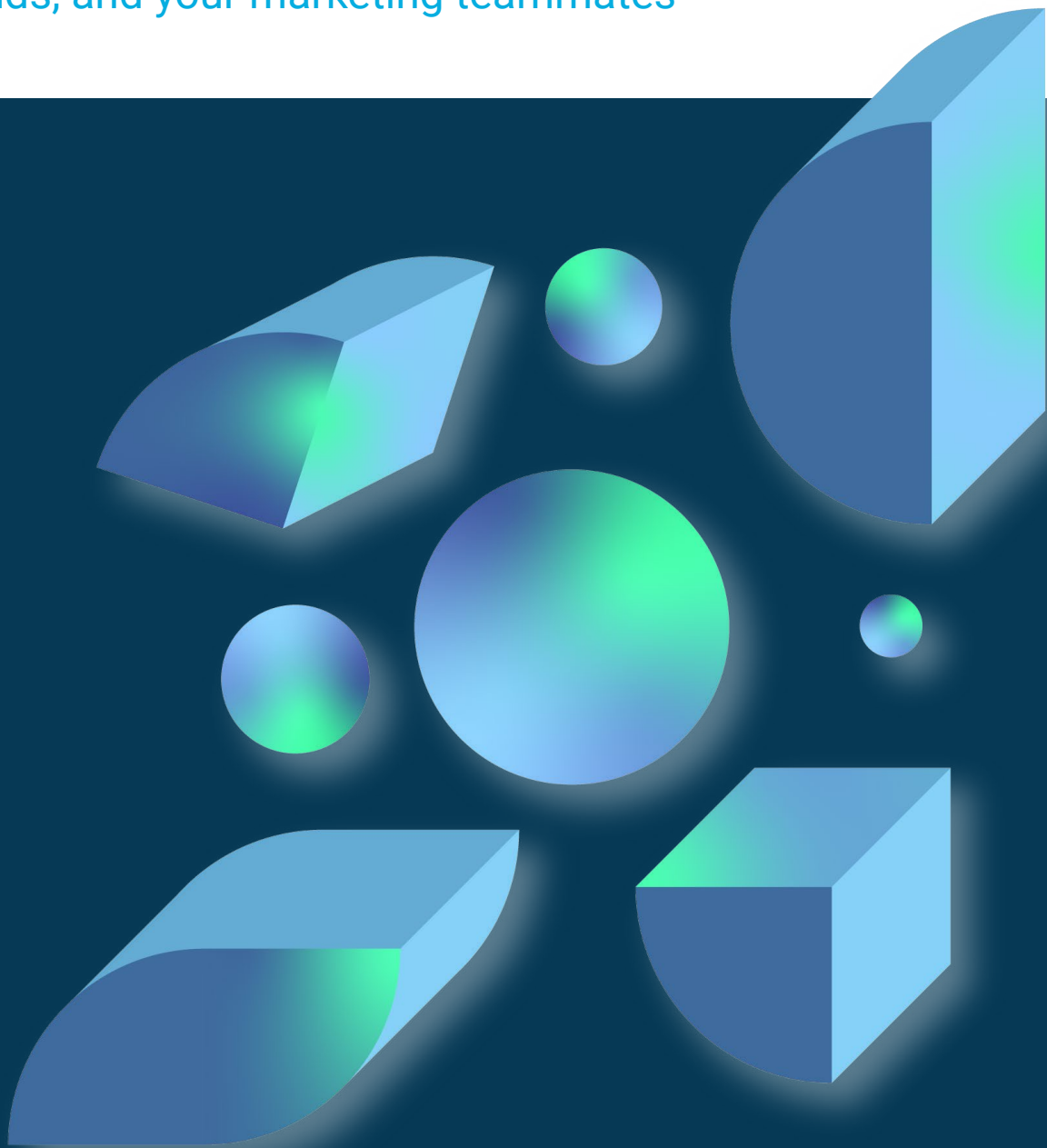


# TECHNICAL SEO FOR NON-TECHNICAL THINKERS

When, why and how technical SEO can help your site

Everything you've always wanted to share with your clients, your friends, and your marketing teammates



Oncrawl

# TABLE OF CONTENTS

<b>03</b>	<b>Introduction</b>
<b>05</b>	<b>Contributors</b>
<b>08</b>	<b>What do search engines need to access to create a search listing?</b>
<b>09</b>	Title tag
<b>10</b>	URL or slug
<b>11</b>	Search listing text
<b>14</b>	Additional rich elements in a search listing
<b>21</b>	Additional elements on a search results page
<b>24</b>	<b>How do search engines find your page?</b>
<b>27</b>	<b>How do search engines evaluate your page?</b>
<b>28</b>	Can Google reach your page?
<b>31</b>	Is your page the right version of the content?
<b>22</b>	Did you give Google any other instructions about your page?
<b>33</b>	Is your site reliable?
<b>35</b>	Is your content authoritative on the web?
<b>38</b>	Is the page important to your website?
<b>40</b>	<b>How do search engines know which search queries to show your page for?</b>
<b>41</b>	Semantic understanding
<b>47</b>	User intent signals
<b>50</b>	Linking structure
<b>50</b>	Search analytics and predictions
<b>52</b>	<b>What some other goals we can accomplish using technical SEO?</b>
<b>63</b>	<b>What are the tools used in technical SEO?</b>
<b>69</b>	<b>OnCrawl: the technical SEO platform</b>

# INTRODUCTION

Technical SEO is the hands-on, under-the-hood version of SEO. If you're reading this, it's because you or someone you're working with is interested in finding SEO solutions based on the how and the why of how search (and websites) work.

“ *Technical SEO is an assurance, from the technical point of view, that your published content is indexable by the search engines that operate ... because they aren't perfect.*

**-Detlef Johnson - SEO in Orbit, Episode 1**

This ebook will look at the type of questions technical SEO asks, the sort of answers it aims to provide, and how these drive a website's SEO performance.

“ *Making technical SEO accessible to non-technical individuals is EVERYTHING!*

*Murky technical knowledge is not only a disservice to your client/boss but it hinders your ability to move the needle and grow professionally.*

*While it can seem intimidating, most concepts are fairly easy to grasp. Understanding how a website and crawlers works help to make everything click and gives you a confidence boost.*

*You owe it to yourself to level up your technical knowledge.*

**-Britney Muller**

**CONTRIBUTORS**



### **Hamlet Batista - CEO at RankSense**

Prior to founding RankSense Inc, Hamlet co-invented Altruik's Page Priority®, an enterprise-level SEO software product that pioneered automated resolution of site architecture issues. On average, clients experienced a double-digit increase in sales year over year from organic search and a patent was filed for this invention. Hamlet also created the patented software RankSense®, an SEO product focused on simplifying SEO for small businesses, which earned WebHost Magazine's prestigious Editor's Choice Award in December 2008.

Additionally, Hamlet is an award-winning blogger and conference speaker and champions of the use of programming and machine learning in SEO.



### **Dave Davies - CEO at Beanstalk Internet Marketing**

Dave has 19 years of SEO experience and an intimate knowledge of search engines, internet marketing, and even a fair bit about web dev and server admin.

He is specialized in SEO, entities, copywriting, stats analysis, and PPC.



### **Marie Haynes - Owner at Marie Haynes Consulting**

Marie's career in SEO began in 2008. She became a recognized leader in tackling Google's Penguin algorithm as very few people understood it. She began giving advice on SEO forums, and her knowledge was in high demand. She began analyzing websites and provided feedback for clients based off of the forums she participated in.

With more demand than she could maintain on her own, she hired and trained a team of SEO's that now comprise Marie Haynes Consulting Inc. She and her team edit a rich monthly SEO newsletter and podcast that has become an industry standard.



### **Britney Muller - Senior SEO Scientist at Moz, Keynote Speaker, Founder at Pryde Marketing**

Britney Muller is a proud Minnesota native with a degree in Strategic Marketing and Public Relations. Britney began her work in multi-platform marketing in 2007 and founded Pryde Marketing in 2012. At Moz, she is currently responsible for optimizing technical SEO, staying up to date on Google algorithm changes.

Britney is interested in machine learning and data science, Python, and snow boarding. She loves getting people excited about big data, Tensorflow, data science, technical SEO and working smarter not harder.



### **Max Prin - Head of Technical SEO at Merkle**

Max Prin is the Head of Technical SEO at Merkle. The SEO technical team ensures the accuracy, feasibility and scalability of the agency's technical recommendations. Max supports Merkle's SEO department by developing SEO best practices, tools and processes through research and testing. Max is also in charge of developing and publishing TechnicalSEO.com.

Max focuses exclusively on technical SEO: from fixing crawling, rendering and indexing issues to optimizing page speed and relevance through structured data markup.



### **Vincent Terrasi - Data Marketing Director at OVH**

At OVH, Vincent has been involved in Datalake, CRM, Analytics, and Data Science, managing machine learning, business intelligence, and CRM teams working on SEO and big data projects.

A fan of computers and the web since he was little, Vincent studied computer science at a French engineering university. Following 7 years as an entrepreneur working on his own sites, he joined Oxygem in 2014 and quickly concentrated on specializing in machine learning, strengthening his SEO skills. With the acquisition of Oxygem by the Groupe M6, he became SEO Director, managing websites with over 15 million monthly visitors.

WHAT DO SEARCH  
ENGINES NEED  
TO ACCESS TO  
CREATE A SEARCH  
LISTING?



## WHAT DO SEARCH ENGINES NEED TO ACCESS TO CREATE A SEARCH LISTING?

Approximately half of on-page SEO -- particularly technical on-page SEO -- aims at making sure Google has the information it needs to create a search listing.

Technical SEO will audit the elements that go into building a search listing, which helps SEOs to check a large number of pages at once, and to identify and resolve strategic problems first.

Here is a basic search listing:

This Is A Title Tag - You Can ChAnGe HoW it lOOks in search

<https://example.com/your-url-goes-here>

Most snippets are limited to two lines, and we still generally recommend a 160-character limit for your Meta descriptions, but Google can allow exceptions in certain cases.

*Source: Britney Muller ( THE BEGINNER'S SEO CHEATSHEET )*

### Title tag

Title tags provide search engines with the title you want to give your page. In many cases, this is the blue title that will end up as the heading of your search listing. This information appears in the head section of the URL of your page:

```
<head>
  <title>Your page title</title>
</head>
```

## WHAT DO SEARCH ENGINES NEED TO ACCESS TO CREATE A SEARCH LISTING?

### What makes a good title?

- Something people want to click on
- A good summary of your page content containing your main keyword
- Contains your brand or company name
- Approximately 50-60 characters long
- Unique titles that aren't used on other pages on your site

### What technical SEO can tell you:

- Whether the theme of the content is reflected in the title tag
- The length of your title tags
- Whether or not your title tags are unique
- Whether any of your pages are missing a title tag

### URL or slug

The green text, the URL, or slug, indicates where on the web your page is located.

Recent information suggests that Google may be phasing out the appearance of green URLs on the search results pages. However, the URL has traditionally had value both for search engines and for humans. Search engines used to (but no longer do) use the structure of the URL to understand how the page fits into your site architecture, or the words in the URL to reinforce the theme of the page content.

“

URL structure has little value, use internal linking instead.

**-John Mueller, Google Webmasters Hangout - 3 May 2019**

Today, the URL reassures users that they're going to the site they think they will. It helps establish content authority, and confirms to users that the page is pertinent to their search.

## WHAT DO SEARCH ENGINES NEED TO ACCESS TO CREATE A SEARCH LISTING?

### What makes a good URL?

- Easy to read for humans. Instead of using numerical categories and page IDs, use words that have meaning for people searching for the page's content
- Avoid "stop words": words that contain little information, such as articles, conjunctions, and prepositions
- Short and simple

### What technical SEO can tell you:

- Whether your URL contains words
- Whether your URL contains stop words
- The length of your URL
- Whether a single page has multiple URLs

### Search listing text

The text of the search listing is usually drawn from or based on your meta description. Like the title tag, the meta description occurs in the head of the HTML of the page:

```
<head>
<meta description="A description of what your
page is about. This summary should accurately
represent the page content and make searchers want
to click.">
</head>
```

Google's algorithms often rewrite meta descriptions using content from the page in order to better adapt the search listing to the search query. Often, Google will use the first line of text on the page, or in the section that appears in the search results.

## WHAT DO SEARCH ENGINES NEED TO ACCESS TO CREATE A SEARCH LISTING?

### What makes a good description?

- An appropriate length. Description lengths have been adjusted multiple times in the last several years. The current suggestion is to write approximately 160 characters of content. Google's advice states that there is no limit, but that descriptions are truncated for "device width".
- Descriptions on sites that have a description for every page.
- Unique descriptions that aren't used on other pages of the site.
- Representative of page content.
- Contain specific facts or structures.

For example, the following meta description provides detailed information about a book.

```
<meta name="Description" content="Written by A.N. Author,  
Illustrated by V. Gogh, Price: $17.99,  
Length: 784 pages">
```

In this example, information is clearly tagged and separated.  
[-Google Webmasters Documentation](#)

## WHAT DO SEARCH ENGINES NEED TO ACCESS TO CREATE A SEARCH LISTING?

### What technical SEO can tell you:

- Whether the theme of the content is reflected in the meta description
- The length of your meta descriptions
- Whether or not your meta descriptions are unique
- Whether any of your pages are missing a meta description

You can, alternatively, prevent snippets from being created and shown for your site in Search results. Use the <meta name=»nosnippet»> tag to prevent Google from displaying a snippet for your page in Search results.  
[-Google Webmasters Documentation](#)

## WHAT DO SEARCH ENGINES NEED TO ACCESS TO CREATE A SEARCH LISTING?

### Additional rich elements in a search listing

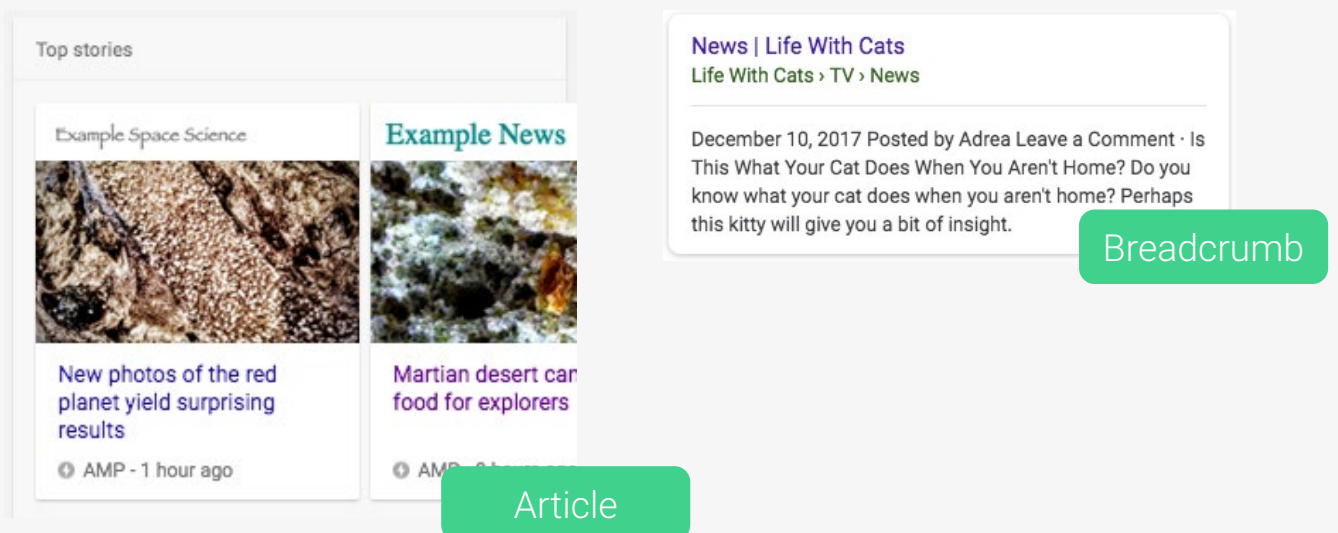
Another significant part of technical SEO aims at providing the technical underpinnings that allow search engines to create a great search listing that attracts clicks by searchers.

Structured data can add additional elements to your search listing. These make your search listing more appealing, and can increase the rate at which searchers click on it.

Structured data provides machine-readable information on your page for search engines. Structured data is often used to refer to Schema.org mark-up for entities, or things with properties that are the subject of pages on your site.

Google provides enrichments for a number of different types of data:

<https://developers.google.com/search/docs/guides/search-gallery>



**To Kill a Mockingbird**  
Novel by Harper Lee

SUMMARY QUOTES **GET BOOK** MORE BY AUTHOR TOP

Google Play Books  
\$12.99 · ebook [Get book](#)

Barnes & Noble  
\$12.99 · ebook [Get book](#)

**BORROW EBOOK**



Check availability at libraries near you

Enter a city or zip code

No results found near you. Try another location.

Book (only available for «broadly available» publications)

Recipes

**Perfect Apple Pie**  
Pillsbury  
4.5 ★★★★★ 276 reviews  
3 hr · 230 calories

**Apple Pie by Grandma Pillsbury**  
Allrecipes

Carousel: a list with multiple items of the same type (article, recette, course)

**Nest Labs**  
Company

Follow

Nest Labs is a home automation company headquartered in the San Francisco Bay Area. The company designs and manufactures smart home products, including smart thermostats and smoke detectors. Wikipedia

Customer service: 1 (855) 469-6378 (Consumer)  
Tony Fadell, Matt Rogers, Tony Fadell

Recent posts

Want a Nest? Get it delivered today if you live in the Bay Area. Nest Learning Thermostat and Nest Protect. Smoke + CO are now available on Google Shopping. Mar 7, 2014

People also search for

Google Honeywell Apple Inc. Control4 ecobee

Nest Customer Service

1 (855) 469-6378 (Consumer)  
Nest Labs, Customer service

**Nest Support**  
<https://support.nest.com/>  
Customers have been asking Thermostat or Nest Protect Smoke + CO customer service. We've created a guide to help you get the most out of your Nest products. Visit our website for more information. Troubleshooting the thermostat. A step-by-step guide to solve your problem.

**Nest Customer Phone Number** (Shortest Wait) | Best Support  
gethuman.com: Find a phone number  
This is the Nest phone number with the shortest wait time and best customer service, as ranked by the millions of customers who have given tips and their scores.

**Contact Nest** (Shortest Wait) | Best Support | GetHuman.com  
gethuman.com: Find Contact Information  
This is the fastest way to contact Nest. Whether by phone, email, chat or otherwise, we'll get you the best support and quality of customer service.

**Nest and Google's Customer Service Problem** - Dan Pallotta  
https://danpallotta.com/2014/03/04/nest-and-googles-customer-service-problem/

Corporate contact

**Android Courses | Coursera**  
Coursera > courses > query=android

**Android App Development**  
This Specialization enables learners to successfully apply co...  
Vanderbilt University

**Java for Android**  
This MOOC teaches you how to program core features and cl...  
Vanderbilt University

**Build Your First Android App (Project-Centered Cour...**  
What you'll achieve: In this project-centered course\*, you'll de...  
CentraleSupélec

Course

Google Dataset Search

weather.snoaa.gov

**Monthly Weather Review**  
data.noaa.gov  
catalog.data.gov  
Updated May 2, 2013

**World Weather Records**  
data.noaa.gov  
Published May 31, 2017

**Mariners Weather Log**  
data.noaa.gov  
catalog.data.gov  
+1more  
Published 1957

**Daily Weather Records**  
data.noaa.gov  
catalog.data.gov  
+1more  
Published Dec 1, 2013

**Surface Weather, Signal Service and Weather Bureau**  
data.noaa.gov  
catalog.data.gov  
+1more  
Published 2011

**Monthly Weather Review**  
gov.noaa.nodc:CO1044  
data.nodc.noaa.gov catalog.data.gov

Dataset created Mar 15, 2011  
Dataset updated Mar 2, 2013  
Dataset published Mar 15, 2011

Dataset provided by  
National Oceanic and Atmospheric Administration

Time period covered 1914 - 1949

Area covered  
United States of America, Pacific Ocean, North Pacific Ocean

Description  
Supplements to the Monthly Weather Review publication. The Weather Bureau published the Monthly weather review Supplement irregularly from 1914 to 1949. The Supplement replaced numerous independent series of bulletins that the Bureau published before 1914. The Supplements featured contributions to the science of meteorology and weather forecasting that were too voluminous to publish in the regular Monthly weather review. The Bureau never published no. 43. The Monthly Weather Review series has which assumed publication in 1974.

Dataset (in beta)

**CRITIC REVIEWS**

Like Taika Waititi before him, Ryan Coogler gives the Marvel template a bold auteurist twist with an African extravaganza that packs a muscular intensity and challenges as much as it exhilarates.

**EMPIRE** Jimi Famurewa  
Empire 4/5

Virtually everything that distinguishes "Black Panther" from past Marvel pics works to this standalone entry's advantage.

**Variety** Peter Debruge  
Variety

**MORE CRITIC REVIEWS**

Critic review (available for local businesses, films, books)



## Google

Glassdoor

★★★★★ 4.4 - 7,167 reviews



Indeed

★★★★★ 4.3 - 2,296 reviews



kununu

★★★★★ 4.1 - 218 reviews



Employer Aggregate Rating  
(will soon replay «Review» for companies  
that have job postings)

## Jazz concerts

● Near you [Edit location](#)

This weekend

Free

Within 0.5 mi

[FILTERS](#)

**24**  
APR  
**Jan Lieberman Concert**  
**Series: Journey in Jazz**  
Sat, 7:30–11 PM  
Central Park Library, Santa C...



FREE

**26**  
APR  
**Anton Schwarz Quintet**  
Sat, 7–10 PM  
Portolo Vineyards, Portola V...



\$90+

**27**  
APR  
**Contemporary Jazz**  
**Orchestra**  
Fri, Apr 27–Sun, Apr 29  
1015 Folsom Street



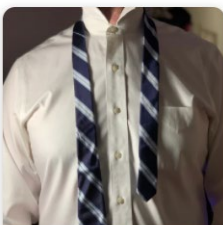
\$84

Event

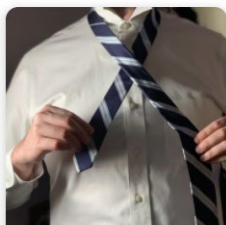
## How to Tie a Tie

[Example.com](#) > [fashion](#) > [neckwear](#)

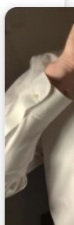
The four-in-hand knot is a great look for any occasion. From formal suits to casual jacket and tie affairs, it's a simple way to quickly look great. Once you have mastered the knot you can tie it in just a minute, so it's also great when you're in a pinch too.



1. Make one end  
about your neck...



2. Cross the short  
end over the...



3. Cro  
your b

How to



the world is flat



[ALL](#)

[NEWS](#)

[IMAGES](#)

[VIDEOS](#)

[MAPS](#)

[BOOKS](#)

**Square World Society Insists the World is Flat - Example.com**

[www.example.com](#) > [square-world-](#)

Claim: "The world is flat; you can sail over the edge"

Claimed By: Square World Society Press Corps

Fact check by Example.com: False

Fact Check  
( = ClaimReview)

## Example Site - Frequently Asked Questions(FAQ)

[Your site title here](#) > [support](#) > [answer](#)

Welcome to our support center. Here you can find the most frequently asked questions. If you cannot find an answer below, please call us at 1-800-555-5555.

What is the return policy?



How long does it take to process a refund?



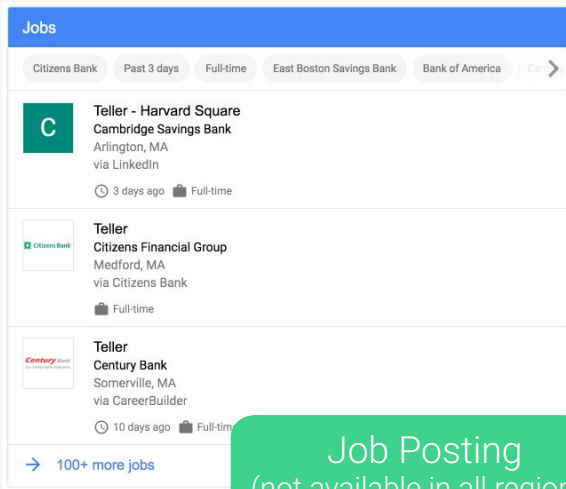
What is the policy for late/non-delivery of items ordered online?



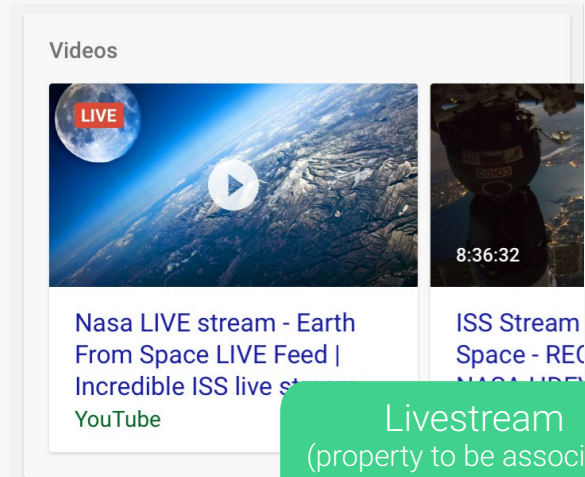
Show more

FAQ

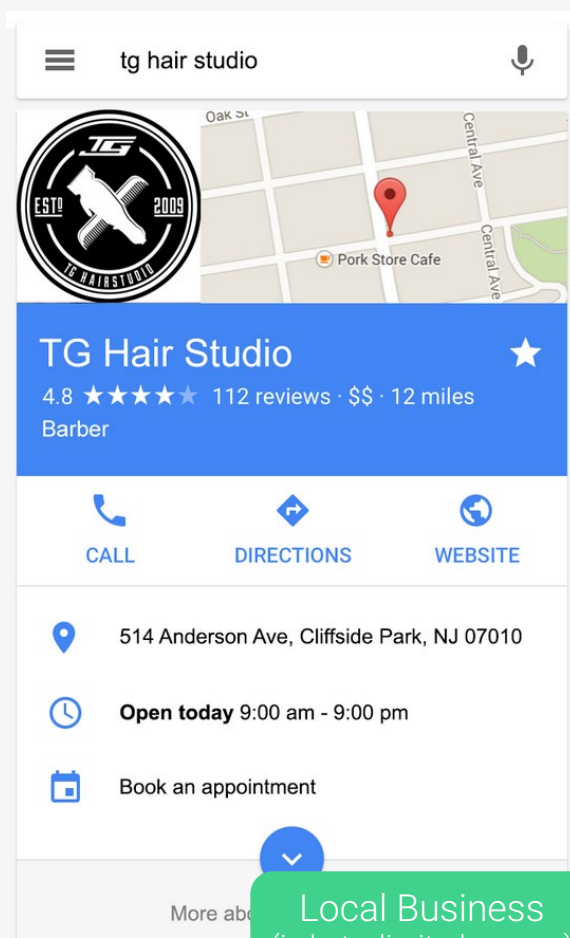




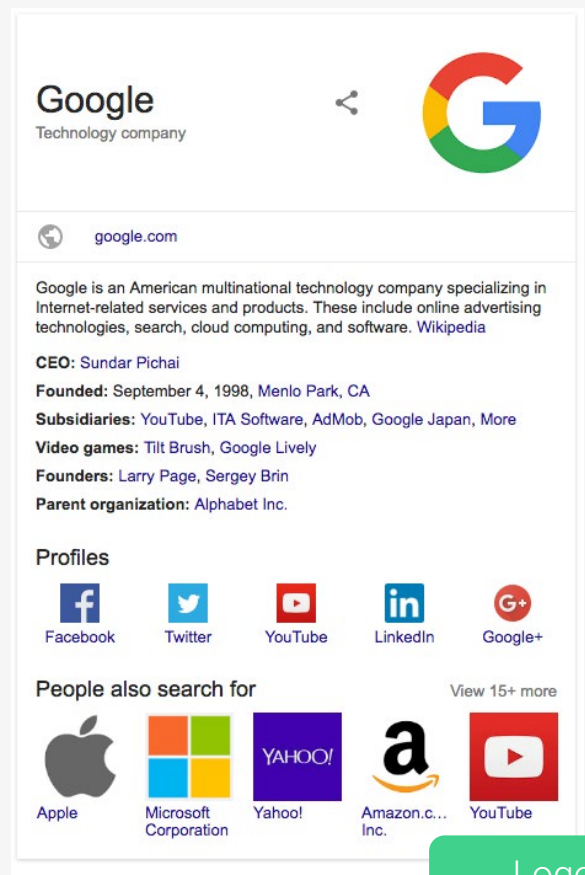
Job Posting  
(not available in all regions)



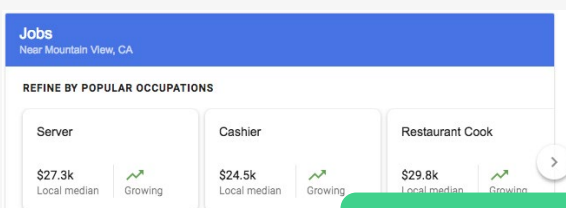
Livestream  
(property to be associated with VideoObject)



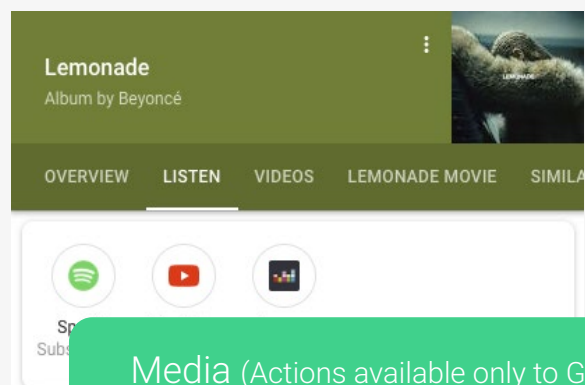
Local Business  
(in beta, limited access)



Logo



Occupation



Media (Actions available only to Google partners that provide Google Search/Google Assistant applications)

How do I remove a cable that is stuck in a USB port ...

Example.com > questions

Nov 28, 2012 · 4 answers

I inserted the plug from my USB cable that came with my phone into a USB hub. When I tried to unplug the cable, it seemed impossible ...

✓ **Top answer** · 2 votes

I had a similar issue recently with my phone. The cable wouldn't come out, so I grabbed my utility multi-tool ... [More](#)

0 votes

I used an eyelash better grip on the sticking out [More](#)

Q&A page

Lava Lite Classic Lava Lamp, Purple/Blue - Walmart.com

Walmart > Lava-2118-Lava-Lite-Classic...

★★★★★ Rating: 3.5 - 60 votes - \$13.17 - In stock

Mobile-friendly - Buy Lava Lite Classic Lava Lamp, Purple/Blue at Walmart.com.

Product  
(price, availability, rating)



The Nightingale by Kristin Hannah – Reviews ...

Goodreads app - Installed

★★★★★ Rating: 4.5 - 159,018 votes

The Nightingale has 159018 ratings and 21386 reviews. Emily May said: "Oh, for heaven's ... [More](#) lists with this book.

[Open on goodreads.com](#)

Review snippet

Recipes



Perfect Apple Pie  
Pillsbury

4.5 ★★★★★ 276 reviews  
3 hr · 230 calories



Apple Pie by Grandma  
Allrecipes

4.8 ★★★★★ 6,931 reviews  
1 hr 30 min · 512 calories

Recipe

Cut the Rope FULL FREE - Android Apps on Google Play

<https://play.google.com/store/apps/details?id=com...ctr...>



★★★★★ Rating: 4.5 - 1,116,042 votes - Free  
Cut the rope to feed candy to  
downloads around the world  
and more to come!

Software App  
(in beta, not globally available)

NASA

<https://www.nasa.gov/>

NASA.gov brings you the latest news, images and videos from America's space agency, pioneering the future in space exploration, scientific discovery and ...

Results from nasa.gov

NASA TV

NASA TV Programming: 11:00 p.m..  
Earth Views High Definition ...

International Space Station

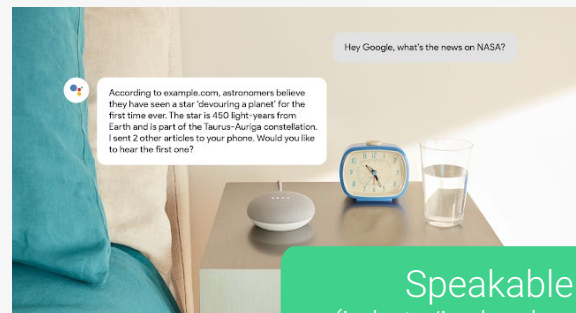
Space Station Images - Overview -  
Crews and Expeditions - HDEV

NASA Live

Watch live broadcasts from NASA  
Television and NASA's ...

For Students

Sitelinks searchbox



Speakable  
(in beta/in development)

ON THESE LISTS

timeout.com

The best Greenwich Village restaurants in NYC

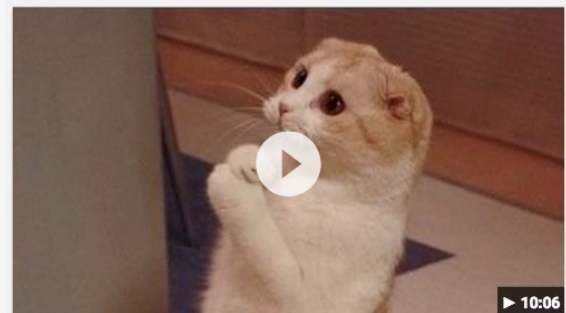
timeout.com

A list of all the Michelin-starred restaurants in NYC

zagat.com

The Best Restaurants Near Washington Square Park

Top places (list of sites, recommendations to be included in the SERPs upon request made to Google)



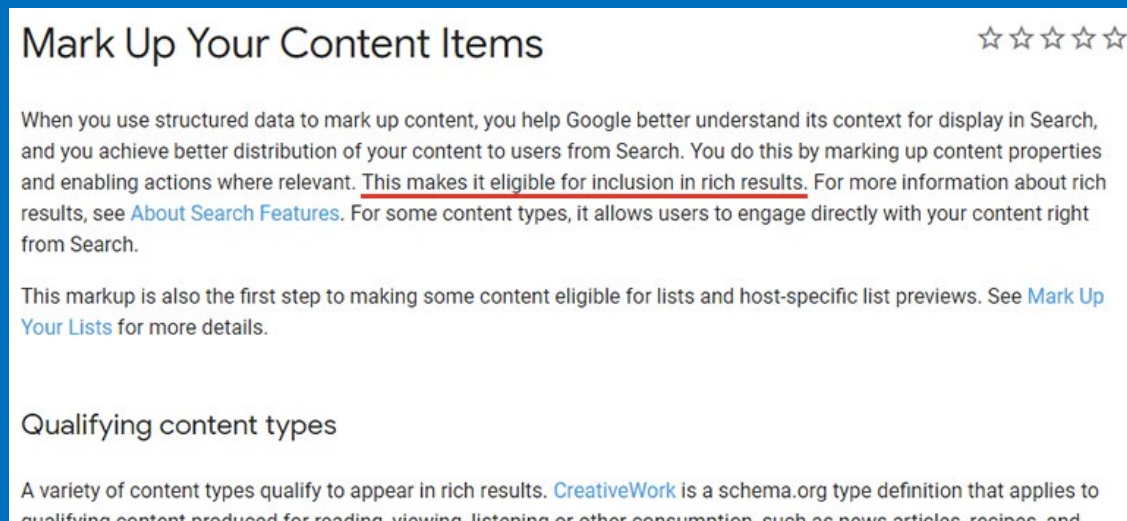
Cats are so funny - Funny cat compilation - ...  
[https://m.youtube.com > watch](https://m.youtube.com/watch)

Video

“

It's often best to go to the horse if you want to hear it from its mouth. So to answer the question as to whether technical SEO assists in attaining rich results one is well-advised to seek the answer from Google.

On their page on markup they write:

A screenshot of a Google search result for 'Mark Up Your Content Items'. The title is 'Mark Up Your Content Items' with a five-star rating to its right. The main text explains that using structured data helps Google understand content context and improves distribution to users. A sentence is underlined: 'This makes it eligible for inclusion in rich results.' It also mentions 'About Search Features' and 'Mark Up Your Lists'. A section titled 'Qualifying content types' follows, stating that various content types qualify for rich results and mentions 'CreativeWork' as a schema.org type definition.

**Mark Up Your Content Items** ☆☆☆☆☆

When you use structured data to mark up content, you help Google better understand its context for display in Search, and you achieve better distribution of your content to users from Search. You do this by marking up content properties and enabling actions where relevant. This makes it eligible for inclusion in rich results. For more information about rich results, see [About Search Features](#). For some content types, it allows users to engage directly with your content right from Search.

This markup is also the first step to making some content eligible for lists and host-specific list previews. See [Mark Up Your Lists](#) for more details.

**Qualifying content types**

A variety of content types qualify to appear in rich results. [CreativeWork](#) is a schema.org type definition that applies to qualifying content produced for reading, viewing, listening or other consumption, such as news articles, recipes, and

To be clear, markup is not a strict necessity in every case but it's rare to see instances where it was not used and rich snippets attained. As the heaviest and most easily wielded weapon in the rich snippets toolbox it is basically become necessary, simply because all of the competitors vying for visibility will be using them.

Beyond that we fall to the hybrid areas of SEO... those areas that bridge the gap between technical SEO and content creation.

Markup only works when the content is created in such a way that the most applicable format is used. A simple example would be a list. If the rich snippet for a query would logically be a list, the content needs to reflect that. It should be structured as a list, with headings or list items used in a way that reinforces what would appear in the markup.

“

Review the search results pages for the query you want a rich results for and other similar queries to assess the best formats - whether the result is a list, paragraph, etc. From there the content can be generated and from the markup created.

To speed up discovery, after the markup is added simply request indexing via search console and if all goes right, you will soon see your efforts pay off.

**-Dave Davies**

### **What makes a good rich snippet?**

- Uses Schema.org markup in the JSON-LD format. (Other formats are supported, but JSON-LD is preferred.)
- Has been properly validated using [the Structured Data Testing Tool](#)

### **What technical SEO can tell you:**

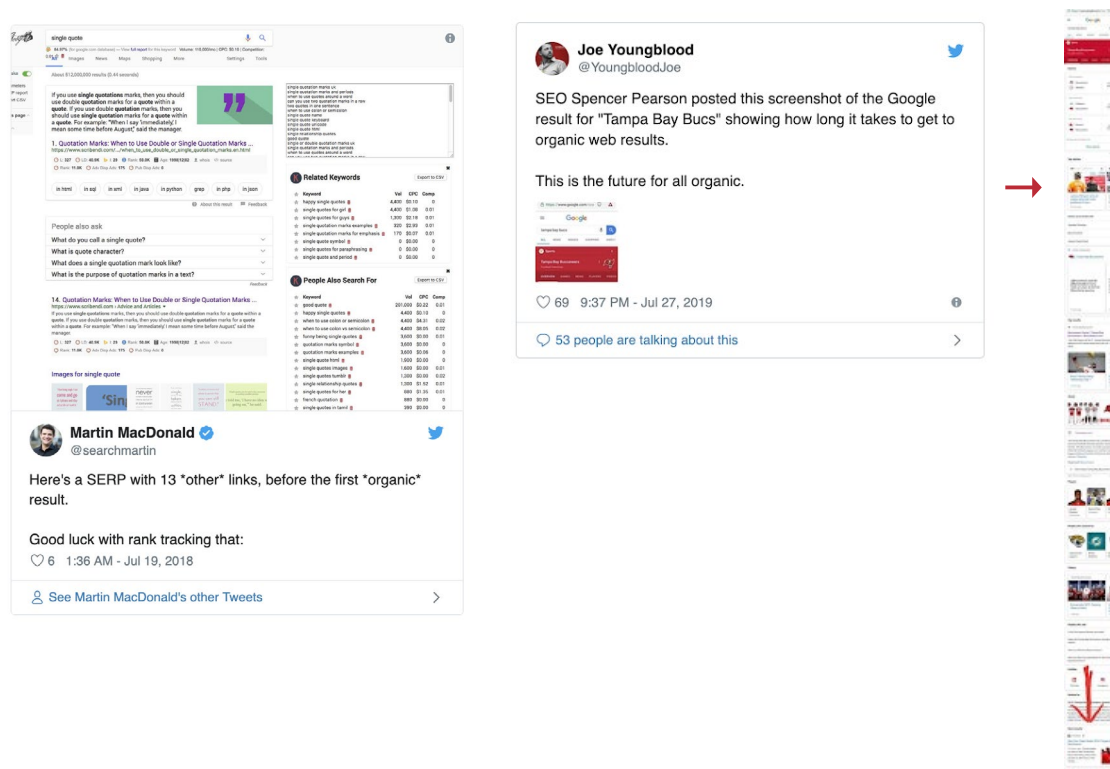
- Whether or not structured data is present on your pages
- What type of markup is used for your structured data
- Whether or not your structured data is valid
- Whether or not there is a difference in searcher click-through-rates for pages with structured data compared to pages without



# WHAT DO SEARCH ENGINES NEED TO ACCESS TO CREATE A SEARCH LISTING?

## Additional elements on a search results page

The search results page is not only composed of listings for URLs that can answer the search query. Since 2016, Google has added more and more elements to the SERPs, pushing the first organic page listings lower and lower on the page.



In response, it has become important for SEOs to secure visibility in additional search results page sections. The value of this strategy, beyond increased brand visibility, is heavily debated: many of these additional features provide information to searchers without encouraging them to click through to your site, resulting in an increasing number of zero-click searches, particularly on mobile devices.

# WHAT DO SEARCH ENGINES NEED TO ACCESS TO CREATE A SEARCH LISTING?

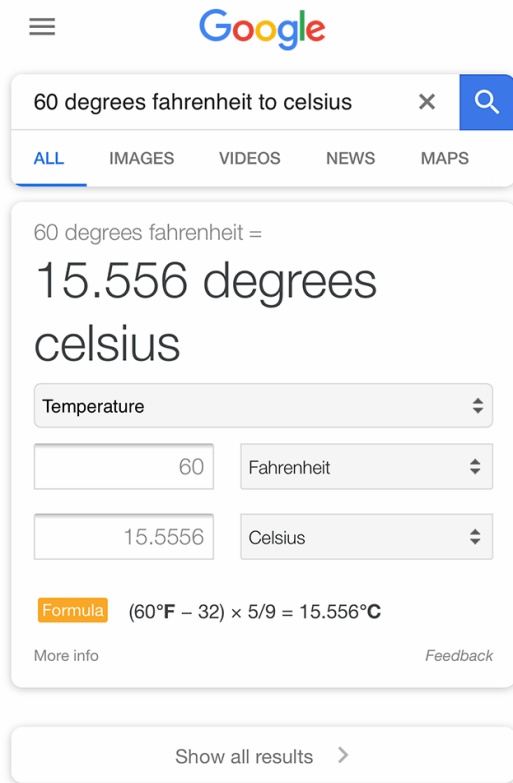
Search results elements include:

1. Knowledge graph
2. Position 0
3. Local pack
4. How-to snippets
5. Video carousel
6. Image carousel
7. News (top stories)
8. People also ask
9. Related searches

The screenshot displays a Google search for "yogurt" with the following elements:

- 3 Local pack:** A map showing locations like "Yogurt Factory" and "Nicaberry" with their ratings and addresses.
- 1 Knowledge graph:** A section titled "Yaourt" providing a definition, ingredients (Lait, Bactéries), and nutrition facts for 100g of yogurt.
- 5 Video carousel:** A row of video thumbnails with titles like "How to Make Fresh Yogurt At Home" and "How to make Yogurt at home? (With Easy Tips)".
- 7 News (top stories):** A list of recent news items, including "Yogurt maker Chobani donates \$50K for school lunch" and "Trader Joe's Launches Vegan Cashew Yogurt Drinks".
- 8 People also ask:** A list of related questions such as "What does plain yogurt do to a woman?" and "Is yogurt good for skin?".
- 9 Searches related to yogurt:** A list of related search terms including "yogurt and card", "yogurt benefits", "yogurt recipe", and "yogurt price".

## WHAT DO SEARCH ENGINES NEED TO ACCESS TO CREATE A SEARCH LISTING?



Since late 2018, there is also a growing number of zero-result searches, in which Google provides a fact or calculation that is the single, undisputed answer to the search query question. To view traditional search results for this sort of query, a mobile searcher must click through to the additional results.

Source

### What technical SEO can tell you:

- Whether your pages have content that can be used in additional search page elements
- Whether this content is correctly formatted so that search engines can use it

**HOW DO SEARCH  
ENGINES FIND  
YOUR PAGE?**



## HOW DO SEARCH ENGINES FIND YOUR PAGE?

In order to propose your page as a search result, and to compose the listing, Google has to be able to find your page.

Google learns a page exists when it finds a reference to your page's URL exists elsewhere on the internet. These references usually take one or more of the following forms:

- XML sitemaps: a list of important pages on your site that is made available to search engine bots.
- Backlinks: links to pages on your site from other sites.
- Site architecture: links to pages on your site from other pages on your site.

The role of technical SEO is to make sure that Google can find your pages.

### **You might use technical SEO to:**

- Create XML sitemaps
- Submit XML sitemaps to search engines
- Check whether or not your pages have backlinks to them
- Examine site architecture to make sure all pages have links to them
- Count the number of incoming and outgoing links per page
- Examine whether the number of links to a page from other pages on your site has an influence on how well a page ranks on search results pages.
- Check whether Google knows about your page using tools like crawls, log file analysis, or tools provided by search engines like the Google Search Console.
- Whether or not your site uses techniques that might accidentally (or intentionally) block Google from discovering your pages and their content. This might include:
  - blocking bots
  - giving instructions to bots to avoid certain pages
  - hiding content from bots
  - using techniques to display content that aren't yet supported by Google
  - using site architecture to delay or hinder page discovery

HOW DO SEARCH  
ENGINES  
EVALUATE YOUR  
PAGE?

Google saves known pages that it believes might be useable in search results in its index. Many algorithms likely use thousands of factors to evaluate a page's usefulness in search results. SEOs often cite "over 200 ranking factors"--but in reality, we can't identify all of the elements used by search engines to evaluate a page.

To demonstrate some transparency, Google has provided its [Search Rater Guidelines](#), which are a mix of qualitative and quantitative criteria for human validators who check the work of Google's algorithms.

Here are some of the main evaluation criteria.

### Can Google reach your page?

Google needs to be able to reach the page and its content. Since we know the type of profile used by Googlebots to discover and evaluate pages, we can use this information to make sure that bots with this profile can access and evaluate the pages on your website.

```
Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X  
Build/MMB29P) AppleWebKit/537.36 (KHTML, like  
Gecko) Chrome/41.0.2272.96 Mobile Safari/537.36  
(compatible; Googlebot/2.1; +http://www.google.  
com/bot.html)
```

*Example of a Smartphone Googlebot User Agent from [Google's official list](#)*

## HOW DO SEARCH ENGINES EVALUATE YOUR PAGE?

---

In 2018 and 2019, Google has been updating its index to a Mobile-First Index. Until now, the version of the page used to respond to search queries was the main version, usually a desktop version. For new sites launched in July 2019 and for sites that have already been migrated to the Mobile-First Index, Google will now use the mobile version, if one is available.

This also means that the main Googlebot used to find and index pages is the smartphone bot cited above. This bot needs to be able to access the full content of your pages.

Finally, in order to make sure Google sees new pages and newly-updated pages while crawling your site, you need to make sure that Google's attention is not already being occupied by low-value pages on your website. The distribution of Google's attention on a website is generally referred to as the crawl budget.

Through crawl budget management, technical SEO undertakes to ensure that the pages Google needs to see aren't being pushed aside by other pages on your site.

### What technical SEO can tell you:

- The distribution of Googlebot hits across your site pages using server log data
- Should you modify site structure to benefit pages you want Google to spend more time and attention on?
- Does your site block search engine bots?
- Your server's response when a page is requested. Server HTTP response codes let bots and visitors know whether a page is available, missing, moved and redirected, or whether the page request resulted in error, which might happen if the server is overloaded or unavailable.
- Does the content of your page show up? Some technologies and techniques can be difficult for search engines and some visitors to see. This can include: JavaScript, Flash, Iframes, content in videos or images...
- Is the content of your page readable/useable for mobile users?
- Does your robots.txt contradict your intentions?
- Does the distribution of crawl budget on your site lead Googlebots to the right pages?

“

If you see ambiguity like this on your site, just fix it. Technical SEO is not about “hope it works” :-).

**-John Mueller, [Twitter in a discussion which Allow/Disallow rules in a robots.txt file take precedence](#) 5 June 2019**

### Is your page the right version of the content?

When content exists at multiple locations, Google wants to index the right version. This criteria might concern your pages if:

- Other sites copy your content, or you copy theirs
- You syndicate (repost or republish) content on other sites, such as Medium or LinkedIn
- Your website's CMS platform produces multiple URLs for the same page
- Some of your pages are very similar to others
- You have multiple versions of the same content that have their own URL, such as printable versions, mobile versions, versions for people in different countries...

#### What technical SEO can tell you:

- Can your page be confused with another page that has similar content on your website?
- How many pages on your website are affected by similar content?
- How similar are similar pages on your website?
- Is your page in the right language?
- Does the page exist in other languages--and, for a site that is translated into multiple languages, is the page translated into all of the other languages?
- Is there a different version for the visitor's device?
- Do you show the same content to all visitors and bots?
- Do you have a strategy in place, such as canonical URLs, to suggest the right version to Google?

### Did you give Google any other instructions about your page?

The technical structure of a website--and a webpage--give website developers and managers many opportunities to signal information to bots and browsers. For example, it's possible to tell Google not to visit a page, or to visit it but not to index it--or to visit it, but not to follow links from it.

Additionally, your server can indicate when a page has been moved to a different URL, and whether this is a temporary or a permanent change.

#### What technical SEO can tell you:

- What instructions a page gives to bots by looking at:
  - Meta robots declarations in the HTML or in the page's HTTP header provided by the server.
  - Site-wide instructions provided in a Robots.txt file or in an htaccess file.
- Which pages return an HTTP status in the 300s range (redirected pages), and the URL they are redirected to.
- Which redirected pages send bots and visitors to pages that are also redirected. This is important because there is a limit to how many redirect jumps browsers and bots will follow. Google has repeatedly indicated that they only follow 5 jumps per crawl session.
- Whether any series of redirections create infinite loops, which result in browser errors.
- How Google treats the difference between temporary and permanent redirects.
- Whether Google interprets instructions to Googlebots as expected.



### Is your site reliable?

A server that can't handle traffic to the website it runs can be a major limiting factor for SEO, particularly as the site-wide experience gains increasing importance in how a page ranks.

Slow servers can slow down Google's discovery of new and updated pages. Google adjusts the number of pages it crawls per second based on your server's bandwidth.

“

Google has sophisticated algorithms that determine the optimal crawl speed for a site. Our goal is to crawl as many pages from your site as we can on each visit without overwhelming your server's bandwidth.

**-Google Webmasters Documentation**

Slow servers can prevent Google from seeing content if your site has frequent overloads, downtime, or other server errors. Content that can't be seen by Google can't be used to rank your site or shown to searchers. When the server cannot provide the content, it returns a 500s series HTTP status for the URL; Google reports these as errors in the Search Console.

Slow servers increase the time it takes to transfer the content and resources for a page to the visitor or the bot. Page speed has always been a factor in how Google evaluates a page. It is an explicit factor for mobile pages since 2018, and starting in 2019, all new sites will be evaluated using a mobile search bot--which means page speed will count.

Servers that are not secure are susceptible to hacking. Google can and does remove sites from their index if it identifies the site as being hacked. This happened (erroneously) to a major SEO news site in 2018.

Servers that do not use the latest security protocols can be penalized in how well their pages rank in the search results. For example, HTTPS, an internet protocol that is used to protect the integrity and confidentiality of data transferred between the website and the user, has been used to evaluate pages since 2016. The current Search Console help on HTTPS states:

“

Users expect a secure and private online experience when using a website. We encourage you to adopt HTTPS in order to protect your users' connections to your website, regardless of the content on the site.

**-Google Webmasters Documentation**

### What technical SEO can tell you:

- Does your server respond when a page is requested?
- How frequent are 5xx status codes, and what pages are affected?
- Is your site safe? (spam/attacks)
- Does your site use HTTPS? If not, what steps need to be taken to migrate your site?
- Do pages load in a reasonable time for users and for Googlebots?
- The total size of the resources (images, styling, scripts...) required to display the page
- The ratio of code (page structure, look, and functionality) to text (content)

Technical SEO can also help when putting server solutions and site-wide parameters in place, whether this includes writing redirects into an htaccess file, or telling your server what information to log whenever it receives a request to provide a URL.

### **Is your content authoritative on the web?**

Content quality and site authority have always been criteria for evaluation for Google and other search engines. For example, it having incoming links from sites on government and higher education domains (.gov and .edu) used to give your site a boost in authority and helped it appear higher in the search results. As Google now uses other measures to determine authoritativeness, this is no longer the case.

Moz, a leading SEO tool and resource center, calculates a Domain Authority (DA) score, which is so often used as a fairly accurate predictor of how the website's pages will rank that many people think DA is an official Google metric. (Spoiler: it's not!)

Since the August 2018 Google algorithm update called «Medic», there has been increased talk about Expertise, Authoritativeness and Trust (EAT). These criteria appear in detail in the Quality Rater Guidelines, and it was confirmed in 2019 that Google's ranking algorithms use EAT metrics to evaluate websites.

## HOW DO SEARCH ENGINES EVALUATE YOUR PAGE?

Google's explanations show that Google's algorithms look for:

- The number of quality or authoritative websites linking to your content
- Mentions, even unlinked, of your site and its content from trustworthy sources
- High quality content
- Trustworthiness of your website
- Information on publishers and authors, particularly on lifestyle (YMYL) websites
- Information about the website, who runs or maintains it, and customer service information on ecommerce websites



### How does Google algorithmically determine E-A-T?

E-A-T, Expertise, Authoritativeness and Trust, has been a hot topic in SEO circles as of late. Google has told us recently that [E-A-T is an important part of core algorithm updates](#). Google's goal is to determine when a particular query is one that indicated a user is looking for information that will affect their lives. If so, then this is determined to be a YMYL, Your Money or Your Life query. For YMYL queries, Google's algorithms are "specifically designed to identify sites with high indicia of expertise, authority, and trustworthiness."

We can gain clues as to how Google would measure this algorithmically by looking at the information in Google's [Search Quality Evaluators guide](#). The Quality Raters use the information in this guide to assess whether a website is high quality or not. The information from the Quality Raters does not directly affect a site, but instead gets fed back to Google's engineers. If they see that their algorithms are not working correctly and allowing low quality sites to rank well, then Google's engineers can tweak the algorithms to make it less likely for that to happen in the future.

“

It is important to know that E-A-T is likely extremely complex. It's not as simple as just adding an author bio to each post. Rather, what Google is looking for is evidence that each of the components of E-A-T is present. Google employees have told us that a big part of E-A-T is determined by seeing what others think of your business, your website and your authors. Gary Illyes from Google has said that E-A-T is largely determined by off site links and mentions from authoritative sites.

While we don't know exactly how Google calculates E-A-T, there are many components that are probably used. It is possible that Google is using sentiment analysis across review sites to determine whether it is obvious that customers have trust issues with a business. Google has also mentioned several times in the Quality Raters' Guidelines that sites that contradict general scientific consensus can be seen as lacking in trust. While we have ideas on how that can be measured, these are all theories at this point.

Google's VP of Search, Ben Gomes, told us the following: "You can view the rater guidelines as where we want the search algorithm to go. They don't tell you how the algorithm is ranking results, but they fundamentally show what the algorithm should do." At MHC, we believe that if something is outlined in the Quality Raters' Guidelines, it is something to which we should pay close attention! While we don't know exactly how Google's algorithms work, these guidelines are an excellent guide to tell us what it is that Google is looking for in terms of quality.

**-Marie Haynes**

### What technical SEO can tell you:

- What sites link to your website (backlinks)
- Which pages receive backlinks
- Whether the authors or editors of content are present on your pages in a Google-readable format

Technical SEO can also offer solutions to reinforce the authoritativeness of your website and its content. This could be things like including reviews and markup for reviews on your pages, programmatically adding author information to article pages, disavowing links from low-quality websites, or establishing threshold metrics to determine what type and length of content yields the best results for your site.

### Is the page important to your website?

Search engines use links to a page to determine its importance. Important pages rank higher in search results, and therefore bring more visitors to websites.

Linking goes beyond the backlinks discussed above. It also includes internal linking. At OnCrawl, we've seen how site architecture created by internal linking can promote or demote pages.

See how a network of job boards used internal links to promote top pages and helped increase conversions by 79%

[Read the case study](#)

### What technical SEO can tell you:

- Which pages receive the most links within your website?
- Which pages and groups of pages are important to your website?
- How popularity flows between groups on your website
- What the popularity score for each page on your website is
- Which pages receive significant popularity from backlinks that can be distributed to other pages in your site using internal links
- Which pages get the most attention from visitors
- Which pages get the most attention from Googlebot

**HOW DO SEARCH  
ENGINES KNOW  
WHICH SEARCH  
QUERIES TO SHOW  
YOUR PAGE FOR?**



## HOW DO SEARCH ENGINES KNOW WHICH SEARCH QUERIES TO SHOW YOUR PAGE FOR?

Google's objective is, for any search query, to return a list of responses ranked in the order of probable usefulness. This means that Google has to select, for each query, URLs in its index that might provide a good response.

Many different factors contribute to Google's understanding of what a page is about, and whether it might be a good response to a searcher's query.

### Semantic understanding

Google uses a semantic understanding of the topic covered on a given URL to decide whether the URL's content matches the search query.

Keywords. These words occur in both the search term and in the content at the URL. Because keywords were, for a long time, the main means of matching URLs to queries, SEOs have had a habit of placing keywords as many times as possible on a page (keyword stuffing). Because this results in content that is hard to read and does not necessarily give an accurate representation of the page's topic, Google's Panda algorithms penalize sites that use this type of practice.

Keywords can still be placed in key sections of the page, at a frequency that feels natural to a human reader.

- Page title
- Headings (H1, H2, H3...) on the page
- First (or "topic") paragraph
- Meta description of the page's content
- Anchor text of links to the page
- Scattered throughout the text itself.

“

Nowadays, search engines are generally able to match search visitors to the most useful pages even if the keywords are not present in the content. They are also getting better at directly answering questions.

At the same time, searchers are growing more comfortable using natural language queries.

Recent research from Google even concludes that we should adapt to unique website visit paths taken by each potential customer.

Considering all these major changes taking place, how do we adapt?

Using machine learning, of course!

Machine learning can help you perform keyword research on steroids. Specifically, you can really understand and predict keyword intent in ways that simply aren't possible manually.

**-Hamlet Batista**

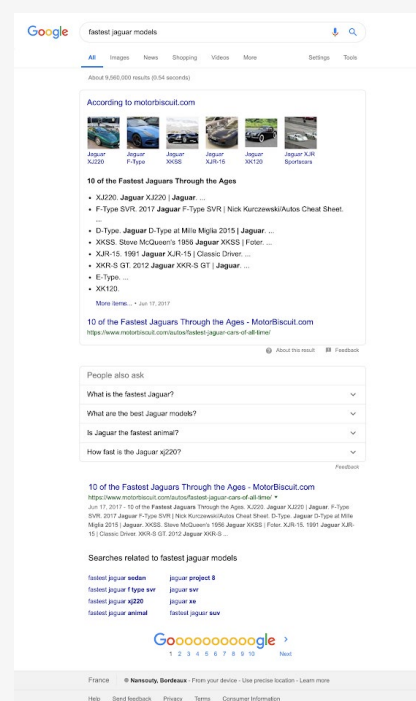
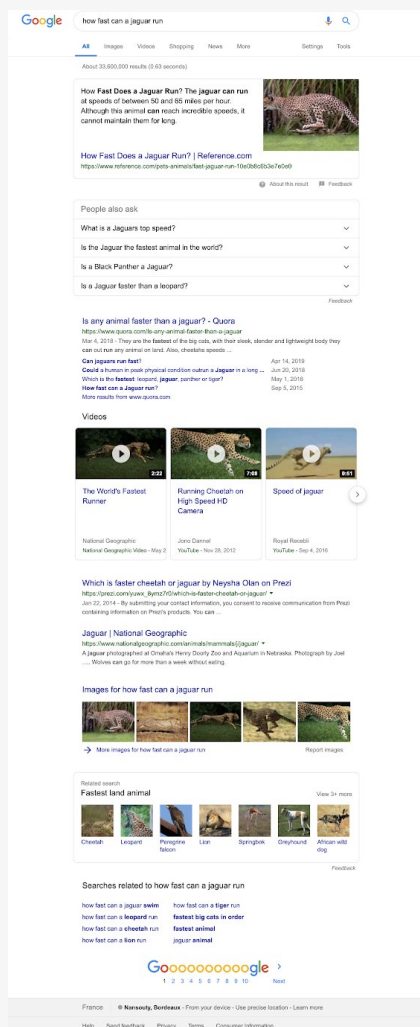
## HOW DO SEARCH ENGINES KNOW WHICH SEARCH QUERIES TO SHOW YOUR PAGE FOR?

**Machine learning.** Google employs machine learning techniques to understand what keywords mean and how topics and keywords are related.

- Latent semantic indexing (LSI) is an old machine learning technique that helps algorithms establish the relationship between words. There is evidence from Google patents and SEO studies that the presence of related words and phrases on a page helps Google better determine the topic and the relevance of the URL, but LSI is not a ranking factor, and Google has confirmed that there is no such thing as LSI keywords.



- Context plays an important role in machine understanding and disambiguation of search terms. Context, in semantic analysis, refers to the words before and after a keyword: in other words, the sentence or sentences in which a keyword occurs. This helps Google understand the difference between search queries like “fastest jaguar models” and “how fast can a jaguar run”--and to pick the right pages to answer the question.



- Structure also helps machines understand which parts of a page and its content should be taken into consideration. You can add structure to content in many ways:
  - Semantic HTML tags such as <head>, <body>, <nav>, <section>, <title>, <table> which identify the purpose and type of content found within them
  - Hierarchical structure such provided by headers such as <H1>, <H2>, which help structure the flow of an article or page
  - Tables and lists, which organize information
  - Structured data markup using Schema.org

## HOW DO SEARCH ENGINES KNOW WHICH SEARCH QUERIES TO SHOW YOUR PAGE FOR?

### Entities.

“

Entities are things that have attributes.

-JP Sherman, [SEO in Orbit, episode 5](#)

Google says that an entity is “a thing or concept that is singular, unique, well-defined, and distinguishable”. In other words, an entity is something specific that can be identified and not confused with something else. Machines can then use entities to understand relationships between ideas by evaluating which entities occur in the same or similar contexts.

Using entities in your content can help Google determine how your pages can answer search queries. In its most simplistic form, “using entities” means being specific:

In this excerpt from the Wikipedia article on SEO:

“

Website owners recognized the value of a high ranking and visibility in search engine results, creating an opportunity for both white hat and black hat SEO practitioners. According to industry analyst Danny Sullivan, the phrase “search engine optimization” probably came into use in 1997. Sullivan credits Bruce Clay as one of the first people to popularize the term. On May 2, 2007, Jason Gambert attempted to trademark the term SEO by convincing the Trademark Office in Arizona that SEO is a “process” involving manipulation of keywords and not a “marketing service.”

[https://en.wikipedia.org/wiki/Search\\_engine\\_optimization](https://en.wikipedia.org/wiki/Search_engine_optimization)

## HOW DO SEARCH ENGINES KNOW WHICH SEARCH QUERIES TO SHOW YOUR PAGE FOR?

Google is able to identify the entities in the paragraph and rank them in order to their pertinence to the topic:

Entities	Sentiment	Syntax	Categories
<p>&lt;Website owners&gt;<sub>2</sub> recognized the &lt;value&gt;<sub>5</sub> of a high &lt;ranking&gt;<sub>3</sub> and &lt;visibility&gt;<sub>4</sub> in &lt;search engine results&gt;<sub>8</sub>, creating an &lt;opportunity&gt;<sub>9</sub> for both white &lt;hat&gt;<sub>7</sub> and &lt;black hat SEO practitioners&gt;<sub>10</sub>. According to &lt;industry analyst&gt;<sub>6</sub> &lt;Danny Sullivan&gt;<sub>6</sub>, the &lt;phrase&gt;<sub>14</sub> " &lt;search engine optimization&gt;<sub>11</sub> " ( &lt;SEO&gt;<sub>1</sub> ) probably came into &lt;use&gt;<sub>13</sub> in &lt;1997&gt;<sub>27</sub> &lt;1997&gt;<sub>30</sub>. &lt;Sullivan&gt;<sub>6</sub> credits &lt;Bruce Clay&gt;<sub>19</sub> as &lt;one&gt;<sub>15</sub> &lt;one&gt;<sub>29</sub> of the first &lt;people&gt;<sub>16</sub> to popularize the &lt;term&gt;<sub>12</sub>. On &lt;May 2, 2007&gt;<sub>28</sub> &lt;2&gt;<sub>31</sub>, &lt;2007&gt;<sub>32</sub>, &lt;Jason Gambert&gt;<sub>26</sub> attempted to trademark the &lt;term&gt;<sub>25</sub> &lt;SEO&gt;<sub>1</sub> by convincing the &lt;Trademark Office&gt;<sub>23</sub> in &lt;Arizona&gt;<sub>24</sub> that &lt;SEO&gt;<sub>18</sub> is a " &lt;process&gt;<sub>17</sub> " involving &lt;manipulation&gt;<sub>20</sub> of &lt;keywords&gt;<sub>21</sub> and not a " &lt;marketing service&gt;<sub>22</sub> ."</p>			
1. SEO Saliency: 0.16	OTHER	2. Website owners Saliency: 0.12	PERSON
3. ranking Saliency: 0.11	OTHER	4. visibility Saliency: 0.11	OTHER
5. value Saliency: 0.10	OTHER	6. Danny Sullivan <a href="#">Wikipedia Article</a> Saliency: 0.05	PERSON
7. hat Saliency: 0.05	OTHER	8. search engine results Saliency: 0.04	OTHER
9. opportunity Saliency: 0.04	OTHER	10. black hat SEO prac... Saliency: 0.04	PERSON
11. search engine opti... Saliency: 0.02	OTHER	12. term Saliency: 0.02	OTHER
13. use Saliency: 0.01	OTHER	14. phrase Saliency: 0.01	OTHER
15. one Saliency: 0.01	OTHER	16. people Saliency: 0.01	PERSON

## HOW DO SEARCH ENGINES KNOW WHICH SEARCH QUERIES TO SHOW YOUR PAGE FOR?

### What technical SEO can tell you:

- What is the density of keywords on your page?
- What keyword search terms does your page rank for?
- What are the entities on your pages?
- What are the most common terms on your page?
- Does your page structure content?
- What pages offer similar content on your website?
- Is any page's content too similar to another page's content?
- What are probable strategies used by Google to understand your content?

### User intent signals

When matching URLs to search queries, search engines also take into consideration what the searcher is trying to do. Based on the context and intent of the search, the results that best meet the searcher's need may be very different.

Google recognizes four different types of search intents:

- Know query, some of which are Know Simple queries
- Do query, some of which are Device Action queries
- Website query, when the user is looking for a specific website or webpage
- Visit-in-person query, some of which are looking for a specific business or organization, some of which are looking for a category of businesses

(Source: Google Search Quality Evaluator Guidelines)

# HOW DO SEARCH ENGINES KNOW WHICH SEARCH QUERIES TO SHOW YOUR PAGE FOR?

Sometimes keywords can indicate the intent of a query:



🔍 **buy** cat food



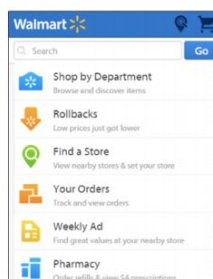
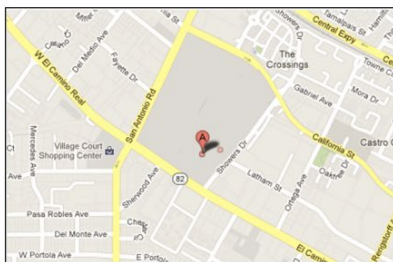
🔍 **What's** the capital of Serbia?



🔍 Ikea **address**



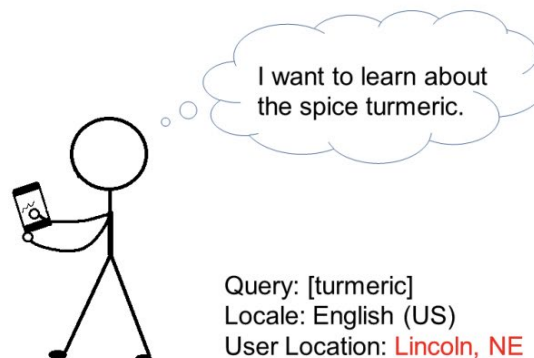
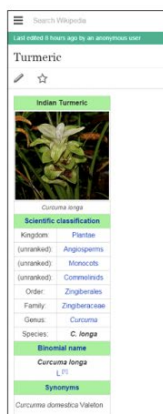
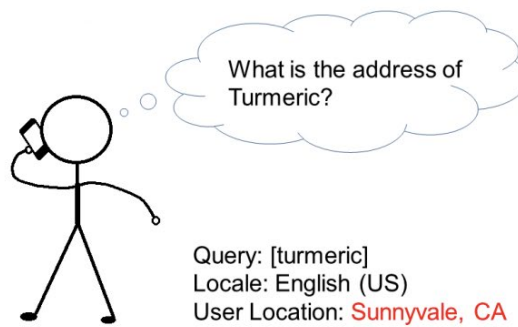
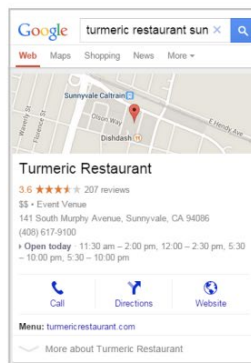
Some queries are ambiguous to Google:





## HOW DO SEARCH ENGINES KNOW WHICH SEARCH QUERIES TO SHOW YOUR PAGE FOR?

The intent of a query and the best results can also vary based on other contextual information. This can produce different search engine results pages based on context.



Since user intent is so closely linked to keyword research, SERP analysis, and content structure, technical SEO will often focus on how to scale these operations. This allows you to address site-wide issues, even on big websites, instead of working on a page-by-page or keyword-by-keyword basis.

### What technical SEO can tell you:

- What are probable user intents for keywords you target
- Whether or not the top pages in results for your keywords have the same objective and structure as yours
- What intent-oriented keywords your competitors are ranking for

### Linking structure

When pages with the same or similar topics reference a page by linking to it, this reinforces a search engine's understanding of the page topic, and helps place it correctly in the results for appropriate queries.

#### What technical SEO can tell you:

- What pages have similar topics on your site
- Whether or not links between these pages exist

### Search analytics and predictions

“

Both Rankbrain and Hummingbird are query rewriting approaches. Hummingbird tried to better understand the context of a query by looking at all of the words in a query.

Previously, Google would only look at words next to one another to understand context; Hummingbird looks beyond the words immediately next to each other. It might even take into account full sentences in conversational queries. Hummingbird tried to use all of the words in the query together to understand the context.

Unlike Hummingbird, Rankbrain uses a word embedding approach. It examines a short textual passage and is able to determine if there are words that are missing. It does this by training on large sets of data (200 billion words).

In addition, we know that it took Google 5 years to develop a language model process that improves results in over 30% of searches.

-Bill Slawski, SEO in Orbit episode 10

## HOW DO SEARCH ENGINES KNOW WHICH SEARCH QUERIES TO SHOW YOUR PAGE FOR?

### User behavior

Google uses user behavior data to validate its guesses at the appropriateness of the URLs it offers as responses and the order in which they appear.

While we don't know exactly what or how Google measures, we can observe mechanisms that indicate Google is trying to re-evaluate and refine a user's search:

- What pages have similar topics on your site
- Whether or not links between these pages exist

### Artificial intelligence

Google uses artificial intelligence--the Rankbrain algorithm--to better respond to queries it has never seen before. This can account for up to 30% of all searches.

This algorithm attempts to match URLs to queries when there's no history of intent or feedback for the query.

### What technical SEO can tell you:

- The click-through-rate (CTR) for your pages
- The number of views and visits that you receive from infrequent queries. These are cases where Rankbrain was likely used to rank your page.
- The search results page positions that earn you the most visits

WHAT SOME  
OTHER GOALS WE  
CAN ACCOMPLISH  
USING TECHNICAL  
SEO?

## WHAT SOME OTHER GOALS WE CAN ACCOMPLISH USING TECHNICAL SEO?

Technical SEO is also about finding technical solutions to SEO problems. Often that means looking for better ways to analyze data and for time-saving ways of implementing fixes.

Here are some examples of technical SEO in use.

**Finding statistically relevant correlations between different elements**, such as crawl rate and rankings.

**Automating repetitive, boring, or time-consuming tasks**, such as generating titles and meta descriptions for product pages on an ecommerce website. This might include custom-developed solutions, or could involve templating, like this example of a step-by-step guide [to bulk updating titles and descriptions in WordPress](#) from Go Fish Digital, which guides you through the necessary steps, or even use machine learning, as [in this series of articles](#) in which Vincent Terrasi automates the search for key metrics affecting indexing and then automatically includes them in a powerpoint presentation.

**Testing SEO implementations and creating predictions.** Technical SEO provides the means to run tests and experiments that SEO, as a “soft” or “inexact” science, can’t easily quantify. This can be useful when estimating the impact of SEO implementations, testing theories and results before investing in a project, or even creating reliable projections and predictions for traffic, conversions, and revenue.





## How to create technical SEO predictions

For this, I'd recommend being familiar with at least one programming language (Python, R, Go, ...). To be honest, programming gives us greater flexibility where technical SEO is concerned.

Then, with machine learning basics, there are a number of strategies that can be used to make predictions.

- Linear or polynomial regression, which allows you to find a connection between a dependent variable (Y) and an independent variable (X), in order to predict Y when X is measured.
- Decision trees, whose goal is to identify all of the factors that lead to a specific event and to determine, in our use case, what is the most likely outcome.
- ARIMA, for time series, is a set of models that can create condensed representations of certain phenomena that vary over time and predict future values with a confidence interval.

**-Vincent Terrasi**

Creating or adapting tools to give us more precise information. Technical SEO aims to make data-based decisions, and to provide technical solutions to obtain data that is otherwise not available.

For example, the Google Search Console's Performance Report table only offers a maximum of 1000 rows. If your site ranks for more than 1000 keywords, or you have more than 1000 ranking pages, it is impossible to view all of them in the Search Console. However, it is possible to develop solutions using the API to get all of your search data. The API will allow you to query as many pages as necessary, with up to 25 000 lines per page.

## WHAT SOME OTHER GOALS WE CAN ACCOMPLISH USING TECHNICAL SEO?

In another example, in local search, it can be essential to view search results from different locations. Technical SEO also creates tools like [Max Prin's local search tool](#) to provide solutions to this sort of need.

At other times, tools like [this FAQPage JSON-LD Schema Generator](#) from Saijo George aim to simplify complex actions that might require advanced skills not everyone has.

“

In the SEO world, the answer to most questions is “it depends”. But in a few cases, mainly on the technical side, answers are Boolean. A URL is crawlable or not, a tag’s implementation is correct or incorrect, etc.

In that particular context testing tools become extremely useful, reliable, and make us save a ton of time. Who can afford to manually verify thousands of hreflang tags across multiple international versions of a site?

And yet, everything needs to be tested and verified before we can move onto the next project or task. The tools available on TechnicalSEO.com have been developed on that basis. Eliminating manual review as much as possible from our processes with efficient and reliable scripts.

Some tools I developed did not exist anywhere else on the web, some others where improvements, in my opinion, of existing resources. Regardless, they’ve been built to fill a gap and meet the specific needs of our team.

However, because the willingness to share knowledge is a cornerstone of our industry, making those tools publicly available just made sense. If we needed them to our job faster, others might too.

**-Max Prin**



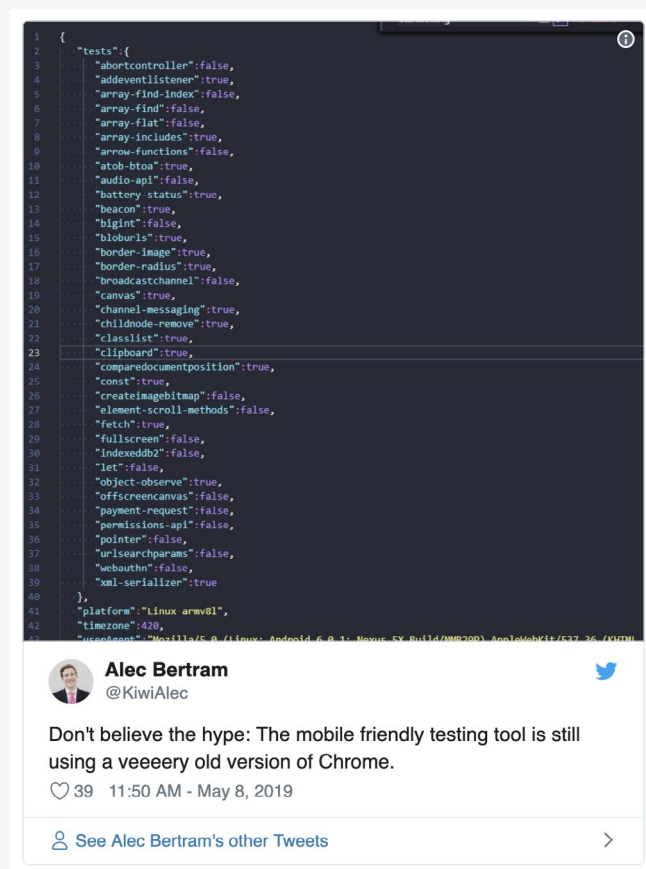
## WHAT SOME OTHER GOALS WE CAN ACCOMPLISH USING TECHNICAL SEO?

**Making sure Google actually does what it says it's doing.** Google's announcements or status quos don't always mean what we think they do. Technical SEO can find discrepancies between what we believe Google should be doing, and how Google actually behaves when examining website.

Recent examples include Max Cyrex's discovery that Googlebot ignores rel=next and rel=prev tags. His discovery came shortly before--but was published at the same time--as Google's announcement to the same effect.

Technical SEOs have also been wary of Google's claim to have updated Googlebot to the most recent version of Chrome, particularly since Google has been vague on the date of the full roll-out. Technical SEO has helped establish that:

- In May, not all Google services use the updated version of Googlebot yet



## WHAT SOME OTHER GOALS WE CAN ACCOMPLISH USING TECHNICAL SEO?

- Googlebots with the old Chrome 41 User-Agent (its identifier) are executing features that weren't supported until later versions of Chrome: Google updated the bot, but did not update the bot's User-Agent.



**Glenn Gabe** @glenngabe · May 8, 2019



Replying to @g33konaut @JohnMu

Great, thanks for the info Martin. Will check out the session for sure. :) The logs contain the version, like this:  
Chrome/41.0.2272.96 Mobile Safari/537.36



**Martin Splitt** @ #athens

@g33konaut

Right, unfortunately the UA is wrong ;-) We wanted to do things step by step and we didn't want to change the UA as people might have hardcoded it somewhere - with enough heads-up, we'll change that!

♥ 19 4:38 PM - May 8, 2019



See Martin Splitt @ #athens's other Tweets



**JR Oakes**

@jroakes



Two things interesting in looking at rendered Googlebot content.  
#1 Googlebot has 112 logical processors, but 8GB of RAM. #2 that the user-agent says Chrome/41.0.2272.96, yet fires deviceorientationabsolute events which were added in Chrome 50. [codeseo.io/assets/html/go...](https://codeseo.io/assets/html/go...)

♥ 7 11:20 PM - Jun 1, 2019



## WHAT SOME OTHER GOALS WE CAN ACCOMPLISH USING TECHNICAL SEO?

- Google runs tests using bots with unofficial User-Agents on “small batches of URLs”

The screenshot shows a Twitter thread starting with a tweet from Lino Uruñuela (@Errioxa) titled "Googlebot Hits: 01-may-2019". The tweet includes a table of user agent strings and their corresponding hit counts. The table lists four entries, all with hit counts of 2, 11, 9, and 6 respectively. The user agent strings are variations of Mozilla/5.0 (Linux; Android 9.0.0; en-us; Pixel 3 XL Build/PD1A.180621.003) AppleWebKit/[WEBKIT\_VERSION] (KHTML, like Gecko) Chrome/[CHROME\_VERSION] Mobile Safari/[WEBKIT\_VERSION] (compatible; Googlebot/2.1; +http://www.google.com/bot.html).

User Agent	Googlebot Hits: 01-may-2019
Mozilla/5.0 (Linux; Android 9.0.0; en-us; Pixel 3 XL Build/PD1A.180621.003) AppleWebKit/[WEBKIT_VERSION] (KHTML, like Gecko) Chrome/[CHROME_VERSION] Mobile Safari/[WEBKIT_VERSION] (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	2
Mozilla/5.0 (Linux; Android 9.0.0; en-us; Pixel 3 XL Build/PD1A.180621.003) AppleWebKit/[WEBKIT_VERSION] (KHTML, like Gecko) Chrome/[CHROME_VERSION] Mobile Safari/[WEBKIT_VERSION] (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	11
Mozilla/5.0 (Linux; Android 9.0.0; en-us; Pixel 3 XL Build/PD1A.180621.003) AppleWebKit/[WEBKIT_VERSION] (KHTML, like Gecko) Chrome/[CHROME_VERSION] Mobile Safari/[WEBKIT_VERSION] (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	9
Mozilla/5.0 (Linux; Android 9.0.0; en-us; Pixel 3 XL Build/PD1A.180621.003) AppleWebKit/[WEBKIT_VERSION] (KHTML, like Gecko) Chrome/[CHROME_VERSION] Mobile Safari/[WEBKIT_VERSION] (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	6

The thread continues with replies from Martin Splitt (@g33konaut) and Lino Uruñuela (@Errioxa) discussing the possibility of these hits being due to a new search engine or an experiment. Martin Splitt explains that the user agent string was an experiment and that Googlebot typically uses a different user agent string. Lino Uruñuela thanks Martin Splitt for the answer.

**User Agent**

**Googlebot Hits: 01-may-2019**

User Agent	Googlebot Hits: 01-may-2019
Mozilla/5.0 (Linux; Android 9.0.0; en-us; Pixel 3 XL Build/PD1A.180621.003) AppleWebKit/[WEBKIT_VERSION] (KHTML, like Gecko) Chrome/[CHROME_VERSION] Mobile Safari/[WEBKIT_VERSION] (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	2
Mozilla/5.0 (Linux; Android 9.0.0; en-us; Pixel 3 XL Build/PD1A.180621.003) AppleWebKit/[WEBKIT_VERSION] (KHTML, like Gecko) Chrome/[CHROME_VERSION] Mobile Safari/[WEBKIT_VERSION] (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	11
Mozilla/5.0 (Linux; Android 9.0.0; en-us; Pixel 3 XL Build/PD1A.180621.003) AppleWebKit/[WEBKIT_VERSION] (KHTML, like Gecko) Chrome/[CHROME_VERSION] Mobile Safari/[WEBKIT_VERSION] (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	9
Mozilla/5.0 (Linux; Android 9.0.0; en-us; Pixel 3 XL Build/PD1A.180621.003) AppleWebKit/[WEBKIT_VERSION] (KHTML, like Gecko) Chrome/[CHROME_VERSION] Mobile Safari/[WEBKIT_VERSION] (compatible; Googlebot/2.1; +http://www.google.com/bot.html)	6

**Lino Uruñuela** @Errioxa

@JohnMu @g33konaut Could these hits be due to the new search engine?

Googlebot hits using User Agent String Android 9.0.0 Pixel 3 XL, 01-may-2019

12 6:42 PM - May 8, 2019

**Lino Uruñuela** @Errioxa · May 8

Replying to @Errioxa

nslookup return "crawl-...ip....googlebot.com."

1 2

**Martin Splitt** @g33konaut · May 14

Replying to @Errioxa and @JohnMu

Regarding the other assumptions as to what is rendering and what isn't - I wouldn't make that assumption because it isn't generally the case.

1 1 1

**Martin Splitt** @g33konaut · May 14

Replying to @Errioxa and @JohnMu

Alright, so the user agent you saw was an experiment we ran. By the way: we do run experiments on small sets of URLs on a regular basis.

1 1 1

**Lino Uruñuela** @Errioxa · May 15

I've seen something like this before, I'm watching you

1 1

**Martin Splitt** @g33konaut · May 14

Replying to @Errioxa and @JohnMu

The assumption about Googlebot you voice might be correct in some cases, but I wouldn't rely on that.

1 1

**Lino Uruñuela** @Errioxa · May 15

Thank you for answering!

1

## WHAT SOME OTHER GOALS WE CAN ACCOMPLISH USING TECHNICAL SEO?

- Until August 7th, Google's testing tools didn't use the same googlebot as its SEO crawls. This means was nearly impossible to test what Google actually sees or does when crawling your pages.



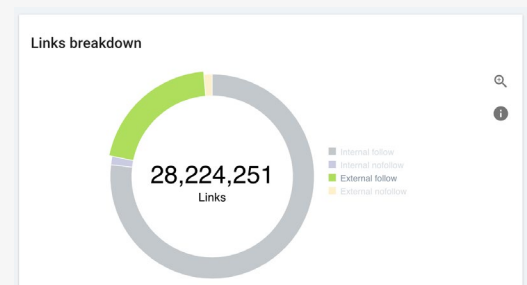
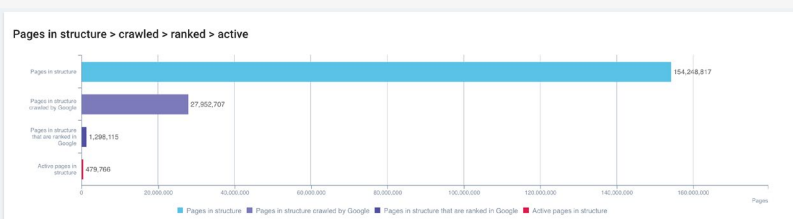
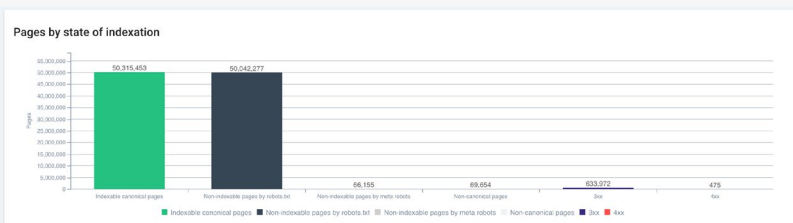
## WHAT SOME OTHER GOALS WE CAN ACCOMPLISH USING TECHNICAL SEO?

**Improving website code and performance** by targeting technical issues.

Technical optimizations can address elements such as:

- **Code bloat.** Pages that contain excessive amounts of code and little content are said to suffer from “code bloat.” These pages can be spotted and corrected. Streamlining code usually increases page speed and reduces the probability that new features, downtime, or unexpected events will break a website.
- **Page speed.** Compression, lazy loading, removing unused scripts and resources, and providing the right content for the right devices are all technical measures that can materially improve the speed of a website.
- **Errors and attacks.** Detecting, correcting and preventing server errors and website attacks helps increase trustworthiness and performance on your website.

**Handling big sets of data on hundreds or thousands of pages** with statistical and algorithmic methods. This makes it possible to audit and implement fixes on sites with millions of pages, including large news and e-commerce sites, where manual actions would be impossible.



**Implementing site-wide solutions** rather than page-by-page adjustments.

By addressing issues at a server level or programmatically via templates and scripts, technical SEO can provide efficient site-wide solutions. This can include tasks like:

- **improving** site speed by updating server software or bandwidth
- **allowing** crawl by search engine bots by removing restrictions at a site configuration level
- **creating** 301 redirects in an htaccess file using regular expressions instead of listing all pages in a directory one by one
- **auto-generating** hreflang declarations in sitemaps
- **using** patterns to automatically create unique descriptions for product pages.

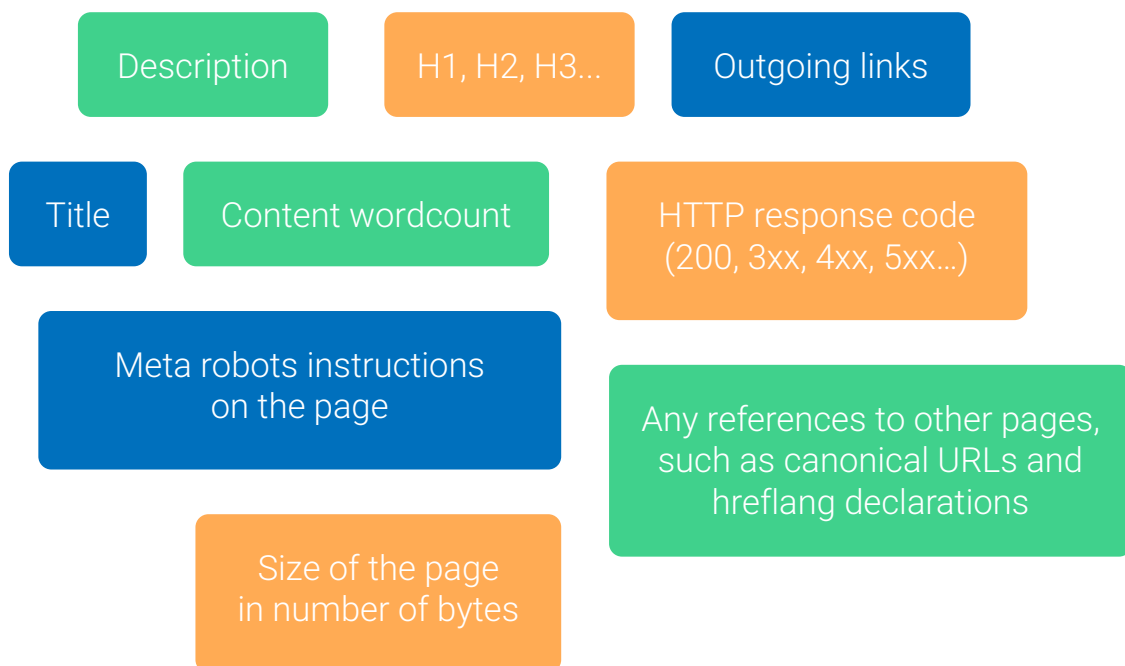
**Offering technical workarounds when SEO standards can't be implemented normally** (Edge SEO). Sometimes legacy website structure or software makes it impossible or impractical for web developers to implement SEO best practices. Edge SEO uses service workers provided by CDNs to make modifications after the page is rendered but before it is served to the bot or visitor that requested it.

**WHAT ARE THE  
TOOLS USED IN  
TECHNICAL SEO?**

## WHAT ARE THE TOOLS USED IN TECHNICAL SEO?

### Crawlers

SEO crawlers examine a website and extract information for every page. Most crawlers will either start from a given page and follow all links they find on that page and subsequent pages, or they'll work from a list of pages you provide, or both. Even the most basic crawl will produce a list of all pages on a site along with information for each page.



Because SEO crawlers function similarly to search engine crawlers, SEO crawls can also be used to predict and to study crawl and indexing issues on a website by allowing you to see how the site is discovered and where the roadblocks for search engine bots are located.

Crawlers exist in both desktop and cloud formats. While desktop crawlers like Screaming Frog or Sitebulb are cheaper, they have their limitations.



“

- Crawls consume your memory and CPU. However, the situation is much better than it used to be in that crawlers are improving in the areas of memory & CPU management.
- You have to use proxies to avoid getting banned.
- Collaboration is limited. You can't just share a report with a client/colleague. You can, however, work around this by sending them a file with a crawl project.
- Unfortunately, desktop crawlers struggle with crawl comparison (Sitebulb is an exception) and scheduling.
- In general, desktop crawlers are more limited than cloud crawlers.

-**Tomasz Rudzki**, [Onely's Ultimate Guide to SEO Crawlers](#)

Cloud crawlers generally offer advantages such as:

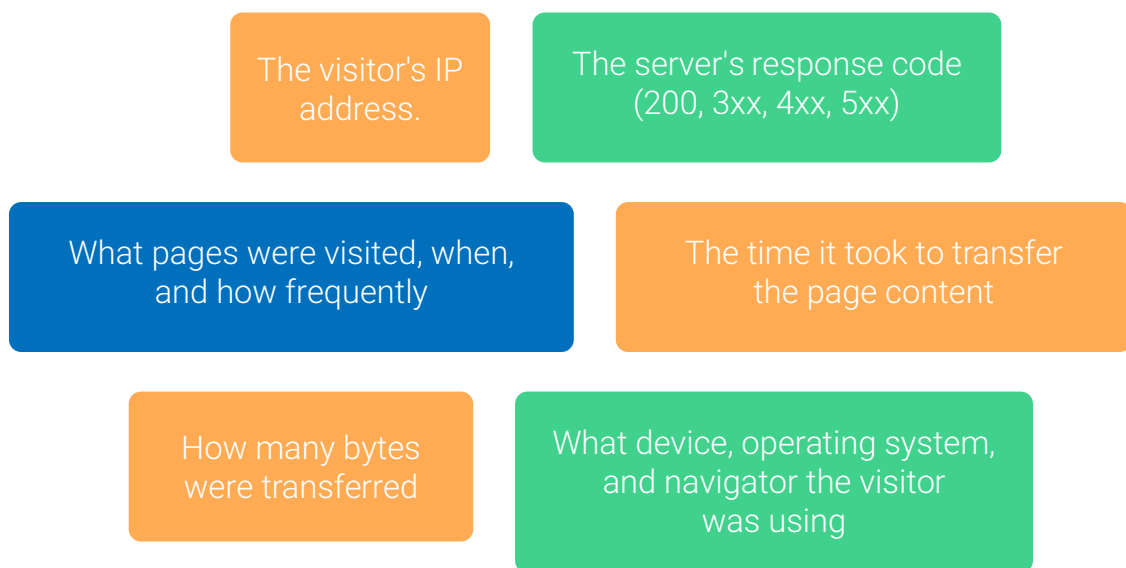
- Collaboration and sharing functions
- Crawl-to-crawl comparison
- Advanced data visualization
- Powerful crawling and customization
- Support for larger sites in both crawl and crawl analysis

Cloud crawlers like OnCrawl's SEO crawler also cross-analyze crawl data with data from other sources, providing site-wide analyses and per-page information on organic traffic, search engine bot visits, social media traffic, structured data, social media markup, page popularity, site architecture, duplicate content, and near-duplicate content.

## WHAT ARE THE TOOLS USED IN TECHNICAL SEO?

### Log analyzers

SEO log analyzers read server logs specifically for information that is of interest to SEOs. This means finding and analyzing records for organic traffic, traffic from social media, and search engine bot traffic. They can tell:



This type of information can be used in deeper analysis:

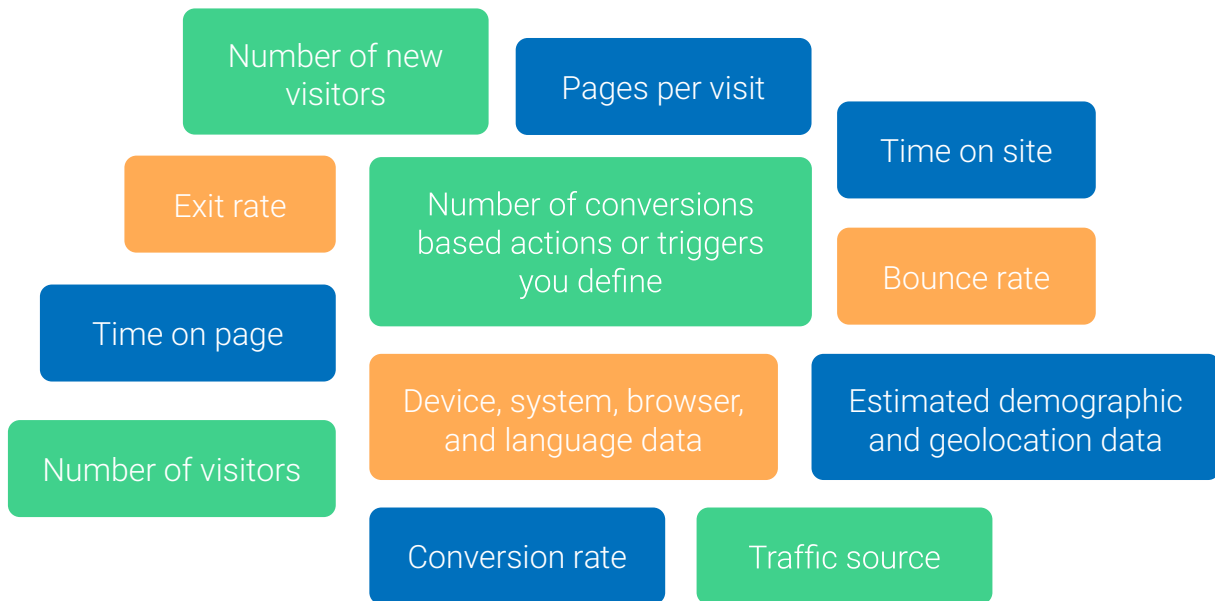
- Provide user navigation profiles
- Identify bots and their behavior
- Monitor site healthy
- Establish the delay between events such as publishing and indexing, indexing and the first organic traffic, exploratory crawls and rendering crawls, first and return visits...

SEO log analysis is critical because it offers the only complete, authoritative data on the activity on a website. Unlike other tools, log analysis does not rely on a third-party service or on tracking scripts or pixels. This means it can't be blocked, and it won't stop working.

## WHAT ARE THE TOOLS USED IN TECHNICAL SEO?

### Analytics solutions

Analytics solutions like Google Analytics use tracking codes, usually lightweight JavaScript, placed on each page of a website in order to monitor and understand user behavior. They will provide information such as:



### Google Search Console

Google Search Console provides SEO tools for people who run websites. These tools allow you to submit sitemaps to Google, monitor and fix errors on your site, submit URLs for indexing or check their indexing status, and more.

It also offers Google's official statistics for:

- Keywords your pages rank for
- Your pages' average position in the search results
- Impressions: number of times your pages are shown in searches
- Clicks: number of times your pages are clicked on from the results pages
- Click-through-rate: the rate at which your pages are clicked in search results pages

## WHAT ARE THE TOOLS USED IN TECHNICAL SEO?

In 2018, Google began a move to a new Search Console interface; some reports and features are still only available in the old version.

Other search engines offer similar tools.

### **Keyword trackers**

Keyword trackers examine how searchers use search engines. They report on how many times a keyword is searched per month (search volume), how many results are provided, and how pages on a given site rank for that keyword. They'll offer estimates on keyword cost and difficulty, whether rich result features are available for the keyword, and other information about the keyword. They may show search result page stability, or even current search trends.

Keyword trackers are based on examining search results and on algorithmic estimations. This accounts for the differences between numbers reported by different tools.

### **Other tools**

Other tools commonly used by technical SEOs include:

- Backlink trackers (Majestic...)
- Spreadsheets (Excel, Google Sheets, Libreoffice...)
- Business Intelligence platforms (Data Studio, Power BI...)
- Data visualization supports (Tableau...)
- APIs, scripts, and programming languages (Python, R...)

# ONCRAWL: THE TECHNICAL SEO PLATFORM

OnCrawl is an award-winning technical SEO platform that helps you make smarter SEO decisions. OnCrawl combines your content, log files and search data at scale so that you can open Google's blackbox and build an SEO strategy with confidence. Backed by a SEO crawler, a log analyzer and third-party integrations, OnCrawl currently works with over 800 clients in 66 countries including e-commerce websites, online publishers and travel websites. OnCrawl produces actionable dashboards and reports to support your entire search engine optimization process and helps you improve your rankings, traffic and revenues.

Start your 14-day free trial

[hello@oncrawl.com](mailto:hello@oncrawl.com)  
[www.oncrawl.com](http://www.oncrawl.com)



**Oncrawl**