(https://www.searchenginewatch.com)

SEO (https://www.searchenginewatch.com/category/seo/) > Matt Cutts on
How Google Tests Its Algorithms

# Matt Cutts on How Google Tests Its Algorithms

Have you ever been curious about how Google decides which algorithm is better than another, when they're pushing out one of the many tweaks they do weekly? Google's Matt Cutts spills the beans on how the search team actually does it.

**Author**
**Jennifer Slegg**
**(https://www.searchenginewatch.com/author/jennifer-slegg/)**

**Date published**
April 3, 2014

**Categories**
**SEO (https://www.searchenginewatch.com/category/seo/)**

(https://www.searchenginewatch.com)

Have you ever been curious about how Google decides which algorithm is better than another, when they're pushing out one of the many tweaks they do weekly? How do they judge which tweak actually produces better results and which produces lots of good results? Or does the spam team just wave a nerf bat over the server before hitting a big red button and hope for the best?

Google's Matt Cutts spills the beans on how the search team actually does it in a webmaster help video, which asks what metrics Google (https://searchenginewatch.com/2019/07/18/how-to-set-up-google-analytics-annotations-to-show-google-updates/) uses to evaluate whether one iteration of the ranking algorithm is delivering better quality results to users than another.

While Cutts starts off saying that he could geek out on this topic for quite some time, and I'm sure many of us would love to hear him do just that, he said he will try and hold back for the sake of video length.

"Whenever an engineer is evaluating a new search quality change, and they want to know whether it's an improvement, one thing that's useful is we have hundreds of quality raters who have previously rated URLs as good or bad, spam, all these sorts of different things.

"So when you make a change, you can see the flux, you can see what moves up and what moves down, and you can look at example searches where the results (https://www.searchenginewatch.com) changed a lot for example," he said. "And you can say OK, given the changed search results, take the URLs that moved up, were those URLs typically higher rated than the URLs that moved down by the search quality raters?"

While Google tries to keep the specifics of their quality rater guidelines secret, they inevitably end up getting leaked. The most recent version (https://www.searchenginewatch.com/article/2308060/Googles-Your-Money-or-Your-Life-Search-Quality-Ratings-What-Webmasters-Need-to-Know) became known in November and detailed exactly what quality raters are looking for when they rate search results.

"Sometimes since these are precomputed numbers, as far as the ratings, we've already got a stored data bank of all those ratings from all the raters that we have, sometimes you'll have question marks or empty areas where things haven't been rated," he said. "So you can also send that out to the raters, get the results either side-by-side, or you can look at the individual URLs, and they say in a side-by-side this set of search results is better, or this set is better or they might say this URL is good in this URL is spam, and you use all of that to assess whether you're making good progress."

While it is good that Google pushes these kinds of things that the quality raters to see what they notice, it doesn't always catch everything. There definitely been times when new tweaks to break something, such as what we saw with entertainment sites that significantly declined in the rankings (https://www.searchenginewatch.com/article/2332225/So-Why-Did-All-Those-Movie-Sites-Lose-Google-Search-Traffic) in February, the quality raters don't always catch.

"If you make it further along, and you're getting close to trying to launch something, often you'll launch what is called a live experiment where you actually take two different algorithms, say the old algorithm and the new algorithm, and you take results that would be generated by one and then the other and then you might interleave them. And then if there are more clicks on the newer set of search results, then you tend to say you know what, this new set of search results generated by this algorithm might be a little bit better than this other algorithm.

This is interesting how he is describing interleaving the two sets of search results, as normally we hear about either full pushes, or pushes to a small percent of
(https://www.searchenginewatch.com) users. However this could be a live experiment limited strictly to Google employees and quality raters.

---

## Resources

**Analytics Searchmetrics Core Web Vitals Study (https://www.searchenginewatch.com/resources/core-web-vitals-study/)**

**Analytics The Ultimate Guide to Forum Link Building in 2020 (https://www.searchenginewatch.com/resources/crowdo-organic-link-building-service/)**

**Analytics Top 5 SEO Mistakes (https://www.searchenginewatch.com/resources/top-5-seo-mistakes/)**

**Development Improving SEO & Developer Relationships (https://www.searchenginewatch.com/resources/improving-seo-developer-relationships/)**

He does say that from within Google, the web spam team is metrics can look quite different from the rest of Google, simply because they like to click on spam and see what's ranking, why it is ranking, and to better figure out how to get rid of it.

"Sometimes our metrics look a little bit worse in web spam because people click on the spam, and we're like we got less spam, and it looks like people don't like the algorithm as much," he said. "So you have to take all those ratings with little bit of a grain of salt, because nothing replaces your judgment, and the judgment of the quality launch committee."

The quality launch committee is actually not that well known, but it is simply a group of the search quality engineers that receives reports and has meetings regarding search quality, something which Matt has mentioned at least once in previous webmaster help videos.

He continues by talking a little bit about what exactly the quality raters are looking for when they're doing their ratings. (https://www.searchenginewatch.com)

"People can rate things as relevant on a scale, they can rate things as spam, they can even rate the quality of the page, which it sort of does it matter based on the query, but how reputable the page is itself," Cutts said. "And then we have metrics that blend all of that together and when we're done we say okay in general we think that the results got a little bit better, and here the kinds of ways they got better or worse. We can even slice and dice and look at different countries or different languages, that sorts of stuff. So in web spam we're not that surprised if users continue to click on the spam, because we can recognize the spam, we have expert raters on those kinds of topics. And we pay special attention to special countries where we know there's more spam and so we can see the sorts of reactions we get there."

Even continues and talks a bit about how every so often they go through and update the quality rating guidelines, something we've seen updated multiple times over the years.

"So we've got it down to a pretty good system," Cutts said. "Every so often we have to revive a process and look at how to improve it, but for the most part things up relatively well in terms of assessing what are the big changes, once you see those big changes it gives you ideas to go back and improve and make things better, and by the time we get to launch committee, normally everybody has a pretty good idea about whether it works and what the strengths and weaknesses of a particular algorithm are.

So if you had visions of Matt Cutts sitting in his office with a big red button on his desk to unleash some new algorithm without any feedback or oversight, you are probably disappointed. There is actually a lot that goes on with testing algorithms, particularly the large ones, and they do get put through the ringer before they go live to remain, to ensure that Google is serving up better search results than the previous algorithm.