

# Google Algorithms

Adithya Sangaraju

---

## I. INTRODUCTION

Basically, according to Computer terminology, an algorithm simply means a certain block of steps employed to perform a particular task. When we enter a search query in the Google search bar, it produces millions of results within the blink of an eye. The process employed to produce these results is best described by an algorithm.

## II. IDEA OF RESEARCH

Google is an imminent part of our everyday life, where in we enter innumerable number of queries into the search bar. The idea behind how Google provides us with so many results in the span of a millisecond is the backbone of this paper.

## III. REQUIREMENT OF RESEARCH

The thirst for knowledge is a rare trait born in the human mind. When we come across any process that challenges the understanding of the mind, we become restless unless we successfully understand what the process is based upon. With this research paper, I hope to illustrate on a brief basis as to how the links to different websites appear when a certain input is entered into the search bar.

### Google Algorithm:

Google employs algorithms in such a way that it holds the capability of sensing what one might want to see, out of the keyword that we have entered. In other words Google's algorithms accept our query as input and produce results in the form of webpages that may count up to many thousands or even millions.

There are certain criteria upon which the algorithm filters our search results. The technical term used to represent such criteria is 'clues'. These clues are unique signals that help in determining what really we are looking for. They include information like the terms and condition of a website, how recent has the content been modified, the region in which the user exists, how good the title of a pages looks and rank of the page.

Google's developers focus on refining the quality of result produced by employing different 'clues', that also determine if a page or website stands at position 1 in search results.

While the overall preview of how the algorithm works has been specified, a Google algorithm is not as easy as it meets the eye. Google algorithms are complicated and increase in the level of complexity as the search for a user's query gets refined from level to level.

Also, Google's algorithms keep changing from time to time to stay updated, to increase the user's satisfaction and the inflow of user traffic.

### Webpage at Position 1 –Then and Now:

At the initial stages of Google's introduction, the algorithms were changed very rarely. This meant that a certain webpage stood at the first position for quite a long amount of time, even for weeks and months in a single row.

When an update of algorithms was released, the changes that occurred would alter the results until the next update was made. This let down the widespread reach of the users and many good webpages were included somewhere unreachable in the result list.

But, with the introduction of “Caffeine” the scenario completely changed. Introduced in the year of 2010, Caffeine allows the search engine to regularly update its results almost every day. With this special inclusion, the results keep getting updated every now and then, to meet the requirements of the users.

### **Algorithm Changes:**

Every year, Google makes almost 600-700 changes to its algorithms. When the change is said to be phenomenal, Google announces a name for the change made. There are three big changes that took place in the last few years, namely, the Panda Algorithm, the Penguin algorithm and the Humming Bird Algorithm.

### **Panda Algorithm:**

Google launched its Panda algorithm on February 23rd in the year of 2011. This algorithm created hype in leading global trends as it was a first of its kind algorithm by Google. Its main function was to concentrate on showing higher quality sites first in the search results compared to low quality ones. When the change in the algorithm was first made, it was unnamed but later on, the name 'Panda Algorithm' was chosen based on one of its creators 'Navneet Panda'.

Due to Panda algorithm on the run, many websites that displayed huge number of advertisements on their webpages were affected immensely. All such type of websites were marked as low quality ones and the results refused to include links of these websites. At the same time, websites that included content that was copied from elsewhere were given high priority. This led to a lot of complaints related to copyright policy, that duplicate websites were receiving more attention.

It was later observed that Google Panda took into consideration a certain block of data and then decides whether a whole website should be given high priority or not. This was improbable and in order to rectify this major issue, Google laid out some guidelines with which a website's priority could be determined. In general, a website was said to get high priority if:

- The information present in the website was trustworthy.
- The information was written by a certified author or an expert.
- The users could provide their credit/debit card details without fear of theft or malpractice.
- The webpage provided better outcome when compared to other webpages.
- The webpage did not display too many of advertisements.
- The content written was well edited and did not contain any errors.
- The owner of a site was a real and trustworthy person.

In general, these guidelines were included to show that the webpage was being real-time user friendly. Apart from these guidelines, there are other factors that are related to the content, which are considered important when Google Panda is being used to determine the reliability of a webpage. These new factors include the strength of the content present in a certain webpage. The strengths of these contents are divided as ‘Thin content’, ‘Low-quality content’ and ‘Duplicate content’.

### **What is Thin Content?**

A page is said to possess thin content if it does not contain information that is quite helpful to the user. In other words, the user may find it hopeless if there are a number of pages in front of him with each page consisting of only one or two sentences of information.

### **What is Low-quality content?**

The content of a page falls under the category of low quality content when the author copies each line of his information from different pages. This in turn creates an impression in the user that the content lacks in better understanding or meaning.

### **What is Duplicate content?**

A page falls under the category of duplicate content when the author copies a whole lot of information and displays it on his website as his own. It is also permissible to fall under duplicate content category when the author copies information from his own page itself.

Google Panda is devised in such a way that no low-quality website is ever displayed in the higher orders of search results.

### **Google Panda Hit:**

When a website is certified as a low quality one, this is known as a hit, otherwise a Panda Hit. Google Panda algorithm is refreshed for every couple of days and each time a refresh occurs, Google rechecks all the websites for re-evaluating their quality. If a site was saved under the category of low quality, and the author had taken necessary measures to improve the content, then after the refresh occurs, this site will recover from the Panda hit. This site will next be transferred to the category of high-priority websites.

### **Panda Update:**

Whenever a refresh to the existing Google Panda occurs, it is simply considered as an update. Since the release of Google Panda in 2011, it has been updated many times with the latest being Panda 4.1 on 21st September, 2014. Panda 4.1 has made many significant changes ever since and this marks the fact that many websites have been converted to high priority category too.

The next important change that took place in the list of Google Algorithms was the Penguin Algorithm.

### **Google Penguin:**

Google Penguin was first released on 24th of April, 2012. It mainly focuses on websites that previously tricked Google by employing duplicate backlinks and appear in the best of results. Google Penguin takes measures and tracks all those websites that possess backlinks and simply marks them as low quality ones.

Google Penguin 1.1 was the last update which was released on May 25th, 2012. It was said to refine the results using language as a criteria too. The languages included English, Chinese, German and Arabic as major ones.

Before explaining further on what and how to recover from a Penguin hit, it is quite important to first know what a backlink actually refers to.

### **Backlinks:**

Links are very important to a website and the more links associated with a website, the more popular or high-priority it becomes. Links from big and well known companies are preferred by Google and when it sees that a website is linked with a multi-national firm, it automatically falls under high-priority rankings. However, links from smaller webiste companies are also accepted but, the number of links has to be quite large.

In recent times, website owners manipulated these links in such a way that it looks as if they are linked with a big company, while in reality it is not true. They create self made links exactly identical to the real ones. This manipulation is simply called Backlinking. Previously, when Google Penguin was not introduced, Backlinking became the favourite option for website owners in order to obtain good rankings.

### **Penguin Hit:**

When Google Penguin verifies your webpage for backlinks (self-made links) and in the cases the algorithm finds one, your webpage will be certified with less priority. Google entrusts a little amount of faith in every webpage and if it finds that a page has been implementing self made links, then the faith breaks leading to Google not showing the name of that webpage in search results.

This phenomennon is called a Penguin Hit. When a Penguin Hit occurs, the global rankings of a website fall down gradually and a recovery method has to be employed to overcome a hit.

### **Penguin Hit recovery:**

Just like a Panda algorithm, Penguin Algorithm gets updated or refreshed many times. Hence, if a website currently contains backlinks that led to the demotion, care has to be taken that these fake links are identified and removed. When a refresh of the algorithm occurs, it re-verifies the webpage and in the situation of not finding any fake links, the website rankings shall increase compared to previous levels. This is the best available option to recover from a Penguin Hit but however, the last known Penguin update, Penguin 5 dates back to the month of October 2013. Eversince, there has been no update and new websites that are free from backlinks are not being judged to alter the current rankings.

The next big change to Algorithms is the Google Humming Bird update.

### **Google Humming Bird:**

Google launched Humming Bird algorithm on 26th September 2013, but it admitted that the algorithm was already in use by a span of a month. Google Humming Bird works in an entirely different way compared to the previous Google Panda or Google Penguin. It does not verify the priority status of a webpage but instead, it dives into the query given by the user. It looks out for keywords in the query and uses them to provide better search experience to the user. For example, when someone enters a query such as 'List of Resorts in Maldives', Google Humming Bird identifies keywords like Resorts, Maldives and works in bringing out different webpages that contain lists of resorts present in Maldives. Though Google Humming Bird and other algorithms work collectively, it is said to be more user-related as it deals with queries entered by the user themselves.

### **Google Humming Bird makes my webpage better?**

Unlike the other algorithms, there exists no Humming Bird hit as it does not verify content on pages. Instead, it focuses on providing a better search experience to the users.

Google Humming bird does make a website better. The authors must take care to include quality rich content that is synonymous to the words that a user might enter in the search bar. In this way, Google Humming Bird readily matches the words in a webpage to the words the user has entered. It increases the visibility of that webpage and in turn fetches high priority.

### **Humming Bird and Importance:**

Humming Bird is said to be the most important and major update to the algorithms since 2001. The reason it gained such recognition is that it deals directly with user interface in providing best results out of the queries. It makes sure that the page content is relevant for the user to oblige.

### **Other Google Algorithms:**

Apart from these three major changes, there are other several updates to Google Algorithms. Some of them are listed below:

PayDay Loan Algorithm

Panda Dance

Phantom update

Venice - a monthly update.

The +1 button

Buffy algorithm

Big Daddy algorithm

Jagger Algorithm

Bourbon

Brandy

Dominic algorithm

Cassandra

Google also launched several other unnamed updates.

#### **IV. ACKNOWLEDGEMENTS**

I am in never ending debt to the above mentioned resources for providing me with a clear insight on the history details and the correct names given to each Google Algorithm. I thank my research paper evaluator, my mentor for reviewing this paper for the first time.

#### **V. CONCLUSION**

Hence a quick insight into the process employed by Google to refine the millennia of websites is introduced and explained. Also, certain recovery mechanisms are provided where a website can overcome the Algorithm hit and re-appear in top notch positions.

#### **REFERENCES**

- [1] Wikipedia
- [2] Moz.com/google-algorithm-change
- [3] Webopedia