**Google Webmaster Tools**

About our stats and data

We're always working to increase the update frequency for your verified sites' data, such as crawl, index, and search query stats. Much of this data depends on the content of your site, and is a close approximation of the status of your site. Our internal systems are always changing, and the web itself is an ever-shifting ecosystem. In addition, there may be a lag between when the numbers are calculated and when they are visible to webmasters - although data gets published in intervals, we are continually collecting it. If your content doesn't change very often, or if you're not getting new links to your site, you may not see updates to your data every time you sign in to Webmaster Tools.

Webmaster Tools provides data based on site visits and links to your pages. You need to get the word out about your web site to get more people visiting and naturally linking to it - that is, of course, after you've built a web site that people will want to visit and link to. The more links to your site on the web, the more likely it is that Googlebot will stop by for a visit. Once Google starts crawling your site more often, you'll notice that Webmaster Tools will begin to show more detailed data, and that this data is updated more often.

Webmaster Tools provides information and data about the sites you have added to your account. You can use this data to improve how search engines crawl and index your site's content.

### Crawl data

The **Crawl** section provides information about potential problems Google may have encountered when crawling and indexing your site.

* [Crawl Errors](https://support.google.com/webmasters/answer/answer.py?answer=35120) lists any pages on your site that Google was unable to access, and why we couldn’t crawl them. Addressing these issues can help ensure that all the content on your site can be successfully crawled by search engines.
* [Crawl Stats](https://support.google.com/webmasters/answer/answer.py?answer=35253) age lists recent Googlebot activity on your site.
* [Blocked URLs](https://support.google.com/webmasters/answer/answer.py?hl=en&answer=2579744) shows you which pages on your site Google was prevented from crawling because of restrictions in your [robots.txt](https://support.google.com/webmasters/answer/answer.py?hl=en&answer=156449) file.
* [Malware](https://support.google.com/webmasters/answer/answer.py?answer=163633) helps you identify any potential problems with hacking and malicious software on your site.

### Traffic data

The **Traffic** section provides details about how users reach your site, and how it appears to search engines and humans.

* [Search Queries](https://support.google.com/webmasters/answer/answer.py?answer=35252) lists the top search queries for which your site appeared in search results, as well as queries for which users actually clicked through to your site.
* [Links to Your Site](https://support.google.com/webmasters/answer/answer.py?answer=55281) provides details about pages on your site linked to from other sites.
* [Internal Links](https://support.google.com/webmasters/answer/answer.py?answer=138752) lists URLs on your site that are linked to from other URLs on the same site.
* +1 Reports provides insight into the effects of the Google +1 button on your site's traffic. Data is available for [Search Impact](https://support.google.com/webmasters/answer/answer.py?answer=1140189), [Activity](https://support.google.com/webmasters/answer/answer.py?answer=1140192), and [Audience](https://support.google.com/webmasters/answer/answer.py?answer=1198127).

### Optimization data

* [Sitemaps](https://support.google.com/webmasters/answer/answer.py?answer=78808) lists details about any Sitemaps submitted for your site, including Sitemaps submitted by others.
* [Remove URLs](https://support.google.com/webmasters/answer/answer.py?answer=1663416) lets you remove your URLs from Google's search results, and lists the status of existing removal requests.
* [HTML Improvements](https://support.google.com/webmasters/answer/answer.py?answer=80407) page lists ways in which you can improve the HTML on your site.
* [Content Keywords](https://support.google.com/webmasters/answer/answer.py?answer=35255) lists the most significant terms Google found on your site. It provides insight into how the content of your site is perceived.

Check your user index status

You can use special searches to see sample results that indicate how your site is indexed by Google. To see a full list of these, use the info: operator, like this: [info:google.com](http://www.google.com/search?&q=info%3Agoogle.com). See below for more details on each search type. Or site:google.com. You can perform the search on a whole domain or limit it to a certain subdomain or subdirectory—for example, [site:google.com/webmasters](http://www.google.com/search?hl=en&q=site%3Agoogle.com%2Fwebmasters).

To exclude pages from your search, use a minus sign before the operator. For example, the search [site:google.com -site:adwords.google.com](http://www.google.com/search?hl=en&q=site%3Agoogle.com+-site%3Aadwords.google.com) gives you all the indexed pages on the google.com domain without the pages from adwords.google.com

If you do a site: search for a page that is redirected, you'll see the redirected URL in the results.

Use [link:google.com](http://www.google.com/search?hl=en&q=link%3Agoogle.com). (You can use link:google.com or link:www.google.com, but the first search will return more complete results.) You can also search for links to specific pages or directories: [link:google.com/webmasters](https://www.google.com/search?hl=en&q=link%3Agoogle.com%2Fwebmasters).

### See Google's cache of your site

Use the cache: operator to see an archived copy of a page indexed by Google. For example, [cache:google.com](https://www.google.com/search?hl=en&q=cache%3Agoogle.com) displays the last indexed version of the Google homepage, along with information about the date the cache was created. You can also view a plain-text version of the page. This is useful because it shows how Googlebot sees the page.

If you don't want searchers to be able to access a cached version of your page, use the [noarchive meta tag](https://support.google.com/webmasters/answer/answer.py?answer=35306) like this:

<meta name="robots" content="noarchive">

The page will still be crawled and indexed by Google, but users will not see a Cached link in the search results.

### See pages that are similar to your site

The related: operator displays websites similar to the site you are looking for. It returns the same results as clicking **Similar pages** next to a result on our search results page.

This search is like searching a bookstore to find books similar to the first Harry Potter novel. The results could include other children's books, a biography of J.K. Rowling, or a non-fiction book on children's literature. In general, use this operator to find resources that overlap. You'll get the best and most useful results if you use sites that cover a broad range of content.

Google uses several factors to determine the similarity of different sites. However, the quality of the sites returned has no impact on your ranking or on how Google indexes your site.

Check your site performance

Google's goal is to return highly relevant results for every query. Search results are returned from our [search index](https://support.google.com/webmasters/answer/answer.py?answer=70897). Our search index is constantly evolving as content is added and modified on the web. The changing content, as well as updates to our ranking algorithms, can cause URLs to change position in search results, and possibly, though less likely, be removed.

If your site isn't appearing in Google search results, or it's performing more poorly than it once did, check out these steps to identify and fix potential causes of the problem.

* [Check your site is in the Google index](https://support.google.com/webmasters/answer/34444?hl=en#1)
* [See if your site has been impacted by a manual spam action](https://support.google.com/webmasters/answer/34444?hl=en#spam)
* [Make sure Google can find and crawl your site](https://support.google.com/webmasters/answer/34444?hl=en#2)
* [Make sure that Google can index your site](https://support.google.com/webmasters/answer/34444?hl=en#3)
* [Make sure your content is useful and relevant](https://support.google.com/webmasters/answer/34444?hl=en#4)

**Verify that your site ranks for your domain name**

Do a Google search for www.[yourdomain].com. If your site doesn't appear in the results, or if it ranks poorly in the results, this is a sign that your site may have a manual spam action for violations of the [Webmaster Guidelines](https://support.google.com/webmasters/answer/answer.py?answer=35769). If we find certain problems with your site—for example, malware—we'll let you know via the [Message Center](https://www.google.com/webmasters/tools/message-list).

### Make sure Google can find and crawl your site

Crawling is how Googlebot discovers new and updated pages to be added to the Google index. Our crawl process begins with a list of web page URLs, generated from previous crawl processes, and augmented with Sitemap data provided by webmasters. As Googlebot visits each of these websites, it detects links on each page and adds them to its list of pages to crawl. New sites, changes to existing sites, and dead links are noted and used to update the Google index.

* Check for [crawl errors](https://support.google.com/webmasters/answer/answer.py?answer=35120). The **Crawl errors** page in Webmaster Tools provides details about the URLs in your site that we tried to crawl and couldn't access. Review these errors, and fix any you can. The next time Googlebot crawls your site, it will note the changes and use them to update the Google index.
* Review your robots.txt file. The [Test robots.txt tool](https://support.google.com/webmasters/answer/answer.py?answer=35237) lets you analyze your robots.txt file to see if you're blocking Googlebot from any URLs or directories on your site.
* Make sure that the URLs haven't been [blocked with meta tags](https://support.google.com/webmasters/answer/answer.py?answer=61050).
* Review your site structure and [make sure that it's easily accessible](https://support.google.com/webmasters/answer/answer.py?answer=72746). Most search engines are text-based. If you use JavaScript, DHTML, images, or rich media such as Silverlight to create navigation and links, Googlebot and other spiders may have trouble crawling your site.
* If you have recently restructured your site or moved to a new domain, pages that previously performed well may now rank poorly. To avoid this, use 301 redirects ("RedirectPermanent") in your .htaccess file to smartly redirect users, Googlebot, and other spiders. (In Apache, you can do this with an .htaccess file; in IIS, you can do this through the administrative console.) For more information about 301 HTTP redirects, please see [http://www.ietf.org/rfc/rfc2616.txt.](http://www.ietf.org/rfc/rfc2616.txt)
* Consider [creating and submitting a Sitemap](https://support.google.com/webmasters/answer/156184). Even if your site is already indexed, Sitemaps are a way to give Google information about your site and the URLs you consider most important. Sitemaps are particularly helpful if your site has dynamic content or other content not easily discoverable by Googlebot, or if your site is new or does not have many links to it.

### Make sure Google can index your site

Googlebot processes each of the pages it crawls in order to compile a massive index of all the words it sees and their location on each page. In addition, we process information included in key content tags and attributes, such as title tags and alt attributes. Google can process many types of content. However, while we can process HTML, PDF, and Flash files, we have a more difficult time understanding (e.g. crawling and indexing) other rich media formats, such as Silverlight.

* Check your site's [index stats](https://support.google.com/webmasters/answer/answer.py?answer=35256). These stats show how your site is represented in the Google index.
* Review your site's structure. Google, like most search engines, is text-based. This means that Googlebot can't read text included in images or in most rich media files other than Flash files, or pages that are hidden behind JavaScript or require sign-in. [Making sure your content is text-based and readable](https://support.google.com/webmasters/answer/answer.py?answer=72746) helps make it more accessible to humans and to Googlebot.

### Make sure your content is relevant and useful

* Understand how users are reaching your site by reviewing the [Search queries](https://support.google.com/webmasters/answer/answer.py?answer=35252) page. The first column shows the Google searches in which your site most often appears. The page also lists the number of impressions, the number of clicks, and the CTR (click-through rate) for each query. This information is particularly useful because it gives you an insight into what users are searching for (the query), and the queries for which users often click on your site. For example, your site may often appear in Google searches for espresso gadgets and coffee widgets, but if your site has a low CTR for this query, it could be because it's not clear to users that your site contains information about coffee widgets. In this case, consider revising your content to make it more compelling and relevant. Avoid [keyword stuffing](https://support.google.com/webmasters/answer/answer.py?answer=66358), though, because this can cause your site's ranking to suffer, as well as degrading the user experience for your readers.
* Understand how Google sees your site. The [Content Keywords](https://support.google.com/webmasters/answer/answer.py?answer=35255) page shows the keywords and phrases other sites use when they link to yours. Understanding how other people see your site can help you figure out how best to target your audience.
* Check the [HTML Improvements](https://support.google.com/webmasters/answer/answer.py?answer=80407) page in Webmaster Tools. Descriptive information in title tags and meta descriptions will give us good information about the content of your site. In addition, this text can appear in search results pages, and useful, descriptive text is more likely to be clicked on by users.
* Tell the world about your site. Incoming links to your site help Google determine your site's relevance to the user's query. Natural links to your site develop as part of the dynamic nature of the web when other sites find your content valuable and think it would be helpful for their visitors.
* Check to see if any of your content has been flagged as adult content by turning off [SafeSearch](http://video.google.com/support/bin/answer.py?answer=66500). Google's SafeSearch filter eliminates sites that contain pornography and explicit sexual content from search results. While no filter is 100% accurate, SafeSearch uses advanced proprietary technology that checks keywords and phrases, URLs, and Open Directory categories.
* Great image content can be an excellent way to generate traffic. We recommend that when publishing images, you think carefully about creating the best user experience you can, and follow our [image guidelines](https://support.google.com/webmasters/answer/answer.py?answer=114016).

There's almost nothing a competitor can do to harm your ranking or have your site removed from our index. If you're concerned about another site linking to yours, we suggest contacting the webmaster of the site in question. Google aggregates and organizes information published on the web; we don't control the content of these pages.

Occasionally, fluctuation in search results is the result of differences in our data centers. When you perform a Google search, your query is sent to a Google data center in order to retrieve search results. There are numerous data centers, and many factors (such as geographic location and search traffic) determine where a query is sent. Because not all of our data centers are updated simultaneously, it's possible to see slightly different search results depending on which data center handles your query.

Make sure googlebot is not blocked

### Intentional webmaster blocking

Some webmasters intentionally prevent Googlebot from reaching their websites, perhaps using a firewall as described above. In these cases, usually the intent is not to entirely block Googlebot, but to control how the site is crawled and indexed. In this case, check the following:

* If you would like to control Googlebot’s crawling of your content, we have detailed help about using the [robots exclusion protocol](http://code.google.com/web/controlcrawlindex/docs/getting_started.html" \t "_blank), including using a [robots.txt file](https://support.google.com/webmasters/answer/answer.py?answer=156449) and [configuring URL parameters](https://support.google.com/webmasters/answer/1235687).
* If you’re worried about rogue bots using the Googlebot user-agent, you can [verify whether a crawler is actually Googlebot](https://support.google.com/webmasters/answer/answer.py?answer=80553).
* If you would like to change how frequently Googlebot crawls your site, you can [verify your website](https://support.google.com/webmasters/answer/answer.py?answer=34592) in Webmaster Tools and [change Googlebot’s crawl rate](https://support.google.com/webmasters/answer/answer.py?answer=48620). Hosting providers can verify ownership of their IP addresses too.

Review your page titles and snippets

Google's generation of page titles and descriptions (or "snippets") is completely automated and takes into account both the content of a page as well as references to it that appear on the web. The goal of the snippet and title is to best represent and describe each result and explain how it relates to the user's query.

The more information you give us, the better your search result snippet can be. With rich snippets, webmasters with sites containing structured content—such as review sites or business listings—can label their content to make it clear that each labeled piece of text represents a certain type of data: for example, a restaurant name, an address, or a rating. [Learn more about how rich snippets can improve your site's listing in search results.](https://support.google.com/webmasters/answer/answer.py?answer=99170)

We use a number of different sources for this information, including descriptive information in the title and meta tags for each page. We may also use publicly available information—for instance, anchor text or listings from the Open Directory Project (DMOZ)—or create [rich snippets](https://support.google.com/webmasters/answer/answer.py?answer=99170) based on markup on the page.

While we can't manually change titles or snippets for individual sites, we're always working to make them as relevant as possible. You can help improve the quality of the title and snippet displayed for your pages by following the general guidelines below.

### Create descriptive page titles

Titles are critical to giving users a quick insight into the content of a result and why it’s relevant to their query. It's often the primary piece of information used to decide which result to click on, so it's important to use high-quality titles on your web pages.

Here are a few tips for managing your titles:

* As explained above, **make sure every page on your site has a title specified in the <title> tag**. If you’ve got a large site and are concerned you may have forgotten a title somewhere, the [HTML suggestions](https://support.google.com/webmasters/answer/answer.py?answer=80407) page in Webmaster Tools lists missing or potentially problematic <title> tags on your site.
* Page titles should be **descriptive and concise**. Avoid vague descriptors like "Home" for your home page, or "Profile" for a specific person's profile. Also avoid unnecessarily long or verbose titles, which are likely to get truncated when they show up in the search results.
* Avoid **keyword stuffing**. It's sometimes helpful to have a few descriptive terms in the title, but there’s no reason to have the same words or phrases appear multiple times. A title like "Foobar, foo bar, foobars, foo bars" doesn't help the user, and this kind of [keyword stuffing](https://support.google.com/webmasters/answer/answer.py?answer=66358) can make your results look spammy to Google and to users.
* Avoid **repeated or boilerplate titles**. It’s important to have distinct, descriptive titles for each page on your site. Titling every page on a commerce site "Cheap products for sale", for example, makes it impossible for users to distinguish one page differs another. Long titles that vary by only a single piece of information ("boilerplate" titles) are also bad; for example, a standardized title like "<band name> - See videos, lyrics, posters, albums, reviews and concerts" contains a lot of uninformative text. One solution is to dynamically update the title to better reflect the actual content of the page: for example, include the words "video", "lyrics", etc., only if that particular page contains video or lyrics. Another option is to just use "<band name>" as a concise title and use the meta description (see below) to describe your site's content. The [HTML suggestions](https://support.google.com/webmasters/answer/answer.py?answer=80407) page in Webmaster Tools lists any duplicate titles Google detected on your pages.
* **Brand your titles**, but concisely. The title of your site’s home page is a reasonable place to include some additional information about your site—for instance, "ExampleSocialSite, a place for people to meet and mingle." But displaying that text in the title of every single page on your site hurts readability and will look particularly repetitive if several pages from your site are returned for the same query. In this case, consider including just your site name at the beginning or end of each page title, separated from the rest of the title with a delimiter such as a hyphen, colon, or pipe, like this:

<title>ExampleSocialSite: Sign up for a new account.</title>

* **Be careful about disallowing search engines** from crawling your pages. Using the [robots.txt](http://code.google.com/web/controlcrawlindex/" \t "_blank) protocol on your site can stop Google from crawling your pages, but it may not always prevent them from being indexed. For example, Google may index your page if we discover it by following a link from someone else's site. To display it in search results, Google will need to display a title of some kind and because we won't have access to any of your page content, we will rely on off-page content such as anchor text from other sites. (To truly block a URL from being indexed, you can use [meta tags](https://support.google.com/webmasters/answer/answer.py?answer=93710).)

If we’ve detected that a particular result has one of the above issues with its title, we may try to generate an improved title from anchors, on-page text, or other sources. However, sometimes even pages with well-formulated, concise, descriptive titles will end up with different titles in our search results to better indicate their relevance to the query. There’s a simple reason for this: the title tag as specified by a webmaster is limited to being static, fixed regardless of the query. Once we know the user’s query, we can often find alternative text from a page that better explains why that result is relevant. Using this alternative text as a title helps the user, and it also can help your site. Users are scanning for their query terms or other signs of relevance in the results, and a title that is tailored for the query can increase the chances that they will click through.

If you’re seeing your pages appear in the search results with modified titles, check whether your titles have one of the problems described above. If not, consider whether the alternate title is a better fit for the query. If you still think the original title would be better, let us know in our [Webmaster Help Forum](https://groups.google.com/a/googleproductforums.com/forum/#%21forum/webmasters).

### Create good meta descriptions

The description attribute within the <meta> tag is a good way to provide a concise, human-readable summary of each page’s content. Google will sometimes use the meta description of a page in search results snippets, if we think it gives users a more accurate description than would be possible purely from the on-page content. Accurate meta descriptions can help improve your clickthrough; here are some guidelines for properly using the meta description.

* Make sure that **every page on your site has a meta description**. The [HTML suggestions](https://support.google.com/webmasters/answer/answer.py?answer=80407) page in Webmaster Tools lists pages where Google has detected missing or problematic meta descriptions.
* **Differentiate the descriptions for different pages.** Identical or similar descriptions on every page of a site aren't helpful when individual pages appear in the web results. In these cases we're less likely to display the boilerplate text. Wherever possible, create descriptions that accurately describe the specific page. Use site-level descriptions on the main home page or other aggregation pages, and use page-level descriptions everywhere else. If you don't have time to create a description for every single page, try to prioritize your content: At the very least, create a description for the critical URLs like your home page and popular pages.
* **Include clearly tagged facts in the description.** The meta description doesn't just have to be in sentence format; it's also a great place to include structured data about the page. For example, news or blog postings can list the author, date of publication, or byline information. This can give potential visitors very relevant information that might not be displayed in the snippet otherwise. Similarly, product pages might have the key bits of information—price, age, manufacturer—scattered throughout a page. A good meta description can bring all this data together. For example, the following meta description provides detailed information about a book.

<meta name="Description" content="Author: A.N. Author, Illustrator: P. Picture, Category: Books, Price: $17.99, Length: 784 pages">

In this example, information is clearly tagged and separated.

* **Programmatically generate descriptions.** For some sites, like news media sources, generating an accurate and unique description for each page is easy: since each article is hand-written, it takes minimal effort to also add a one-sentence description. For larger database-driven sites, like product aggregators, hand-written descriptions can be impossible. In the latter case, however, programmatic generation of the descriptions can be appropriate and are encouraged. Good descriptions are human-readable and diverse, as we talked about in the first point above. The page-specific data we mentioned in the second point is a good candidate for programmatic generation. Keep in mind that meta descriptions comprised of long strings of keywords don't give users a clear idea of the page's content, and are less likely to be displayed in place of a regular snippet.
* **Use quality descriptions.** Finally, make sure your descriptions are truly descriptive. Because the meta descriptions aren't displayed in the pages the user sees, it's easy to let this content slide. But high-quality descriptions can be displayed in Google's search results, and can go a long way to improving the quality and quantity of your search traffic.

### Prevent search engines from displaying DMOZ data in search results for your site

One source Google uses to generate snippets is the [Open Directory Project](http://www.dmoz.org/). You can direct us not to use this as a source by adding a meta tag to your pages.

To prevent all search engines (that support the meta tag) from using this information for the page's description, use the following:

<meta name="robots" content="NOODP">

To specifically prevent Google from using this information for a page's description, use the following:

<meta name="googlebot" content="NOODP">

If you use the robots meta tag for other directives, you can combine those. For instance:

<meta name="googlebot" content="NOODP, nofollow">

Note that once you add this meta tag to your pages, it may take some time for changes to your snippets to appear in the index.

If you're concerned about content in your title or snippet, you may want to double-check that this content doesn't appear on your site. If it does, changing it may affect your Google snippet after we next crawl your site. If it doesn't, try searching [Google.com](http://www.google.com) for the title or snippet enclosed in quotation marks. This will display pages on the web that refer to your site using this text. If you contact these webmasters to request that they change their information about your site, any changes to their sites will be recognized by our crawler after we next crawl their pages.

The snippet text is bolded when it’s exact to the search query, the texts snippet can come from your meta tag description, the body of your page or the open directory Project.

Sitelinks (under the snippet)

### Security checklist

In addition to monitoring your site regularly, we also recommend the following:

#### All webmasters

* **Choose good passwords.** The Gmail [guidelines](http://mail.google.com/support/bin/answer.py?answer=29409) are helpful.
* **Pick third-party content providers very carefully.** If you're considering installing an application provided by a third party, such as a widget, counter, or ad network, be sure to exercise due diligence. Ad space is often syndicated to other parties who are not known to the website owner. While there are many great third-party content on the web, it's also possible for providers to use these applications to push exploits, such as dangerous scripts, towards your visitors. Make sure the application is from a reputable source. Do they have a legitimate website with support and contact information? Have other webmasters used the service.
* **Contact your hosting company or publishing platform for support.** Most companies have helpful and responsive support groups and/or security pages. If a security page or site has an RSS feed, subscribe to it to make sure you stay up to date.
* Keep all of your computers safe. Especially when working on a website, make sure that your local workstation has up-to-date software, is clean from viruses, trojans or similar malware and has recently updated anti-virus software installed.

#### Webmasters with server access

* **Check your server configuration.** Apache has some [security configuration tips](http://httpd.apache.org/docs/1.3/misc/security_tips.html) on their site and Microsoft has some [tech center resources for IIS](http://technet2.microsoft.com/windowsserver/en/library/354f4539-982a-418c-bfe7-4d5155b83f4a1033.mspx?mfr=true) on theirs. Some of these tips include information on directory permissions, server-side includes, authentication and encryption.
* **Make a backup copy of your .htaccess file** (or other access control mechanisms depending on your website platform). Use your backup file to recover if the following fails. Be sure to delete the backup file once you are finished.
* **Stay up-to-date with the latest software updates and patches.** There are lots of tools that make building a website easy, but each one adds some risk of being exploited. A common pitfall for many webmasters is to install a forum or blog on their website and then forget about it. Much like taking your car in for a tune-up, it's important to make sure you have all the latest updates for any software program you have installed. Make a list of all the software and plug-ins used for your website, and keep track of the version numbers and updates. Even if you're diligent and keep all your website components updated, you may still be vulnerable if your web hoster has not installed the most recent operating system patches. This is not a problem only for smaller sites; there have been warnings on the websites of banks, sports teams, and corporate and government websites.
* **Keep an eye on your log files.** Making this a habit has many great benefits, one of which is added security. For example, unfamiliar URL parameters (like "=http:" or "=//") or spikes in traffic to redirect URLs on your site may indicate that a hacker is exploiting [open redirects](https://support.google.com/webmasters/answer/answer.py?answer=171297). Also, bear in mind that hackers often try to alter log files. Take measures to protect these files from attack. For example, you can move these files from their default location, making it harder for hackers to find them.
* **Check your site for common vulnerabilities.** Avoid having directories with open permissions. This is like leaving the front door to your home wide open.

Also check for any [XSS](http://www.owasp.org/index.php/Cross_Site_Scripting) (cross-site scripting) and [SQL injection](http://www.owasp.org/index.php/SQL_injection) vulnerabilities.

* **Use secure protocols.** Google recommends using SSH and SFTP for data transfer, rather than plain text protocols such as telnet or FTP. SSH and SFTP use encryption and are much safer. For this and many other useful tips, check out StopBadware.org's [Tips for Cleaning and Securing Your Website](http://www.stopbadware.org/home/security).
* **Keep up to date on the latest security news.** The [Google Online Security Blog](http://googleonlinesecurity.blogspot.com/) provides useful information about online security and safety, as well as pointers to other resources. The government site [US-CERT](http://www.us-cert.gov/) (United States Computer Emergency Readiness Team) provides technical security alerts and tips.

### Learn how Google discovers, crawls, and serves web pages

When you sit down at your computer and do a Google search, you're almost instantly presented with a list of results from all over the web. How does Google find web pages matching your query, and determine the order of search results?

In the simplest terms, you could think of searching the web as looking in a very large book with an impressive index telling you exactly where everything is located. When you perform a Google search, our programs check our index to determine the most relevant search results to be returned ("served") to you.

The three key processes in delivering search results to you are:

* [Crawling: Does Google know about your site? Can we find it?](https://support.google.com/webmasters/answer/70897?hl=en#1)
* [Indexing: Can Google index your site?](https://support.google.com/webmasters/answer/70897?hl=en#2)
* [Serving: Does the site have good and useful content that is relevant to the user's search?](https://support.google.com/webmasters/answer/70897?hl=en#3)

## Crawling

Crawling is the process by which [Googlebot](https://support.google.com/webmasters/answer/answer.py?answer=182072) discovers new and updated pages to be added to the Google index.

We use a huge set of computers to fetch (or "crawl") billions of pages on the web. The program that does the fetching is called Googlebot (also known as a robot, bot, or spider). Googlebot uses an algorithmic process: computer programs determine which sites to crawl, how often, and how many pages to fetch from each site.

Google's crawl process begins with a list of web page URLs, generated from previous crawl processes, and augmented with Sitemap data provided by webmasters. As Googlebot visits each of these websites it detects links on each page and adds them to its list of pages to crawl. New sites, changes to existing sites, and dead links are noted and used to update the Google index.

Google doesn't accept payment to crawl a site more frequently, and we keep the search side of our business separate from our revenue-generating AdWords service.

## Indexing

Googlebot processes each of the pages it crawls in order to compile a massive index of all the words it sees and their location on each page. In addition, we process information included in key content tags and attributes, such as Title tags and ALT attributes. Googlebot can process many, but not all, content types. For example, we cannot process the content of some rich media files or dynamic pages.

## Serving results

When a user enters a query, our machines search the index for matching pages and return the results we believe are the most relevant to the user. Relevancy is determined by over 200 factors, one of which is the [PageRank](http://infolab.stanford.edu/%7Ebackrub/google.html) for a given page. PageRank is the measure of the importance of a page based on the incoming links from other pages. In simple terms, each link to a page on your site from another site adds to your site's PageRank. Not all links are equal: Google works hard to improve the user experience by identifying spam links and other practices that negatively impact search results. The best types of links are those that are given based on the quality of your content.

In order for your site to rank well in search results pages, it's important to make sure that Google can crawl and index your site correctly. Our [Webmaster Guidelines](https://support.google.com/webmasters/answer/answer.py?answer=35769) outline some best practices that can help you avoid common pitfalls and improve your site's ranking.

Google's [Did you mean](http://www.google.com/support/websearch/bin/answer.py?answer=1723) and [Google Autocomplete](http://www.google.com/support/websearch/bin/answer.py?answer=106230) features are designed to help users save time by displaying related terms, common misspellings, and popular queries. Like our [google.com](http://www.google.com) search results, the keywords used by these features are automatically generated by our web crawlers and search algorithms. We display these predictions only when we think they might save the user time. If a site ranks well for a keyword, it's because we've algorithmically determined that its content is more relevant to the user's query.

### Make sure your content is relevant and useful

* Understand how users are reaching your site by reviewing the [Search queries](https://support.google.com/webmasters/answer/answer.py?answer=35252) page. The first column shows the Google searches in which your site most often appears. The page also lists the number of impressions, the number of clicks, and the CTR (click-through rate) for each query. This information is particularly useful because it gives you an insight into what users are searching for (the query), and the queries for which users often click on your site. For example, your site may often appear in Google searches for espresso gadgets and coffee widgets, but if your site has a low CTR for this query, it could be because it's not clear to users that your site contains information about coffee widgets. In this case, consider revising your content to make it more compelling and relevant. Avoid [keyword stuffing](https://support.google.com/webmasters/answer/answer.py?answer=66358), though, because this can cause your site's ranking to suffer, as well as degrading the user experience for your readers.
* Understand how Google sees your site. The [Content Keywords](https://support.google.com/webmasters/answer/answer.py?answer=35255) page shows the keywords and phrases other sites use when they link to yours. Understanding how other people see your site can help you figure out how best to target your audience.
* Check the [HTML Improvements](https://support.google.com/webmasters/answer/answer.py?answer=80407) page in Webmaster Tools. Descriptive information in title tags and meta descriptions will give us good information about the content of your site. In addition, this text can appear in search results pages, and useful, descriptive text is more likely to be clicked on by users.
* Tell the world about your site. Incoming links to your site help Google determine your site's relevance to the user's query. Natural links to your site develop as part of the dynamic nature of the web when other sites find your content valuable and think it would be helpful for their visitors.
* Check to see if any of your content has been flagged as adult content by turning off [SafeSearch](http://video.google.com/support/bin/answer.py?answer=66500). Google's SafeSearch filter eliminates sites that contain pornography and explicit sexual content from search results. While no filter is 100% accurate, SafeSearch uses advanced proprietary technology that checks keywords and phrases, URLs, and Open Directory categories.
* Great image content can be an excellent way to generate traffic. We recommend that when publishing images, you think carefully about creating the best user experience you can, and follow our [image guidelines](https://support.google.com/webmasters/answer/answer.py?answer=114016).

If you don’t want sth to appear in a search engine then better contact the webmaster

* If a photo or information shows up in Google search results, it just means that the information exists on the Internet and it doesn’t mean that Google endoreses it.

# Learn about Fetch as Google

Fetch as Google is a diagnostic tool that allows you to simulate how Google crawls or renders a URL on your site.

Googlebot constantly crawls and renders URLs across the entire web. Fetching is an operation you run as a user when you use our Fetch as Google tool to retrieve a specific URL. It’s a small-scale simulation of the real deal.

The Fetch as Google tool has two modes of operation that you can use to test Google crawling and rendering for URL on your site:

### ****Fetch—quick check****

When the Fetch as Google tool is in fetch mode, Googlebot crawls any URL that corresponds to the path that you requested. If Googlebot is able to successfully crawl your requested URL, you can review the response your site sent to Googlebot. This is a relatively quick, low-level operation that you can use to check or debug suspected network connectivity or security issues with your site.

### ****Fetch and render—deeper view****

The fetch and render mode tells Googlebot to crawl and display your page as browsers would display it to your audience. First, Googlebot gets all the resources referenced by your URL such as picture, CSS, and JavaScript files, running any code. to render or capture the visual layout of your page as an image. You can use the rendered image to detect differences between how Googlebot sees your page, and how your browser renders it.

**Use fetch as google**

Here are the steps to test the crawling and rendering of your site with Fetch as Google:

**Example textbox entries:**

* “/” fetches your homepage.
* “/images/puppies/” fetches all the images files in that directory of your domain.

1. On the **Webmaster Tools** home page, select the site you would like to fetch.
2. On the left dashboard, click **Crawl** and then click **Fetch as Google**, which appears as a child menu item below **Crawl**.
3. In the textbox, enter the path component of a URL on your site that you want Googlebot to retrieve.
4. From the grey drop-down that reads **Googlebot**, choose the type of Googlebot you wish to perform a fetch (or fetch and render) with.

**Choose a Type of Googlebot**   
  
Your site users see your content differently when they view your web pages from different devices. You can select a Googlebot from the dropdown list to simulate different browsers types that members of your site audience might be using. As an example, for a fetch and render operation, the size of the rendered image typically varies based on the Googlebot you choose.

* To show the version of your site your browser sees, select **Desktop**.
* To see what our mobile crawler for smartphones sees, select **Mobile Smartphone**.
* To see what our mobile crawler for feature phones sees, select **Mobile cHTML** (mainly for Japanese websites) or **Mobile XHTML/WML**.

Note that, for the **WML** and **cHTML** options for mobile feature phones, the rendering function is not available.

 5.  Click **fetch** to see the HTTP response Google received from your page, or click **fetch and render** to see the HTTP response as well as a rendered image of your page.

Following this action, a row is added to the Fetches Table. New rows or fetches are initially given a pending status until Googlebot has finished its operation. The fetch operation is usually faster than the fetch and render operation.

 6.  Click any record in the Fetches Table in order to open the Fetch Details page to learn more about the outcome of your fetch (or fetch and render).

The Fetch Details page can have up to two tabs (Fetching and Rendering) that appear near the top of the page, depending on whether you use the fetch or fetch and renderoperation on your URL.

* Click **Fetching** to see how your site communicates with Google during the crawling process in the form of an HTTP response (learn more in the section below).
* Click **Rendering** to view a rendered image that displays how your site looks to the Googlebot you selected. Beneath the rendered image, you can see a list of links to missing resources (if any) that your page referenced and that Google was unable to retrieve.

You have a weekly quota of **500 fetches**. Once you have 100 fetches remaining, a notification appears below the fetch form to indicate you are approaching your fetch quota limit.

## ****Analyze your Fetch results****

When Googlebot asks your site to crawl a page, your site server responds with an **HTTP response**. You can find the HTTP response from your site for a completed fetch in Fetch as Google: First click on the corresponding row in the Fetches Table to see the Fetch Details page, and then view click **Fetching** to open the Fetching Tab and see the HTTP response text file. The text file can hold data of different types: HTML code, Javascript code, CSS code, image data, and more. For example, if you fetch an image, the Fetching Tab still shows an HTTP response, but that response holds image data symbols.

The following example scenario shows you how to use the Fetching Tab to troubleshoot a problem with the redirection of a URL.

### Example

A URL www.example.com/pageX.php redirects to www.example.com/pageY.php in your web browser but, when you fetch the first path /pageX.php with Fetch as Google, Google isn’t able to follow the redirect to the second page.

To observe more closely the redirection setup, you can read the HTTP response in the Fetching Tab to find exactly what your server said to Google when Google asked to crawl your page. You can scan the HTTP code if you know what to look for, or you can copy-paste your HTTP response code from Fetch as Google, and compare it against what you’d expect to get in a response for a successful fetch.

Note that an HTTP response in the Fetching Tab is truncated if it exceeds a certain length.

Diagnose your fetch as google results

Analyze the outcome of a fetch orfetch & render operation that you performed with Fetch as Google tool by locating the **Fetch Status** of the operation in the Fetches Table on the Fetch as Google page, and looking up this status in either the Fetch Statuses table or the Fetch & Render Statuses table, below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Status** | | **Fetch** | | **Notes & Next Steps** |
| Complete | | Google successfully contacted your site and crawled your page. | | On the Fetch as Google page, click the table row that corresponds to your fetch to see more details about the fetch results. |
| Redirected | | When Google contacts the site server for the URL you want to fetch, the server tells Google to visit a different URL. | | While the actual Google mechanism that crawls URLs on the web follows redirects, the Fetch as Google simulation tool instead tells you about redirects. You can then choose to follow redirects manually, in one of two ways:   * If a page on your site redirects to another URL on your site, the tool displays a button that allows to quickly follow the redirect. * If the URL redirects to a page that Google considers to be on another site, you must add that domain of the other site to the sites associated with your Webmaster Tools account. Once the new site is set up with Webmaster Tools, enter the path for the redirect's destination URL into the textbox on the Fetch as Google page to follow the redirect.   You can inspect the HTTP response on the Fetch as Google details page to look see what is causing the redirect. Locate the HTTP error code to [learn more](https://support.google.com/webmasters/topic/6001951" \t "_blank). |
| Not found | | Google can contact the server, but the server can't find the URL you want to fetch. | | This error indicates that you might see the HTTP 404 error code when you access your page using a web browser. |
| Not authorized | | Google contacted the server, but the server replied that access to your URL is restricted or blocked from Google. | | This error indicates that you might see the HTTP 403 error code when you access your page using a web browser. |
| DNS not found | | Google couldn’t retrieve the URL you entered because the domain name hasn’t been registered. | | Make sure you typed in your domain name (for example, www.example.com) properly so that Google can find your site server. |
| Blocked | | Your robots.txt file is blocking Google from retrieving the URL you want to fetch. | | This error is a common problem that you can fix by updating your [robots.txt file](https://support.google.com/webmasters/answer/6062608" \t "_blank). |
| Unreachable robots.txt | | Google can’t reach your robots.txt file. | | To resolve this issue, read our Help Center articles on how to create and test [robots.txt files](https://support.google.com/webmasters/answer/6062608" \t "_blank). |
| Unreachable | | 1) The site server is taking too long to reply to the request sent by Googlebot.  OR  2) Google can reach your server, but your server won't allow Google to retrieve the URL Googlebot requested. | | Check to see that your server is up and running. |
| Temporarily unreachable | | 1) Fetch as Google can’t currently fetch your URL because the server took too long to reply.  OR  2) Fetch as Google cancelled your fetch because too many consecutive requests were made to the server for different URLs. | | Note the URL is not unreachable for all of Google-- it is just unreachable for the Fetch as Google simulation tool. |
| Error | | An error prevented Google from completing the fetch. | | If this error happens again, we ask that you contact Webmaster Tools product support. |
| **Status** | **Fetch and Render** | | **Notes & Next Steps** | |
| Complete | Google can successfully contact your site, crawl your page, and get every resource referenced by that page. | | On the Fetch as Google page, click the on the table row that corresponds to your fetch to see more details about the results of the operation. | |
| Partial | Google got a response from your site and fetched URL, but we could not reach some of the resources on that URL for rendering. | | You can assess the gravity of the situation by clicking through the missing resources. A fetch with the Partial status could mean a small problem with the page (e.g. few or insignificant resources could not be retrieved) or a more serious one (e.g. many or important resources which include a large part of the page content are missing).  Note that occasionally, blocked resources reside on other domains, which you might not be able to control. If important page content is missing due a blocked resource on another site, you might want to contact the resource owner to allow Google to access that resource. It is also possible that the rendered URL references some resources that simply do not exist. | |
| Redirected | 1) At the fetching stage, when Google contacted the host server for your URL, the server told Google to visit a different URL.  OR  2) At the rendering stage, HTML meta tags and Javascript code might tell the Google to visit other sites. | | While the actual Google mechanism that crawls URLs on the web follows redirects, the Fetch as Google simulation tool instead tells you about redirects. You can then follow redirects manually in one of two ways:   * If a page on your site redirects to another URL on your site, the tool displays a button that allows to follow the redirect. * If the URL redirects to a page on another site, you must manually perform another fetch by entering that site into your sites on webmaster tools, and then entering the path into Fetch as Google to follow that redirect.   You can inspect the HTTP response on the Fetch as Google details page to look see what is causing the redirect. Locate the HTTP error code to [learn more](https://support.google.com/webmasters/topic/6001951" \t "_blank).  You can also search through your meta tags and Javascript code if the redirects are happening at the rendering level. | |
| Not found | Google can contact your site, but it is telling Google that the URL you entered wasn’t found. | | This error indicates that, when you access your page using a web browser, you might see the HTTP 404 error code. | |
| Not authorized | Google can contact your site, but it tells Google that access to your URL is restricted or has been blocked from crawling. | | This error indicates that, when you access your page using a web browser, you might see the HTTP 403 error code. | |
| DNS not found | Google couldn’t retrieve the URL you entered because the domain name hasn’t been registered. | | Make sure you typed in your domain name (for example, www.example.com) properly so that Google can find your site server. | |
| Blocked | Your robots.txt file blocks Google from retrieving this URL. | | This is a common problem that you can fix by updating your [robots.txt file](https://support.google.com/webmasters/answer/6062608" \t "_blank).  If your site address is at the root of a domain (for example, www.example.com and not www.example.com/my\_site/), you can use the robots.txt Tester tool to diagnose why the URL is blocked from Google. | |
| Unreachable robots.txt | Google can’t reach your robots.txt file. | | Learn more about how to create and test [robots.txt files](https://support.google.com/webmasters/answer/6062608" \t "_blank). | |
| Unreachable | 1) Googlebot can’t reach your host server because the wait was too long.  OR  2) The server was reached but replied that it could not allow the request for the URL. | | Check to see that your server is up and running. | |
| Temporarily unreachable | 1) Fetch as Google can’t currently fetch your URL because the server took too long to reply.  OR  2) Fetch as Google cancelled your fetch because too many consecutive requests were made to the server for different URLs. | | Note the URL is not unreachable for all of Google: it is just unreachable for the Fetch as Google simulation tool. | |
| Error | An error prevented Google from completing the fetch. | | If this happens again, we ask that you contact Webmaster Tools product support. | |

**What's next?** Once you are able to successfully fetch and render your URL, you have the option to let Google know if you made changes! Learn more in the next article of this guide.