

Check Under the Hood of your Website

by Katie Ayres
katie@1happyplace.com

①HappyPlace

Introduction

This checklist was created to support the video [Check Under the Hood of your Website](#), created by Katie Ayres, web developer, to help artists better understand their website.

We will be checking three areas of your website to ensure they are supporting the search engines when they come to index your website, to ensure better search results.

Robots.txt

This file lets the legitimate bots know if there are areas of your website that are best not indexed and thus not served up as search results. The main thing you want to check is that anything that is not allowed makes sense for your website.

Sitemap.xml

Sitemaps direct the search engine bots to all the pages on the website. You want to ensure that list is comprehensive.

404 page

When a visitor is on a website, if they accidentally request a page that doesn't exist, a 404 page appears. That page might be a default page, but we want to see a custom page that gently leads the visitor back to the website.

On the next page is instructions on how to check each one and how to ensure they are healthy and supporting your website.

Robots.txt

To see this file, type your domain and add /robots.txt.

Here is an example shown with 1happyplace.com:

1happyplace.com/robots.txt

```
User-agent: *  
Allow: /  
Sitemap: https://1happyplace.com/sitemap.xml  
Sitemap: https://sproutvideo.com/sitemaps/a69bdfb41e1be1c028.xml
```

1happysite.com/robots.txt

```
User-agent: *  
Allow: /wp-admin/admin-ajax.php  
Disallow: /wp-admin/  
  
Sitemap: https://1happysite.com/sitemap.xml
```

Notice there are the words **Allow** and **Disallow**. You want to make sure it makes sense for your website.

In the first example, 1happyplace.com, it is a static site, so all of it is public and I have no reason to block off a section, therefore you see with “Allow: /” – the slash means all files.

The second example, 1HappySite.com, is a WordPress website, and you see Disallow: This is listed to ask the bots to leave the admin area alone. Note, this is not for security, the nefarious bots can still crawl anything not protected. In WordPress case, there is also extensive code protecting the admin area.

You want to ensure that you do not see “Disallow: /” – which means don’t crawl any of the website. This would prevent your website from being listed in search results.

SiteMap.xml

Notice under the allow/disallow section of the **robots.txt** is a Sitemap listing. Your robots.txt may or may not have this listing, but if they do, then you can cut and paste those links to test yours. For most websites, the name of the file is **sitemap.xml**.

To see this file, type your domain (example shown with 1happyplace.com, change it for your domain) and add /sitemap.xml or the link you saw in the robots.txt file.

1happyplace.com/sitemap.xml

```
<url>
  <loc>https://1happyplace.com/about</loc>
  <lastmod>2021-03-28</lastmod>
</url>
<url>
  <loc>https://1happyplace.com/contact</loc>
  <lastmod>2021-05-01</lastmod>
</url>
<url>
  <loc>https://1happyplace.com</loc>
  <lastmod>2021-05-01</lastmod>
</url>
```

The sitemap listing is shown in a bit of a crazy format. Remember, this is software talking to software, and computers often have their own language, like in this case, XML.

You do not have to understand XML to understand the file. If you look closely, you see a listing of mainly web addresses (surrounded by <loc>), and dates (surrounded by <lastmod>).

The main thing look for is that all your pages are listed, and the dates are properly updated.

404 Page

The 404 page is found when a user clicks or types in a web address that does not exist. To test this out, you need to use a web address that is invalid. When I am testing this, I always use /junk.

1happyplace.com/junk

What do you see? Is it friendly? Is it styled like the rest of the website? Does it include links to important areas of your website?

Hopefully you see a pleasant page that helps the visitor back from the wrong corridor.