Try Hack Me

Access Machines   1

● ..ıl Easy   ⏱ 0 min

💻 Start AttackBox   ▼   Help   ▼   🔖 Save Room   👍 4802   👎

⚙ Options   ▼

Room progress (25%)

Google Dorking | DarkStar • Jan 16, 2021   Source: YouTube

TryHackMe Google Dorking Official Walkthrough

Task 1	✅	Ye Ol' Search Engine	⌄

Task 2	✅	Let's Learn About Crawlers	⌄

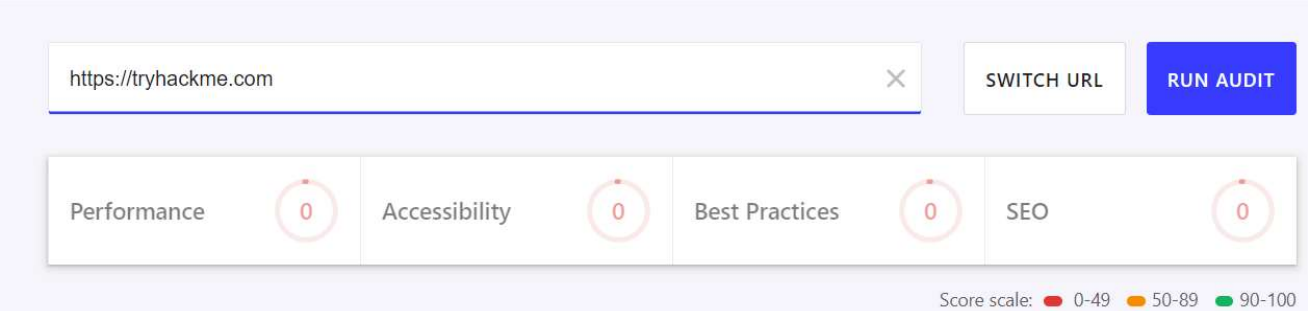| Task 3	○	Enter: Search Engine Optimisation	⌃

Search Engine Optimisation

Search Engine Optimisation or SEO is a prevalent and lucrative topic in modern-day search engines. In fact, so much so, that entire businesses capitalise on improving a domains SEO "ranking". At an abstract view, search engines will "prioritise" those domains that are easier to index. There are many factors in how "optimal" a domain is - resulting in something similar to a point-scoring system.

To highlight a few influences on how these points are scored, factors such as:

• How responsive your website is to the different browser types I.e. Google Chrome, Firefox and Internet Explorer - this includes Mobile phones!

• How easy it is to crawl your website (or if crawling is even allowed ...but we'll come to this later) through the use of "Sitemaps"

• What kind of keywords your website has (i.e. In our examples if the user was to search for a query like "Colours" no domain will be returned - as the search engine has not (yet) crawled a domain that has any keywords to do with "Colours"

There is a lot of complexity in how the various search engines individually "point-score" or rank these domains - including vast algorithms. Naturally, the companies running these search engines such as Google don't share exactly how the hierarchic view of domains ultimately ends up. Although, as these are businesses at the end of the day, you can pay to advertise/boost the order of which your domain is displayed.
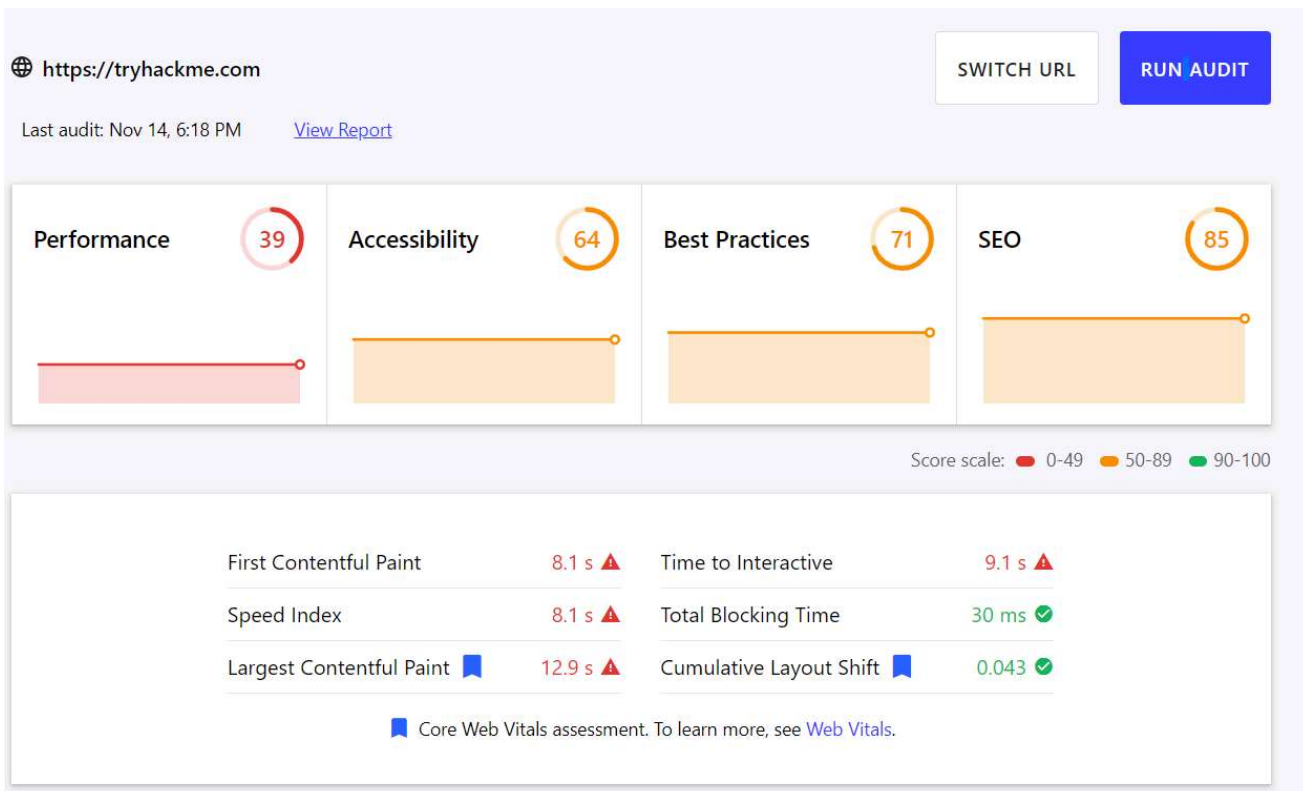
There are various online tools - sometimes provided by the search engine providers themselves that will show you just how optimised your domain is. For example, let's use Google's Site Analyser to check the rating of TryHackMe:

According to this tool, TryHackMe has an SEO rating of **85/100** (as of 14/11/2020). That's not too bad and it'll show the justifications as to how this score was calculated below on the page.

But...Who or What Regulates these "Crawlers"?

Aside from the search engines who provide these "Crawlers", website/web-server owners themselves ultimately stipulate what content "Crawlers" can scrape. Search engines will want to retrieve **everything** from a website - but there are a few cases where we wouldn't want **all** of the contents of our website to be indexed! Can you think of any...? How about a secret administrator login page? We don't want **everyone** to be able to find that directory - especially through a google search.

Introducing Robots.txt...

Use the same SEO checkup tool and other online alternatives to see how their results compare for https://tryhackme.com and http://googledorking.cmnatic.co.uk

| No answer needed | ✈ Complete |

Task 4 ○ Beepboop - Robots.txt ⌄

Task 5 ○ Sitemaps ⌄

Task 6 ○ What is Google Dorking? ⌄

Created by	**Room Type**	**Users in Room**	**Created**
cmnatic	Free Room. Anyone can deploy virtual machines in the room (without being subscribed)!	114,695	1621 days ago