

# Búsqueda y Gestión de Información

---

Christopher Expósito Izquierdo

Airam Expósito Márquez

Israel López Plata

{cexposit, aexposim, ilopezpl}@ull.edu.es



1. Introducción
2. Navegador Web
3. Internet Profunda
4. Estrategias de Búsqueda

# Introducción

---

A día de hoy en Internet se pueden hallar una gran cantidad de información. Sin embargo, este gran volumen hace necesario que cuando nos ponemos a tratar de encontrar información referente a un tema particular, sea esencial saber emplear herramientas y metodologías adecuadas y destinadas al acceso eficaz.

Se hace imprescindible consecuentemente que conozcamos las herramientas destinadas a la búsqueda de información en la red así como su adaptación al tipo información de interés.

Otro de los aspectos esenciales a la hora de identificar información en la red es aprender la manera en que Internet se organiza y cómo las páginas web están construidas.

Esto permite comprender la manera en que la información se analiza e indexa por medio de motores de búsqueda. A su vez, esto nos ayuda a crear nuevas estrategias de búsqueda efectivas.

El gran volumen de información existente en Internet da lugar a que sea indispensable la formación del profesorado así como del alumnado. Esta formación debe estar relacionada con estrategias y competencias de búsqueda de información en contextos virtuales. En este sentido, se hace necesario formarse en la gestión de la información, cómo poder localizarla, cómo filtrarla, cómo organizarla y cómo usarla de una manera eficaz y eficiente.

## Introducción:

De acuerdo con Monereo y Fuentes (2008), actualmente se presentan los siguientes desafíos a la hora de tratar información:

---

Cantidad de información	Es sencillo perderse en Internet debido a la cantidad de información. Se presenta el desafío de poder realizar búsquedas eficaces y llevar a cabo un análisis profundo y selectivo de la información.
Caducidad de la información	Aquella información que no es actualizada de manera regular termina por desaparecer de la red.
Certidumbre de la información	Uno de los grandes problemas de la información en la red es su fiabilidad. Se emplea el término de <i>infoxicación</i> para hacer mención al exceso de información en Internet.

---

## Introducción:

---

Calidad de la información	Se hace especialmente relevante que el usuario tenga capacidad crítica para determinar la precisión, rigor, originalidad, etc. de la información.
Comprensibilidad de la información	La estructura de la información demanda estrategias específicas que permitan acceder y comprender la información en la red.
Consumibilidad o grado de usabilidad de la información	Se debe tener en cuenta las características y contexto de los usuarios que acceden a contenidos en la red.

---

# Navegador Web

---



# Navegador Web:

---

Un navegador web es un software informático cuya finalidad principal es poder visualizar páginas web. Con esta finalidad en mente, un navegador web lee e interpreta el código fuente de las páginas web a las que accede el usuario y lo transforma en los elementos visuales e interactivos que ve el usuario durante la navegación.

## Navegador Web:

En la actualidad cualquier dispositivo electrónico como smartphones, ordenadores o tabletas incluyen navegadores web de forma gratuita. A su vez, es posible incorporar otros navegadores web sin coste alguno.

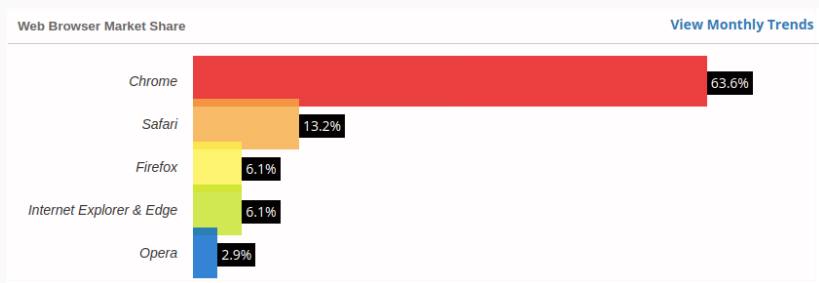
Los diferentes navegadores web no proporcionan las mismas funcionalidades ni rendimiento. Por ello, es imprescindible realizar una reflexión sobre éstos para poder maximizar la experiencia de usuario y no restringir las posibles acciones a llevar a cabo.

Por otro lado, el navegador web es la herramienta empleada habitualmente cuando se navega por la red. Esto hace que sea una herramienta imprescindible para la gestión de nuestra actividad online. Esto inevitablemente hace que sea necesario vehículos de protección y garantía de la seguridad.

A día de hoy, cuatro son los principales navegadores web que existen en el mercado. Éstos son Google Chrome, Internet Explorer, Safari y Firefox.

Lamentablemente, no es posible conocer las cuotas de mercado precisas, pero sí se realizan estimaciones que persiguen medir el porcentaje de visitas realizadas a través de cada navegador a una serie de páginas web preestablecidas. Hay que tener en cuenta que las estimaciones realizadas fluctúan notablemente en base a la región considerada y el período de tiempo bajo análisis.

La siguiente imagen muestra las cuotas de mercado según W3Counter:



<https://www.w3counter.com/globalstats.php>

Es importante comprender que cuando realizamos una búsqueda a través de un navegador web realmente se realiza la búsqueda sobre el índice de páginas web creado por el navegador web y no en todo Internet.

La indexación de la web es llevada a cabo mediante programas software denominados **arañas**. Estos programas están destinados a la visita y carga de páginas web, lectura de contenido para ubicar enlaces a otras páginas webs, etc.

La labor de las arañas es llevada a cabo de manera continua para poder localizar el mayor número de páginas web posibles. Al mismo tiempo, el contenido de las páginas web es almacenado en servidores ubicados en centros de datos que se encuentran dispersos por diferentes partes del mundo.

En el caso de Google, las páginas web que son indexadas por la empresa son clasificadas en base a su relevancia.

La relevancia de una página web viene establecida en base a un algoritmo denominado Page Rank. Éste analiza las relaciones existentes entre páginas web, las cuales vienen determinadas por el número y calidad de los enlaces entre páginas.

Cuando se realiza una búsqueda a través de Google, éste revisa las páginas web indexadas donde se incluye los términos de búsqueda establecidos. A su vez emplea Page Rank con el objetivo de establecer qué páginas web poder proporcionar al usuario.

# Internet Profunda

---

Habitualmente se tiene la concepción de que las páginas accesibles a través de un buscador como Google es todo lo existente en Internet.

Sin embargo, actualmente los navegadores web no son capaces de indexar todo el contenido existente en Internet. En concreto, solo son capaces de indexar una pequeña parte conocida como **Internet superficial**.

Existe un amplio abanico de contenidos en Internet que no está indexado, que se conoce como **Internet profunda** o **deep web** y es mucho más grande que la anterior.

En la Internet profunda se pueden encontrar páginas web accesibles mediante contraseña o bases de datos que solicitan consultas para el acceso a su contenido. Este tipo de sitios web no son indexables mediante el uso de arañas, por lo que los buscadores no los recogen.



Se hace referencia a la Internet profunda como aquel repertorio de contenidos en la red que se encuentran por debajo de la Internet superficial.

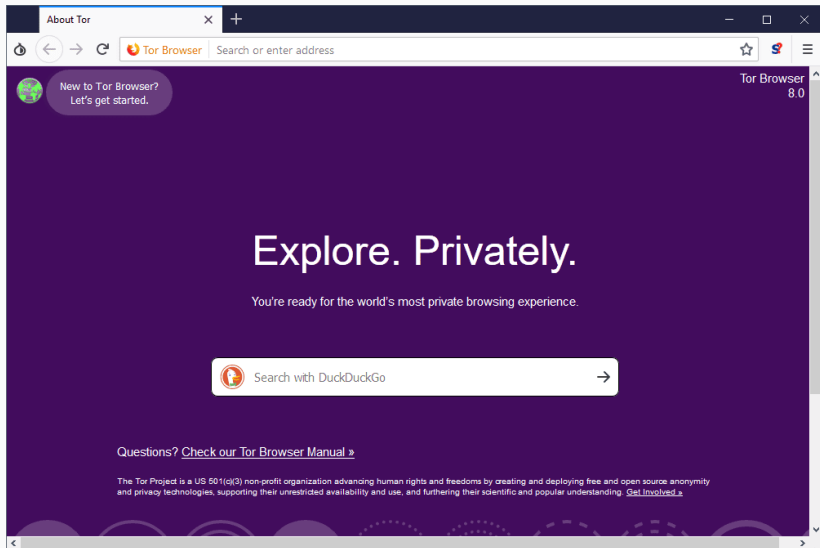
Una de las características fundamentales asociadas a la Internet profunda es que no se cuenta con ninguna entidad que regule cómo se gestiona o accede a los datos existentes. En su lugar, la información existente se intercambia de forma libre entre los usuarios de la red, habitualmente mediante conexiones **peer to peer** (P2P).

Debido a que las páginas incluidas en la Internet profunda no pueden ser indexadas, habitualmente se ha vinculado su contenido con **actividades no lícitas** y con contenido no permitido.

Esto hace que las páginas de la Internet profunda queden en un limbo con respecto a los buscadores web más populares y por tanto se requiera de otro tipo de herramientas para su acceso.

Sin embargo, en cierta medida, la falta de controles adecuados así como el **anonimato** han posibilitado la aparición de contenidos ilegales.

# Internet Profunda:



<https://www.torproject.org/>

# **Estrategias de Búsqueda**

---

## Estrategias de Búsqueda:

Cuando se emplean **marcadores** de búsqueda en Internet, es siempre recomendable crear carpetas y subcarpetas que permitan organizar por área temática el contenido accedido.

Es también posible importar y exportar marcadores para que el contenido esté accesible desde cualquier dispositivo.

Una buena estrategia de búsqueda además debe basarse en el **conocimiento del navegador empleado**. En la mayoría de los casos se pueden crear estrategias muy elaboradas y con resultados eficaces.

# Estrategias de Búsqueda:

## Google

Google

Advanced Image Search

Find images with...

all of these words:	<input type="text"/>	To do this in the search box:	Type the important words: winter heartfruit
this exact word or phrase:	<input type="text"/>		Put exact words in quotes: "fruit flower"
any of these words:	<input type="text"/>		Type OR between all the words you want: fruit OR seeds OR grasses
none of these words:	<input type="text"/>		Put a minus sign just before words that you don't want: -windows

Then narrow your results by...

image size:	<input type="text" value="any size"/>	Find images in any size you need.
aspect ratio:	<input type="text" value="any aspect ratio"/>	Specify the shape of images.
colours in the image:	<input checked="" type="radio"/> any colour <input type="radio"/> full colour <input type="radio"/> black & white <input type="radio"/> transparent <input type="radio"/> this colour <input checked="" type="checkbox"/>	Find images in your preferred colours.
type of image:	<input type="text" value="any type"/>	Limit the kind of images that you find.
region:	<input type="text" value="any region"/>	Find images published in a particular region.
site or domain:	<input type="text"/>	Search one site (like <code>news.ny.gov</code> ) or limit your results to a domain like <code>edu</code> , <code>org</code> or <code>gov</code> .
SafeSearch:	<input type="text" value="Show most relevant results"/>	Set <b>SafeSearch</b> whether to filter sexually explicit content.
file type:	<input type="text" value="any format"/>	Find images in the format that you prefer.
usage rights:	<input type="text" value="not filtered by license"/>	Find images that you are free to use.

[https://support.google.com/websearch/answer/134479?hl=es&ref\\_topic=3081620](https://support.google.com/websearch/answer/134479?hl=es&ref_topic=3081620)

## Firefox



The screenshot shows the Mozilla Support website in Spanish. The main heading is "Consejos de Búsqueda" (Search Tips). The page content includes:

- A paragraph explaining that search suggestions are provided by search engines like Google, Yahoo, etc., and Firefox only recovers suggestions from those engines.
- A bullet point: "Para más información sobre la barra de búsqueda, puedes consultar el artículo [La barra de búsqueda: elige fácilmente tus buscadores favoritos.](#)"
- A "Tabla de contenidos" (Table of contents) section with the following links:
  - [Habilitar o deshabilitar las sugerencias de búsqueda](#)
  - [A través de la barra de búsqueda](#)
  - [A través del Administrador de motores de búsqueda](#)
  - [Utilizar sugerencias de búsqueda](#)
  - [¿Cómo diferenciar entre las sugerencias de búsqueda y resultados del historial de búsqueda?](#)
  - [¿Por qué las sugerencias de búsqueda pueden ser inapropiadas?](#)

[https:](https://support.mozilla.org/es/kb/Consejos%20de%20B%C3%BAsqueda)

[//support.mozilla.org/es/kb/Consejos%20de%20B%C3%BAsqueda](https://support.mozilla.org/es/kb/Consejos%20de%20B%C3%BAsqueda)

## Safari

---



[http://support.apple.com/kb/HT5364?viewlocale=es\\_ES](http://support.apple.com/kb/HT5364?viewlocale=es_ES)



# Estrategias de Búsqueda:

## Bing

