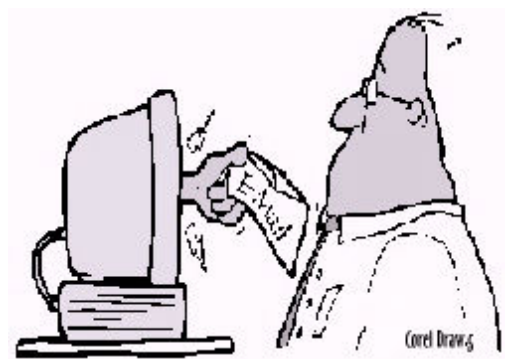


# Internet

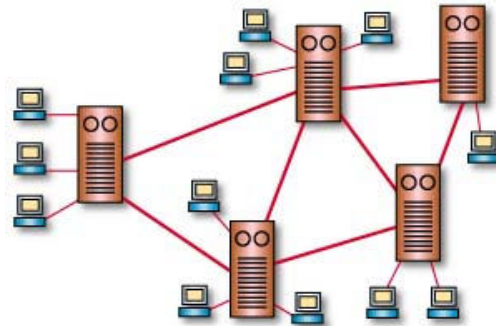


## Internet

### Historia de Internet

Aunque se pueda pensar que Internet es algo que ha surgido en estos últimos tiempos debido a su auge, no es así: Internet ya lleva con nosotros unas cuantas décadas.

Los inicios de Internet se remontan a los años '60. En plena guerra fría, Estados Unidos crea una red exclusivamente militar, con el objetivo de que, en el hipotético caso de un ataque soviético, se pudiera tener acceso a la información militar desde cualquier punto del país. Este red se creó en 1969 y se llamó ARPANET.



En principio, la red contaba con 4 computadoras distribuidas entre distintas universidades del país. Dos años después, ya contaba con unas 40 computadoras conectadas. Tanto fue el crecimiento de la red que su sistema de comunicación se quedó obsoleto. Entonces dos investigadores crearon el Protocolo TCP/IP, que se convirtió en el estándar de comunicaciones dentro de las redes informáticas (actualmente seguimos utilizando dicho protocolo).

ARPANET siguió creciendo y abriéndose al mundo, y cualquier persona con fines académicos o de investigación podía tener acceso a la red. Las funciones militares se desligaron de ARPANET y fueron a parar a MILNET, una nueva red creada por los Estados Unidos. La NSF (National Science Foundation) crea su propia red informática llamada NSFNET, que más tarde absorbe a ARPANET, creando así una gran red con propósitos científicos y académicos. El desarrollo de las redes fue abismal, y se crean nuevas redes de libre acceso que más tarde se unen a NSFNET, formando el embrión de lo que hoy conocemos como INTERNET.

El desarrollo de NSFNET fue tal que hacia el año 1990 ya contaba con alrededor de 100.000 servidores.

En este contexto, el CERN (Centro Europeo de Investigación de Partículas) crea las *páginas web*, con el objetivo de comunicarse con otros científicos europeos. En 1993 un estudiante norteamericano escribió el código del primer explorador web, el Mosaic, que se distribuía de forma gratuita por la red y permitía tener acceso a gráficos y documentos de texto dentro de Internet. Esto supuso una auténtica revolución, y a partir de ese momento, Internet no ha parado de crecer. Es así como en el año 1996 existían cerca de 90.000 sitios web, mientras que en la actualidad se superan los millones de sitios web.

Contenido extraído de: <http://www.iespana.es/mundointernet/historia.htm>

## Funcionamiento de Internet

Internet es, básicamente, una gran red mundial donde millones de computadoras están conectadas entre sí, siendo independientes unas de otras. Para que todas estas computadoras puedan coexistir y comunicarse, deben ponerse de acuerdo. Con este motivo fueron creados los Protocolos, que son reglas de comunicación que han de adoptarse para que una computadora sea entendida por las otras computadoras de la red. Los dos protocolos más importantes son el *Protocolo de Control de Transmisión* (**Transfer Control Protocol**) y el *Protocolo de Internet* (**Internet Protocol**). Usualmente se trata a estos dos protocolos como si fueran uno solo, **TCP/IP**. Una computadora, si maneja estos dos protocolos, no tendrá ningún problema para ser entendida por las demás computadoras de Internet.

Todas las computadoras que componen Internet no están conectadas unas con otras: sólo se está conectado con las más cercanas. Si se realiza un pedido de cierta información a una computadora que está ubicada a cientos de kilómetros de donde uno se encuentra, se accede a esta computadora a través de otras, formando una especie de cadena. Para saber donde está la computadora a la que se quiere acceder, se utilizan las Direcciones IP (son como una especie de DNI de las máquinas) las cuales están compuestas por una combinación de números entre 0 y 255 y se utilizan para identificar a cada computadora dentro de una red. Una vez sabido esto, la información que hemos solicitado a la computadora lejana, nos llega a nosotros en pequeños paquetes que una vez en nuestras máquinas, se enlazan mostrándonos la información pedida en pantalla.

## Servicios de Internet

Cuando un usuario se conecta a Internet, puede acceder a miles de computadoras de todo el mundo. Ahora bien, ¿cómo interpretar la información de todas esas computadoras? Existen varios **servicios** de Internet que definen normas para presentar y enviar la información que circula por la red, de forma que el usuario pueda recuperarla. Los más importantes son el correo electrónico (**E-MAIL**), la World Wide Web (**WWW**), los grupos de noticias (**NEWS**), conversación en tiempo real (**IRC o CHAT**) y acceder y transferir archivos (**FTP**).

Otros servicios de Internet, son Gopher, Archie, o Telnet. Todos ellos son distintas formas de acceder a la información de las computadoras de Internet, pero se utilizan con menos frecuencia.

### World Wide Web (WWW)

La World Wide Web es, por decirlo de alguna manera, el servicio más popular de Internet. Cuando ya se haya pasado cierto tiempo en la Web, se empezará a sentir que no hay límites a lo que se puede descubrir en ella. La Web permite una comunicación rica y variada gracias a que puede presentar información mediante distintos medios como son el texto, gráficos, animaciones, fotos, sonidos y videos.

La Web es conocida como un sistema cliente-servidor. La computadora que estamos utilizando es el cliente y la computadora remota que alberga los archivos electrónicos a los que queremos acceder es el servidor. A continuación explicamos cómo funciona:

Digamos que se quiere visitar el sitio Web del museo del Louvre. Primero se escribe la dirección URL del sitio web en el explorador Web o Browser (hablaremos de esto enseguida). Entonces el explorador solicita la página Web del servidor que alberga el sitio del Louvre. El servidor envía de vuelta los datos a nuestra computadora a través de Internet. El explorador interpreta los datos, mostrándolos en la pantalla.

El sitio de Louvre también tiene vínculos a sitios web de otros museos, como el Museo del Vaticano. Cuando se hace clic con el mouse sobre ese vínculo, se puede acceder al servidor web del Museo del Vaticano.

Esta manera de recorrer los sitios es posible gracias al uso del hipertexto. Theodor H. Nelson señala... "Con hipertexto, me refiero a una escritura no secuencial, a un texto que bifurca, que permite que el lector elija y que se lea mejor en una pantalla interactiva. De acuerdo con la noción popular, se trata de una serie de bloques de texto conectados entre sí por nexos, que forman diferentes itinerarios para el usuario". Los nexos o también conocidos como hipervínculos, son el elemento que permite vincular los archivos electrónicos en la Web, de forma que se pueda saltar fácilmente de unos a otros. En la Web, uno navega a través de páginas de información -es lo comúnmente conocido como explorar o navegar- conforme a lo que le interese en un determinado momento.



Como comentábamos, para acceder a la Web se necesita de un explorador Web o browser, como el Netscape Navigator, Microsoft Internet Explorer, Mozilla Firefox u Opera software, entre otros. ¿Cómo hace el explorador para distinguir entre páginas Web y otros tipos de datos que hay en Internet? Las páginas Web están escritas en un lenguaje de computación llamado Lenguaje de Marcación de Hipertexto o HTML (del inglés HyperText Markup Language).

Entonces uno puede imaginarse la World Wide Web como una red de archivos electrónicos (archivos HTML) o como mencionábamos anteriormente páginas web, las cuales están organizadas en forma de sitios y están almacenados en computadoras de todas partes del mundo. El **Hipertexto** vincula estos recursos unos a otros. Los **Localizadores Uniformes de Recursos** o **URLs** (del inglés Uniform Resource Locator) son las direcciones que se utilizan para localizar estos archivos. La información que contiene una URL proporciona la capacidad de saltar de una página web a otra simplemente con un clic del mouse. Cuando se teclea una URL en el explorador o se hace clic en un vínculo (no solamente el texto puede servir de vínculo o enlace, un objeto como una imagen también puede serlo), el explorador envía una petición

para descargar uno o más archivos a una computadora remota, llamada un **servidor web**.

### ¿Cómo está formada una URL típica?

Veamos algunos ejemplos:

<http://www.unrc.edu.ar> La página de inicio (Home Page) del sitio de la Universidad Nacional de Río Cuarto.

<http://www.learnthenet.com> La página de inicio de Aprenda la Red.

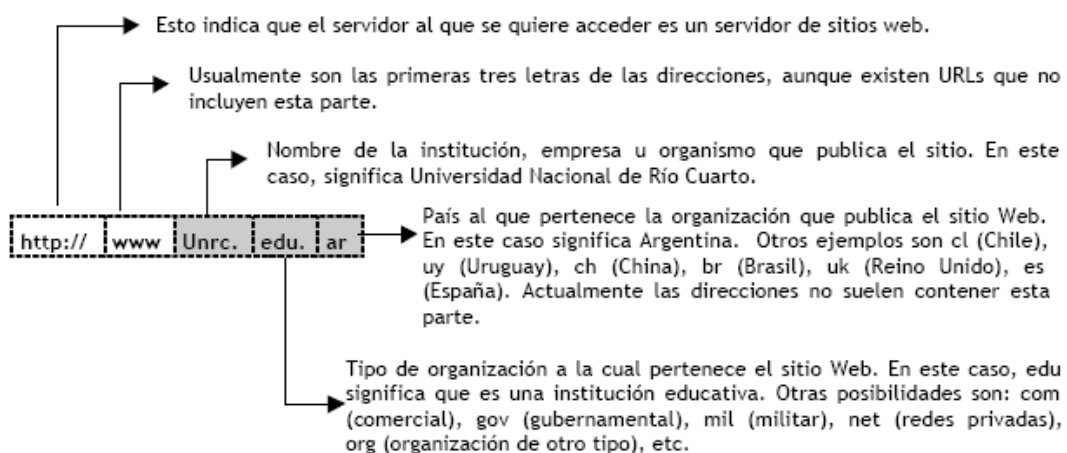
<ftp://rtfm.mit.edu/pub/> Un directorio de archivos en el MIT, disponibles para ser descargados.

La primera parte de una URL (antes de las dos barras inclinadas) indica el tipo de recurso o método de acceso a esa dirección. Por ejemplo:

- **http:** Un documento de hipertexto.
- **ftp:** Un archivo, o un directorio de archivos, disponibles para ser descargados.
- **news:** Un grupo de noticias.
- **telnet:** Un sistema informático en el que se puede abrir una sesión a través de Internet.
- **file:** Un archivo situado en un disco local, por ejemplo su disco duro.

La segunda parte es, por lo general, la dirección del computadora donde se encuentran los datos o el servicio. Puede haber otros elementos adicionales para especificar nombres de archivos, el puerto al que conectarse, o un texto a buscar en una base de datos.

En síntesis una URL puede estar formada por las siguientes partes:



La URL de un sitio se introduce tecleándola en la barra *Dirección* del explorador Web, situada justo debajo de la barra de herramientas.



Muchos exploradores almacenan las URL's que se pretenden utilizar de nuevo, añadiéndolas a una lista especial. En Netscape Navigator se llaman **Marcadores**; en Microsoft Explorer, **Favoritos**. Una vez que se agregue una URL a la lista, se puede regresar a esa página web simplemente haciendo clic en el nombre correspondiente de la lista, en vez de tener que volver a teclear toda la URL.

La mayoría de las URL's que se utilizarán empezarán con **http**. A continuación se detallan algunas otras cosas a saber sobre las URL's:

- Generalmente, una URL no contiene espacios.
- Una URL siempre usa barras diagonales (//).
- Si se introduce una URL incorrectamente, el explorador no será capaz de localizar el sitio o recurso deseado. Si se recibe un mensaje de error o se accede a un sitio equivocado, hay que asegurarse de que se ha tipeado correctamente la dirección.
- Se puede saber cuál es la URL que se oculta tras cualquier vínculo colocando el cursor sobre el mismo. El puntero se convierte en una mano, y la URL aparecerá en la **barra de estado**, normalmente situada en la parte inferior de la pantalla.

Contenido de base: <http://www.learnthenet.com>

### Guardando elementos de una página Web

Muchas veces es útil poder grabar en nuestro disco rígido (o en un diskette) el contenido de una página Web. Si vamos al menú *Archivo* y luego a la opción *Guardar Como*, en la ventana del explorador Web, nos aparecerá el cuadro de diálogo mediante el cual elegiremos en qué carpeta, de qué disco y con qué nombre queremos grabar el archivo.

Si queremos almacenar todos los archivos necesarios para mostrar la página (gráficos, imágenes, sonidos, etc) debemos elegir guardar con el tipo de archivo *Página Web, completa (\*.htm; \*.html)*. En cambio, si necesitamos guardar únicamente la página HTML actual, elegimos el tipo de archivo *Página Web, sólo HTML (\*.htm; \*.html)*. Esta opción almacena la información de la página Web, pero no los gráficos, sonidos ni otros archivos. También si queremos guardar únicamente el texto de la página Web actual, seleccionamos *Archivo de texto (\*.txt)*. Esta opción graba la información de la página Web como texto sin formato.

Si el documento fue guardado como página web, puede ser abierto utilizando un editor de páginas web o un procesador de textos, que además nos permitirá modificarlo si es necesario.

Otra forma de rescatar el texto de una página, es aplicando los comandos de copiar y pegar. El procedimiento es bastante sencillo. Primero se selecciona el texto de nuestro interés dentro de la página actual. En el menú *Edición* de la ventana del browser seleccionamos la opción *Copiar*. Luego vamos al editor, donde pegaremos el texto y del menú *Edición* de la aplicación seleccionamos la opción *Pegar*.

Finalmente, si lo que deseamos es guardar una imagen de una página Web que estamos viendo en el navegador Web, el procedimiento es simple.



Debemos hacer clic con el botón secundario sobre la imagen, apareciendo un menú contextual como el que sigue:

La opción *Guardar imagen como...* **(a)** nos permite grabar la imagen en el disco. *Establecer como papel tapiz* **(b)** coloca la imagen como fondo del escritorio y *Copiar* **(c)** nos permite enviarla al portapapeles, para luego pegarla en un

programa de edición de gráficos o un procesador de textos.

### Algunas direcciones útiles

[www.unrc.edu.ar](http://www.unrc.edu.ar)  
[intra.exa.unrc.edu.ar](http://intra.exa.unrc.edu.ar)

[dc.exa.unrc.edu.ar](http://dc.exa.unrc.edu.ar)  
[www.mcye.gov.ar](http://www.mcye.gov.ar)  
 Nación.

[www.museosargentinos.org.ar](http://www.museosargentinos.org.ar)  
[www.bibnal.edu.ar](http://www.bibnal.edu.ar)  
[www.nationalgeographic.com](http://www.nationalgeographic.com)  
[www.rae.es](http://www.rae.es)  
[www.diccionarios.com](http://www.diccionarios.com)  
[www.oma.org.ar](http://www.oma.org.ar)  
[www.lingolex.com/espan.htm](http://www.lingolex.com/espan.htm)  
[www.clarin.com.ar](http://www.clarin.com.ar)  
[www.intervoz.com.ar](http://www.intervoz.com.ar)  
[www.uba.ar](http://www.uba.ar)  
[www.unlp.edu.ar](http://www.unlp.edu.ar)  
[www.algoritmica.com.ar](http://www.algoritmica.com.ar)

[www.monografias.com](http://www.monografias.com)  
[www.google.com.ar](http://www.google.com.ar)  
[www.yahoo.com.ar](http://www.yahoo.com.ar)

Universidad Nacional de Río Cuarto  
 Facultad de Ciencias Exactas, Físico-Químicas y Naturales  
 Departamento de Computación  
 Ministerio de Cultura y Educación de la Nación.  
 Museos argentinos.  
 Biblioteca Nacional de la República Argentina.  
 National Geographic, en inglés.  
 Real Academia Española.  
 Diccionarios en la red.  
 Olimpiada Matemática Argentina.  
 Aprender-inglés.Net.  
 Diario Clarín.  
 Diario La Voz del Interior.  
 Universidad de Buenos Aires.  
 Universidad Nacional de La Plata.  
 Consulta y materiales para estudiantes de programación.  
 Descarga y publicaciones de trabajos.  
 Buscador Google.  
 Portal Yahoo!.

## Búsqueda de información en la Web

Con la llegada de Internet toda la información de la Red existe en formato digital y miles de computadoras de todo el mundo la almacenan públicamente. Se puede buscar cualquier palabra o concepto y acceder a ella en cuestión de segundos. Lo mejor de todo es que no hace falta ningún programa especial: existen **buscadores de Internet**, en forma de sitios de la WWW a los que se puede acceder con cualquier explorador desde una computadora.



Los buscadores se dividen en dos categorías: los **directorios de búsqueda** y los **motores de búsqueda**. Sin embargo, se puede encontrar una combinación de ellos ya que por ejemplo el buscador *Google*, es un motor de búsqueda que entre sus opciones ofrece el acceso a un directorio.

A continuación, se analizan que son los motores de búsqueda y los directorios y de que manera se busca en ellos, para culminar con sus ventajas y desventajas.

### Motores de búsqueda

En estos sitios la búsqueda se realiza ingresando una o más palabras que tienen que ver con el tema del cual se busca información; éstas son las llamadas "palabras claves". Los resultados de la búsqueda se visualizan en una página con la lista de sitios que contienen los términos solicitados. Esto es lo que se puede llamar una **búsqueda por palabras claves**.

Para realizar este proceso de búsqueda automática, los buscadores utilizan unos programas llamados "robots" o "spiders" que exploran la red siguiendo los enlaces de las páginas (o sea saltando de una página a otra, aprovechando la característica de hipertexto que poseen los sitios web). Cuando llegan a un sitio (por ejemplo el de la UNRC), recogen la mayor parte de las palabras que aparecen en las partes públicas del sitio y en una base de datos las almacenan junto con un índice.

Es así que cuando se hace una consulta a un motor, lo que se está pidiendo es que busque en su índice de sitios y compare la consulta (realizada mediante palabras claves) con los términos recogidos y almacenados en la base de datos del buscador.

Se pueden mencionar como motores de búsqueda a: **Google**, **Altavista**, entre otros.

### Directorios de búsqueda

Estos sitios ofrecen una lista de categorías donde al seleccionar la que es de nuestro interés, se abren nuevas opciones o subcategorías y así en pasos sucesivos, se va refinando la búsqueda hasta llegar a la información requerida, obteniendo al final un listado de páginas incorporadas al directorio que ofrecen la información buscada. Es lo que se puede llamar una *búsqueda por categoría*.

Para ello, en los directorios hay personas que examinan las páginas web y los recursos de todo el mundo, clasificándolos en una categoría

predeterminada, haciendo un pequeño resumen y publicándolos en un catálogo general.

Se puede mencionar como directorios a **Google, Yahoo!, Altavista**, etc.

**Para tener en cuenta:** Como se dijo anteriormente se puede encontrar que un motor de búsqueda también incluya un directorio, lo que también genera una combinación de técnicas de búsqueda. Por ejemplo, Google permite hacer una búsqueda por palabras clave dentro de una categoría del directorio. Es decir, si primero se acota parcialmente el tema de búsqueda (localizando la categoría que se necesita) y se usa palabras claves relacionadas con el tema de búsqueda se obtendrán las páginas que, en esa categoría, hacen referencia a lo buscado.

#### *Ventajas y desventajas de los mecanismos de búsqueda*

En primer lugar, se plantean ventajas y desventajas de un motor de búsqueda:

##### *Ventajas*

- Exhaustividad** en los resultados: Se proporciona una gran cantidad de información.
- Amplitud** de resultados: Cuando no se tienen criterios muy definidos de lo que se busca, este tipo de búsqueda brinda una visión muy amplia.
- Aciertos** en los resultados: Cuando se sabe acotar la búsqueda y se emplean las opciones avanzadas, los resultados pueden ser muy certeros y completos.

##### *Desventajas*

- Falta de precisión** en los resultados: Debido a que la información no ha sido seleccionada por un humano especialista sino por programas informáticos, es probable que la ambigüedad propia de las lenguas naturales dé listas enormes de posibles páginas, dando resultados descontextualizados. La precisión es el punto débil de estos sistemas.

Ahora se plantean ventajas y desventajas de un directorio de búsqueda:

##### *Ventajas*

- Selectividad** de los resultados: Como la búsqueda se acota a los sitios contenidos en una categoría o subcategoría, los sitios devueltos tratan más específicamente lo que buscado.

##### *Desventajas*

- Filtro** en los resultados: De la misma manera en que la selectividad de los resultados es una ventaja, puede verse como desventaja si se

considera que puede haber páginas que son de interés y que los editores del directorio temático hayan decidido excluir.

Los dos tipos de buscadores comparten la desventaja de **desactualización** de los resultados, esto se debe a que cuando se utiliza un buscador no se consulta la información que en ese momento existe en la red, sino la porción de la red capturada en una fecha anterior, puede entonces que el enlace del sitio al que se quiere acceder haya cambiado. Por ello, es que al intentar acceder al resultado de la búsqueda da el mensaje de “Página no encontrada”.

*Contenido extraído de: Manual del Seminario-Taller “Buscar y validar información en Internet” – Programa IRC (Secretaría de Extensión y Desarrollo de la UNRC)*