FOREST PATHWAYS

Your accessibility journey starts here.

Google robot.txt

Index

Description

Google seems to rule the world. Years ago, you uploaded your robot.txt document to your server. Now you do it through google.

Why

Robot.txt files play a crucial role in managing how web crawlers, like those from search engines, interact with your website. Here are some key reasons why it's beneficial to use a robots.txt file:

Control Crawler Access

- Block crawling of specific content - You can prevent crawlers from accessing and indexing certain pages or directories on your website. This might be useful for things like internal search results pages, test pages, or content under development.
- Focus crawler attention on important pages - By allowing access to specific pages or sections, you can encourage crawlers to prioritize more relevant content for indexing, potentially improving your search ranking.

Manage Crawl Rate

- Prevent overloading your server - If your website receives a lot of traffic from crawlers, you can use robots.txt to limit the number of requests they can make within a certain timeframe. This helps ensure your server doesn't get overwhelmed and slows down.
- Optimise crawler behaviour - Depending on your website's size and structure, you can guide crawlers through specific paths for efficient indexing.

Improve Search Engine Optimization (SEO)

- Prevent indexing unnecessary or duplicate content - Excluding duplicate or low-quality content from indexing can improve your website's overall search performance by focusing search engines on valuable and unique content.
- Maintain control over search results - By blocking indexing of certain pages, you can prevent irrelevant or outdated content from appearing in search results associated with your website.

How

The process is relatively simple but is an important SEO step.

Creating and submitting a Sitemap helps make sure that Google knows about all the pages on your site, including URLs that may not be discoverable by Google's normal crawling process such as blog entries.
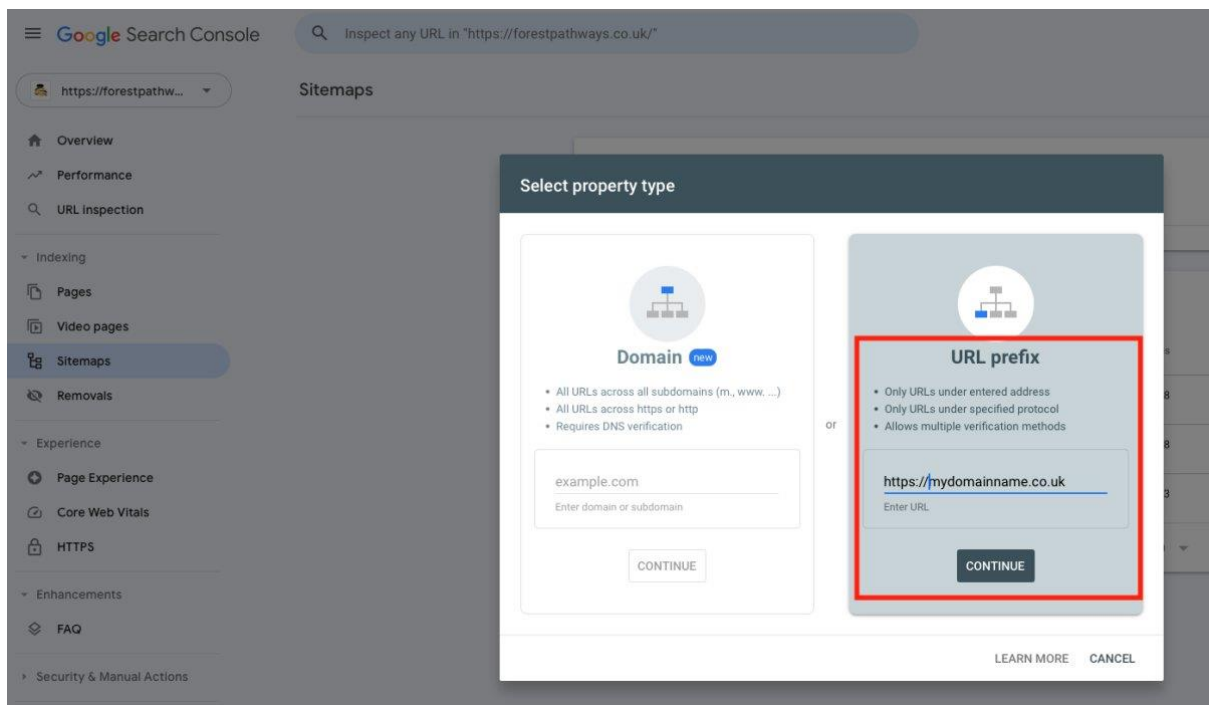
Google adheres to Sitemap Protocol 0.9 as defined by sitemaps.org. The Sitemap protocol uses an XML schema to define the elements and attributes that can appear in your Sitemap file.

Open a Google Search console page here >>>>>>

Next you need to add a domain name by clicking this button.
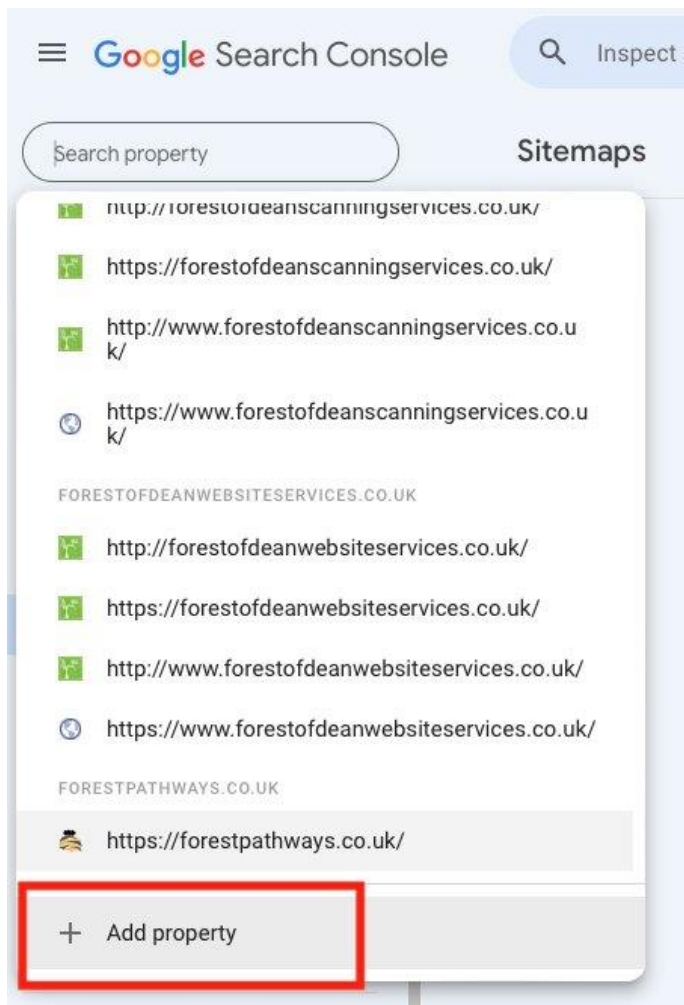
This brings you to a popup menu. Simply add the four instances of the domain name:

For this site they are:

- http://forestpathways.co.uk
- https://forestpathways.co.uk
- http://www.forestpathways.co.uk
- https://www.forestpathways.co.uk
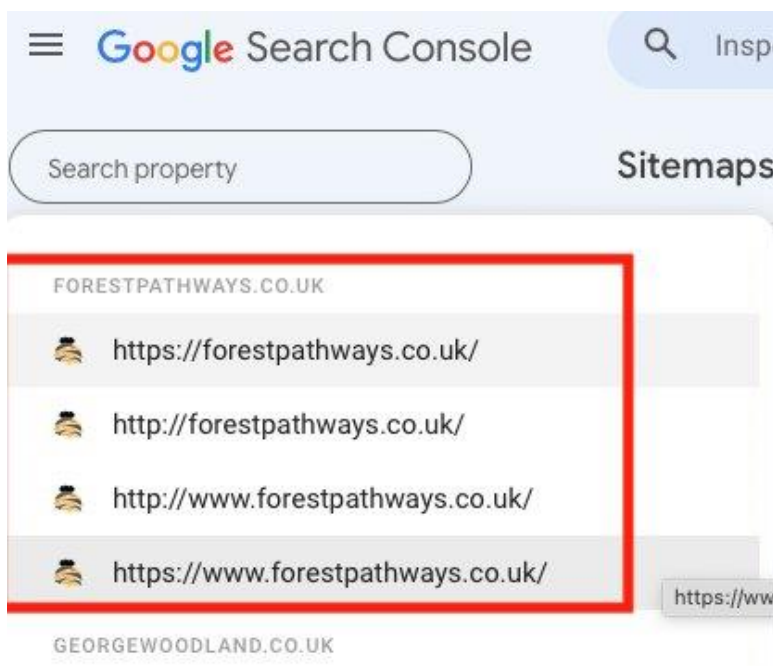
Next you need to submit the robot.txt file and there may be more than one location. On this site we have two principal XML file locations as illustrated here.

Once added Google will verify as illustrated here.

Now you are good to go. This helps Google understand your website.