SEO Basics: Is Translation an SEO Duplicate Content Issue for Google?

WYNNEWOOD, PENNSYLVANIA, UNITED STATES, July 30, 2019 /EINPresswire.com/ -- Though people have tried myth-busting the idea of a penalty associated with duplicate content, blatantly using content that has been copied and pasted from somewhere else still steers your website away from being optimized. There are exceptions to the rules associated with penalties and optimization when it comes to using duplicate content or syndication of content, such as in situations where:

- You're "duplicating" content on LinkedIn or Medium
- You designate a canonical tag for the original source/master page
- You syndicate content on other websites with a canonical and/or meta noindex tag
- Product pages

Even by knowing these exceptions, it can be hard to be sure of what's "allowed"—and what isn't.

A 2015 Raven Tools study revealed that about 29% of pages crawled showed up with duplicate content. Of course, if you're confident with the originality of your content (a tool like Copyscape or Grammarly Premium can help you run a plagiarism check), there's really not much to worry about.

However, if you have multilingual websites, you might find yourself wondering if translations of content across different websites will be flagged. Let's set the record straight right now—translation is not a duplicate content issue.

Let's Dive Deeper into the Question of Duplicate Content:
But that's the short answer. Having a better idea of duplicate content, how it impacts SEO, and its causes can help you further optimize the websites or web pages you work with.

Before we delve into the duplicate content, let's talk about websites in general. Did you know the first website was created on August 6, 1991, by Tim Berners-Lee? And the first website address was http://info.cern.ch/hypertext/WWW/TheProject.html. Can you imagine typing in this address? Most people wouldn't because it too long. Today, you will find non-www web addresses,

such as http://example.com. And you can have different versions of an URL, such as http://www.example.com, to increase traffic results.

Domains have more power than ever before. URL parameters now give a page the power to have limitless views. A canonical URL tells a search engine to treat it as authoritative. And a canonical tag (also known as use rel = canonical tag) tells search engines that a specific URL is the master copy of a web page(s) or the preferred version of a web page via a canonical link. It also prevents duplicate or copied content from appearing. And when creating a domain, you want to create a preferred domain. This tells search engines which domain (www. vs. non-www) to crawl, which produces better search results. The preferred version of your domain should also contain HTTPS. This tells visitors to your website that its secure, especially if you have a store or sensitive information. HTTPS also gets you a lock symbol on search engines, such as Chrome and Google. Plus, Google loves secure websites, which can provide a boost in rankings. Domains aren't the only piece of a website that can build strong SEO or cause duplicate content. Authoritative content is another piece.

When websites became popular, the focus was primarily on the design and a paragraph or two about the business. This is when duplicate content issues became rampant. People didn't know what to write and they were simply told to put a piece of content on separate pages to allow for better search results (search engine rankings). Thus, they just duplicated content or duplicate versions of a page (e.g. substantive blocks of content), which resulted in a duplicate content penalty from Google. It not only hurts a website's ranking, it damages the company's brand and user experience. Identical content of the original article can show up not only on a website, but on different URLs from the same company. To avoid the temptation of duplicate content, it's beneficial to have unique, high-quality, or authoritative content. It's easy to write this type of content if the author is an expert on the subject (or company). The expert is the original source of the content. When a website is published on the internet or a page is updated, a Googlebot will crawl over the page(s), indexing keywords as well as alt tags and meta tags (meta robots). Duplicate content can easily appear in these tags. When a Googlebot sees the robots.txt, file it will not crawl a page, which can prevent duplicate copy. However, some could use this to get around using duplicate unique content by blocking the page from the bot.

Today, anyone can build a website using a content management system, such as WordPress and Drupal. Content has become "king" on websites today. There is more pressure for it to be authentic without duplication. Visitors do not want to see duplicate pages; they want to see content that is customized for each page, such as the About and Blog pages. The homepage does not give everything away about the company anymore; it merely gives an overview of the site.

However, web content writers should not use the same phrases or sentences on their corresponding pages. This can lead to duplication. Also, visitors want to see updated and new content, which keeps them coming back to a site. Each user has a session id, which can be used to measure traffic. When creating page titles and descriptions, you don't want to copy and paste from the website. This is an opportunity to use new keywords as SERP content, which can help

build SEO. Another piece is an internal link, which links one webpage to another. You should also have external links, which links a page on your own site to another website. This is a great way to make your website credible.

Better SEO means more traffic, and that can lead to more conversions and revenue. A Google webmaster can measure SEO through webmaster tools, such as the Google search console. Just enter the keywords that would be used in the search query for your website, product or service. For more SEO tips, check out articles by Matt Cutts and John Mueller.

What is Duplicate Content?
Duplicate content is content that appears exactly the same, or almost exactly the same, on more than one website. Google explains that "duplicate content generally refers to substantive blocks of content within or across domains that either completely matches other content or are appreciably similar." The company adds, "Mostly, this is not deceptive in origin."

How Does Duplicate Content Affect Your SEO?
Though some SEOs argue that you're not going to be directly penalized by Google for accidentally having duplicate content, it does impact your search engine ranking—especially when it comes to ranking on Google.

If there are multiple sources of similar content across the internet, Google can struggle to identify the most relevant result for any given query. Not knowing which content offering to rank higher, the search engine might not rank any of the pages with the same content (though this is an extreme example).

Not showing up on the first page of a search can be detrimental to a business. Not showing up anywhere is a surefire way to go down in flames.

According to Google, you're only really going to be getting yourself in trouble with the search engine if you "engaged in deceptive practices." If flagged, this can result in your website being removed from the search engine results completely.

Google explains, "Once you've made your changes and are confident that your site no longer violates our guidelines, submit your site for reconsideration."

There is also a distinction between accidental duplicate content and obviously plagiarized content. If you believe someone else has stolen your content, you can request Google to remove it from its search results. On a similar note, you can also petition Google to disavow spammy backlinks through Search Console. Trailing slashes is another cause of duplicate content to keep in mind.

Managing Translated Content:
If you think about Google's goal—to provide the most relevant information to any given query—it should become immediately clear that translated content would not be considered duplicate content. Someone searching for information on coconut water in English is not going to find an answer in Spanish as similarly relevant to an answer in English.

According to Google's former head of web spam, content in different languages—although identical in context—is still quite different, so it is not considered as duplicate content. However, if the original content was simply dumped into Google Translate and then copied and pasted, it could trigger spam flags.

This kind of flagging is due to the automated nature of the content translation process on tools such as Google Translate. Without being reviewed by a human, such content can be low-quality due to many grammatical issues. By curating the Google Translate content before publishing (perhaps by hiring a freelance writer who's proficient in the languages you're translating to and from), you can easily avoid this issue and ensure a better experience for those visiting your site.

What Can Cause Duplicate Content Issues?
The majority of duplicate content cases are not intentional. In fact, it's very possible that you already have some duplication on your website.

One common issue where people end up creating duplication is if they are actively running and maintaining both http:// and https:// versions of a site with identical content.

If both sites are live, active, and visible to search engines, search engines will see the pages as duplication.

In the same way, if your site has a "regular" and "print" version of each article, they can be categorized as duplicate content. In such cases, it's best to block crawlers from one with a noindex meta tag. If not, Google will choose one of them to list.

Content created by web scraping or scrapers, most often described as an automated process of extracting data from a website, is prone to be seen as a duplicate content issue or pagination. This is often the case for e-commerce websites, as many of them sell multiple versions of the same products, where the product descriptions are scraped from somewhere else online (i.e., the original supplier) and are added to new e-commerce stores without change.

Localized domains can also be a source of duplicate content. When these geo-targeted websites—such as .co.uk for the United Kingdom or .ca for Canada—are all owned by the same company and are dialed into separate English speaking locations, it's easy to make the mistake of thinking that posting the same content won't be recognized as duplication (spoiler alert, it will).

However, if the content has been translated into a different language and curated for a geo-targeted website, you should have no issues.

Final Thoughts: Is Translation an SEO Duplicate Content Issue?
The bottom line is that publishing good translations of the same content and information elsewhere on your site will not adversely impact your SEO. However, you are setting yourself up to be flagged as spam if you are relying on automated translations and putting no additional work into curating the content you're publishing—even if it is on a geo-targeted website.

This does not mean that you don't have to worry about duplicate content. As noted, there are many different situations where you can accidentally end up publishing duplicate content that you may not have even realized are live on your website! Though this content may not cause you to be directly penalized, it moves your website away from being completely optimized.

At the very least, you need to ensure that you are not creating duplicate content out of laziness or in an attempt to deceive search engines. Being flagged for doing so can cause serious damage to your company and brand by your website being removed from search engine results until you have proved that you have remedied the situation. It's simply not worth the risk.

Thankfully, professionally translating content and publishing it in multiple languages is not going to put you in that situation.

About [Results Driven Marketing, LLC](#):
Results Driven Marketing, LLC is located at 300 E Lancaster Ave, Suite 202, Wynnewood, PA 19096 and can be reached by phone at 215-393-8700. Founded by Janeene High in 2013, High is a twelve-year veteran of the industry. She has been a leader in the Greater Philadelphia Area in the retention of clients and forecasting new opportunities for her client using highly advanced digital marketing strategies for her firm's clientele. For a complimentary review of your website, SEO or marketing strategy, book a 45-minute review session with Janeene today!

This release was drafted by Results Driven Marketing, LLC: a full-service digital marketing, public relations, advertising, and content marketing firm located at 300 E. Lancaster Ave., Wynnewood, PA 19096.

Mike Bannan
Results Driven Marketing LLC
2156003540
email us here