

# HOW TO USE CANONICAL TAGS

Posted on August 23, 2020 by Fabian Pott



Categories: [Level: Beginner](#), [On-Page SEO](#), [SEO Tips](#)



Since February 2009 Google supports a [particular way](#) to specify your preferred version of your URL and its content. The goal behind it was to solve the eternal problem of duplicate content once and for all.

Once you know how to use Canonical Tags correctly, you can benefit from it, but mistakes should be avoided at all costs. So what exactly do you need to pay attention to here?

I have put this together for you and I'm sure that you will be confident enough with Canonical Tags after reading this article and you will know how to use them properly.

## What is a canonical tag?

So what are canonical tags and how do you use them to avoid duplicate content?

After a page has grown continuously (generally also when errors occur during content creation), it often happens that very similar or even the same content exists on different URLs. However, Google naturally wants to know which of these contents is the "original" one. That means: Which of these is the content to be considered and indexed in the end.

## How does a canonical tag look like?

Canonical tags are basically defined in this form in the HTML code of a website: `rel="canonical"`. This canonical tag is set in the `<head>` section of a website.

The code looks like this: `<link rel="canonical" href="https://your website.com/your-page/" />`

It therefore consists of 2 central components:

- `link rel="canonical"`: This link in this tag is the main version on this page.
- `href="https://your website.com/your-page/"`: The canonical version of it is on this URL.

## Why are canonical tags so important for SEO?

Every SEO should know that Google hates duplicate content. But why? The answer is quite simple: Google doesn't know what is most important and what should be chosen.

Since only one version of a page is getting indexed, it must be explained which one of them should be. Also: which URL is relevant for which search term?

Furthermore, the so-called "crawl budget" of a website is very important here. And if you have too much duplicate content on a website, this can lead to an enormous negative effect.

Finally, it's best not to waste the Google Bot's time and make sure that multiple versions of a website

are not crawled if they are the same or very similar to each other.

Once the crawl budget is saved, the Google Bot will have enough resources to discover all the other content on the page and is ultimately able to index it as well.

And this is exactly where Canonical Tags come into action. These tags can be used to tell Google very precisely which page should be indexed and possibly ranked.

**But be aware:**

*"If you don't explicitly tell Google which URL is canonical, Google will make the choice for you, or might consider them both of equal weight, which might lead to unwanted behavior".*

(Source [click here](#))

So if you don't care enough, Google will make that decision for you - which does not always mean that the right site will be chosen.

Below I will explain why Canonical Tags are only a recommendation in front of Google and can be ignored in rare cases. How to avoid this is explained further below.

## **Other reasons for using canonical tags**

A/B testing is used to see which elements work best on a page. E.g. small details such as the color of different buttons, page layout, or content. Canonical tags are designed to keep things in order and point to the appropriate, original content.

Canonical URLs are also frequently used on e-commerce websites. For example, if the category for shoes is changed to divide the category into men's and women's shoes. Therefore, the URLs for these product pages change, even though the pages themselves remain the same. As a result, there are two URLs with the same content.

## **Content that doesn't appear to be duplicate content**

Search engines crawl URLs, not web pages. That means that they see example.com/product and example.com/product?color=red as unique pages, even though they're the same web page with identical or similar content. These are called parameterized URLs, and they're a common cause of duplicate content, especially on e-commerce sites with faceted/filtered navigation.

Other reasons are SEO meta descriptions and titles that are the same for multiple pages or maybe there are technical issues such as pagination problems or multiple versions of printable and text-only pages.

Below are some other frequent causes of duplicate content that affect all types of websites:

- **Having parameterized URLs for search parameters** (e.g., *example.com?q=search-term*)
- **Having parameterized URLs for session IDs** (e.g., *https://example.com?sessionid=3*)
- **Having separate printable versions of pages** (e.g., *example.com/page* and *example.com/print/page*)
- **Having unique URLs for posts under different categories** (e.g., *example.com/services/SEO/* and *example.com/specials/SEO/*)
- **Having pages for different device types** (e.g., *example.com* and *m.example.com*)
- **Having AMP and non-AMP versions of a page** (e.g., *example.com/page* and *amp.example.com/page*)
- **Serving the same content at non-www and www variants** (e.g., *http://example.com* and *http://www.example.com*)
- **Serving the same content at non-https and https variants** (e.g., *http://www.example.com* and *https://www.example.com*)
- **Serving the same content with and without trailing slashes** (e.g., *https://example.com/page/* and *http://www.example.com/page*)
- **Serving the same content at default versions of the page such as index pages** (e.g., *https://www.example.com/*, *https://www.example.com/index.htm*, *https://www.example.com/index.html*, *https://www.example.com/index.php*, *https://www.example.com/default.htm*, etc.)
- **Serving the same content with and without capital letters** (e.g., *https://example.com/page/* and *http://www.example.com/Page/*)

Source click [here](#).

In all these cases it is essential to set a Canonical tag!

## How to use canonical tags in the best way?

It is quite easy to set canonical tags. However, there are some points that need to be considered:

### Only use one canonical tag per page

If you use more than one canonical tag, then Google won't use any of them and rather will ignore those. Learn more about this [here](#).

### Check your domain version and make sure the URL is correct

The classical case is, as soon as you switch to SSL, that then the corresponding "s" after HTTP often gets lost. This error generally confuses the bot, because it doesn't know which URL to use exactly.

If you use SSL - always put in the "s".

## Use so-called self-referential canonical tags

A self-referential canonical tag (as is said: self-referential) points to itself.

E.g. your URL is: `https://yourwebsite.com/your-page` then use `<link rel="canonical" href="https://yourwebsite.com/your-page" />`.

### The reason why it's good to use them:

Google's John Mueller recently stated that self-referencing canonical tags are not absolutely necessary, but they do help.

In Mueller's words: *"It's a great practice to have a self-referencing canonical but it's not critical."*

This topic came up during a recent Google Webmaster Central hangout when a site owner asked about the importance of using self-referencing canonicals.

Canonicals are typically used to link a non-canonical page to the canonical version, but they can also be used to link a page to itself.

Self-referencing canonicals are beneficial because URLs may get linked to parameters and UTM tags.

Click [here](#) to read the source.

## Do not use upper case letters in the URL

Google treats letters differently. So always make sure that only lowercase letters are used in your canonical tags. How this can be done on the server-side, you can [find out here](#).

## Always use the complete URL

Relative: `<link rel="canonical" href="https://yourwebsite.com/your-page/" />`

Absolute - complete: `<link rel="canonical" href="https://example.com/sample-page/" />`

## How should you implement canonical tags?

There are several reasons for specifying canonical URLs [with advantages and disadvantages](#).

- 301 redirect
- HTTP header
- Internal links
- HTML tag (rel=canonical)
- Sitemap

[Find out more](#) here.

## Canonicals with rel="canonical" HTML tags

This canonical tag is the easiest way to specify a canonical URL. All you have to do is to add the code (below) into the <head> section of the duplicate page.

```
<link rel="canonical" href="https://yourwebsite.com/canonical-page/" />
```

If you use WordPress then just get an SEO plugin where you can set this **canonical URL**. Otherwise, you have to hardcode it. Yoast, rankMath, or SEOultimate Pro will give you the opportunity to change those settings accordingly.

Canonical URL:

The Canonical URL that this page should point to. Leave field empty to default to existing permalink.

## Custom HTTP header customized from images or PDF to HTML page

When the rel="canonical" tag was introduced in 2009 it was quickly adopted by SEOs. Unfortunately, because the canonical tag resides in the HTML head you cannot insert it into non-HTML pages.

Why? If you have images or PDF documents that play an important role in your website, they can outrank HTML pages on your site. If you created a redirect no one could read the document or see the image.

The solution is to create a rel="canonical" for the image or document. Since you cannot place the canonical tag in the HTML head on non-HTML documents, search engines provided the option to provide it as an HTTP header

See the source [here](#) or learn more about this [here](#).

## Canonical tags vs. 301 redirects

You can use 301 redirects if you want to redirect traffic coming from a duplicate URL and redirect it to the canonical version.

Assume that your page can be reached at these URLs:

- example.com
- example.com/index.php
- example.com/home/

Select one URL as the canonical URL and then redirect the other URLs to it.

You may want to do the same for secure HTTPS/HTTP and www/non-www versions of your

website. Select a canonical version and redirect the other URLs there.

E.g. the canonical version of seogeeklab.com is the HTTPS-non-www URL (<https://seogeeklab.com>). Any other URLs that follow will be redirected to those URLs:

- <http://seogeeklab.com/>
- <http://www.seogeeklab.com/>
- <https://www.seogeeklab.com/>

## What to do when using internal links

The way in which one links from one page to the next on the website is an important signal of canonicalization.

The more thoroughly you handle all these signals, the easier it will be for search engines to identify your preferred canonical URL. John Mueller created a [video](#) for exactly this purpose.

## Mistakes when implementing canonical tags

Setting canonical tags isn't always easy. Try to avoid those mistakes here when you canonicalize your URLs.

### Don't block the canonicalized URL via robots.txt

Everything you block via robots.txt will keep Google from crawling this content. The bot is therefore no longer able to recognize the Canonical tag. So the sense and power of a Canonical Tag will not work out.

### Don't set the canonicalized URL to noindex

Never combine rel= canonical with noindex. Both are two completely conflicting directives.

Normally Google will follow the canonical tag and ignore the noindex tag. This was also [stated](#) accordingly by John Mueller. You should avoid this process in all cases. As soon as a URL is set to noindex and you want to canonicalize it at the same time, it is very advisable to use a 301 redirect. Otherwise, simply use rel=canonical.

### Don't use more than one rel=canonical tag

As soon as more than one rel=canonical tag is set, they are completely ignored by Google. Often this happens through the CMS itself, a theme, or a plugin.

Also via Javascript, you have to take care that HTML and JavaScript are not mixed up. How to insert canonical tags via Javascript, you can [read here](#).

## **Don't put the canonical tag in the <body>**

The rel=canonical tag must ALWAYS be included in the <head> section of a document or set there. However, if it ends up in the <body>, it will be completely ignored.

## **Don't canonicalize all paginated page to the main page**

Use only self-referencing canonicals that are used on all paginated pages. So if the page no. 2 is not the same as page no. 1, using a canonical tag in this way would be wrong. More about the statement from John Mueller you can [read here](#).

## **Always use canonical tags with hreflang**

Hreflang tags are commonly used to define both the language and the geographical focus of a website.

According to Google, when using Hreflang, "a canonical page in the same language, or the best possible alternative language as long as no canonical page exists for the same language".

## **4XX Set HTTP status codes for the canonical URL**

Using a 4XX HTTP status code for a canonical URL will have the same effect as using a noindex-tag. Google will not be able to recognize the canonical tag and cannot transfer the "link equity" to the canonical version.

## **Summary**

Setting Canonical tags is not as complicated as it may have felt in some parts. You just have to get used to it at first - once the learning curve is a bit more advanced, the easier the process is.

But remember: Canonical tags are still leading the way, but can be ignored by Google or Google decides itself which content is taken as the original.

To find out where any errors are or if you just feel unsure, you can use the [URL Inspection Tool](#) in Google Search Console to find out what Google and you have set as canonical.



**There are no comments yet.**