

Overview of Google crawlers (user agents)

See which robots Google uses to crawl the web

"Crawler" is a generic term for any program (such as a robot or spider) that is used to automatically discover and scan websites by following links from one webpage to another. Google's main crawler is called [Googlebot](#). This table lists information about the common Google crawlers you may see in your referrer logs, and how they should be specified in [robots.txt](#), the [robots](#) meta tags, and the X-Robots-Tag HTTP directives.

The following table shows the crawlers used by various products and services at Google:

- **User agent token** is used in the **User-agent:** line in robots.txt to match a crawler type when writing crawl rules for your site. Some crawlers have more than one token, as shown in the table; you need to match only one crawler token for a rule to apply. This list is not complete, but covers most of the crawlers you might see on your website.
- **Full user agent string** is a full description of the crawler, and appears in the request and your web logs.

These values can be spoofed. If you need to verify that the visitor is Googlebot, you should [use reverse DNS lookup](#).

Crawler	User agent token (product token)	Full user agent string
APIs-Google	APIs-Google	APIs-Google (+https://developers.google.com/webmasters/APIs-Google.html)
AdSense	Mediapartners-Google	Mediapartners-Google
AdsBot Mobile Web Android (Checks Android web page ad quality)	AdsBot-Google-Mobile	Mozilla/5.0 (Linux; Android 5.0; SM-G920A) AppleWebKit (KHTML, like Gecko) Chrome Mobile Safari (compatible; AdsBot-Google-Mobile; +http://www.google.com/mobile/adsbot.html)
AdsBot Mobile Web (Checks iPhone web page ad quality)	AdsBot-Google-Mobile	Mozilla/5.0 (iPhone; CPU iPhone OS 9_1 like Mac OS AppleWebKit/601.1.46 (KHTML, like Gecko) Version/9 Mobile/13B143 Safari/601.1 (compatible; AdsBot-Google-Mobile; +http://www.google.com/mobile/adsbot.html)
AdsBot	AdsBot-Google	AdsBot-Google (+http://www.google.com/adsbot.html)

(Checks desktop web page ad quality)		
Googlebot Image	<ul style="list-style-type: none"> • Googlebot-Image • Googlebot 	Googlebot-Image/1.0
Googlebot News	<ul style="list-style-type: none"> • Googlebot-News • Googlebot 	Googlebot-News
Googlebot Video	<ul style="list-style-type: none"> • Googlebot-Video • Googlebot 	Googlebot-Video/1.0
Googlebot (Desktop)	Googlebot	<ul style="list-style-type: none"> • Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html) • Mozilla/5.0 AppleWebKit/537.36 (KHTML, like Gecko; compatible; Googlebot/2.1; +http://www.google.com/bot.html) Chrome/W.X.Y.Z[‡] Safari/537.36 <p>or (rarely used):</p> <ul style="list-style-type: none"> • Googlebot/2.1 (+http://www.google.com/bot.html)
Googlebot (Smartphone)	Googlebot	Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/W.X.Y.Z[‡] Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)
Mobile AdSense	Mediapartners-Google	<i>(Various mobile device types)</i> (compatible; Mediapartners-Google/2.1; +http://www.google.com/bot.html)
Mobile Apps Android (Checks Android app page ad quality. Obeys AdsBot-Google robots rules.)	AdsBot-Google-Mobile-Apps	AdsBot-Google-Mobile-Apps

Feedfetcher	FeedFetcher-Google Does not respect robots.txt rules - here's why	FeedFetcher-Google; (+http://www.google.com/feedfetcher.html)
Google Read Aloud	Google-Read-Aloud Does not respect robots.txt rules - here's why	<ul style="list-style-type: none"> Current agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/41.0.2272.118 Safari/537.36 (+https://support.google.com/webmasters/answer/10046302) Former agent (<i>deprecated</i>): google-speakr
Duplex on the Web	DuplexWeb-Google May ignore the * user-agent wildcard - here's why	Mozilla/5.0 (Linux; Android 8.0; Pixel 2 Build/OPD3.170816.012; DuplexWeb-Google/1.0) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/74.0.3729.131 Mobile Safari/537.36
Google Favicon (Retrieves favicons for various services)	Google Favicon For user-initiated requests, ignores robots.txt rules	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/49.0.2623.75 Safari/537.36 Google Favicon

‡ Chrome/*W.X.Y.Z* in user agents

Wherever you see the string **Chrome/*W.X.Y.Z*** in the user agent strings in the table, *W.X.Y.Z* is actually a placeholder that represents the version of the Chrome browser used by that user agent: for example, 41.0.2272.96. This version number will increase over time to [match the latest Chromium release version used by Googlebot](#).

If you are searching your logs or filtering your server for a user agent with this pattern, you probably should use wildcards for the version number rather than specifying an exact version number.

User agents in robots.txt

Where several user-agents are recognized in the robots.txt file, Google will follow the most specific. If you want all of Google to be able to crawl your pages, you don't need a robots.txt file at all. If you want to block or allow all

of Google's crawlers from accessing some of your content, you can do this by specifying Googlebot as the user-agent. For example, if you want all your pages to appear in Google search, and if you want AdSense ads to appear on your pages, you don't need a robots.txt file. Similarly, if you want to block some pages from Google altogether, blocking the user-agent Googlebot will also block all Google's other user-agents.

But if you want more fine-grained control, you can get more specific. For example, you might want all your pages to appear in Google Search, but you don't want images in your personal directory to be crawled. In this case, use robots.txt to disallow the user-agent Googlebot-image from crawling the files in your /personal directory (while allowing Googlebot to crawl all files), like this:

```
User-agent: Googlebot  
Disallow:
```

```
User-agent: Googlebot-Image  
Disallow: /personal
```

To take another example, say that you want ads on all your pages, but you don't want those pages to appear in Google Search. Here, you'd block Googlebot, but allow Mediapartners-Google, like this:

```
User-agent: Googlebot  
Disallow: /
```

```
User-agent: Mediapartners-Google  
Disallow:
```

User agents in robots meta tags

Some pages use multiple robots **meta** tags to specify directives for different crawlers, like this:

```
<meta name="robots" content="nofollow"><meta name="googlebot" content="noindex">
```

In this case, Google will use the sum of the negative directives, and Googlebot will follow both the **noindex** and **nofollow** directives. [More detailed information about controlling how Google crawls and indexes your site.](#)

Was this helpful?

Yes

No

中文 (简体)

