УДК 004.77

D.S.Kolesnyk, student
Svitlana Shcherbyna, Candidate of Pedagogical Sciences
Central Ukrainian National Technical University, Kropyvnytskyi

ALGORITHM, SEARCH, RANKING FACTORS, WEBSITES, WEBPAGES, UPDATES, GOOGLE

This article is dedicated to Google's search algorithm, definition of this algorithm, how does it work, stages of the search process, how many and what ranking factors the algorithm uses and for what reasons this information is not public. This article also raises the issue of algorithm updates, their differences and the principle of operation.

Using Google Search Algorithm for Internet search

Google's algorithms are a complex system used to retrieve data from its search index and instantly deliver the best possible results for a query. The search engine uses a combination of algorithms and numerous ranking factors to deliver webpages ranked by relevance on its search engine results pages (SERPs) [1].

The search process consists of three stages:

1. Crawling. The first stage involves Google's bots, also known as "spiders", crawling the web and looking for new or updated web pages. Pages need to be crawled and indexed in order to rank.

2. Indexing. This step is to analyze the content contained in URLs and try to figure out what each page is about. It does this by looking closely at the content, images, and other media files on the page, and then stores this information in a huge database known as the Google index.

3. Searching and ranking. The user enters a query, and the search engine ranks and returns content in relation to the query.

Exact information about how Google algorithms work is not public information and nobody knows outside of Google's inner circle. There are several compelling reasons for this. One of them is that if the algorithm becomes publicly available, everyone will be able to exploit it and doctor the system in their favor. As a result, this will generate useless search results for users and certainly create a worse internet. Also, one of the reasons it would greatly diminish the company's value, since the algorithm is a closely-guarded business secret, and releasing it would lead to this consequence [2].

Google keeps the details of its search ranking algorithm private but uses a lot of specific criteria to rank content. It's believed that there are well in excess of 200 ranking factors —and nobody knows them all. Even if someone knows all these factors, they will not remain relevant for long, since the algorithm is always changing. On average, Google releases updates for its update six times per day. Despite all of the above, Google still provides some information [3].

The algorithm categorizes information based on many factors, but some of the key factors, which can be found on Google's "How Search Works" page, include the following:

1. Meaning of query. The meaning of the query relates to the intent of the searcher. For Google to provide the most helpful answers, they need to fully understand what you're looking for. Google uses their proprietary language models to decipher the words in a query to match up to useful content. This includes steps such as recognizing and correcting spelling errors, and using a complex system of synonyms to allow relevant documents to be acted upon even if they do not contain the exact words being used.

2. Relevance of content. Google's systems analyze the content to assess whether it contains information that might be relevant to what is being searched for. The most basic signal for this is matching keywords in content to keywords in the query. If the page contains an exact match, that

sends a strong relevance signal to the system. In addition to assessing the types and number of keywords, ranking systems also analyze content relevance in other ways. The data is converted into signals that allow Google's machine learning algorithms to more accurately predict future relevance. The only drawback of this system as of now is that as it uses objective ranking signals, it fails to analyse subjective concepts such as the viewpoint or gist of a page's information [4].

3. Quality of content. The system looks for signals of expertise, authoritativeness and trustworthiness. One of several factors is understanding if other prominent websites link or refer to the content.

4. Usability of webpages. The algorithm prioritizes more user-friendly websites. Google also considers factors such as page loading speed, mobile responsiveness, and the overall user experience when evaluating the ranking of a web page.

5. Context and settings. The algorithm looks at information such as location, search history and search settings to return content based on the specific user profile. Search also personalizes results based on previous activity on the Google account [8].

Within these 5 categories are the aforementioned ranking factors. In addition to their complete list, the weights of each ranking factor is also unknown, but each aims to validate one of the above five principles in content. It's better for content creators to optimize for many of the factors rather than just a few. Some specific factors that they can directly measure include content organization, content length, website structure, backlinks, domain authority, meta descriptions, keywords etc. Five most prominent ranking factors can be highlighted:

1. Backlinks – backlinks, or links from other sites, show Google that other sites trust your content;

2. Freshness - refers to how "fresh," or recent, the content on webpage is;

3. Keyword mentions - keyword variations in the headings of posts, the title of posts, at least one subheading, the intro paragraph, the page's URL;

4. User experience - encompasses a lot of different things, including the page load speed, mobile-friendliness, website design etc;

5. Topical authority - means sites that have additional, valued content about queries relevant to the one being searched.

As mentioned before, Google updates its algorithm regularly. Daily updates make minor tweaks. Algorithm updates that do significantly affect the search engine results page are called core algorithm updates. Google usually doesn't announce upcoming updates. A sign that Google is doing something on its end is inexplicable drops in traffic and conversion rates.

Updates to the core algorithm usually occur once a year. Some have been given memorable names by the SEO community or by Google itself to help alert people to them, prepare for them, and make them easier to reach. The eight main algorithm updates are: Panda, Penguin, Hummingbird, Mobilegeddon, RankBrain, Medic, Bert and Helpful Content.

Panda was first introduced on February 23, 2011 and is the most popular algorithm now. It works on a permanent basis and is a full-fledged component of the Google search algorithm. The most pervasive myth about Panda is that it is about duplicate content. John Mueller (Google Search Advocate) has clarified that duplicate content is independent of Panda. Google employees have stressed that Panda encourages unique content, but this goes deeper than avoiding duplication. What Panda is looking for is genuinely unique information that provides outstanding value to users [5].

The penguin algorithm can be termed as a focused version of the panda but it relatively checks for the micromanagement which is done on the website. In 2012, Google officially launched the "webspam algorithm update," which specifically targeted link spam and manipulative link-building practices. The webspam algorithm later became known (officially) as the Penguin algorithm update via a tweet from Matt Cutts, who was then head of the Google webspam team. Google Penguin's objective is to down-rank sites whose backlinks look unnatural. This update put an end to low-effort link building, like buying links from link farms and PBNs.

Hummingbird is a significant update to Google's search algorithm introduced in 2013. It emphasises the meaning behind a user's search query rather than just the individual keywords. It

uses this information to return more relevant and accurate search results. Hummingbird is considered to be the beginning of Google's transition from keyword to topic. This update led to the fact that Google can show the user content relevant to his query, even if the text does not have an exact key, but uses, for example, synonyms or describes the necessary concept in other words. The mechanism of machine learning and artificial intelligence analyzes the semantic content of pages, the search behavior of users and gradually begins to rank higher those sites that are better suited to users [6].

Mobilegeddon, also has nicknames such as mobilepocalyse, mopocalypse, mobocalypse, it's a fairly simple but very important update to Google's algorithm. It is part of the search engine algorithm that evaluates the mobile version of websites. This update provided no gray area. Your pages were either mobile-friendly, or they weren't. There was no in-between.

RankBrain is a system by which Google can better understand the likely user intent of a search query. This can be taken as an extension of the Hummingbird algorithm which will be providing more focus on the delivery of relevant content and it is a major factor when the page ranking is considered.

Medic update made expertise and topical authority a strong ranking factor for health, wellness and personal finance sites. Many sites that had non-experts share advice on these topics dropped in SERP ranking following this update.

BERT added natural language processing capabilities to help the algorithm understand complex language in search queries. Unlike RankBrain, BERT doesn't rely on past search data to interpret intent. Since more than 15% of queries every day were new, Google needed BERT to solve the problems associated with the growing number of natural language queries. BERT is needed to understand the context and meaning of the query, regardless of the spelling and phrases used. It works by processing words in relation to other words in a sentence, rather than in order. This means BERT can understand the full meaning of a word in a search query by looking at the words that come before and after it. [4].

Helpful Content implemented to further prioritize high-quality content that is more likely to meet readers' needs. It penalizes websites that produce low-quality content or rely heavily on automation tools to create content.

The recurring theme in Google's algorithm updates is that Google wants to provide the most in-depth, authoritative, and high-quality content possible. As AI chatbots become more competent, Google can incorporate them into search and update its algorithm accordingly to handle the types of queries coming through that interface [11].

References

1. https://www.searchenginejournal.com/google-algorithm-history/
2. https://www.semrush.com/blog/google-search-algorithm/
3. https://ahrefs.com/blog/google-search-algorithm/#3-ranking-factors
4. https://intuji.com/how-does-google-search-work/
5. https://www.searchenginejournal.com/google-algorithm-history/panda-update/
6. https://searchengineland.com/8-major-google-algorithm-updates-explained-282627
7. https://project-seo.net/blog-uk/yak-pratcuyut-filtry-poshulovyh-system/
8. https://www.google.com/intl/en_us/search/howsearchworks/how-search-works/ranking-results/#meaning
9. https://web-promo.ua/ua/blog/shho-take-filtry-google-za-shho-mozhna-pid-nyh-potrapyty/#
10. https://www.pixelproductionsinc.com/how-the-evolution-of-googles-algorithms-have-filtered-the-internet/
11. https://www.techtarget.com/whatis/feature/Google-algorithms-explained-Everything-you-need-to-know