



Facts about Google and Competition

- [Overview](#)
- About Search
- [About Ads](#)
- [Better Answers](#)
- [Transparency](#)
- [Supporting a Competitive Web](#)
- How Google Search Works
- [Algorithms Rank Relevant Results Higher](#)
- [Testing and Evaluation](#)
- [Manual Control and the Human Element](#)
- [Fighting Spam](#)

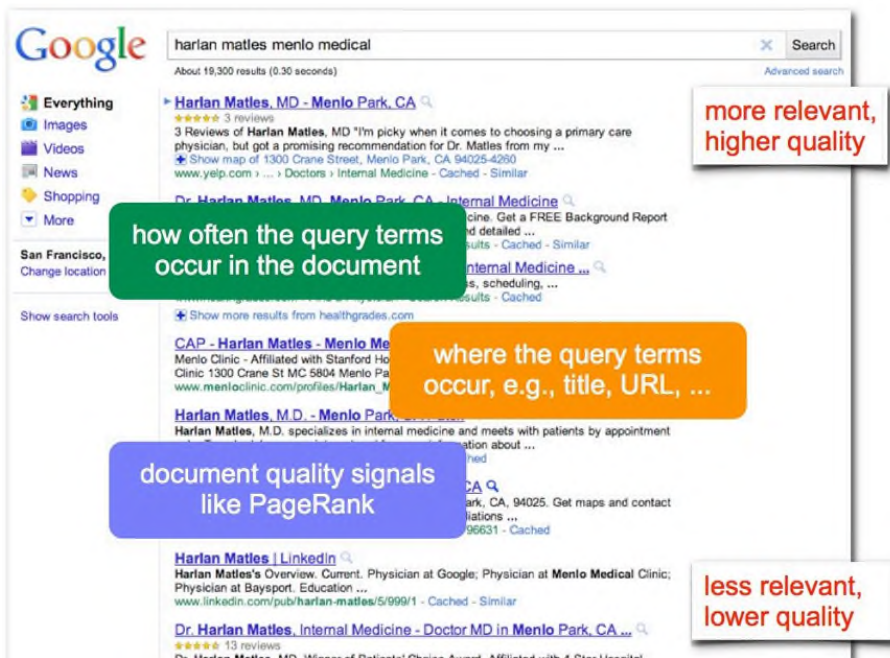


Matt Cutts explains the basics of how Google Search works.

About Search

Every day Google answers more than one billion questions from people around the globe in [181 countries and 146 languages](#). 15% of the searches we see everyday we've never seen before. Technology makes this possible because we can create computing programs, called "algorithms", that can handle the immense volume and breadth of search requests. We're just at the beginning of what's possible, and we are constantly looking to find better solutions. We have more engineers working on search today than at any time in the past.

Search relies on human ingenuity, persistence and hard work. Just as an automobile engineer designs an engine with good torque, fuel efficiency, road noise and other qualities – Google's search engineers design algorithms to return timely, high-quality, on-topic, answers to people's questions.



Our algorithms attempt to rank the most relevant search results towards the top of the page, and less relevant search results lower down the page.

Algorithms Rank Relevant Results Higher

For every search query performed on Google, whether it's [\[hotels in Tulsa\]](#) or [\[New York Yankees scores\]](#), there are thousands, if not millions of web pages with helpful information. Our challenge in search is to return only the most relevant results at the top of the page, sparing people from combing through the less relevant results below. Not every website can come out at the top of the page, or even appear on the first page of our search results.

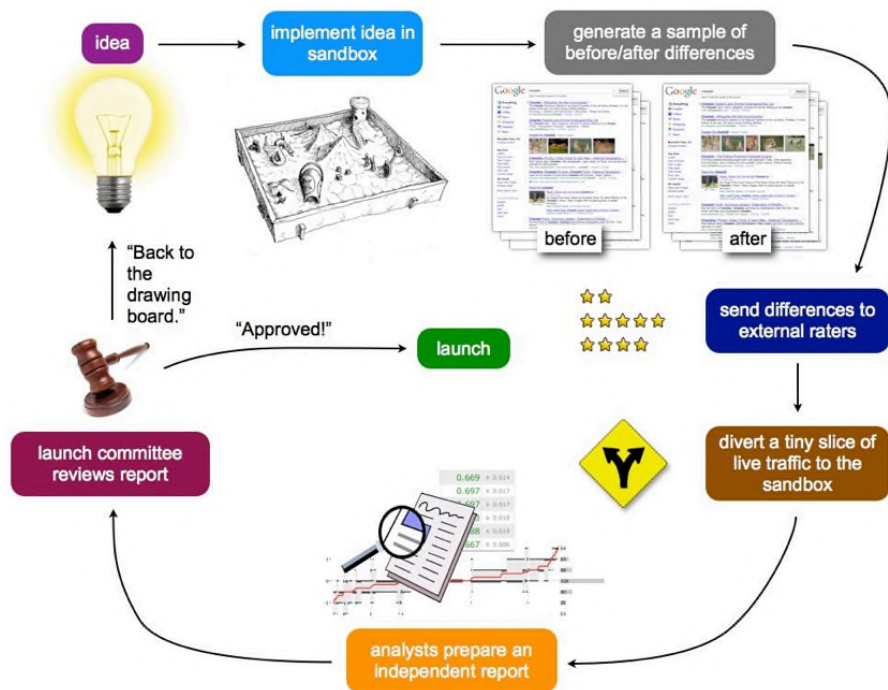
Today our algorithms rely on more than 200 unique signals, some of which you'd expect, like how often the search terms occur on the webpage, if they appear in the title or whether synonyms of the search terms occur on the page. Google has invented many innovations in search to improve the answers you find. The first and most well known is PageRank, named for Larry Page (Google's co-founder and CEO). [PageRank](#) works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The underlying assumption is that more important websites are likely to receive more links from other websites.

Panda: Helping People Find More High-Quality Sites

To give you an example of the changes we make, recently we launched a pretty big algorithmic improvement to our ranking—a change that noticeably impacts 11.8% of Google searches. This change came to be known as “Panda,” and while it's one of hundreds of changes we make in a given year, it illustrates some of the problems we tackle in search. The [Panda update](#) was designed to improve the user experience by catching and demoting low-quality sites that did not provide useful original content or otherwise add much value. At the same time, it provided better rankings for high-quality sites—sites with original content and information such as research, in-depth reports, thoughtful analysis and so on.

Market Pressure to Innovate

“[Google] has every reason to do whatever it takes to preserve its algorithm's long-standing reputation for excellence. If consumers start to regard it as anything less than good, it won't be good for anybody—except other search engines.” Harry McCracken, TIME, [3/3/2011](#)



We rely on rigorous testing and evaluation methods to rapidly and efficiently make improvements to our algorithms.

A Peek Inside

“At any moment, dozens of these changes are going through a well-oiled testing process...Every time engineers want to test a tweak, they run the new algorithm on a tiny percentage of random users, letting the rest of the site’s searchers serve as a massive control group.” – Read more from Steven Levy’s in-depth story in [Wired](#), 02/22/10

Testing and Evaluation

Google is constantly working to improve search. We take a data-driven approach and employ analysts, researchers and statisticians to evaluate search quality on a full-time basis. Changes to our algorithms undergo extensive quality evaluation before being released.

A typical algorithmic change begins as an idea from one of our engineers. We then implement that idea on a test version of Google and generate before and after results pages. We typically present these before and after results pages to “raters,” people who are trained to evaluate search quality. Assuming the feedback is positive, we may run what’s called a “live experiment” where we try out the updated algorithm on a very small percentage of Google users, so we can see data on how people seem to be interacting with the new results. For example, do searchers click the new result #1 more often? If so, that’s generally a good sign. Despite all the work we put into our evaluations, the process is so efficient at this point that in 2010 alone we ran:

- 13,311 precision evaluations: To test whether potential algorithm changes had a positive or negative impact on the precision of our results
- 8,157 side-by-side experiments: Where we show a set of raters two different pages of results and ask them to evaluate which ones are better
- 2,800 click evaluations: To see how a small sample (typically less than 1% of our users) respond to a change

Based on all of this experimentation, evaluation and analysis, in 2010 we launched 516 improvements to search.

Manual Control and the Human Element

In very limited cases, manual controls are necessary to improve the user experience:

1. [Security Concerns](#): We take aggressive manual action to protect people from security threats online, including malware and viruses. This includes removing pages from our index (including pages with credit card numbers and other personal information that can compromise security), putting up [interstitial warning pages](#) and adding notices to our results page to indicate that, “this site may harm your computer.”
2. [Legal Issues](#): We will also manually intervene in our search results for legal reasons, for example to remove child sexual-abuse content (child pornography) or copyright infringing material (when notified through valid legal process such as a [DMCA takedown request](#) in the United States).
3. Exception Lists: [Like the vast majority of search engines](#), in some cases our algorithms falsely identify sites and we sometimes make limited exceptions to improve our search quality. For example, our SafeSearch algorithms are designed to protect kids from sexual content online. When one of these algorithms mistakenly catches websites, such as essex.edu, we can make manual exceptions to prevent these sites from being classified as pornography.
4. [Spam](#): Google and other search engines publish and enforce guidelines to prevent unscrupulous actors from trying to game their way to the top of the results. For example, our guidelines state that websites should not repeat the same keyword over and over again on the page, a technique known as “keyword stuffing.” While we use many automated ways of detecting these behaviors, we also take manual action to remove spam.

The Engineers Behind Search

“So behind every algorithm, and therefore behind every search result, is a team of people responsible for making sure Google search makes the right decisions when responding to your query. Obviously, there’s no other way it could have happened: Google is a living example of what’s possible when brilliant people devise a smart algorithm and marry it to limitless computing resources.” – Tom Krazit, The human process behind Google’s algorithm, [CNET](#), [09/07/10](#)



Matt Cutts explains how Google deals with spam through a combination of algorithms and manual action, and how websites can request reconsideration of their sites.

Fighting Spam

Ever since there have been search engines, there have been people dedicated to tricking their way to the top of the results page. Common tactics include:

- Cloaking: In this practice a website shows different information to search engine crawlers than users. For example, a spammer might put the words “Sony Television” on his site in white text on a white background, even though the page is actually an advertisement for Viagra.
- Keyword Stuffing: In this practice a website packs a page full of keywords over and over again to try and get a search engine to think the page is especially relevant for that topic. Long ago, this could mean simply

ST Reproductive Technologies, LLC EX2019

Smartec, Inc. v. ST Reproductive Technologies, LLC; No. IPR2024-00875

repeating a phrase like “tax preparation advice” hundreds of times at the bottom of a site selling used cars, but today spammers have gotten more sophisticated.

- Paid Links: In this practice one website pays another website to link to his site in hopes it will improve rankings based on PageRank. PageRank looks at links to try and determine the authoritativeness of a site.

Today, we estimate more than one million spam pages are created each hour. This is bad for searchers because it means more relevant websites get buried under irrelevant results, and it’s bad for legitimate website owners because their sites become harder to find. For these reasons, we’ve been working since the earliest days of Google to fight spammers, helping people find the answers they’re looking for, and helping legitimate websites get traffic from search.

- More:
- [Questions and Answers](#)
- [What People Are Saying](#)
- [Google’s Economic Impact](#)

© Google