

# Duplicate Content

## Background

Web pages with unique, rich information and keywords about the subject will help you rank in search engines and earn more site traffic. Search engines access web page content, like text, images and videos, through a web address or a URL. If multiple web pages, or URLs, on a site have the same or very similar content, then it's considered duplicate content.

Internal duplicate content is caused by web pages within a single site. Google distinguishes between two types of internal duplicate content: malicious and non-malicious.

1. **Malicious Internal Duplicates:** When website deliberately creates multiple versions of pages with largely similar content to manipulate search engines. Google may adjust indexing and ranking of a site, making it less likely to rank in search.
2. **Non-malicious Internal Duplicates:** Sometimes multiple versions of the same page on a website are indexed accidentally. This often happens with forums and product pages, where parameters like `rel="prev"` and `rel="next"`, as well as product options can result in duplicate content.

## Issues

Identify, solving and preventing duplicate content is possible. Below is a list of duplicate content issues found while auditing the website.

1. Missing Canonical Element

### Missing Canonical Element

You can indicate to search engines what version of your webpage is the most authoritative version, or which one to rank in search by using a canonical element.

```
<link rel="canonical"
href="https://www.samplesite.com/most-authoritative-page" />
```

If you don't declare a canonical element on the page, then search engines will decide which one is the most authoritative.

**Problem:** None of the pages on the website have canonical tags, which means search engines are choosing which pages to rank as the authoritative page.

**Solution:** In order to explicitly communicate to Google which pages to rank in search, add canonical tags to the <head> section of all pages on the site. For different page types, follow the logic below:

Page Type	Example(s)	Canonical Logic	Canonical Example
Category Page	<ol style="list-style-type: none"> <li>1. <a href="https://www.sample-site.com/category/shoes">https://www.sample-site.com/category/shoes</a></li> <li>2. <a href="https://www.sample-site.com/category/shirts">https://www.sample-site.com/category/shirts</a></li> </ol>	<ol style="list-style-type: none"> <li>1. Always self referential canonical</li> </ol>	<pre>&lt;link rel="canonical" href="https://www.sample-site.com/category/shoes" /&gt;</pre>
Product Page	<ol style="list-style-type: none"> <li>1. <a href="https://www.sample-site.com/product/blue-shoes">https://www.sample-site.com/product/blue-shoes</a></li> <li>2. <a href="https://www.sample-site.com/product/blue-shirts">https://www.sample-site.com/product/blue-shirts</a></li> </ol>	<ol style="list-style-type: none"> <li>1. Always self referential canonical</li> </ol>	<pre>&lt;link rel="canonical" href="https://www.sample-site.com/product/blue-shoes" /&gt;</pre>
Product Option	<ol style="list-style-type: none"> <li>1. <a href="https://www.sample-site.com/product/blue-shoes?pon=1">https://www.sample-site.com/product/blue-shoes?pon=1</a></li> <li>2. <a href="https://www.sample-site.com/product/blue-shoes?pon=2">https://www.sample-site.com/product/blue-shoes?pon=2</a></li> </ol>	<ol style="list-style-type: none"> <li>1. Always canonical to base product URL</li> </ol>	<pre>&lt;link rel="canonical" href="https://www.sample-site.com/product/blue-shoes" /&gt;</pre>

**Resource Data:** [All Pages Missing Canonical URLs](#)