

# HUMMING BIRD SEARCH ALGORITHM

Sanjana Singh

Computer Science Department  
Dronacharya College of Engineering  
Khentawas, Farukhnagar, Gurgaon  
Email-Ssingh2294@gmail.com

Surbhi Bhardwaj

Computer Science Department  
Dronacharya College of Engineering  
Khentawas, Farukhnagar, Gurgaon  
Email-surbhi.bhardwaj094@gmail.com

Shivansh Mudgil

Computer Science Department  
Dronacharya College of Engineering  
Khentawas, Farukhnagar, Gurgaon  
Email-shivanshmudgil@gmail.com

## INTRODUCTION

Google Hummingbird is a search algorithm used by Google.

Google started using Hummingbird about August 30, 2013, and announced the change on September 26 on the eve of the company's 15th anniversary.

Gianluca Fiorelli said Hummingbird is about synonyms but also about context. Google always had synonyms, he writes, but with Hummingbird it is also able to judge context - thereby judging the intent of a person carrying out a search, to determine what they are trying to find out. Danny Sullivan said of Hummingbird, "Google said that Hummingbird is paying more attention to each word in a query, ensuring that the whole query — the whole sentence or conversation or meaning — is taken into account." Michelle Hill said Hummingbird is about "understanding intent". Steve Masters wrote, "The Hummingbird approach should be inspirational to anyone managing and planning content — if you aren't already thinking like Hummingbird, you should be. In a nutshell, think about why people are looking for something rather than what they are looking for. A content strategy should be designed to answer their needs, not just provide them with facts." "The age of Semantic Search.

"Hummingbird" is the name of the new search platform that Google is using as of September 2013, the name comes from being "precise and fast" and is designed

to better focus on the meaning behind the words. Read our Google Hummingbird FAQ [here](#).

Hummingbird is paying more attention to each word in a query, ensuring that the whole query — the whole sentence or conversation or meaning — is taken into account, rather than particular words. The goal is that pages matching the meaning do better, rather than pages matching just a few words.

Google Hummingbird is designed to apply the meaning technology to billions of pages from across the web, in addition to Knowledge Graph facts, which may bring back better results.

The name of Google's new search algorithm is called "Hummingbird". Derived from being "precise and fast", it's been the biggest change in Google's search algorithm since the [Caffeine update](#) in 2010.

There were also some major changes before like Penguin and Panda, but what sets Hummingbird apart from them is that two previous ones were just changes done to some parts of the old algorithm, while Hummingbird is an entire replacement of the old one.

One of the new search activities Google features is "Conversational Search", which is meant for smart phone users who may find having conversations more convenient when doing research.

For example, you could ask, "What's the nearest taco store from my home?" The traditional search engine will most probably focus on finding matches for your keywords like "taco" or "store", and then lead you to a website owned by a restaurant that sells tacos but may not be necessarily close to your home.

Hummingbird makes Google almost human with the way it responds to queries. Instead of looking at keywords, Hummingbird makes Google look deeper and focus on the whole statement or question.

In relation to the example above, the new update helps the search engine understand that you're looking for a physical store near your area of residence that sells tacos, provided that you use the Google domain based in your country.

In a nutshell, Google responds to whole statements and questions in a way similar to how a human being would by looking at the query as a whole, and not segmenting its keywords. This then makes the results to match the query better since it looks for concepts, not words.

## **HUMMINGBIRD SPEEDS US INTO THE FUTURE**

Hummingbird is a great move for search results and could be a great way for websites to gain more visibility if they focus on the user and the content first.

It may actually be a relief for some SEOs to know that with Hummingbird and some of the other changes we've seen Google putting out, it's a clear message that site owners should stop obsessing over keywords *only* and start focusing on creating a great experience.

Google quietly rolled out Hummingbird with little or no fanfare. That should not be taken as an indication that Google sees Hummingbird as a minor change to their search engine algorithm. Just the opposite.

With Hummingbird, Google hopes to significantly alter the way users interact with their search engine. All search engines seek to provide relevant search results quickly and efficiently. Google is working toward a new level of speed and precision in their search result

## **WHAT DOES HUMMINGBIRD MEAN FOR SEO?**

Does Hummingbird mean the end for SEO? Definitely not. While any change to Google's search engine protocols is routinely answered with a cry of frustration from SEOs and webmasters, Hummingbird shows no signs of significantly changing the search engine optimization landscape. If you are following best SEO practices, there should be little or no adverse affects on your sites.

In my opinion, this search update gives us SEOs some advantages. How you ask? Simple.

The title of this post was purposely written in the form of a question. Using this concept and the 5 W's (who, what, when, where, why, and of course how) will get your content one step close to rising to the top of search.

However, Hummingbird will also place a renewed emphasis on authoritative quality content. If your SEO strategy is too heavily weighted towards keyword deployment, at the expense of creating authoritative content and quality links, your sites will continue to lose traction and rankings.

SEO's have long dreaded Google's updates since they believe that these would mess with the ranks they've worked so hard for.

But Google said that as long as you have been following their age-old rule to make original and high quality content, then there's nothing really to worry about since the Hummingbird was just meant to process information in a different way.

In a recent blog post, Krystian Włodarczyk suggests that Google is intent on reducing the importance of keywords. What will always matter is the quality of the content that an author produces. Hence, the new algorithm only has massive effect on those sites which use dirty tactics in order to achieve their ends.

Although SEO's purpose is to get traffic, it's not about manipulating search engines in devious ways, sending out spams, or killing web design. It's all about making quality content that's friendly with search engines and people alike, in a way that both would have legitimate reasons to rank it high and share it.

## **NITTY-GRITTY OF THE ALGORITHM**

The Hummingbird algorithm responds to end users by means of processing the whole statements or questions, and then looks for websites or webpages that has the most relevant concept to the query. But how does Hummingbird do this exactly?

Hummingbird applies a statistical language approach in a way that it rewrites search queries to make it simpler. This can be done by browsing the large database of synonyms or other related terms that might be substituted in place of one or more terms in the query.

If a query indicates, "Where is the best place to buy burgers?", Google will most likely try to find a substitute for the terms "best", "place", "buy", and "burgers". The words "where", "is", "the", and "to" would require no substitutes since they hold no great significance – these words are called "skip words".

The synonyms that are stored in the database were created through observation of what end users search for in their queries. If an end user, for example, submitted two consecutive queries, the words he sent will give the hint that the words are related to the ones sent in the previous one.

To illustrate, when an end user searches for "Korean idol" and then "Korean star", Google will understand that the word "idol" and "star" are of the same context.

## **THE PROCESS**

So, how does this work? What is the process involved? In a nutshell, Google needs to understand the user queries. They take the query, leverage grammatical structure and boil the query down to one of these forms. They leverage user intent (and various other implicit signals to assist with determining this intent). They can then determine which form to map to.

In short, a simplified version of the process is approximately as follows, and this process may well change or be altered by leveraging machine learning. I would state this merely as an educated guess:

- Parse the grammar of the query
- Identity the form from the user intent
- Identify entity(s) involved
- Attribute synonyms
- Qualify what the user is looking for (refined intent)
- Determine what entities to retrieve
- Determine what properties those entities have
- Determine what to show the user in a meaningful way (the latter of which is device-dependent and needs to be attractive and engaging to the user)

Leveraging context and other implicit factors to understand the query is clearly critical, as understanding user intent is paramount for correct disambiguation and relevant synonym expansion throughout.

Google's goal is clearly to expand on the number of forms and types of forms they can handle. And Google, as I have stated many times, is a master at big data. Google is merely taking older technology and making it work at scale. They can add new *forms* based on the combination of *search volume* (for queries from incoming sources, namely user demand) and *lowest-hanging fruit* (best bang for the buck, the latter being from a computational cost as well as the possible Question Space they can cover). And so, they keep adding to the types of questions that Google can answer.

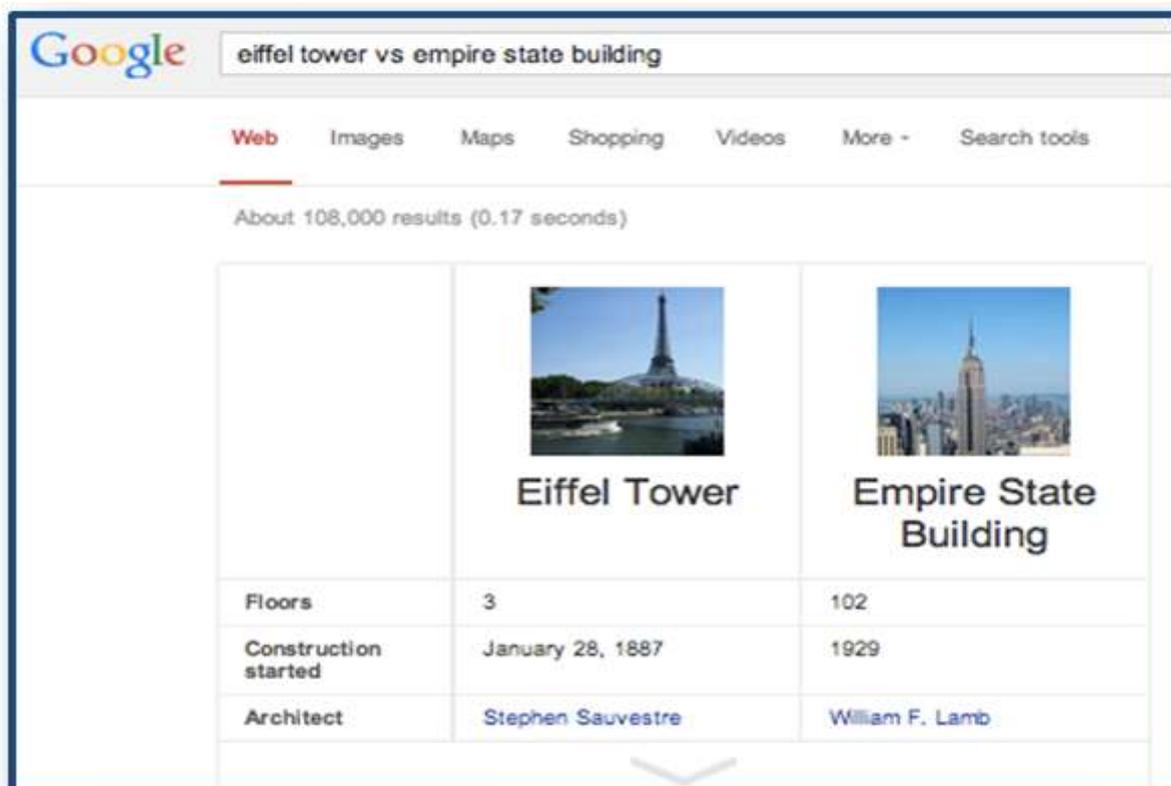
You can also see here where Google's acquisition of Metaweb and the concept of "entity equivalence" are enormously critical, as they now have the Knowledge Graph as a single source to which they can map all related information for an entity, and from which they can pull the required answers (entities and associated attributes or properties of those entities).

In light of the above, let us now re-look at Google's latest announcement on its 15<sup>th</sup> birthday.

Let's look at the first new feature announced in [their blog post](#): “Comparisons and filters in the Knowledge Graph.” This means they added at least one new form which deals with the concept (and I am approximating here):

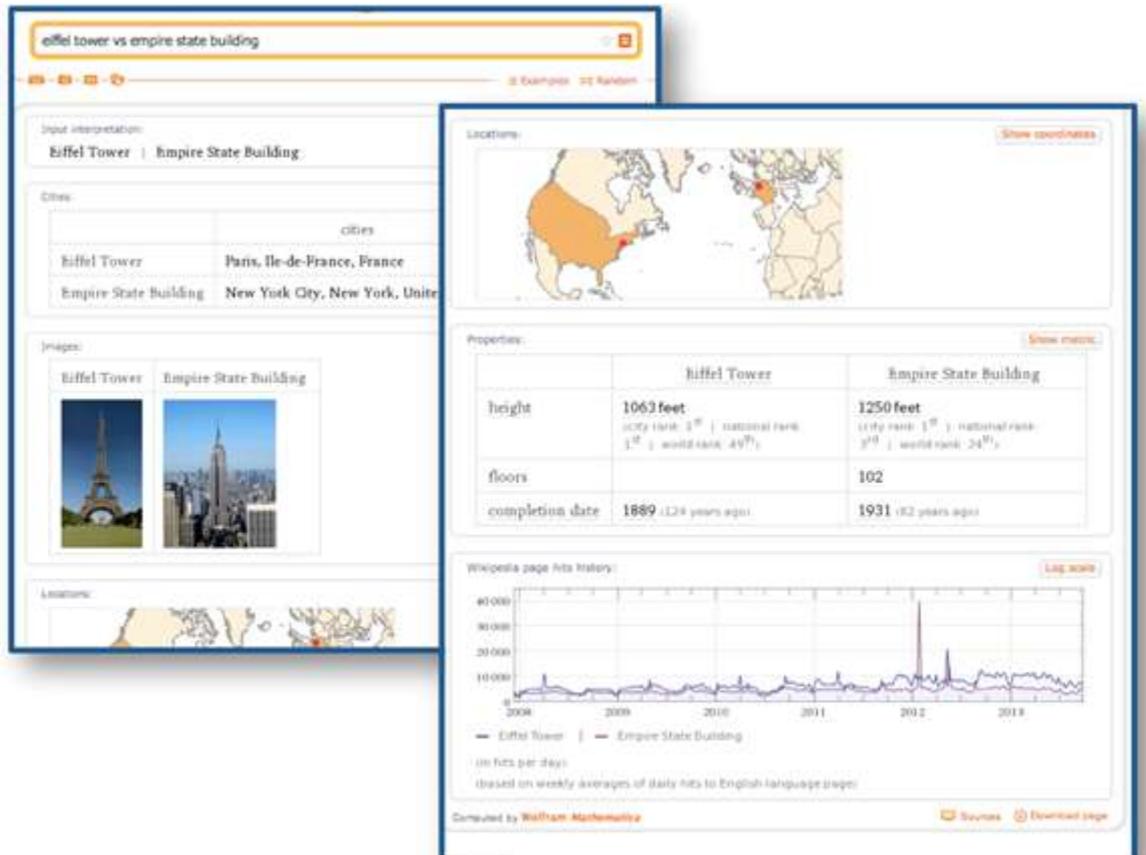
*?entity1 vs. ?entity2* where both ?entity1 and ?entity2 are of the same type

Knowing what you do now of the Knowledge Graph, you can see it is just a matter of pulling up the required entities from the Knowledge Graph, looking at the associated attributes/properties, and figuring out the best means to attractively and engagingly display the answer to the user. An example in Google is depicted below.



Example of “Eiffel Tower vs. Empire State Building”

You can see where other search engines leverage the technology in a similar fashion; for example, if you type the same query into [Wolfram Alpha](#), you get the example below:



Example of “Eiffel Tower vs. Empire State Building” using Wolfram Alpha. The other two items in Google’s announcement also relate to optimizing displays across devices and leveraging the predictive capabilities of Google Now. These are also critical components of the Hummingbird ecosystem, necessary to maintain, support and give traction to the power embodied in the Knowledge Graph and Google’s ability to leverage Semantics at Scale and across devices.

The proliferation of new query patterns and different device types clearly require a paradigm shift, such as the move to a pragmatic semantic model, which was needed for the survival of search engines.

## REFERENCES

- <http://www.sitepoint.com/what-googles-hummingbird-update-means-for-seo/>
- <https://zadroweb.com/hummingbird-affect-seo/>
- <http://searchengineland.com/google-hummingbird-the-keyword-what-you-need-to-know-to-stay-ahead-175531>
- <http://searchengineland.com/google-hummingbird-when-evolutionary-becomes-revolutionary-173740>