

# Módulo 2. Algoritmos de búsqueda de Google

## Unidad 2.1 Algoritmos de búsqueda de Google

Es casi imposible hablar de SEO sin mencionar a Google; y casi tan difícil es también hacerlo sin referirse a los algoritmos que dan vida a este motor de búsqueda. Los algoritmos constituyen el criterio determinante a la hora de definir y asumir las optimizaciones a realizar en cada etapa de una estrategia de posicionamiento.

Para comprender cómo funcionan los distintos algoritmos de búsqueda de Google será de utilidad primero definirlos. Un algoritmo de búsqueda es una serie de pasos y secuencias lógicas que permiten mejorar la toma de decisiones en función de determinar si el elemento que se desea analizar cumple con el criterio programado o no. A partir de esta evaluación, el mismo algoritmo asignará un puntaje, conocido como *page-rank*. Este puntaje es el resultado de la evaluación y ponderación de una gran variedad de factores distintos en una ecuación compleja, que se simplifica al entender lo que se está buscando.

En este módulo se analizarán las versiones más importantes de los últimos años del algoritmo de Google y qué implicancias tienen para el trabajo práctico a realizar en una estrategia de SEO.

### 2.1.1 Qué es un algoritmo y por qué cambian

Para poder comprender acabadamente el impacto de cada actualización individual del algoritmo de búsqueda, es necesario entender qué es un algoritmo de motor de búsqueda: ***La palabra "algoritmo" se refiere al procedimiento paso a paso basado en la lógica para resolver un problema particular.***

En el caso de un motor de búsqueda, el problema es "cómo encontrar las páginas web más relevantes para este conjunto particular de palabras clave (o términos de búsqueda)". El algoritmo es cómo Google encuentra, clasifica y devuelve los resultados relevantes.

Google es el motor de búsqueda número uno en la web debido a su enfoque en ofrecer los mejores resultados para cada búsqueda. Como dijo Ben Gomes, vicepresidente de ingeniería de Google, "nuestro objetivo es brindarle la respuesta exacta que busca más rápido" (Carrero, s. f., párr. 20).

Desde el principio, en un intento por mejorar su capacidad para devolver esas respuestas correctas rápidamente, Google comenzó a actualizar su algoritmo de búsqueda. Este fue cambiando, a veces drásticamente, la forma en que brindaba resultados relevantes a los usuarios de búsqueda.

Como resultado de estos cambios en el algoritmo, muchos sitios fueron penalizados con clasificaciones más bajas mientras que otros experimentaron un aumento en el tráfico orgánico y clasificaciones mejoradas.

### 2.1.2 Historia de los algoritmos de Google

Google ha utilizado algoritmos de evaluación de resultados desde sus orígenes a finales de los años '90. Estos algoritmos recopilaban una serie de criterios programados para evaluar de forma automatizada grandes volúmenes de información y a partir de allí, brindar resultados a las búsquedas hechas por los usuarios.

Google inició en 1997 con un buscador sencillo pero poderoso, el cual realizaba un *match* entre las palabras colocadas por los usuarios en una barra de búsqueda y el contenido en las distintas páginas web agregadas en el índice; dicho índice estaba creado por la compañía en base al descubrimiento y rastreo de páginas. El mecanismo de rastreo, desarrollado por Google, consistía en tener robots que leían y rastreaban sitios web constantemente para descubrir contenidos y ampliar la base de opciones.

En los años posteriores se fueron agregando algunos elementos adicionales como tipos de resultados de búsqueda distintos, como lo fue la incorporación de imágenes en 2001, de noticias en 2002 y la comprensión de sinónimos lanzado poco antes de 2003.

En el 2003 entró en vigor Florida que consistió en modificaciones en el algoritmo de clasificación; se implementaron mecanismos de filtrado para palabras clave generando un cambio masivo en los resultados de búsqueda, por lo que muchos sitios cayeron del ranking de Google. En el 2003 había tantos problemas de spam, especialmente en el link building y manipulaciones con técnicas Blackhat, que Google cambió su algoritmo PageRank y creó Florida para mejorar los resultados de búsqueda. Respecto a este algoritmo, Google nunca dio muchas especificaciones sobre los factores que se evaluaban para “pulir” los resultados, pero se puede sostener que fue el pionero de las actualizaciones que vendrían después.

Para el año 2005 ya se habían agregado resultados que incluían libros, información de la bolsa de valores y se comenzaba a experimentar con búsquedas geolocalizadas. A inicios de 2005 se hace el lanzamiento del *auto-complete* de Google, el cual consistía en completar automáticamente la frase que el usuario estaba buscando en función de lo que detectaba que podría ser su intención de búsqueda. Esto permitió estandarizar las búsquedas realizadas y agilizar la velocidad y efectividad con que se podía entregar resultados de calidad.

Para 2008 ya se habían lanzado las búsquedas de vuelos comerciales, deportes, el clima y contenido multimedia o información de películas; aunque el lanzamiento más importante de ese año fue, sin duda, la búsqueda universal que estandarizaba el criterio con el que se generaba la *relevancia* en los resultados de búsqueda, al mismo tiempo que permitía separar la búsqueda general de indexación, introduciendo la posibilidad de generar modificaciones en cómo se presentaban los resultados (o la posición en las que se presentaban) según la personalización de lo que podía ser relevante para cada usuario de forma específica.

En 2009 se comienzan a realizar pruebas con las búsquedas por medio de comandos de voz (es decir, mediante preguntas al buscador hablando normalmente), aunque estas no despegaron del todo debido a la necesidad de mejorar la interpretación de cada idioma y la pronunciación de cada persona.

En conclusión, Google ha utilizado una serie de algoritmos distintos a través de su historia; y, en la actualidad, continúa mejorándolo y agregando otros complementarios, con el fin de optimizar las experiencias y las respuestas de búsquedas que se realizan alrededor del mundo.

A continuación, explicaremos algunos de los principales cambios e incorporaciones existentes para comprender de dónde vienen los criterios que se explicarán más adelante para optimizar un sitio web.

### 2.1.3 Evolución de los algoritmos y principales cambios

#### Google Caffeine

Google Caffeine fue el comienzo de todo, con su lanzamiento en agosto del 2009. Desde su incorporación ha agregado significativas mejoras de velocidad al buscador y una nueva infraestructura de indexación más relevante.

Figura 1: Google Caffeine



Fuente: Imagen extraída de "Google Caffeine Is Not An Algorithm Change", por A. J. Kohn, 2010.  
<http://www.blindfiveyearold.com/google-caffeine-is-not-an-algorithm-change>

#### Google Panda

En 2010 se lanzó Google Instant, que permitía realizar búsquedas de sucesos en tiempo real. En el mismo año se lanzó la primera versión de Google Panda, el cual buscaba pulir la calidad de los resultados de búsqueda en base a la calidad de enlaces que apuntaban a cada sitio, tratando de identificar cuáles eran de baja calidad según los sitios web que los estaban referenciando. Google Panda atacaba principalmente a las *granjas de contenido* que se caracterizaban por ser de baja calidad y contenido duplicado.

De Google Panda se han hecho una gran cantidad de actualizaciones; al día de hoy algunas actualizaciones siguen vigentes, apuntando a ser cada vez más certeras en la identificación de contenidos de calidad y de aquellos que no lo son, premiando y penalizando los contenidos duplicados o publicados en base a plagio.

**Figura 2: Google Panda**



Fuente: Imagen extraída de "The Ultimate Google Algorithm Cheat Sheet", por N. Patel, s.f. <https://neilpatel.com/blog/the-ultimate-google-algorithm-cheat-sheet/>

### **Google Penguin y Pirate**

En 2012 se incorporaron elementos más visuales y gráficos para presentar los resultados de búsqueda como consecuencia de la incorporación de *Knowledge Graph*. Así mismo, en este año, también se realizó el lanzamiento de uno de los algoritmos más importantes en la historia reciente del buscador: Google Penguin.

Penguin consiste en identificar los enlaces no-naturales que fueron creados masivamente en base a un robot y en detectar que los sitios donde aparecían dichos enlaces también eran de baja calidad, por lo que se consideraba que el sitio que generaba una estrategia basada en estos principios también era de baja calidad y se lo penalizaba por ello. Este lanzamiento afectó fuertemente a una gran cantidad de sitios web en el mundo e incluso hizo quebrar a algunas compañías que basaban su modelo de negocio en conseguir tráfico de esta manera. Ahora bien, los sitios que sí seguían las buenas prácticas y tenían un crecimiento natural, en su mayoría, se vieron beneficiados por estas actualizaciones.

**Figura 3: Google Penguin**



Fuente: Imagen extraída de “The Ultimate Google Algorithm Cheat Sheet”, por N. Patel, s.f. <https://neilpatel.com/blog/the-ultimate-google-algorithm-cheat-sheet/>

Penguin continúa más vigente que nunca, filtrando enlaces de baja calidad y permitiendo, a través de herramientas como Google Search Console, que los usuarios puedan gestionar los enlaces que efectivamente desean permitir que sean parte de su ecosistema de *link-building*.

En el 2012, Google también lanzó otra actualización: Google Pirate. Este algoritmo está enfocado en justamente lo que dice la palabra: penalizar aquellas páginas web que tengan sanciones de derechos de autor o que sigan la línea de la piratería.

### **Google Hummingbird**

Hacia 2013 se realizó el lanzamiento de Hummingbird, un algoritmo de búsqueda bastante disruptivo que intentaba tratar a las páginas web como elementos semánticos que debían responder preguntas de un ser humano y no estar solo optimizados para elementos técnicos, robots o *crawlers*.

La actualización Hummingbird de Google es probablemente la más significativa de todas, después de Panda. Desde este lanzamiento el posicionamiento en Google está directamente relacionado a la semántica -lo que conceptualmente se conoce como la "Indexación Semántica Latente" (LSI por sus siglas en inglés)- y la geolocalización.

Ya no es más meramente la palabra clave (keyword) lo que marca diferencias al momento de optimizar una web. Ahora hay que considerar la semántica, la sinonimia, el contexto de uso y la geolocalización.

**Figura 4: Google Hummingbird**



Fuente: Imagen extraída de "Google Algorithm: 8 Fun Facts You Should Know", por Silver Edge, 2017. <http://www.silveredgemedial.com/google-algorithm-8-fun-facts-know/>

### **Google Pigeon**

Google Pigeon se lanzó en el 2014 con el objetivo de depurar las búsquedas locales, de manera de crear una fórmula matemática que se encarga de priorizar los contenidos y backlinks locales (mediante geolocalización), por sobre los globales.

Para el SEO local es esencial. Es la causante de que un negocio local aparezca en las primeras posiciones. Si bien no es un algoritmo que haya causado tanto revuelo como Panda y Hummingbird, porque no presenta ningún tipo de penalización, sin dudas marcó un antes y un después en la experiencia del usuario.

### **Google MobileGeddon**

De peculiar nombre, pero a la vez muy acertado (mobile + armageddon), en el 2015 surge este código que incide en los algoritmos de Google para dar prioridad en la clasificación de resultados de búsqueda a sitios web que sean compatible con dispositivos móviles, aquellos que se optimicen adecuadamente. La tecnología AMP de Google (por sus siglas en inglés “Páginas Móviles Aceleradas”), diseñada para mejorar la velocidad de apertura de páginas web en celulares, también es parte de esta actualización.

### **Google RankBrain**

En el 2015 también nace RankBrain. Se trata de la primera versión de algoritmo basado enteramente en inteligencia artificial, la cual cambiaba el criterio de presentar un contenido que se asemejara a la pregunta que se escribía en la casilla de búsqueda a premiar a las páginas que más bien brindaban la respuesta a la pregunta formulada. Este algoritmo busca detectar páginas web de UX de baja calidad, con contenido superficial y poco relevante.

De acuerdo a fuentes de Google, es parte del algoritmo de Google Hummingbird.

Hummingbird y RankBrain fueron la primera camada de una serie de actualizaciones del algoritmo de búsqueda que trataban de entender las necesidades, contexto y motivaciones de las personas más allá de las palabras escritas en los términos buscados. Intentaban dar respuestas prácticas y en el

momento preciso, incluso incorporando dispositivos móviles para detectar dónde debía ser relevante brindar una respuesta según la ubicación de la persona que preguntaba.

**Figura 5: Google Hummingbird y RankBrain**



*Fuente:* Imagen extraída de “How to get the most from RankBrain, Google’s new ranking robot”, por Braffton, 2015.  
<https://www.braffton.com/news/seo-1/how-to-get-the-most-from-rankbrain-googles-new-ranking-robot/>

### **Google Possum**

Google Possum se lanzó para dar prioridad a los resultados de búsqueda locales. Le indica al motor de búsqueda de Google la distancia entre una persona y el negocio que está buscando. También indica todos aquellos establecimientos cercanos de productos o servicios relacionados a lo que se está buscando. Trabaja en directa retroalimentación con el programa Google My Business.

### **Google Fred**

En el 2017 Google Fred se lanzó netamente para “ordenar” el gigante mercado de publicidad en páginas web.

Bajo Fred, los algoritmos de Google están castigando a todas aquellas páginas web que buscan ganar dinero a través de la publicidad sin importarles la experiencia del usuario y la calidad de los contenidos.

### **Google Medic y Google EAT**

El denominado “Google Medic” surge en el 2018 con exclusivo enfoque sobre los sitios web relacionados con la salud y la medicina, penalizando los que no son “user friendly” y los que no brindan una buena experiencia de usuario.

El algoritmo trabaja bajo el concepto de Google “YMYL, your money or your life”. Para Google, la información errónea puede afectar a los usuarios financiera, física o emocionalmente. Por lo tanto, la actualización está dirigida especialmente a aquellas páginas YMYL que tienen contenido de baja calidad e información errónea.

### **Figura 6: Google Medic**



*Fuente:* Imagen extraída de “Google Medic Update: Things You Need to Know”, por P. Singh, 2020. <https://cyfuture.com/blog/google-medic-update-things-you-need-to-know/>

Con una directa relación a Google Medic, se desarrolla el algoritmo de EAT, cuyas siglas significan: experticia, autoridad y confiabilidad (en inglés “Expertise, Authoritativeness, Trustworthiness”). Con este algoritmo Google muestra los requisitos y recomendaciones que utiliza el robot para evaluar la relevancia social de una web y clasificar contenidos en su ranking de resultados.

Vamos a profundizar brevemente en cada criterio:

- Expertise – Experticia: es muy recomendable que los autores firmen los textos y que Google detecte que éstos tengan una buena reputación online y que son expertos en su campo.
- Authoritativeness – Autoridad: la autoridad se consigue gracias a menciones de otras páginas de renombre hacia nuestra página o hacia el autor de contenidos de la misma.
- Trustworthiness – Confianza: en este aspecto hemos de demostrar a través de nuestro sitio web que el usuario puede confiar en que el contenido que se expone es cierto y legal. Para ello, un imprescindible es instalar el certificado SSL, solicitar que los usuarios creen contraseñas complejas cuando se requieran, e incluir cualquier certificado relevante en nuestra web.

Hablamos de una especial relación de Google EAT con Medic porque no todos los sitios están obligados a dar votos de fe sobre la autoría de los contenidos que publican, pero sí lo están los sitios YMYL, como webs del sector de la salud, financiero, educación, entre otros.

## **Google Florida 2**

En marzo del 2019 Google actualizó el mítico “Florida”, que como se mencionó en párrafos anteriores fue el pionero de todos los algoritmos que se vendrían. De esta actualización no se sabe demasiado, como tampoco nunca se reveló información concreta del funcionamiento de Florida. Lo que sí es cierto es que impactó en gran medida en webs de nichos médicos, de salud, de entretenimiento y juegos, y sitios YMYL. Todos ellos tienen en común que se caracterizan por:

- Solicitar información personal como datos bancarios, que pueden utilizarse para comprar cosas por internet, pero que también pueden ser blancos del robo de identidad.
- Ofrecer información médica o de salud, que pueda afectar el bienestar físico de los usuarios.

- Ofrecer consejos sobre decisiones importantes sobre la calidad de vida, como la compra de seguros, de vehículos, de casas, de viajes; también sobre la crianza de los hijos, el divorcio, etc.
- Información y consejos financieros, de bienes raíces o legales.

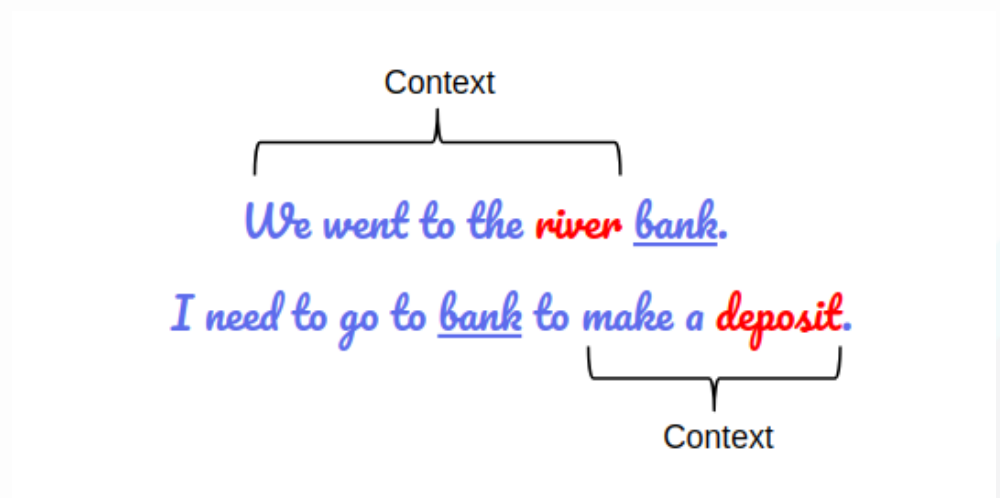
### Google BERT

A finales del 2019 surge BERT. Con este algoritmo Google empieza a utilizar el machine learning para la interpretación más natural del lenguaje. Con esto cambia toda la forma de uso del lenguaje para el potenciamiento del SEO.

BERT es particularmente interesante y se trata de uno de los cambios más importantes de los últimos tiempos. BERT es el acrónimo para Bidirectional Encoder Representations from Transformers (Representaciones de Codificador Bidireccional de Transformadores). En otras palabras, es un sistema basado en inteligencia artificial (IA) para ayudar a los algoritmos de Google Search a entender mejor el lenguaje que utilizamos los usuarios en el momento de realizar una búsqueda mediante oraciones.

BERT posee una característica que se llama “bidireccionalidad”, esta consiste en analizar una oración en dos direcciones. Es decir, analiza las palabras que se encuentran tanto a la izquierda como a la derecha de una palabra clave, y esto le permite entender en profundidad el contexto y la temática de toda la frase que introduce un usuario para la búsqueda en Google.

**Figura 7: Funcionamiento del análisis bidireccional de BERT**



Fuente: Imagen extraída de “Demystifying BERT: A Comprehensive Guide to the Groundbreaking NLP Framework”, por Mohd Sanad Zaki Rizvi, 2019.  
<https://www.analyticsvidhya.com/blog/2019/09/demystifying-bert-groundbreaking-nlp-framework/>

## 2.1.4 Cuando somos víctimas del algoritmo

A pesar de la creencia popular, el algoritmo de Google cambia constantemente. Si bien a menudo discutimos los principales hitos del algoritmo, como Panda y Penguin, Google implementa actualizaciones y sanciones más pequeñas para complementar las más importantes.

### **Señales de que ha sido penalizado por un cambio de algoritmo de Google:**

Si el sitio que maneja sufre alguno de los siguientes síntomas, es posible que Google haya aplicado una penalización.

**1-Degradación de rango y caída de tráfico orgánico:** Quizás el signo más obvio de una penalización es una caída repentina en la clasificación de su sitio en las SERP.

Si su sitio es expulsado de la primera página de Google durante un período, por ejemplo, verá una caída abrupta de visitas.

**2- Cambio de clasificación manual:** en la mayoría de los casos, el algoritmo de Google hace la mayor parte del trabajo preliminar al penalizar las páginas. Sin embargo, en ocasiones Google toma acciones manuales si sospechan que el sitio está haciendo acciones forzosas para el posicionamiento. De todas formas, antes de ejecutar alguna penalización, Google le notificará si cree que algo está sospechoso (y si es que el sitio está registrado en Search Console). Se da un margen al webmaster para que corrija los errores, o bien, apele en contra de dicha notificación.

**3- Google Banned:** esto es cuando Google decide incluir el sitio en una lista negra, es decir, se le prohíbe ser indexado. Ocurre porque probablemente se

haya detectado algo muy sospechoso, como comprar enlaces, enviar spam a directorios, hacer uso excesivo de palabras clave o distribuir malware a través del sitio.

### Conceptos clave

- 1) Google realiza actualizaciones de su algoritmo constantemente, por lo que es necesario mantenerse al día con las implicancias de cada uno y apuntar a realizar las optimizaciones necesarias para ganar relevancia.
- 2) Conocer la historia de los algoritmos recientes de Google permite entender hacia dónde va el criterio de lo que consideran *relevante* y evitar posibles penalizaciones que acaben, en un segundo, con todo el esfuerzo de meses.
- 3) Google Panda se enfoca en contenido de calidad, Google Penguin en enlaces, y Google Hummingbird utiliza la base de RankBrain para darle inteligencia artificial a la forma de entender las preguntas de los seres humanos; y BERT es el complemento de ellos a través de una interpretación semántica del contexto en el que se encuentran las keywords. Mobile First les da mayor importancia a los sitios desde dispositivos móviles, y Possum se enfoca en búsquedas con contenido local. Cada uno de los algoritmos aportó un importante granito de arena en lo que hoy se traduce como el mayor motor de búsquedas del mundo, con una experiencia de usuario incomparable que nunca antes habíamos presenciado.

## Referencias

**Aarav Infotech.** (2019, 25 febrero). *9 Best SEO Trends You Must Know in 2019.* Aarav Infotech. <https://www.aaravinfotech.com/blog/9-best-seo-trends-must-know-2019/>

**Brafton.** (2015, 29 octubre). *How to get the most from RankBrain, Google's new ranking robot.* Brafton Fuel your brand. <https://www.brafton.com/news/seo-1/how-to-get-the-most-from-rankbrain-googles-new-ranking-robot/>

**Carrero, J.** (s. f.). *El Algoritmo de Google y tu Estrategia SEO*. 2IM Marketing. <https://2immarketing.com/algoritmo-de-google-estrategia-seo/>

**Doantam Phan.** (2016, 4 noviembre). *Mobile-first Indexing*. Google Search Central Blog. <https://webmasters.googleblog.com/2016/11/mobile-first-indexing.html>

**Kohn, A. J.** (2010, 5 febrero). *Google Caffeine Is Not An Algorithm Change*. Blind Five Year Old. <http://www.blindfiveyearold.com/google-caffeine-is-not-an-algorithm-change>

**Mohd Sanad Zaki Rizvi.** (2019, 25 septiembre). *Demystifying BERT: A Comprehensive Guide to the Groundbreaking NLP Framework*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2019/09/demystifying-bert-groundbreaking-nlp-framework/>

**Singh, P.** (2020, 2 marzo). *Google Medic Update: Things You Need to Know*. Cyfuture. <https://cyfuture.com/blog/google-medic-update-things-you-need-to-know/>

**Patel, N.** (s. f.). *The Ultimate Google Algorithm Cheat Sheet*. Neil Patel. <https://neilpatel.com/blog/the-ultimate-google-algorithm-cheat-sheet/>

**Silver Edge Media.** (2017, 2 julio). *Google Algorithm: 8 Fun Facts You Should Know*. Silver Edge. <http://www.silveredgemedia.com/google-algorithm-8-fun-facts-know/>