

## Investigar en la Web Searching

*Por Sandra Crucianelli*  
*@spcrucianelli*

Internet acaparó mi atención desde el mismo momento en que supe de su existencia y está absolutamente ligada a mi pasión por el descubrimiento.

Encontrar información desconocida y poder utilizarla como soporte para la documentación de un reportaje, debo confesar, es una tarea que ha marcado mis 30 años de trabajo como periodista.

La creciente potencialidad de la Web y la aparición de las herramientas digitales funcionales al periodismo hicieron posible demostrar, empíricamente, que sin demasiados recursos económicos y sólo a fuerza de voluntad, no solamente se pueden hacer descubrimientos aislados, sino también, crear un medio digital e imponer esa impronta: la de contener **información no conocida**, ni difundida, sumergida en las entrañas de la Web, casi como esperando ser expuesta bajo la luz de la opinión pública.

Muchos me preguntan si el *Searching* es un método. No lo creo. Si fuera un método, podría ser descrito y replicado de idéntica manera por otros, con similares resultados.

Más bien creo que el *Searching* es una *actitud*, marcada profundamente por el uso del entorno digital, pero dominada por la resistencia a creer que en un día no encontraremos nada nuevo.

### Internet como vía de recuperación documental para Texto

Para todo documento hallado en la Web, sea cual fuera su naturaleza, el documento debe cumplir con 4 principios básicos:

- 1 Principio **verificativo** o comprobatorio
- 2 Principio **explicativo** (relacionado positiva o negativamente con la hipótesis)
- 3 Principio **editorial** o de crédito

## 4 Principio de **perdurabilidad**

### **Buscar texto en Google**

#### **Métodos de Filtrado mediante Operadores**

En general, los periodistas de investigación usamos buscadores y bases de datos como vía de recuperación documental en la Web.

**La mejor búsqueda no es la que ofrece más resultados.** Más bien, cuando eso sucede probablemente el proceso no fue lo suficientemente preciso y sería imposible analizar gran cantidad de enlaces. Por el contrario, una buena búsqueda es aquella que devuelve no más de 50 a 100 resultados y en lo posible, menos de 50 si el proceso se ajustó lo suficiente.

Hay que tener cuidado con la restricción porque si ese ajuste se agudiza demasiado, el buscador podría no recuperar digitalmente lo que se busca. Ya veremos cómo hacerlo.

En todos los casos, estas búsquedas responden a la Teoría de Conjuntos y usan operadores (los más comunes son los llamados “*booleanos*”), que combinan determinados signos y palabras para búsquedas más eficientes.

Los denominados “*Métodos de Filtrado*” no son más que técnicas de búsqueda mediante las cuales se usan operadores.

Los operadores lógicos u *operadores booleanos* surgieron a partir del Álgebra del matemático inglés del siglo XIX George Boole y permiten combinar los términos de búsqueda de acuerdo con nuestras necesidades. La combinación de palabras o frases, permite obtener universos más acotados.

En el uso de la Web, los operadores booleanos AND (+) y NOT (-) son operadores de restricción (permiten obtener un menor número de resultados), mientras que OR es un operador de expansión (permite obtener un mayor número de resultados).

Estos elementos facilitan la búsqueda vinculando términos de búsqueda y definiendo la relación entre ellos.

- 1 Operador AND (su traducción del inglés = y) para localizar registros que contengan todos los términos de búsqueda especificados. **La misma función (y más usada de todas), cumple el operador SUMA (+), es decir el signo +**
- 2 Operador OR localiza registros que contengan cualquiera o todos los términos especificados.
- 3 Operador NOT, representado por el signo (-), logra localizar registros que contengan uno o varios términos, pero excluyendo otros.
- 4 Operador COMILLAS, devolverá las páginas que contengan la **exacta frase** que se escribió y se incluyó dentro de ellas.

### ¿Qué es la *Búsqueda Combinatoria Especial*?

La denominada *búsqueda combinatoria especial* permite combinar un dato y una frase o varias frases, relacionándolas con una o varias sumas.

El reportero puede intentar, mediante el uso de sinónimos, variar sus opciones de búsqueda para optimizar sus resultados. Una estrategia muy usada es realizar un filtrado mediante fechas, más específicamente años.

Por ejemplo, en el cuadro de búsqueda se puede escribir:

**“corrupción en América Latina” + 2020**

El resultado será diferente si se escribe:

**“Corrupción en América Latina” + “octubre de 2020”**

Como se podrá ver en la práctica, el resultado en número de enlaces disminuye drásticamente.

Otros ejemplos de combinatoria especial:

**“corrupción en América Latina” + presidente + 2019 + 2021**

**“corrupción en Ecuador” + “corrupción en Venezuela”**

## “análisis presupuestario” + “auditoría social”

La combinación de comillas y el signo + (o el operador AND), es una búsqueda bastante específica dentro de las que podemos hacer en la ventana principal. Pero por cierto, no la mejor de todas.

Cuando se busca información específica, sobre la que se conocen algunos datos aislados, la mejor opción es la herramienta que se conoce como **Búsqueda Avanzada**.

Para poder usar estas Opciones Avanzadas, identifique primero cada elemento de la Página Principal de Google y luego ajuste sus preferencias.

### **Ajuste de Preferencias:**

Arriba, a la derecha del cuadro de búsqueda, vaya a **Preferencias** y configure su buscador de modo que le devuelva 100 resultados; este sencillo truco le evitará 9 de cada 10 clics y ahorrará tiempo. Dentro de esta sección decida el idioma en el que realizará la búsqueda. Es mejor tener todas las opciones habilitadas, porque para la búsqueda de nombres propios o compañías, es probable que haya documentos en otros países en otros idiomas, que podría no localizar si solamente escoge la opción “español”.

Por otra parte, al usar palabras en español, Google excluirá los resultados en otro idioma o usted podrá ajustar la búsqueda a la opción español desde la ventana principal, sin que eso afecte la selección hecha en la sección **Preferencias**.

### **Aprenda a Usar la Búsqueda Avanzada de Google**

Dominar las búsquedas utilizando las opciones avanzadas es lo que marca la diferencia entre un resultado malo o mediocre, de uno bueno o excelente.

Ubique en la página principal de Google, “**Búsqueda Avanzada**” o acceda mediante su acceso directo: [http://www.google.com/advanced\\_search](http://www.google.com/advanced_search)

Tal como se ve en la captura de pantalla, en la parte superior del formulario hay diferentes campos que representan a *los operadores booleanos*

### Búsqueda avanzada

| Mostrar páginas que contengan... |                                              | Para hacer esto en el cuadro de búsqueda                                                                     |
|----------------------------------|----------------------------------------------|--------------------------------------------------------------------------------------------------------------|
| todas estas palabras:            | <input type="text"/>                         | Ingresar las palabras importantes: Terrier ratonero tricolor.                                                |
| esta palabra o frase exactas:    | <input type="text"/>                         | Ingresar las palabras exactas entre comillas: "Terrier ratonero".                                            |
| cualquiera de estas palabras:    | <input type="text"/>                         | Ingresar OR entre las palabras que desees:<br>En miniatura OR estándar.                                      |
| ninguna de estas palabras:       | <input type="text"/>                         | Ingresar un signo menos justo delante de las palabras que no desees que aparezcan: -Roedor, -"Jack Russell". |
| números del:                     | <input type="text"/> al <input type="text"/> | Ingresar dos puntos entre los números y agrega una unidad de medida:<br>10..35 lb, \$300..\$500, 2010..2011. |

---

Luego restringe tus resultados por...

|                                  |                                                                        |                                                                                                                                 |
|----------------------------------|------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| idioma:                          | <input type="text" value="cualquier idioma"/>                          | Busca páginas en el idioma que seleccionas.                                                                                     |
| región:                          | <input type="text" value="cualquier región"/>                          | Busca páginas publicadas en una región determinada.                                                                             |
| última actualización:            | <input type="text" value="en cualquier momento"/>                      | Busca páginas actualizadas en el transcurso del período que especificas.                                                        |
| sitio o dominio:                 | <input type="text"/>                                                   | Realiza búsquedas en un sitio (como wikipedia.org) o restringe los resultados a un dominio como .edu, .org o .gov.              |
| términos que aparecen:           | <input type="text" value="En cualquier parte de la página"/>           | Busca términos en toda la página, en su título o en su dirección web, o vínculos que te dirijan a la página que estás buscando. |
| <a href="#">SafeSearch</a>       | <input type="text" value="Ocultar resultados explícitos"/>             | Indica a <a href="#">SafeSearch</a> si quieres que filtre contenido sexualmente explícito.                                      |
| tipo de archivo:                 | <input type="text" value="Cualquier formato"/>                         | Busca páginas del formato que prefieras.                                                                                        |
| <a href="#">derechos de uso:</a> | <input type="text" value="Páginas cuyo uso no requiera de licencias"/> | Busca páginas que puedas usar libremente.                                                                                       |

Búsqueda avanzada

**Primer Campo:** Allí se deben colocar palabras sin operadores, ya que este campo representa al operador AND, o lo que es igual, el signo +. Por lo tanto, **todas las palabras** que incluya en este campo estarán incluidas en las páginas que devuelva el buscador. Debe quedar en claro que en este cuadro de búsqueda se colocan palabras sueltas. Ejemplo: corrupción Ecuador. Vea el ejemplo en la captura de pantalla:

## Búsqueda avanzada

Mostrar páginas que  
contengan...

Pe

todas estas palabras:

corrupción Ecuador

esta palabra o frase exactas:

cualquiera de estas palabras:

ninguna de estas palabras:

números del:

al

**Segundo Campo:** Representa **la frase exacta**, por lo tanto, ***no debe colocar allí comillas***, ya que este campo funciona como el operador comillas. Ejemplo: corrupción en Ecuador (vea el ejemplo en la captura de pantalla de la imagen inferior). Utilizar este cuadro actúa como un modo de restringir la búsqueda; es una de las opciones de filtrado más seguras.

## Búsqueda avanzada

Mostrar páginas que  
contengan...

Pe

todas estas palabras:

esta palabra o frase exactas:

corrupción en Ecuador

cualquiera de estas palabras:

ninguna de estas palabras:

números del:

al

**Tercer Campo:** Si no tiene muchos datos relacionados con la información que busca, puede escribir en este cuadro palabras sueltas o sinónimos, ya que opera como un selector de **algunas palabras** y no todas.

### Búsqueda avanzada

---

Mostrar páginas que contengan...

todas estas palabras:

esta palabra o frase exactas:

cualquiera de estas palabras:

ninguna de estas palabras:

números del:  al

**Cuarto Campo:** Generalmente páginas comerciales interfieren en la meta de obtener pocos pero buenos resultados. Por lo tanto, este cuadro permite excluir términos que usted no desea estén incluidos en sus resultados.

**Inmediatamente después de estas cuatro ventanas, dentro de la búsqueda avanzada, encontrará otros cuadros (*o campos*), de selección que permiten un filtrado más preciso. Veamos:**

**Campo Idioma:** Puede ser que requiera información que se encuentre en otro país, por ejemplo, Brasil, por lo que será necesario ajustar el idioma al portugués.

**Campo Región:** Permite remitir la búsqueda a un determinado país. Por ejemplo, suponiendo que usted intuye que cierto documento fue generado en la embajada de España en su país, probablemente en este país haya un correlato digital del mismo. Vea el ejemplo en la captura de pantalla.

**Campo Tipo de Archivo (formato):** Es quizá, uno de los campos de selección más importantes. Los documentos relevantes, generalmente información gubernamental, suelen estar en PDF, acrónimo del inglés *Portable Document Format*, desarrollado por la empresa Adobe Systems. Pliegos de licitación, fallos judiciales y otros documentos públicos, frecuentemente se suben a la Web en PDF. Seleccione esta opción como una manera de restringir entre los resultados de la búsqueda aquellos sitios comerciales o de fuentes poco confiables que no suelen utilizar este formato. Si busca estadísticas, es más probable que el documento se encuentre en Excel, aunque no necesariamente, ya que muchos organismos y gobiernos, con la excusa de proteger sus documentos, convierten los archivos Excel a formato fotográfico, con lo cual se imposibilita el traslado de la información a una Hoja de Cálculo. Gobiernos que utilicen formatos fotográficos para almacenar documentos contribuyen a que sus sitios operen de manera más lenta y dificultan el proceso de acceso a la información pública que debería estar garantizado por una ley. También puede elegir otras opciones, como el formato PPT (*PowerPoint*), si está buscando información académica, presentaciones en congresos, etc.

**Campo Última actualización o Fecha:** Resulta importante que ajuste la fecha, ya que permite acotar el plazo de manera bastante precisa, aunque no exacta. Si decide buscar dejando la opción “en cualquier momento”, probablemente los resultados serán muchos. Vaya ajustando este parámetro gradualmente. Suponga que necesita un documento del 2009. Comience seleccionando la opción “*durante la última semana*”, y si no encontró lo que buscaba, vaya ampliando la selección, a “*en el último mes*” hasta finalmente “*en el último año*”. La opción “*en las últimas 24 horas*” es bastante restrictiva cuando previamente se seleccionó un formato, por lo que podría excluir resultados valiosos. En la captura de pantalla anterior, se añadió la restricción utilizando este campo, solicitando a Google que sólo devuelva los resultados en PDF, indexados durante el último mes.

**Campo sitio o dominios:** Hace posible ubicar un documento, dato o información dentro de un sitio. Es como si Google “*se metiera*” dentro de un sitio Web. No deberá colocar las 3 www.

Ejemplo:



Luego restringe tus resultados por...

|                        |                                                                        |        |
|------------------------|------------------------------------------------------------------------|--------|
| idioma:                | <input type="text" value="cualquier idioma"/>                          | E      |
| región:                | <input type="text" value="cualquier región"/>                          | E      |
| última actualización:  | <input type="text" value="el último año"/>                             | E      |
| sitio o dominio:       | <input type="text" value="ministeriodegobierno.gob.ec"/>               | F<br>r |
| términos que aparecen: | <input type="text" value="En cualquier parte de la página"/>           | E<br>V |
| SafeSearch             | <input type="text" value="Ocultar resultados explícitos"/>             | I      |
| tipo de archivo:       | <input type="text" value="Cualquier formato"/>                         | E      |
| derechos de uso::      | <input type="text" value="Páginas cuyo uso no requiera de licencias"/> | E      |

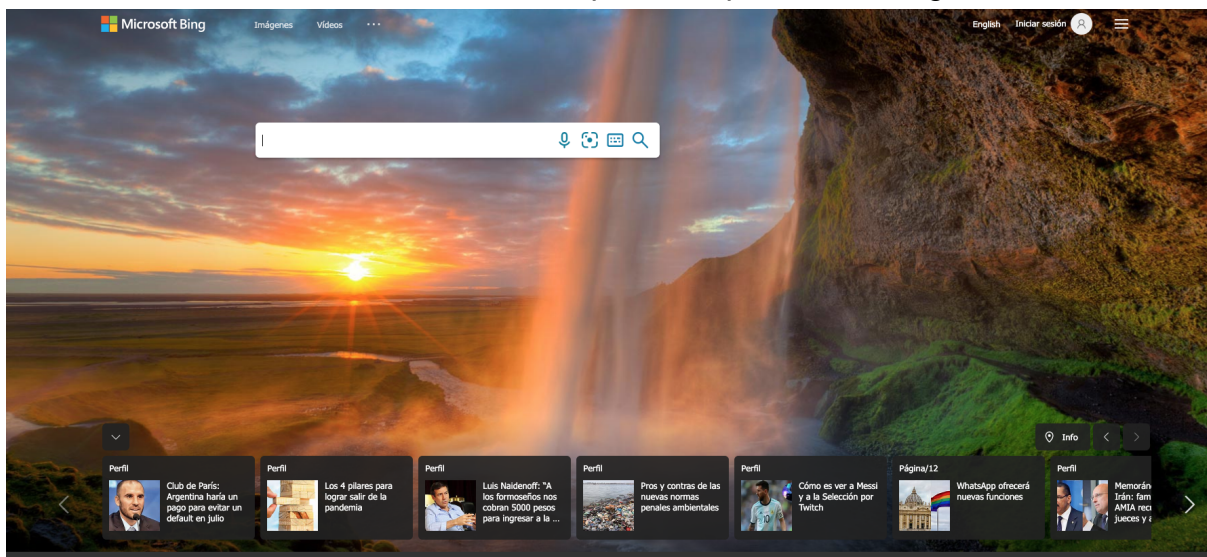
## Bing

Microsoft Bing es un buscador web de Microsoft.

<https://www.bing.com/?setlang=es>

Bing es un excelente buscador, que ofrece alternativas muy interesante a la hora de completar las búsquedas Web

Permite buscar con la voz y buscar con una imagen, entre múltiples utilidades que lo convierten en una excelente herramienta para el reportero investigador.



Permite la búsqueda por palabras pero también por frases exactas (mediante el uso de comillas) y filtrar por fecha, incluso personalizando el intervalo.

Para acceder a la **búsqueda avanzada de Bing**, colocamos el cursor al final de las palabras claves y hacemos clic en el vínculo Búsqueda avanzada que aparece.

Puedes sumar palabras con el operador +

Ejemplo: Corrupción + Latinoamérica

Puedes excluir palabras o frases con el operador -

Ejemplo: Argentina - "Buenos Aires"

Puedes decidir que las palabras solo aparezcan en el título con el operador: intitle

Ejemplo: intitle:Muro de Berlín

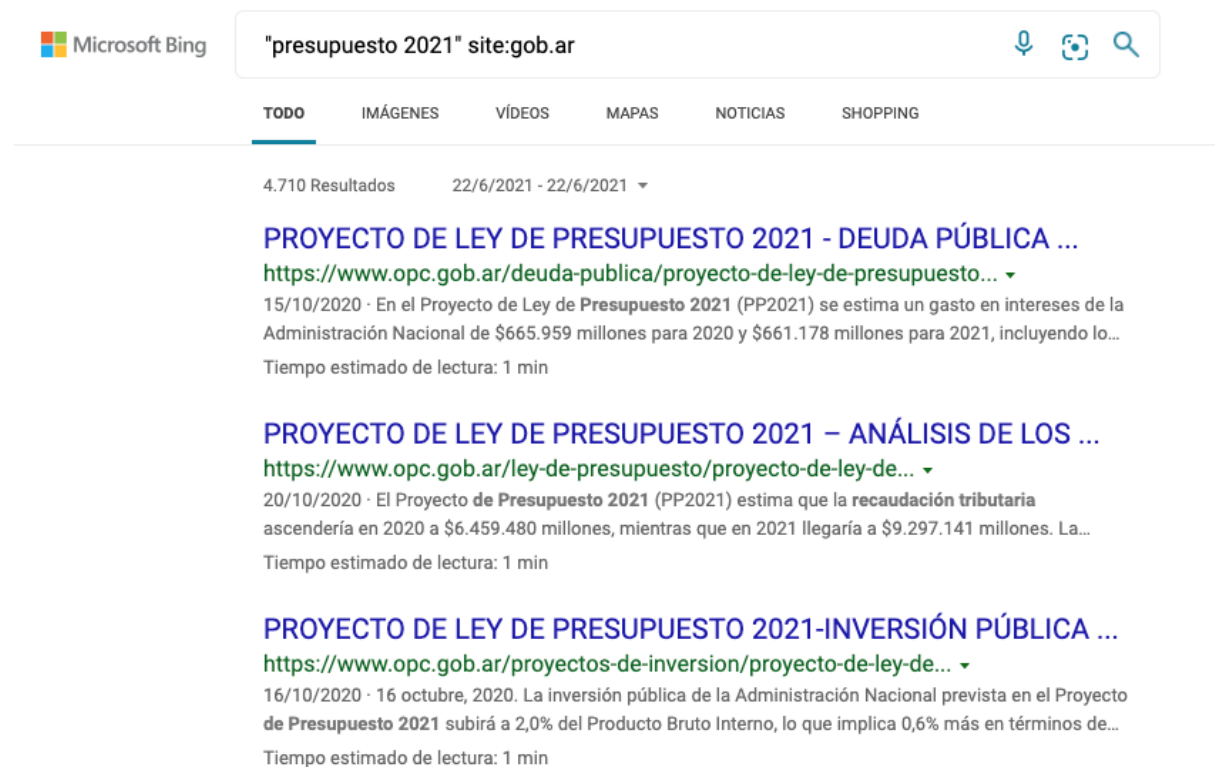
Puedes buscar dentro de una página con el operador site

Ejemplo: música site:[www.wikipedia.org](http://www.wikipedia.org)

Y así podrás ir refinando tus búsquedas.

Ejemplo:

[https://www.bing.com/search?q=%22presupuesto+2021%22+site%3agob.ar&filters=ex1%3a%22ez5\\_18435\\_18800%22&qs=n&sp=-1&pq=%22presupuesto+2021%22+site%3agob.a&sc=0-29&cvid=B553A3B63BEA43D7AA75172EE5AEE3ED](https://www.bing.com/search?q=%22presupuesto+2021%22+site%3agob.ar&filters=ex1%3a%22ez5_18435_18800%22&qs=n&sp=-1&pq=%22presupuesto+2021%22+site%3agob.a&sc=0-29&cvid=B553A3B63BEA43D7AA75172EE5AEE3ED)



Microsoft Bing

"presupuesto 2021" site:gob.ar

TODO IMÁGENES VÍDEOS MAPAS NOTICIAS SHOPPING

4.710 Resultados 22/6/2021 - 22/6/2021 ▼

**PROYECTO DE LEY DE PRESUPUESTO 2021 - DEUDA PÚBLICA ...**  
<https://www.opc.gob.ar/deuda-publica/proyecto-de-ley-de-presupuesto...> ▼  
15/10/2020 · En el Proyecto de Ley de Presupuesto 2021 (PP2021) se estima un gasto en intereses de la Administración Nacional de \$665.959 millones para 2020 y \$661.178 millones para 2021, incluyendo lo...  
Tiempo estimado de lectura: 1 min

**PROYECTO DE LEY DE PRESUPUESTO 2021 – ANÁLISIS DE LOS ...**  
<https://www.opc.gob.ar/ley-de-presupuesto/proyecto-de-ley-de...> ▼  
20/10/2020 · El Proyecto de Presupuesto 2021 (PP2021) estima que la recaudación tributaria ascendería en 2020 a \$6.459.480 millones, mientras que en 2021 llegaría a \$9.297.141 millones. La...  
Tiempo estimado de lectura: 1 min

**PROYECTO DE LEY DE PRESUPUESTO 2021-INVERSIÓN PÚBLICA ...**  
<https://www.opc.gob.ar/proyectos-de-inversion/proyecto-de-ley-de...> ▼  
16/10/2020 · 16 octubre, 2020. La inversión pública de la Administración Nacional prevista en el Proyecto de Presupuesto 2021 subirá a 2,0% del Producto Bruto Interno, lo que implica 0,6% más en términos de...  
Tiempo estimado de lectura: 1 min

## Recuperación Documental

A veces, los enlaces no abren el documento y nos reportan *error 404*, lo que nos indica que el link no funciona o literalmente, “se rompió”. Por ejemplo, para el caso de este resultado:

En general, los enlaces que devuelven las búsquedas proporcionan bastante información, de modo tal que usted puede ir leyendo la página de resultados (con 100 por página, como explicamos antes, para ahorrar tiempo entre página y página), de modo tal que con el tiempo desarrollará, seguramente, la capacidad para “*entender*” intuitivamente qué hay detrás de cada enlace.

No obstante, si el sitio Web desapareció o fue modificado; o para el caso de que usted necesite seguir el registro de un sitio a lo largo del tiempo, la herramienta más poderosa es la “*Way Back Machine*” (*Máquina del Tiempo*), alojada en **Archive:** [www.archive.org](http://www.archive.org)

Archive es una organización sin fines de lucro, fundada en 1996 con la misión de crear una biblioteca en Internet. El sitio es de consulta obligada para investigadores, periodistas, historiadores y académicos, entre otros, ya que contiene varias colecciones históricas en formato digital compuestas por textos, audio, imágenes y *software*, así como páginas Web inactivas. El sitio dispone de bases de datos dentro de las que se destaca la colección de filmes y videos, acaso la más completa que existe en su tipo.

## **Buscadores de Información Académica**

Dentro de las utilidades de Google, si observa con atención la imagen anterior, verá una solapa (*extremo superior derecho*), que dice **Más**, con una flecha hacia abajo, que se despliega con otras opciones y al final se lee “**todavía más**”. Haga una prueba y acceda al panel central de Herramientas de Google haciendo clic en “**todavía más**”. Si observa a la izquierda, arriba, la primera que se ve es la denominada “**Académico**”, cuyo enlace directo es <http://scholar.google.com>,

Se trata del buscador de información académica de Google, un recurso aliado del periodista, en especial para quienes cubren noticias de fondo o hacen periodismo de investigación.

## ¿Dónde encontrar fotografías?

### Google Imágenes

<http://images.google.com/>

Es el buscador de fotografías de Google. Permite seleccionar el rastreo al tamaño de la fotografía o a su categoría: rostros, bocetos, etc.

### Flickr

[www.flickr.com](http://www.flickr.com)

Es una plataforma para alojar imágenes fotográficas (también videos) y funciona de manera similar a Youtube, es decir como una red social donde los usuarios pueden subir contenidos y compartirlos con todo el mundo o solamente con los contactos que deseen.

### Fotografías de libre uso

La siguiente es una lista de repositorios con fotografías que pueden ser usadas con permiso del autor (libre uso)

<https://unsplash.com/>

<https://yandex.com/>

<https://www.istockphoto.com/es>

<https://www.pexels.com/es-es/>

## Validación de Datos

Sin duda, una de las principales metas del reportero frente a la Web es conseguir la mejor calidad de información en el menor tiempo posible. Pero que un dato o documento aparezca en Internet no implica que el mismo sea fidedigno. Así como el reportero debe distinguir, muy especialmente en el campo del periodismo de investigación, entre los distintos grados de confiabilidad de sus fuentes, este proceso se agudiza aún más en Internet. Las sugerencias, para analizar la confiabilidad de los contenidos de un sitio Web, podrían resumirse en:

- 1) Analice la confiabilidad del recurso, utilice un buscador para recuperar las menciones de la URL, en otras fuentes.

- 2) Encuentre información sobre el responsable del sitio. Una sección denominada “*¿Quiénes somos?*” ó “*Acerca de Nosotros*”, debería visualizarse con facilidad. Dentro de la misma deben datos básicos tales como nombre y dirección de correo electrónico. El curriculum de los responsables en línea es un buen indicador de calidad.
- 3) Observe la fecha de actualización. Si no es diaria al menos debería ser semanal.
- 4) Analice si el sitio contiene errores de ortografía y si la información está presentada de manera clara y precisa.
- 5) Busque información dentro del sitio sobre la fuente de financiamiento del mismo.

## Consejos Finales

Los motores de búsqueda pueden (o no) contener opciones para optimizar el rastreo, tales como Configuración, Preferencias y Búsqueda avanzada.

Usar un buscador sin conocer sus características es algo así como conducir un auto que nunca se manejó.

Recuerde que, salvo excepciones, **la búsqueda avanzada o la restricción por dominios o tipo de archivos siempre será la mejor opción**, ya que permite recuperar información de manera más eficiente.