GOOGLE ALGORITHM UPDATES

PANDA UPDATE

The Panda algorithm update addressed a number of problematic phenomena in Google SERPs, including:

- **Thin content** - Weak pages with very little relevant or substantive text and resources, such as a set of pages describing a variety of health conditions with only a few sentences present on each page.
- **Duplicate content -** Copied content that appears on the Internet in more than one place. Duplicate content issues can also happen on your own website when you have multiple pages featuring the same text with little or no variation. For example, a chimney sweep company might create 10 pages, one for each city the business serves, with content that is nearly identical on all of the pages with only the city names swapped out (e.g. “We clean chimneys in Denver” on one page and “We clean chimneys in Boulder” on the next, and “We clean chimneys in Aspen” on the next).
- **Low-quality content -** Pages that provide little value to human readers because they lack in-depth information.
- **Lack of authority/trustworthiness -** Content produced by sources that are not considered definitive or verified. A Google rep stated that sites aiming to avoid Panda’s impact should work to become recognized as authorities on their topic and entities to which a human user would feel comfortable giving their credit card information.
- **Content farming -** Large numbers of low-quality pages, often aggregated from other websites. For example, of a content farm might be a website that employs large numbers of writers at a low wage to create short articles covering a vast variety of search engine queries, producing a body

of content that lacks authority and value to readers because its core purpose is simply to gain search engine rankings for every conceivable term.

- **Low-quality user-generated content (UGC) -** An example of this type of low-value User Generated Content would be a blog that publishes guest blog posts that are short, full of spelling and grammatical errors and lacking in authoritative information.

PENGUIN UPDATE

The Penguin update is an update to the Google ranking algorithm, first carried out in 2012. Back then, the rollout of the update had massive effects on many websites. Google intended the update to fight website spam

- *Bought links*
- *Backlinks from untrustworthy sources*
- *Backlinks that have the same anchor text*
- *Bot-built backlinks*
- *Backlinks from irrelevant sites*

HUMMINGBIRD UPDATE

The Google Hummingbird update was put into place in August 2013 and announced one month later, in September 2013.

The Hummingbird update has been described by Google as the biggest change to the algorithm since 2001.

It was also described by multiple Googlers as a total rewrite of the core algorithm.

Semantic search attempts to match appropriate SERP results to the language of Internet users' queries **beyond** the meanings of individual keywords, taking a broader context into account, even when the intent of the user is implicit rather than explicit.

Google's increasing mastery of semantic search enables them, in their own words, to understand "real-world entities and their relationships to one another." Hummingbird's focus on matching query context to results relies on the intelligence of Google's technology — *the ability to parse intent*.

GOOGLE PIGEON

One of Google's biggest-impact algorithm updates was implemented for the sake of local search results, and local businesses typically saw the effects of the update in their website's analytics data.

To ultimately improve its local search capabilities, Google enhanced hundreds of ranking signals for both Google Search and Google Maps. This also meant Google's location and distance ranking parameters were also improved to better provide local, relevant results to users based on proximity.

RANKBRAIN UPDATE

RankBrain is a system by which Google can better understand the likely user intent of a search query. It was rolled out in the spring of 2015, but not announced until October 26 of that year.

At inception, RankBrain was applied to queries that Google had not previously encountered which accounted then and still does, for about 15% of all searches. It was expanded from there to impact all search results.

At its core, RankBrain is a machine learning system that builds off Hummingbird, which took Google from a “strings” to “things” environment.

This is to say, it took it from “reading” literal characters, and instead “seeing” the entity they represented.

To clearly conceptualize RankBrain, it can help to put yourself in Google’s shoes, trying to understand the intent of a search engine query like “Olympics location.”

What is the true intent of this search? Does the searcher want to know about the Summer or Winter Olympic Games? Are they referring to an Olympics that just concluded, or one that will take place four years from now? Is the searcher attending the Olympics right now, sitting in a hotel and looking for directions to the venue for the opening ceremonies? Could they even be looking for historic information about the location of the very first Olympics in ancient Greece?

Now, imagine that in trying to answer this query, all you have is simplistic algorithm signals like the quality of content or the number of links a piece of content has earned to rank results for this searcher. Imagine that the Winter Games in Sochi, Russia just concluded last month and the official Sochi Olympics website has earned millions of links for its content about this past event. If your algorithm is simplistic, it may only show results about the Sochi Games, because they have earned the most links… *even if the searcher was actually hoping to learn the location of the next Winter Olympics in Pyeongchang, South Korea*.

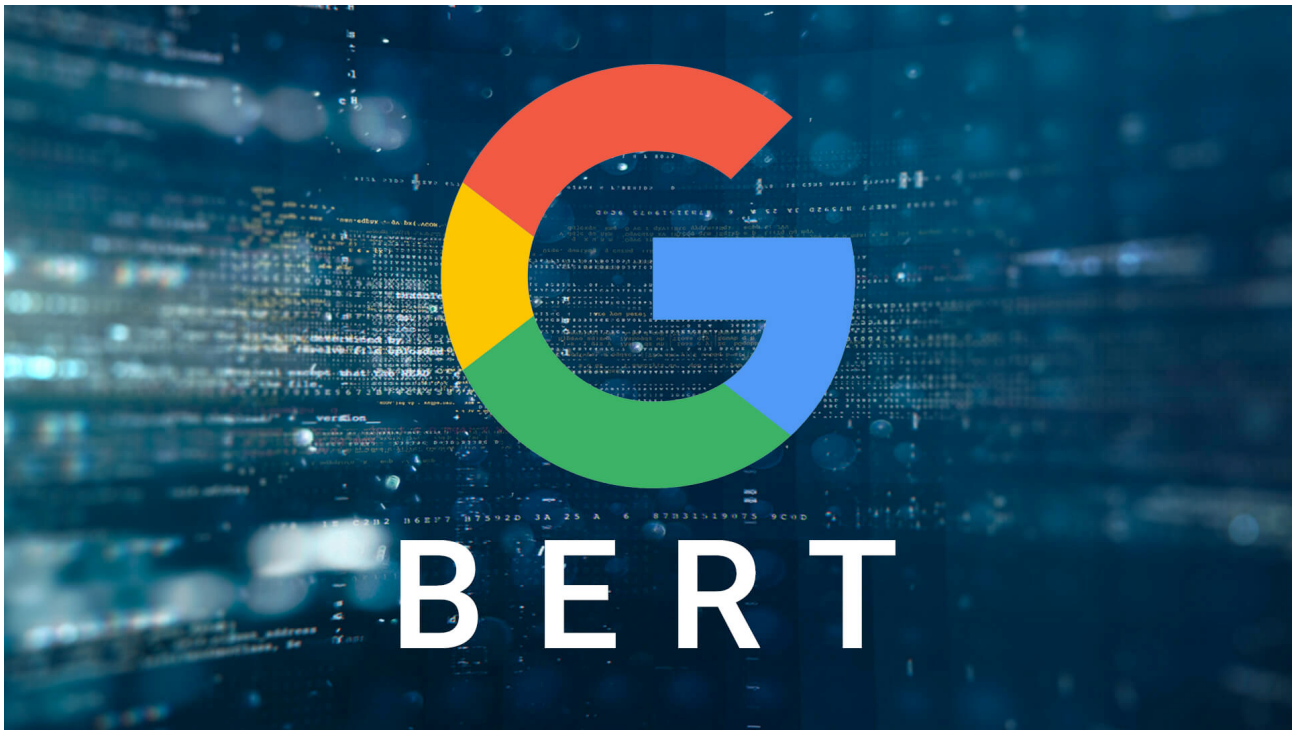
It’s within this complicated but common situation that the capacity of RankBrain emerges as essential. It’s only by being able to mathematically calculate results based on patterns the machine learning algorithm has “noticed” in searcher behavior that Google can determine that, for example, the majority of people looking up “Olympics location” want to know where the *very next* Games (be they Summer or Winter) will be held. So, in this case, a Google answer box with the upcoming Games’ location in it will serve the majority of searchers’ needs.

While that answer box may address the intent behind most “Olympics location” searches, there are notable exceptions Google must address. For instance, if the search is being performed by a user within an Olympic city (like Pyeongchang) the week of the games, Google might instead provide driving directions to the pavilion where the opening ceremonies will be held. In other words, signals like user location and content freshness must be taken into account to **interpret intent** and deliver the results most likely to **satisfy searchers**.

**RankBrain is a work in progress*, with the goal of machine learning perfecting Google’s interpretation of searcher intent over time. Interestingly, our hypothetical query, “Olympics location,” performed in the United States in April of 2017 is returning this Google answer box result:

Does this indicate that the machine believes most people searching for this term are still more interested in the 2016 Rio de Janeiro Summer Games than they are in the next event, the 2018 Pyeongchang Winter Games? Is RankBrain succeeding here, based on patterns it has calculated, or is it still "in the works," not sure from the vagueness of our query whether we want an older, popular answer, or a fresher one that looks to the future? And what would this query return if we could perform it in January of 2018? Would the answer box show Pyeongchang because the signals surrounding the event have intensified by that time?

Because the extent and nuances of RankBrain's influence on how Google's core search algorithm works hasn't yet been fully fleshed out, one of the best ways of learning more about how RankBrain works may come from observing how often Google responds to a variety of your own queries with satisfying answers. How often to they correctly interpret *your intent*?



BERT UPDATE

BERT identifies sentences or phrases within web page content, which, although missing individual search keywords, are more relevant to the search query. BERT utilizes Natural Language Processing (NLP) to understand search terms better. NLP understands and analyzes how we interact as humans in terms of terminology. This way, BERT gets a better grasp of content searched by the user. It is important to note that BERT analyses search queries and not web pages.

Many On-Page SEO blogs are centered around keywords rather than content. BERT's unique natural language processing ability will help Google easily distinguish such content. For consumers, this will be beneficial as the quality of search query results will increase, and better content will be displayed.

There are ways to avoid your business suffering from the implementation of BERT. First ensure that your content is reliable, while keywords shouldn't be your priority. Generate top-quality content that caters to your industry's SEO needs. It helps you to revamp your copy if you are worried about BERT's effect on your visibility.