

# How To Rank In Google News

Style and Content Guide



**Date:** December 11, 2009

**Prepared By:**

Dan Nedelko

# **Contents**

**Why Google News** **Page 03**

**Technical Concerns** **Page 04**

<b>Article URLs</b>	<b>04</b>
<b>Different domains</b>	<b>04</b>
<b>Drop-down menus</b>	<b>04</b>
<b>Embedded JavaScript text</b>	<b>05</b>
<b>Flash, graphic, JavaScript links</b>	<b>05</b>
<b>Frames</b>	<b>05</b>
<b>Languages and encoding</b>	<b>05</b>
<b>Multimedia content</b>	<b>05</b>
<b>Non-permanent section pages</b>	<b>05</b>
<b>PDF and non-HTML formats</b>	<b>06</b>
<b>Subscription sites</b>	<b>07</b>
<b>URL errors</b>	<b>08</b>
<b>Article text Requirements</b>	<b>08</b>

**News Sitemap Formatting** **Page 14**

<b>News Sitemaps overview</b>	<b>09</b>
<b>Creating a News Sitemap</b>	<b>09</b>
<b>Submitting a News Sitemap</b>	<b>12</b>
<b>News-specific errors</b>	<b>13</b>
<b>News Sitemap status</b>	<b>16</b>
<b>Sitemap updates frequency</b>	<b>17</b>
<b>Content types</b>	<b>17</b>

**How to Rank in Google News** **Page 20**

# Why Google News

Google news is one of the premier news portals on the web, and is driven by aggregated news sources. Google News sees a very impressive amount of traffic and is a great opportunity for websites with an aggressive content creation strategy to put their best content in front of the eyes of millions of viewers. Successful inclusion and ranking in Google News would clearly be a boon to traffic and PR.

explicitly by the Google News team, which is in stark contrast to their policies regarding the organic index. In the following pages we'll outline the technical The methods for inclusion and ranking in the Google News stream have been outlined fairly requirements for inclusion in Google News, the factors that cause articles to rank well and display prominently, and outline a content strategy that can be used to build on this framework.

Submitting a site to Google News for consideration to be added to their aggregator is relatively easy, simply visit

[http://www.google.com/support/news\\_pub/bin/request.py?contact\\_type=suggest\\_content](http://www.google.com/support/news_pub/bin/request.py?contact_type=suggest_content) and fill out the form.

However certain technical and formatting requirements must first be met. These are outlined in detail in this document, taken from various Google documents, videos, and blog posts.

# Technical Concerns

## Technical Requirements: Article URLs

In order to be included in Google News, your articles URLs should meet the following guidelines:

**Be unique.** Each of your pages that display an article's full text needs to have a unique URL. Google can't include sites in Google News that display multiple articles under one URL, or that do not have links to pages dedicated solely to each article.

**Be permanent.** For example, Google wouldn't be able to crawl the page [www.yoursite.com/news1.html](http://www.yoursite.com/news1.html) if it displayed a different story every day. In order to ensure that Google News' links to articles function properly, each article on a news site needs to be associated with one unique URL, and that URL must be permanent (i.e., it can't be recycled).

**Display a three-digit number.** The URL for each article must contain a unique number consisting of at least three digits. For example, Google can't crawl an article with this URL: <http://www.google.com/news/article23.html>. Google can, however, crawl an article with this URL: <http://www.google.com/news/article234.html>. Keep in mind that if the only number in the article consists of an isolated four-digit number that resembles a year, such as <http://www.google.com/news/article2006.html>, Google won't be able to crawl it. Please note, this rule is waived with News Sitemaps.

## Technical Requirements: Different domains

Google won't be able to crawl your articles if they are hosted on a domain that's separate from your main site. For instance, if your news page is located at [www.domain.com](http://www.domain.com), your articles must also be hosted on this domain. Any articles located at [www.differentdomain.com](http://www.differentdomain.com) would not be included in Google News.

## Technical Requirements: Drop-down menus

If your articles are located in a drop down box, Google won't be able to crawl them. Google News is unable to crawl articles only accessible through a drop down menu.

## Technical Requirements: Embedded JavaScript text

Google News doesn't accept articles embedded in JavaScript because they sometimes display different content for users (who see the JavaScript-based text) than for search engines (which see the no script-based text). If your articles are embedded in JavaScript, you need to provide the textual contents of JavaScript in a no script tag, in order to be included in Google News.

## Technical Requirements: Flash, graphic, JavaScript links

Google News does not recognize or follow Flash, graphic/image or JavaScript links which link to articles. The Google News crawler is best able to crawl plain text HTML links.

## Technical Requirements: Frames

If your site uses frames, Google may not be able to crawl your content properly. Frames can cause problems for search engines because they usually display several URLs (one for each frame) within a single page.

## Technical Requirements: Languages and encoding

Google News is unable to include sites with articles that display more than one language within the text of the article. The Google News system is best able to crawl articles written in one language.

Additionally, please note that sites encoded in UTF-8 are optimal for Google News.

## Technical Requirements: Multimedia content

Our news crawler gathers content by extracting text articles. Google currently don't accept audio files or podcasts, but are accepting video via [YouTube](#). If your site displays multimedia content, and Google doesn't crawl your articles, it maybe due to the fact that your articles do not contain enough words.

## Technical Requirements: Non-permanent section pages

If the URLs of your main news sections change on a daily, weekly, monthly or quarterly basis, Google may not be able to include your site in Google News. Non-permanent URLs prevent Google News from crawling your new content, as their crawlers would be unable to detect the most current URL to be crawled. The Google News automated crawler is best able to crawl sites when the URLs of their main news sections don't change.

## Technical Requirements: PDF and non-HTML formats

Google News doesn't crawl articles in PDF format, although this content is included on Google Web Search. The Google News automated crawler is currently best able to crawl plain text HTML sites.

## Technical Requirements: Subscription sites

In order to add news articles to Google News, the crawler must be able to access the content on your site. Since crawlers can't currently fill out registration forms, nor do they support cookies, Google need to be able to circumvent those pages in order to successfully crawl those sites.

The easiest way to do this is to configure your webserver to not serve the registration when crawlers visit your pages (when the User-Agent is "Googlebot-News"). It is equally important that your robots.txt file allows access by Googlebot-News.

Google is also committed to providing an excellent user experience. Generally, encountering a registration page after following a link on Google News constitutes a poor user experience for users (who may also become your users). There are three possible solutions to this problem:

1. **First click free:** We've worked with other subscription-based news services to arrange that the very first article view by a Google News user (identifiable by referrer) doesn't require subscription. While the first article can be seen without subscribing, all clicks on the article page are "trapped." This means that if users click anywhere else on that page, they'll be prompted to sign up. This allows users to view the article of interest while also exposing them to your site, encouraging an actual subscription.

Publishers have the option to limit the number of free visits to their articles to 5 articles (or more) per day per Google News user. There is some evidence that this model yields more subscriptions and more return visits. This is the preferred solution since it benefits both you and your users. Make sure to return full articles to Google users whose

referrer is one of [`*.google.*`].

2. **Subscription designation:** If the above (preferred) solution isn't feasible, Google's policy is to add a "(subscription)" tag to the publication name of all sources that greet users with a subscription or registration form. This signals to users that they may be required to register or subscribe on your site in order to access the article in question. This setting will only apply to Google News results.
3. **Subscription designation, snippets only:** If you do not want the full content of your articles to be indexed by Google News, you can display a snippet of your article that is at least 80 words long. This content, which is the only text that would appear in Google News, could be an abstract of the full article. As with the above solution, we'll add a "(subscription)" tag to your publication name when your articles appear in search results. If you prefer this option, there's no need to configure your servers as indicated at the beginning of this message.

To summarize, Google News first must be able to crawl your site. Secondly, Google News needs to know how you'd like to handle registration for visitors.

## Technical Requirements: URL errors

Certain URLs may cause technical issues. For instance, articles may be rejected if the automated crawler doesn't accept their URLs. The list of rejected URLs includes but isn't limited to the following:

URLs can't start with: `https` or `javascript`.

URLs can't end with: `.gif`, `.hqx`, `.jpeg`, `.jpg` or `.zip`.

If you're able to change your serving protocol or top level domain and follow the other guidelines, Google News should be able to crawl your site.

## Article text Requirements

Here are some of the most common **extraction issues** related to the article text, and how you can solve them. If the article content:

1. If the article content appears to be **too long to be a news article**, the Google News crawler may not recognize it as an article. This may happen with news articles that contain user-

contributed comments below the article, or HTML layouts that contain other material besides the news article itself.

2. If the article content **doesn't have punctuated sequences of contiguous words**, Google News won't be able to include it. Make sure that the text of your articles is made up of sentences, and that you don't use frequent tags within your paragraphs.

3. If the article content appears to **consist only of isolated sentences** not grouped into paragraphs, Google News won't be able to crawl it. Try formatting your articles into text paragraphs of a few sentences each.

4. If the article content constitutes a **small fraction of the text** on the page, Google News won't be able to include it in the News index. Consider removing some of the non-article text on the page.

5. If the article content appears to contain **too few words to be a news article**, Google News won't be able to include it. This applies to most links that would lead to news briefs or multimedia content, rather than full news articles.

6. If the article content **appears to be empty**, Google News won't be able to crawl it. Make sure that the full text of each of your articles is available in the source code of your article pages (and not embedded in a JavaScript file, for example).

7. If the article content is **prevented from being crawled** by a robots.txt file or a robots Meta tag, Googlebot won't be able to access your article. Try removing the "nofollow" Meta tag or checking that your robots.txt file allows "User-agent: Googlebot" access to the file where your news articles are stored.



# News Sitemap Formatting

## News Sitemaps: News Sitemaps overview

A Google News Sitemap is a file that allows you to control which content you submit to Google News. By creating and submitting a Google News Sitemap, you're able to help Google News discover and crawl your site's articles:

- **Discover news articles faster:** Sitemaps allow Google News to quickly find all of the news articles on a site.
- **Crawl and index all news articles:** Sitemaps point Google News crawler directly to each news articles URL, ensuring full coverage of the content on your site.
- **Extract and display article information more accurately:** Sitemaps identify the article titles, as well as the publication date for each article.
- **Characterize article content more accurately:** Sitemaps specify the different types of content in your articles by using <access> and <genres> tags.
- **Annotate articles with metadata:** Sitemaps clearly identify each article's content based on specific description such as related keywords or stock tickers.

**Google especially recommends using a News Sitemap if your site is new, has dynamic content, or requires users to follow several links to reach your news articles.**

Please note that Google News doesn't favor sites that use News Sitemaps in ranking results. Whether or not there's a News Sitemap available, Google News will use normal crawl methods to search and index all news sites (homepage and sections).

## News Sitemaps: Creating a News Sitemap

Before you begin, ensure that your site is included in Google News. If it's not, you can [manually request inclusion](#).

A News Sitemap uses the [Sitemap protocol](#), with additional News-specific tags as defined below. Here is an example of a News Sitemap entry using News-specific tags:

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
        xmlns:n="http://www.google.com/schemas/sitemap-news/0.9">
  <url>
    <loc>http://www.example.org/business/article55.html</loc>
    <n:news>
      <n:publication>
        <n:name>The Example Times</n:name>
        <n:language>en</n:language>
      </n:publication>
      <n:access>subscription</n:access>
      <n:genres>pressrelease, blog</n:genres>
      <n:publication_date>2008-12-23</n:publication_date>
      <n:title>Companies A, B in Merger Talks</n:title>
      <n:keywords>business, merger, acquisition, A, B</n:keywords>
      <n:stock_tickers>NASDAQ:A, NASDAQ:B</n:stock_tickers>
    </n:news>
  </url>
</urlset>
```

### News-specific tag definitions

Tag	Required?	Description
<publication_name>	Yes	The <publication> tag specifies the publication in which the article appears. It has two required child tags: <name> and <language>. The <name> is the name of the news publication. It must exactly match the name as it appears on your articles in <a href="http://news.google.com">news.google.com</a> , omitting any trailing parentheses. For example, if the name appears in Google News as "The Example Times (subscription)", you should use the name, "The Example Times". The <language> is the language of your publication. It should be an <a href="#">ISO 639 Language Code</a> (either 2 or 3 letters). Exception: For Chinese, please use zh-cn for Simplified Chinese or zh-tw for Traditional Chinese.
<access>	Yes, if access is not open, else should be	Possible values include "Subscription" or "Registration", describing the accessibility of the article. If the article is accessible to Google News readers without a registration or subscription, this tag should be omitted.

	omitted	
<genres>	Yes, if genres apply, else can be omitted	A comma-separated list of properties characterizing the content of the article, such as "PressRelease" or "UserGenerated." See <a href="#">Google News content properties</a> for a list of possible values. Your content must be labeled accurately, in order to provide a consistent experience for users.
<publication_date>	Yes	Article publication date in <a href="#">W3C format</a> , using either the "complete date" (YYYY-MM-DD) or "complete date plus hours, minutes and seconds" (YYYY-MM-DDThh:mm:ss) format, with optional fraction and time zone suffixes. Please ensure that you give the original date and time at which the article was published on your site; do not give the time at which the article was added to your Sitemap.
<title>	Yes	The title of the news article. Note: The title may be truncated for space reasons when shown on Google News.
<keywords>	No	A comma-separated list of keywords describing the topic of the article. Keywords may be drawn from, but are not limited to, the list of <a href="#">existing Google News keywords</a> .
<stock_tickers>	No	A comma-separated list of up to 5 stock tickers of the companies, mutual funds, or other financial entities that are the main subject of the article. Relevant primarily for business articles. Each ticker must be prefixed by the name of its stock exchange, and must match its entry in <a href="#">Google Finance</a> . For example, "NASDAQ:AMAT" (but not "NASD:AMAT"), or "BOM:500325" (but not "BOM:RIL").

When creating your News Sitemap, please keep in mind the following:

- Your News Sitemap should contain only URLs for your articles published in the last two days.
- You're encouraged to update your News Sitemap continually with fresh articles as they're published. Google News crawls News Sitemaps as often as it crawls the rest of your site.

- A News Sitemap can contain no more than 50,000 URLs. If you want to include more, you can break these URLs into multiple Sitemaps, and use a [Sitemap index file](#) to manage them. Use the XML format provided in the [Sitemap protocol](#). Your Sitemap index file shouldn't list more than 1,000 Sitemaps. These limits help ensure that your web server isn't overloaded by serving large files to Google News.

Once you've created your Sitemap, upload it to the highest-level directory that contains your news articles. Please see [this page](#) for further instructions on submitting your Sitemap.

## News Sitemaps: Submitting a News Sitemap

You can submit your news articles via a News Sitemap created using the Sitemap protocol. Google adheres to Sitemap Protocol 0.9 as defined by [sitemaps.org](#) and Sitemaps created for Google using Sitemap Protocol 0.9 are therefore compatible with other search engines that adopt the standards of sitemaps.org.

Before you submit your Google News Sitemap, please check the following:

1. Check if your site is included in Google News. If not, you can [request inclusion](#).
2. Create a news sitemap and place it in the root location of your site.

Once you've listed your new articles in a Sitemap, follow the steps below to submit your Sitemap:

1. Sign into [Google Webmaster Tools](#).
2. Click **Add a site**, and type the URL of the site you want to add to your account.
3. Click **Continue**. The **Site verification** page will open.
4. Select the verification method you want.
  - [Meta tag](#): Google will ask you to add a meta tag with a unique value to your site's home page. This is the easiest solution if editing your home page's HTML is easier than uploading new files.
  - [HTML file](#): Google will ask you to create a file with a specific name and upload it to a specific directory on your webserver. The file can be empty—Google cares only about the file's location, not about its content.
5. Once you've applied one of the verification methods to your website, click **Verify**.
6. Once verified, go to your **Dashboard** page.
7. Under **Site configuration**, click **Sitemaps** in the left navigation bar.
8. In the Sitemaps section, click **Submit a Sitemap**.
9. In the text box, complete the path to your Sitemap.
10. Click **Submit Sitemap**.

Once you've submitted your News sitemap the first time through your Webmaster Tools account, you can re-submit it through one of the following options:

### Re-submitting Sitemaps using Google Webmaster Tools

Before you begin, make sure you have the following sites added and verified in your Webmaster Tools account:

- the site on which the Sitemap is located
- the site(s) whose article URLs are referenced in the Sitemap

Once you have verified these details, please follow the steps below:

1. Upload your Sitemap to your site.
2. On the Webmaster Tools home page, click the site you want.
3. Under **Site configuration**, click **Sitemaps**.
4. In the text box, complete the path to your Sitemap (for example, if your Sitemap is at <http://www.example.com/sitemap.xml>, type **sitemap.xml**).
5. Click **Submit Sitemap**.

### Re-submitting Sitemaps using your robots.txt file

You can tell Google News about your Sitemap by adding the following line to your [robots.txt file](#) (updating the sample URL with the complete path to your own Sitemap):

Sitemap: [http://example.com/sitemap\\_location.xml](http://example.com/sitemap_location.xml)

This directive is independent of the user-agent line, so it does not matter where you place it in your file. If you have a Sitemap index file, you can include the location of just that file. You do not need to list each individual Sitemap listed in the index file.

## News Sitemaps: News-specific errors

In order to view error reports specific to Google News, news publishers need to include their site in Google News, have created a [Webmaster Tools](#) account and added their site to it.

- On the **Home** page, click the site's URL.
- On the **Dashboard**, click **Diagnostics > Crawl** errors.
- Click the **News** tab.
- Click the **News-specific** link.

News-specific errors include:

Article disproportionately short	The article body that Google extracted from the HTML page appears to contain too few words to be a news article. Google generated this error to avoid including what might be an incorrect piece of text.
Article fragmented	The article body that Google extracted from the HTML page appears to consist of isolated sentences not grouped together into paragraphs. Google generated this error to avoid including what might be an incorrect piece of text.
Article too long	The article body that Google extracted from the HTML page appears to be too long to be a news article. Google generated this error to avoid including what might be an incorrect piece of text. Common causes include news articles that contain user-contributed comments below the article, or HTML layouts that contain other material besides the news article itself.
Article too short	The article body that Google extracted from the HTML page appears to contain too few words to be a news article. Google generated this error to avoid including what might be an incorrect piece of text.
Date not found	Google were unable to determine the publication date of the article. Adding a <publication_date> tag to the Sitemap entry for this article will fix the problem. Please include a timestamp in the date if possible, as this will help us detect fresh content on your site.
Date too old	The date that Google determined for this article, either from a <publication_date> tag in the Sitemap, or from a date in the page HTML itself, is too old. Currently Google is only collecting articles that are 2 days old or less from Sitemaps.
Empty article	The article body that Google extracted from the HTML page appears to be empty.
Extraction failed	Google was unable to extract the article from the page. Extractions fail when Google is unable to identify a valid title, body, and timestamp for the article. Google lists URLs with this error to

	<p>provide you with information regarding why some articles may not appear in Google News. You can help us extract your articles by making sure that your title, body, and timestamp are easily crawlable (are available as text and not as images, for instance), but at this time, this error is primarily for informational purposes. Google is actively working to improve extraction methods so that you'll see this error less often.</p>
Invalid date meta tag	<p>The HTML page contains a date meta tag that Google was unable to parse. Date meta tags should be of the form: <code>&lt;meta name="DC.date.issued" content="YYYY-MM-DD"&gt;</code>, where the date is in <a href="#">W3C</a> format, using either the "complete date" (YYYY-MM-DD) or "complete date plus hours, minutes and seconds" (YYYY-MM-DDThh:mm:ss) format, with optional fraction and time zone suffixes. The date should specify when the article was first published.</p>
No links found	<p>Googlebot didn't find any links to valid news articles on the page. This error applies only to news section pages. Check this page and make sure it includes crawlable links to news articles. Googlebot is best able to crawl HTML links and is unable to crawl image links or linked embedded in JavaScript.</p>
No sentences found	<p>The article body that Google extracted from the HTML page appears not to contain punctuated sequences of contiguous words. Google generated this error to avoid including what might be an incorrect section of text.</p>
Noindex tag found	<p>The HTML page of the article contains a meta noindex tag, prohibiting Google from indexing the page. Try removing the meta noindex tag.</p>
Off-site redirect	<p>The section or article page redirects to a URL on a different domain. All section pages and articles must be located within the domain of the site included in Google News.</p>
Page too large	<p>The hub or article page length exceeds the maximum allowed. Pages can be 256KB long.</p>

Title not allowed	The title that Google extracted from the HTML page suggests that it is not a news article. Often this problem can be fixed by setting the TITLE tag on the HTML page to the title of the article, and repeating the title in a prominent place on the HTML page, such as in an H1 tag.
Title not found	Google was unable to extract a title for the article from the HTML page. Try setting the <title> tag on the HTML page to the title of the article, and repeating the title in a prominent place on the HTML page, such as in an <h1> tag. Also ensure that the title is not too long or too short. Currently, the title is required to be between 2 and 22 words, inclusive.
Uncompression failed	Googlebot detected that the page was compressed, but was unable to uncompress it.
Unsupported content type	The page had an HTTP content-type that is not supported by Google News. Articles must have a content-type of text/h

## News Sitemaps: News Sitemap status

If your site is included in Google News, and you have verified ownership of the site using Google Webmaster Tools, you can view the status of all News Sitemaps added for each site. Your sitemap's details include:

- Sitemap **Filename**
- Sitemap **Status**
- **Format** (News, Sitemap Index, etc.)
- Last **Downloaded** date. Although it's not reflected in the Last Downloaded Date, Google download News Sitemaps continually.

To view details about a particular News Sitemap or site, including Sitemap errors and warnings:

1. Sign into [Google Webmaster Tools](#) with your Google Account.
2. On your Home page, click the URL for the site you want.
3. On the Dashboard page, in the Sitemaps section, click the More » link.
4. On the Sitemaps page, click the URL for the Sitemap of interest.

From the Sitemaps page, you can also:



- Delete a sitemap you're no longer using.
- Submit a sitemap to Google.

If you submit your Google News Sitemap and receive an error stating that your news site is unknown, it means that your Sitemap is on a site which is not in the Google News database. Please check that your sitemap URL matches the URLs of your articles as they appear on Google News. For example, if the URLs of your articles on Google News start with:

`http://mynewssite.com/`

## News Sitemaps: Sitemap updates frequency

You should update your News Sitemap as often as you update your news site with new news articles. Google News crawls your News Sitemap and your regular site with the same frequency.

Please keep in mind that you should only list articles in your News Sitemap which have been published in the last two days. You can remove articles older than two days from the News sitemap, but they will remain in the News index for the regular 30 day period.

## News Sitemaps: Content types

You can use the following tags to provide Google more information about the content of your articles:

- The `<access>` tag specifies whether an article is available to all readers, or only to those with a free or paid membership to your site.
- The `<genres>` tag specifies one or more properties for an article, namely, whether it is a press release, a blog post, an opinion, an op-ed piece, user-generated content, or satire.

Because these properties can directly affect the experience of a user, it is essential to accurately characterize your articles in Google News.

### **Tag values:**

*Note: Values for the tags `<access>` and `<genres>` are required when applicable and restricted to the list provided below. Please make sure to include these values in English only. The Google News system will automatically display the visible designations in your site's language when displaying your articles on Google News.*

- The `<access>` tag takes one of the following values:

- **Subscription** (visible): an article which prompts users to pay to view content.
- **Registration** (visible): an article which prompts users to sign up for an unpaid account to view content.

Omitting the <access> tag means that the full article is accessible to all users for at least thirty days.

The <genres> tag takes one or more of the following values, separated by commas:

- **PressRelease** (visible): an official press release.
- **Satire** (visible): an article which ridicules its subject for didactic purposes.
- **Blog** (visible): any article published on a blog, or in a blog format.
- **OpEd**: an opinion-based article which comes specifically from the Op-Ed section of your site.
- **Opinion**: any other opinion-based article not appearing on an Op-Ed page, i.e., reviews, interviews, etc.
- **UserGenerated**: newsworthy user-generated content which has already gone through a formal editorial review process on your site.

### Below is an example to clarify the use of these tags:

```
<xml version="1.0" encoding="UTF-8">
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
        xmlns:n="http://www.google.com/schemas/sitemap-news/0.9">
  <url>
    <loc>http://www.example.org/business/article55.html</loc>
    <n:news>
      <n:publication>
        <n:name>The Example Times</n:name>
        <n:language>en</n:language>
      </n:publication>
      <n:access>subscription</n:access>
      <n:genres>pressrelease, blog</n:genres>
      <n:publication_date>2008-12-23</n:publication_date>
      <n:title>Companies A, B in Merger Talks</n:title>
      <n:keywords>business, merger, acquisition, A, B</n:keywords>
      <n:stock_tickers>NASDAQ:A, NASDAQ:B</n:stock_tickers>
    </n:news>
  </url>
</urlset>
```

Please note that although you can specify <access> and <genres> in your Sitemap, Google News may also add such designations to certain articles as necessary.

The <keywords> tag is used to help classify the articles you submit to Google News by topic. Although there isn't a maximum number of keywords you can add per article, it's unlikely to be

useful to specify more than a dozen keywords per article. Please separate your keywords with commas, as in the example below:

```
<n:keywords>sports, college basketball</n:keywords>
```

### **Suggested keywords:**

Please note: all terms listed below are possible keywords, so you could use "Entertainment," "Books," or both of these keywords to describe a single article. You may use other keywords not listed here; keywords should be chosen based on careful keyword research. These keywords affect what section of Google News that an article would appear in, in combination with semantic content analysis. For example an article containing words like "financial" "CFO" and "stock market" would be likely filed in the financial section of Google news.

- 
- |  |  |
|--|--|
| <ul style="list-style-type: none"><li>• <b>Business</b><ul style="list-style-type: none"><li>○ Companies</li><li>○ Economy</li><li>○ Industry</li><li>○ Markets</li></ul></li><li>• <b>Education</b></li><li>• <b>Entertainment</b><ul style="list-style-type: none"><li>○ Arts</li><li>○ Books</li><li>○ Celebrities</li><li>○ Movies</li><li>○ Music</li><li>○ TV</li></ul></li><li>• <b>Headlines</b></li><li>• <b>Health</b></li><li>• <b>Humor</b></li><li>• <b>Legal</b></li><li>• <b>Lifestyle</b><ul style="list-style-type: none"><li>○ Automotive</li><li>○ Culture</li><li>○ Food and Beverage</li><li>○ Home and Garden</li><li>○ Theater</li><li>○ Travel</li></ul></li><li>• <b>Nation</b></li><li>• <b>Politics</b></li><li>• <b>Religion</b></li><li>• <b>Science</b><ul style="list-style-type: none"><li>○ Environment</li><li>○ Geography</li><li>○ Space</li></ul></li></ul> | <ul style="list-style-type: none"><li>• <b>Sports</b><ul style="list-style-type: none"><li>○ American Football</li><li>○ Athletics</li><li>○ Badminton</li><li>○ Baseball</li><li>○ Basketball</li><li>○ Boxing</li><li>○ College Baseball</li><li>○ College Basketball</li><li>○ College Football</li><li>○ College Sports</li><li>○ Cricket</li><li>○ Cycling</li><li>○ Field Hockey</li><li>○ Golf</li><li>○ Handball</li><li>○ High School Sports</li><li>○ Ice Hockey</li><li>○ Olympics</li><li>○ Racing</li><li>○ Rugby</li><li>○ Soccer</li><li>○ Table Tennis</li><li>○ Tennis</li></ul></li><li>• <b>Technology</b><ul style="list-style-type: none"><li>○ Computing</li><li>○ Internet</li><li>○ Personal Technology</li><li>○ Video Games</li></ul></li><li>• <b>Weird News</b></li><li>• <b>World</b></li></ul> |
|--|--|
-

# How to Rank in Google News

Once content is being crawled and indexed into the Google news feed, the obvious next step is to make sure that it ranks well. Fortunately Google, in a series of papers and videos, has been very forthcoming with the factors which cause articles to rank highly in Google News.

## **Click-Through Rate / Authority**

As in the organic index, "authority" of a source is an important determinant of ranking. Authority in Google News is determined largely by click-through rates. Sites with larger click-through rates are determined to be more authoritative sources, as judged by user behavior. This click-through authority is aggregated over time and assigned not to a site as a whole, but based on topic. For example a site with high click-through rates on it's sports-related stories would not necessarily rank well for finance-related stories.

This added importance of the click-through rate shows how important attractive titling and story leads are.

## **Editorial Interest (going viral)**

Google News also pays close attention to what kind of editorial interest a story gets. For example, a story originally posted on one site, but then is referenced or re-posted by several other sites, is more likely to appear prominently in Google News than a story which was not 'picked up' by other news sites.

In the case of a similar story being reposted by several sites, these may be grouped together into a cluster of links. In order to be the lead article in this cluster, articles are given preference based on article novelty as determined by Google's novelty detection algorithm, and informativeness (op-eds won't lead an existing cluster).

## **Recency**

As having up-to-date information is a critical component of news in general, recently written articles appear more prominently in Google News searches. This dictates that news sites wishing to rank well and draw traffic from Google News, should contribute articles often.

## **Local and personal relevancy**

Locale is also a factor in Google News rankings, in two ways. Firstly, users from Canada will likely be shown news articles from Canadian sources, and about topics relevant to Canadians. Secondly, if a story is about an event in New York, a locally-written article by an institution like

the New York Times would be considered more relevant than one written by the Chicago Tribune, and would therefore rank higher.

## Keywords

Keyword content is also a major factor in rankings. Keyword content in articles determines ranking for specific keyphrases, as well as categorization. Article titles and content should be optimized carefully around popular search terms. Page title tags and article titles should both be optimized carefully.

## PageRank

PageRank, while very important for ranking in the organic index, is less so for Google News. This stands to reason as the newest, most relevant articles that would make sense to rank well would not have the time to accumulate links. The fact that PageRank should not be a concern is a boon to younger websites, as they can take advantage of this to push public awareness of, and traffic to, their sites.

## Image Inclusion

Images can be included in the Google News stream alongside articles, and provide added prominence and a greatly increased click-through rate.

- Images should be relatively high resolution with a relatively squared aspect ratio. Images in news stories should be at least 350px by 350px.
- Images should have descriptive, keyword containing captions **and** alt text.
- Images should be featured near the title of the story, ideally immediately below it.
- Images should be inline (not clickable).
- All images for news stories should be in JPG format. Formats like PNG or GIF are less likely to be included in the Google News results.

## Video Inclusion

Videos can also be included in the Google News Stream, and provided an added level of prominence to articles. Articles which contain a video segment should host that video on YouTube, provide a relevant caption, and ideally link to a transcript.

## Other Best Practices for Google News Articles

- Articles should contain all of their text on one page, it is not advisable to break up the article body with user comments, or page breaks. These breaks make it more difficult for Google to determine where an article starts and stops, and makes the article less likely to be included in Google News.
- As stated, recency is very important in news stories. As such, articles submitted to Google News should be dated, and have the date located between the article title, and the article body.
- Non-news content such as press releases, should not be posted in the section as actual news articles. It is suggested that a directory system be used to separate these topics, for example:  
[www.example.com/news/](http://www.example.com/news/) and [www.example.com/pr/](http://www.example.com/pr/)
- As always, it is suggested that articles be written containing unique and informational content which will be interesting and useful to readers.
- It's worth noting that the Google News crawlers will re-crawl an article within 12hours to see if information has been updated. It is worth updating articles about evolving situations within that time period in an attempt to keep the information relevant and useful, but also to 'bump' the article up in recency and (hopefully) ranking.