecas del sector de las finanzas, del comercio y de los negocios están entre las más afectadas por esta situación. Ellas han estado en la primera fila de los cambios tecnológicos y administrativos en el mundo de la información. Dentro de este sector, el de las finanzas es el más influyente, porque cualquier cambio en el repercute rápidamente en otros sectores de la economía y la sociedad. Este emplea en Inglaterra, por ejemplo, más profesionales de la información que otros sectores. Sin embargo, la carencia de información confiable en la red es uno de los factores que influye decisivamente en su consideración como un medio secundario, cuando de hallar información autorizada se trata.

Internet promete revolucionar la dinámica del comercio internacional y al igual que el teléfono y los equipos de fax, podría ser una fuerza principal en la democratización del capitalismo. Las pequeñas compañías serán capaces de competir más fácilmente en el mercado global y los consumidores en los mercados emergentes en particular, se beneficiarán del un número crecido de productos, servicios e información a los cuales la Internet les dará acceso. Internet desaparece muchas barreras de la comunicación entre los clientes y los empleados al eliminar los obstáculos creados por la geografía, las zonas de tiempo y la localización y crea un ambiente de negocios sin fricciones.

Es el momento en que se han hecho realidad las ideas, las concepciones de hace décadas y se pasa del papel al soporte electrónico. La biblioteca como institución ha alcanzado su clímax con el paso a una nueva etapa de su desarrollo: la biblioteca virtual, sin papel y sin paredes.

Internet no es solo una red para la transmisión de información científica sino un medio de difusión masiva por el que una gran parte de la información que transita no es académica.

Con el desarrollo de aplicaciones que conciben a la red como una gran base de datos distribuida, el acceso a la información se ha visto más favorecido que la comunicación. Sin embargo, nuevos progresos podrían restablecer el dominio de las comunicaciones sobre el acceso a la información. Internet puede estar transformándose en un nuevo dispositivo para la publicación, en la medida en que responde cada vez más a la agenda de las diferentes compañías establecidas en el ciberespacio y cada vez menos a las necesidades de los individuos que desean mantenerse en contacto. El advenimiento de Java, un software para la publicación, refuerza esta tendencia. No obstante, nada es seguro. Suceda lo que suceda, siempre existirá espacio suficiente para aquellos que prefieren comunicarse en lugar vez de recibir información según las estrategias de los medios de radiodifusión. De cualquier forma, Internet es un medio inadecuado hasta el momento para la difusión, hecho el cual podría salvarla de caer en la trampa de la radio o la televisión.

Fuera del contexto de las aplicaciones, Internet se enfrenta a una batalla recurrente con los grupos o compañías que intentan crear redes que funcionan como circuitos y que, por lo tanto, facilitan un mayor control, sea este político o económico, en comparación con el TCP/IP. El debate actual sobre el ATM difiere de las primeras batallas entre el X.25 y el TCP/IP que ya no lucha por la supervivencia. Sin embargo, se puede concebir un despliegue inicial del ATM que utilice TCP/IP y, en forma gradual, introducir aplicaciones propias de esta nueva tecnología. Finalmente, ATM podría dominar la transferencia de información en Internet y el TCP/IP constituirse en un sector del ATM.

De igual forma que Internet revolucionó las comunicaciones a nivel mundial, el Asynchronous Transfer Mode (ATM) aporta actualmente un nuevo sentido a las redes de alta velocidad. El ATM representa a una nueva generación de redes. Promete ser un medio clave en las redes para el empleo tanto de las aplicaciones existentes como emergentes en los campos de la ciencia, la medicina y la educación.

La fusión de las compañías telefónicas y de televisión por cable, el empleo de una tecnología como el ATM para la conmutación y el transporte de la información, así como de sistemas de comunicación, cada vez más potentes, posibilitan el envío de datos, voz y vídeo mediante un solo medio físico en tiempo real de forma tal que exista una sola red capaz de proporcionar a los usuarios - pequeños, medianos y grandes- los servicios que demandan.

Se ha entrado a una nueva era cultural, la era del acceso instantáneo a cualquier lugar del mundo, la era en la que existe, por primera vez en la historia, una total libertad para la comunicación y el flujo sin limitaciones de información, la era en que todas las culturas podrán enlazarse por medio de las computadoras para conformar una inteligencia colectiva que llevará a cambios fundamentales en las respuestas culturales, políticas y sociales de los pueblos.

El futuro de la Internet Society dependerá, en gran medida, de su capacidad de crear y perfeccionar sus protocolos rápida y eficientemente. La solución del problema del direccionamiento y el aumento de la seguridad de las transacciones en la red son las primeras prioridades del perfeccionamiento de Internet. En los años 90, el establecimiento de una Internet Society fue un paso muy importante en esta dirección que pronostica un buen futuro a un sistema mundial que debe permanecer aislado de influencias políticas, económicas o ideológicas.

El sector de las telecomunicaciones experimenta actualmente un proceso de desregulación y privatización. La necesidad de atraer inversiones foráneas directas para incrementar el número de líneas telefónicas de las economías en desarrollo, así como de modernizar y potenciar subsectores como la telefonía local y de larga distancia, la televisión por cable y la transmisión por satélite, han generado una convergencia de las telecomunicaciones, la visión y la inversiones mundiales durante la década de los años 90. En un tiempo relativamente corto, los satélites se han convertido en una parte esencial de la comunicación global. El primer satélite de televisión, el Echo, se lanzó al espacio en 1960.

De cierta forma, el TCP/IP es el idioma que emplean como vehículo para comunicarse las computadoras

en el mundo, razón por la cual, al igual que cualquier otro idioma, debe situarse por encima de las actividades humanas, de forma tal que facilite y mejore su eficiencia. Nadie desearía que el español, el inglés o el francés pasara a propiedad de una entidad política o comercial. De igual forma, nadie desea que el TCP/IP se convierta en una herramienta propiedad de un grupo, compañía o país en particular. El pertenece a toda la humanidad; por lo tanto, debe preservarse, valorarse y perfeccionarse continuamente. Las telecomunicaciones, la tecnología de las computadoras personales y la televisión convergen, tanto a nivel de la red como de las aplicaciones. La tecnología y las aplicaciones de las computadoras personales cada vez soportan mejor el audio y al video mientras que la televisión es cada vez más interactiva. Las redes de telecomunicaciones y los servicios de radiodifusión progresan en la esfera de las aplicaciones multimedias interactivas. Un aspecto clave para el rápido crecimiento de estos nuevos servicios y de las aplicaciones es la disminución del costo de las comunicaciones.

En estas condiciones, los métodos bibliotecarios y de la ciencia de la información, que se desarrollaron cuidadosamente durante décadas, han perdido su habilidad, elegancia, precisión y aplicabilidad como consecuencia del rápido, caótico y espontáneo de desarrollo de Internet. Los documentos, resultantes de la búsqueda a la que se hizo referencia con anterioridad pueden ser diferentes en carácter, extensión y calidad. Por lo tanto, se carece de la posibilidad de realizar comparaciones entre ellos. Esta característica de Internet no sólo afecta a los estudios de bibliotecas y de información sino que también cambian los términos de conceptos profundamente enraizados y estructurados como son las técnicas de indización y la lógica booleana en la cual se sustenta el poder de los operadores lógicos.

El desarrollo descontrolado de las ofertas para realizar búsquedas de información en el web, como la que posibilitan los denominados motores, con frecuencia, genera que una misma solicitud se interprete e indice de forma diferente por cada uno de estos sistemas, la cual no siempre se explica con claridad. Si a esto se une, la forma específica de la aplicación de los operadores lógicos en las búsquedas, se puede decir que existen motivos suficientes como para desconfiar de lo que se busca y lo que se encuentra en Internet. Se conoce que en Internet en materia de información hay "oro" por descubrir; el problema radica en que es casi imposible estimar, por anticipado, cuántos de ellos pueden distinguirse con este calificativo. En la práctica, los usuarios pueden perder una gran cantidad de tiempo en la búsqueda de lo que realmente necesitan.

## Referencias bibliográficas

1. Organización Panamericana de la Salud. Internet, Telematics and Health. Washington D.C: IOS Press, 1997.
2. Chamero J. Historia de Internet y el Internet histórico. Disponible en: [http://www.aunmas.com/future/internet\\_historia/](http://www.aunmas.com/future/internet_historia/)
3. Leiner BM, Cerf VG, Clark DD, Kahn RE, Kleinrock L, Lynch DC, Postel J, Roberts LG, Wolff S. Una breve historia de Internet. Disponible en: <http://www.ati.es/DOCS/internet/histint/>
4. Lynch CA, Preston CM. Internet access to information resources. ARIST 1990;25:263-312.
5. Urra González P. Taller Internet. La Habana, CNICM, 1998.
6. Quelch JA, Klein LR. The Internet and international marketing. Sloan Management Review

1996; (spring):60-75.

7. Quaterman JS, Phillips G, Carl-Michell S. Internet services and access. Public Access to the Internet Symposium. (Harvard University, 1993).
8. Leiby C, Konkol M. Security issues on the Internet. Aslib Proceedings 1996;48(5):123-127.
9. Barrueco Cruz JM, García Testal C, Gimeno MJ. Una aproximación a las revistas científicas en formato electrónico. Rev Esp Doc Cient 1996;19(3):304-313.

Recibido: 16 de diciembre del 2003 Aprobado: 7 de enero del 2004

Lic. *Rubén Cañedo Andalia*

Red Telemática de Salud en Cuba (Infomed). Centro Nacional de Información de Ciencias Médicas.

Calle 27 No. 110 entre N y M, El Vedado. C P 10 400, Ciudad de La Habana, Cuba. Correo electrónico:

[ruben@infomed.sld.cu](mailto:ruben@infomed.sld.cu)

<sup>1</sup> **Licenciado en Información Científico-Técnica y Bibliotecología. Red Telemática de Salud en Cuba (Infomed) Centro Nacional de Información de Ciencias Médicas.**

**[Indice Anterior Siguiente](#)**