Google Dorking

Now it might be rather patronising explaining how these “Search Engines” work, but there’s a lot more going on behind the scenes then what we see.

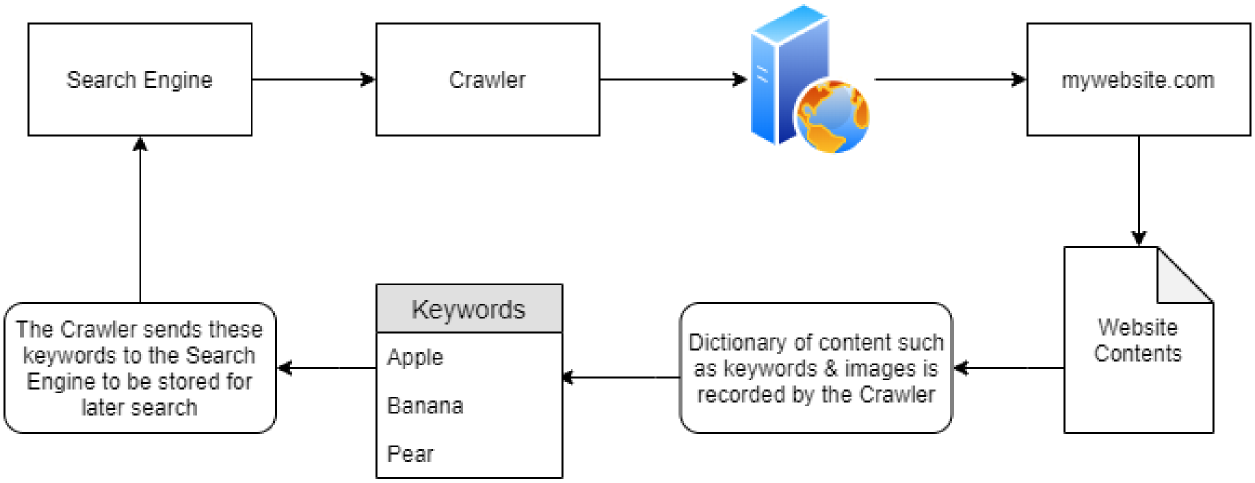
“Search Engines” such as google are huge indexers specifically, indexers of content spread across the world wide web.

What are Crawlers and how do They Work?

these crawlers discover content through various means, one begin by pure discovery, where a url is visited by the crawler and information regarding the content type of the website is returned to the search engine,

The diagram below is a high-level abstraction of how these web crawlers work.

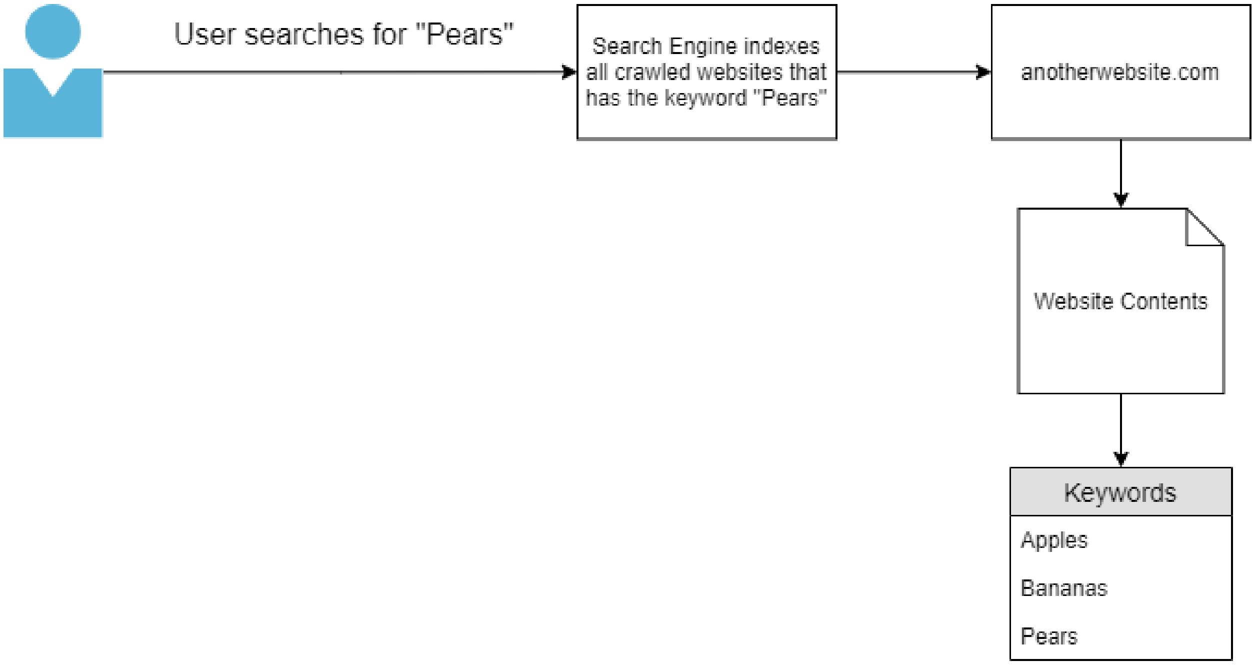
once a web crawler discovers a domain such as mywebsite.com, it will index the entire contents of the domain, looking for keywords and other miscellaneous information


Search Engine
Crawler
mywebsite.com
Website Contents
Dictionary of content such as keywords & images is recorded by the Crawler
Keywords
Apple
Banana
Pear
The Crawler sends these keywords to the Search Engine to be stored for later search

mywebsite.com has been scraped as having the keywords as “Apple” “Banana” and “pear”. these keywords are stored in a dictionary by the crawler. who then returns these to the search engine i.e **Google.** Because of the persistence, google now knows that the domain mywebsite.com has the keywords “Apple", “Banana” and “Pear”.

as only one website has been crawled, if a user was to search for apple mywebsite.com would appear. This would result in the same behaviour if the user was to search for banana. as the indexed contents from the crawler report the domain as having banana it will be displayed to the user.

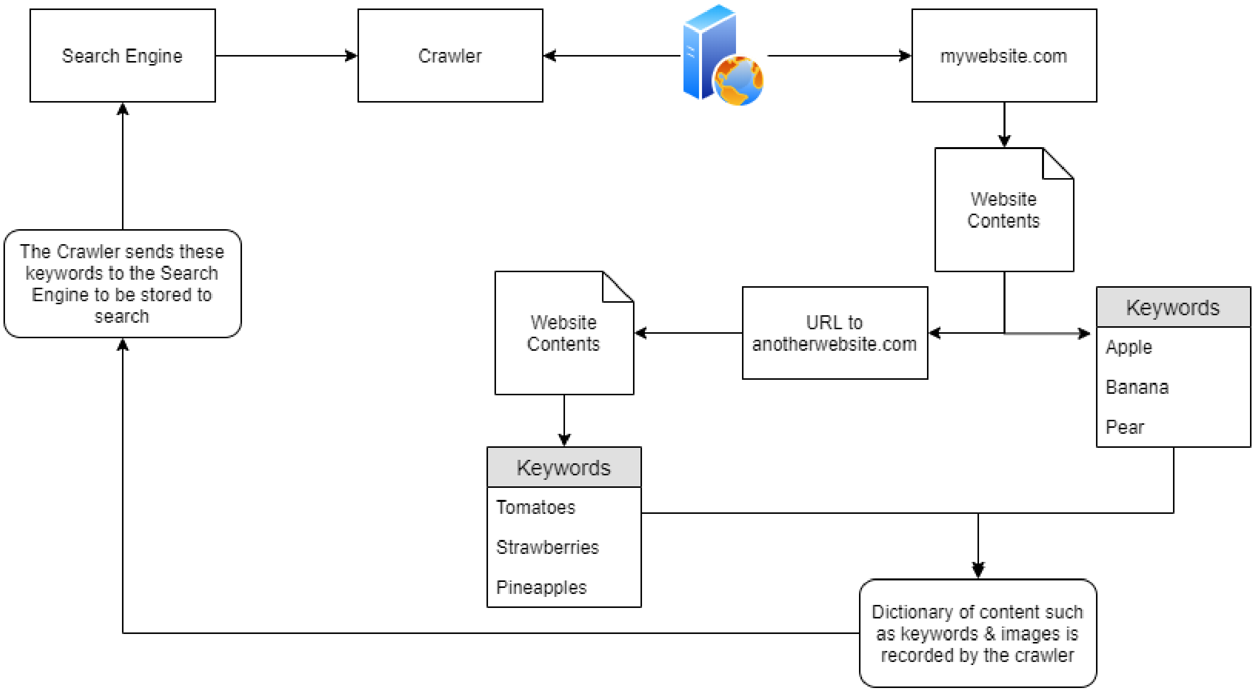
As illustrated below, a user submits a query to the search engine of “Pears". Because the search engine only has the contents of one website that has been crawled with the keyword of “Pears” it will be the only domain that is presented to the user.


User searches for "Pears"
Search Engine indexes all crawled websites that has the keyword "Pears"
anotherwebsite.com
Website Contents
Keywords
Apples
Bananas
Pears

However, as we previously mentioned, **crawlers attempt to traverse, termed as crawling, every URL and file that they can find!** Say if "**mywebsite.com**" had the same keywords as before ("Apple", "Banana" and "Pear"), but also had a URL to another website "**anotherwebsite.com**", the crawler will then attempt to traverse everything on that URL (**anotherwebsite.com**) and retrieve the contents of everything within that domain respectively.

This is illustrated in the diagram below. The crawler initially finds "**mywebsite.com**", where it crawls the contents of the website - finding the same keywords ("Apple", "Banana" and "Pear") as before, but it has additionally found an external URL. Once the crawler is complete on "**mywebsite.com**", it'll proceed to crawl the contents of the website "**anotherwebsite.com**",

where the keywords ("Tomatoes", "Strawberries" and "Pineapples") are found on it. The crawler's dictionary now contains the contents of both "**mywebsite.com**" and "**anotherwebsite.com**", which is then stored and saved within the search engine.


Search Engine
Crawler
mywebsite.com
Website Contents
The Crawler sends these keywords to the Search Engine to be stored to search
Website Contents
URL to anotherwebsite.com
Keywords
Apple
Banana
Pear
Keywords
Tomatoes
Strawberries
Pineapples
Dictionary of content such as keywords & images is recorded by the crawler

Recapping

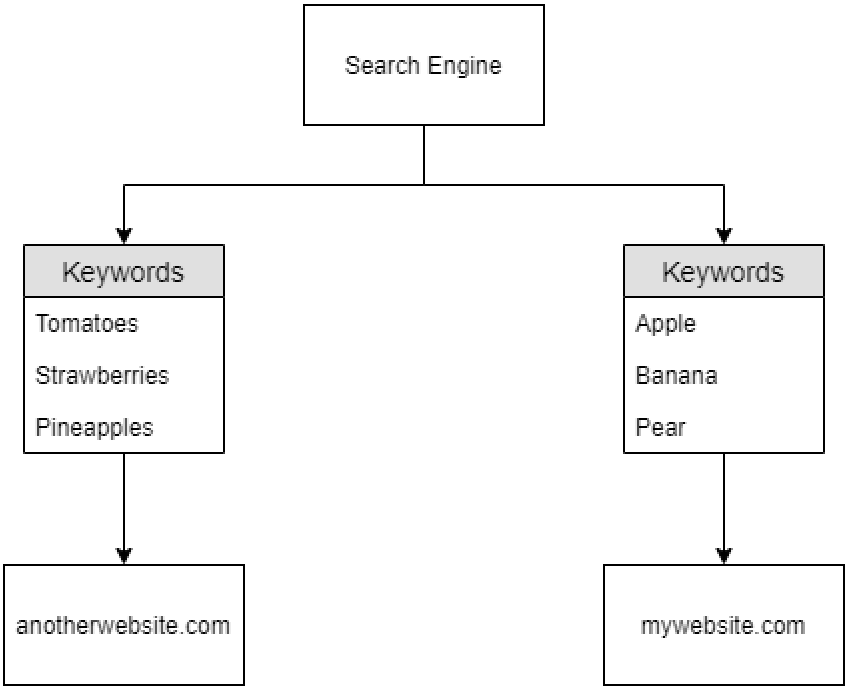
So to recap, the search engine now has knowledge of two domains that have been crawled:

1. mywebsite.com
2. anotherwebsite.com

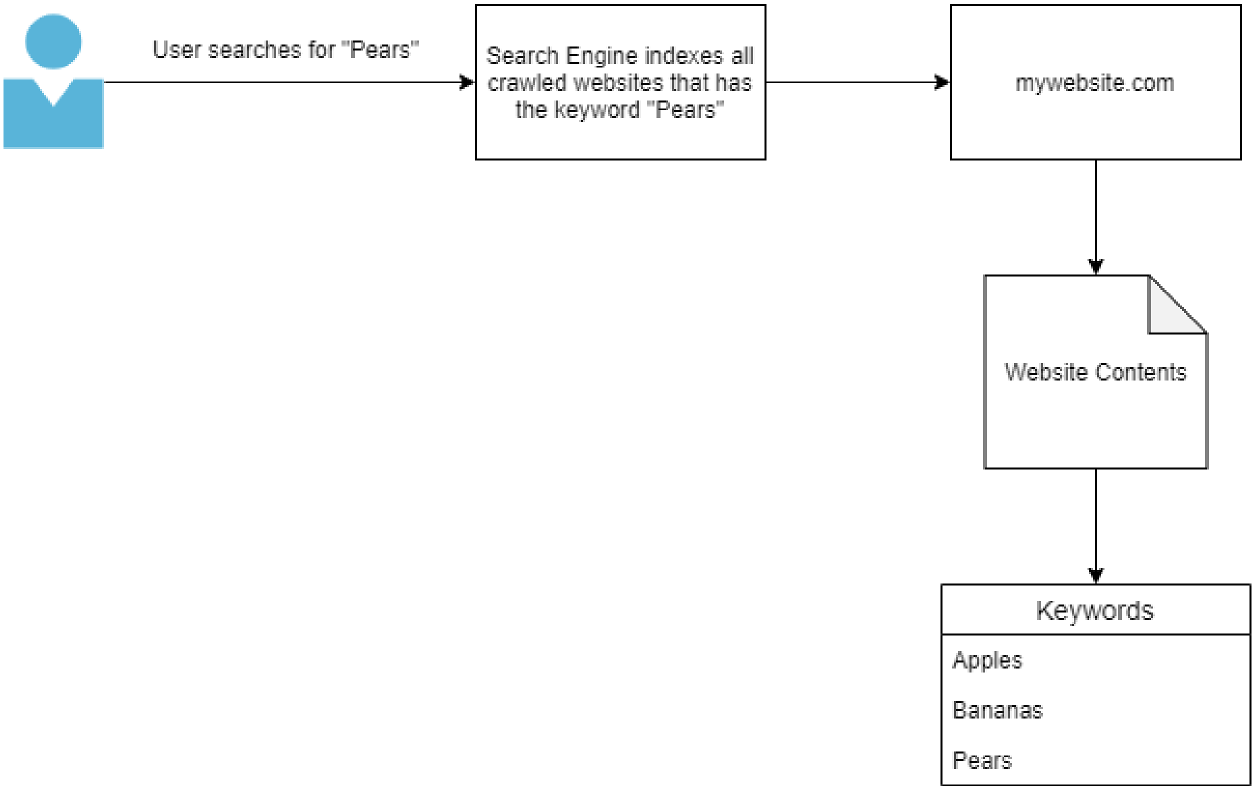
Although note that “**anotherwebsite.com**” was only crawled because it was referenced by the first domain “**mywebsite.com**”. Because of this reference, the search engine knows the following about the two domains:

Domain Name	**Keyword**
mywebsite.com	Apples
mywebsite.com	Bananas
mywebsite.com	Pears
anotherwebsite.com	Tomatoes
anotherwebsite.com	Strawberries
anotherwebsite.com	Pineapples

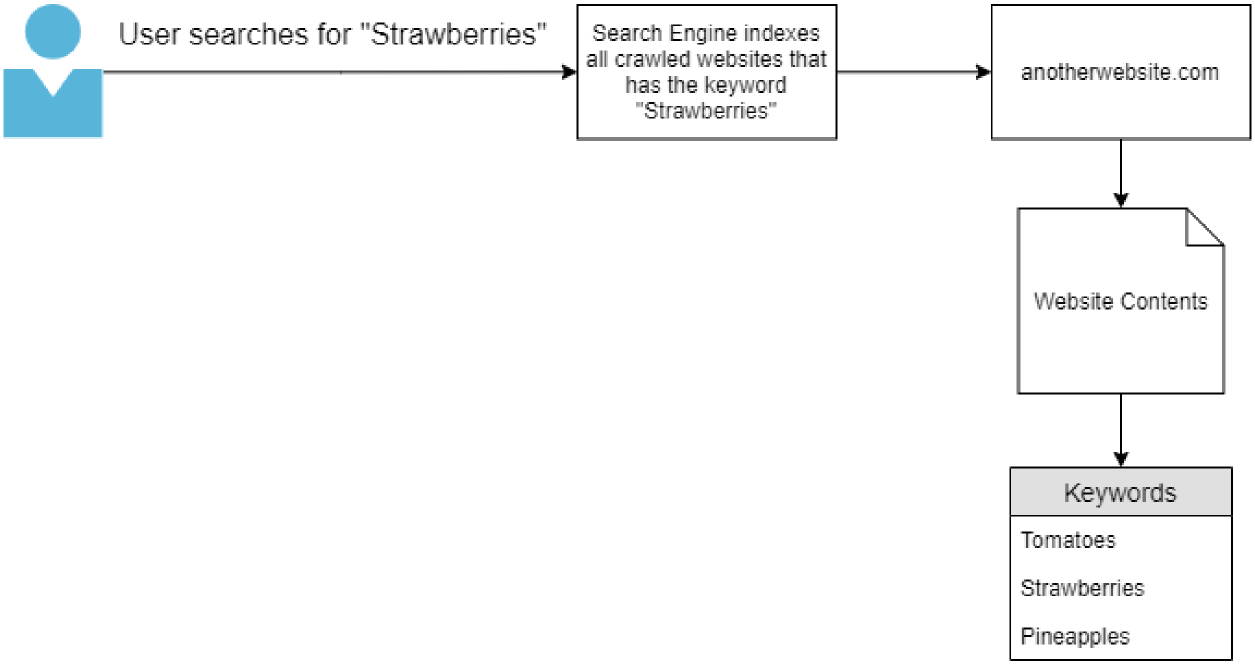
Or as illustrated below:



Now that the search engine has some knowledge about keywords, say if a user was to search for “Pears” the domain “**mywebsite.com**” will be displayed - as it is the only crawled domain containing "Pears":

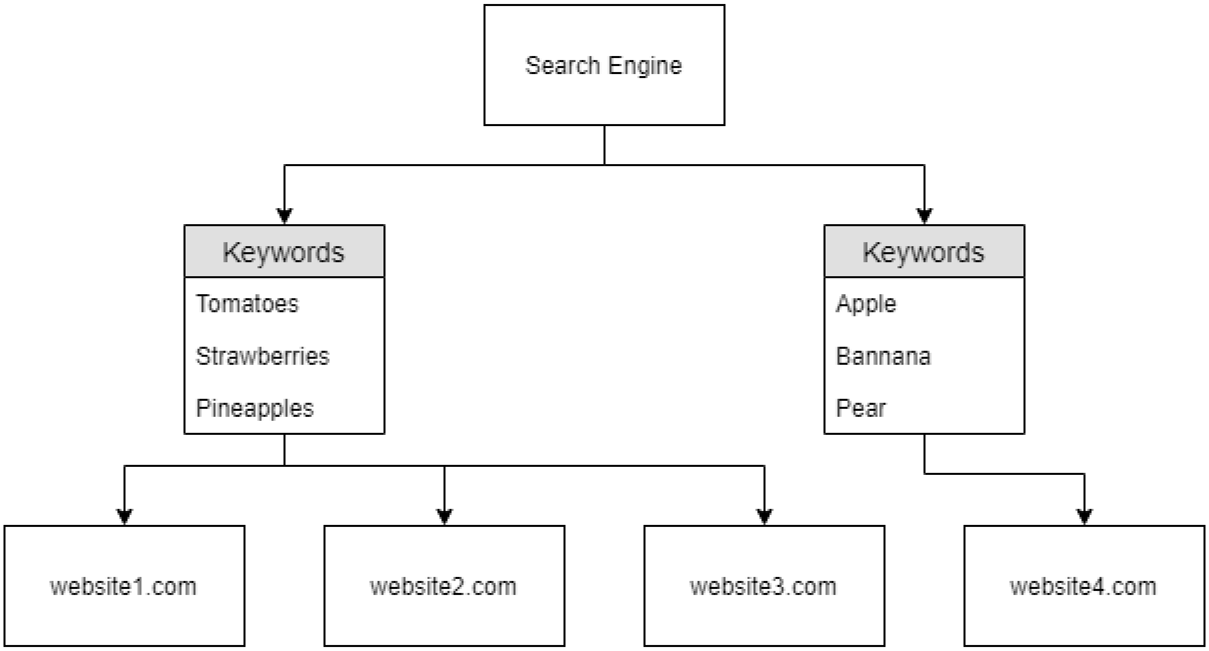

User searches for "Pears"
Search Engine indexes all crawled websites that has the keyword "Pears"
mywebsite.com
Website Contents
Keywords
Apples
Bananas
Pears

Likewise, say in this case the user now searches for "Strawberries". The domain "**anotherwebsite.com**" will be displayed, as it is the only domain that has been crawled by
the search engine that contains the keyword "Strawberries":


User searches for "Strawberries"
Search Engine indexes all crawled websites that has the keyword "Strawberries"
anotherwebsite.com
Website Contents
Keywords
Tomatoes
Strawberries
Pineapples

This is great...But imagine if a website had multiple external URL's (as they often do!) That'll require a lot of crawling to take place. There's always the chance that another website might have similar information as of that another website crawled - right? So how does the "Search Engine" decide on the hierarchy of the domains that are displayed to the user?

In the diagram below in this instance, if the user was to search for a keyword such as "Tomatoes" (which websites 1-3 contain) who decides what website gets displayed in what order?



A logical presumption would be that website 1 -> 3 would be displayed...But that's not how real-world domains work and/or are named.

1. **Name the key term of what a "Crawler" is used to do**

ans: **index**

2. **What is the name of the technique that "Search Engines" use to retrieve this information about websites?**

ans: **crawling**

3. **What is an example of the type of contents that could be gathered from a website?**

ans: **keywords**

Task 3

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”.

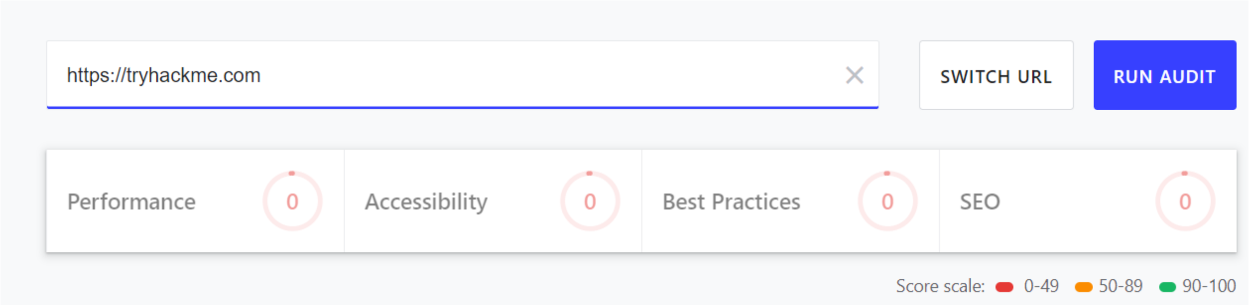
At an abstract view, search engines will “prioritise” those domains that are easier to index. There are many factors in how “optimal” a domain is - resulting in something similar to a point-scoring system.

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



https://tryhackme.com

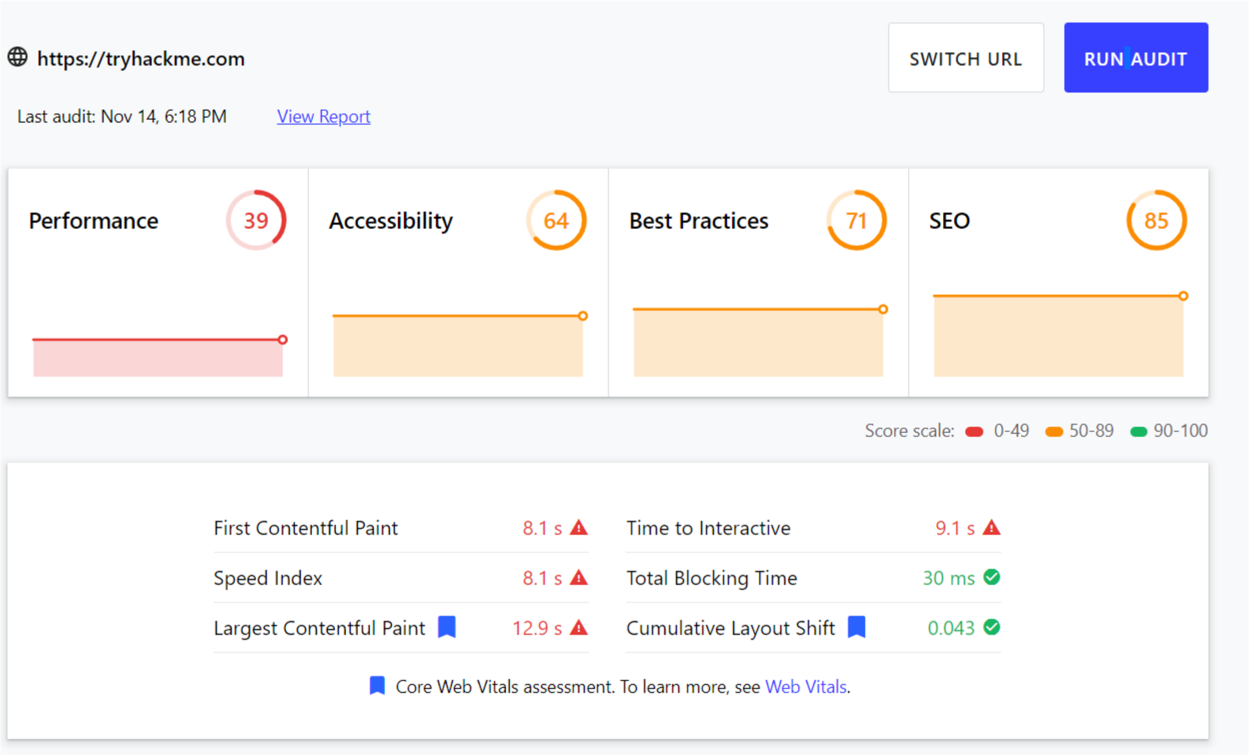
Performance 0

Accessibility 0

Best Practices 0

SEO 0

Score scale: 🟥 0-49 🟧 50-89 🟩 90-100



https://tryhackme.com

Last audit: Nov 14, 6:18 PM View Report

Performance	Accessibility	Best Practices	SEO
39	64	71	85

Score scale: 🟥 0-49 🟧 50-89 🟩 90-100

First Contentful Paint	8.1 s ▲	Time to Interactive	9.1 s ▲
Speed Index	8.1 s ▲	Total Blocking Time	30 ms ✓
Largest Contentful Paint 🔖	12.9 s ▲	Cumulative Layout Shift 🔖	0.043 ✓

🔖 Core Web Vitals assessment. To learn more, see Web Vitals.

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

Similar to "Sitemaps" which we will later discuss, this file is the first thing indexed by "Crawlers" when visiting a website.

Task 4

Beepboop-Robots.txt

Robots.txt

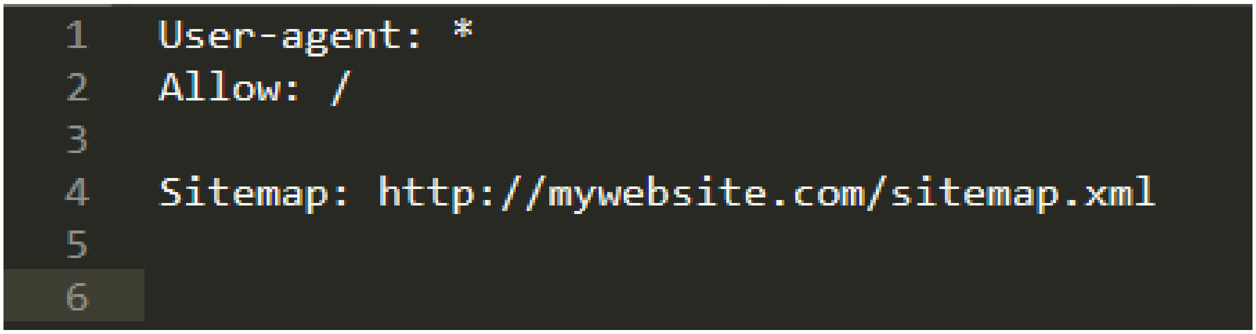
Similar to "Sitemaps" which we will later discuss, this file is the first thing indexed by "Crawlers" when visiting a website.

But what is it?

This file must be served at the root directory - specified by the webserver itself. Looking at this files extension of **.txt**, its fairly safe to assume that it is a text file.

The text file defines the permissions the "Crawler" has to the website. For example, what type of "Crawler" is allowed (I.e. You only want Google's "Crawler" to index your site and not MSN's). Moreover, Robots.txt can specify what files and directories that we do or don't want to be indexed by the "Crawler".

A very basic markup of a Robots.txt is like the following:



```
User-agent: *
Allow: /

Sitemap: http://mywebsite.com/sitemap.xml


```

Here we have a few keywords...

Keyword	Function
User-agent	Specify the type of "Crawler" that can index your site (the asterisk being a

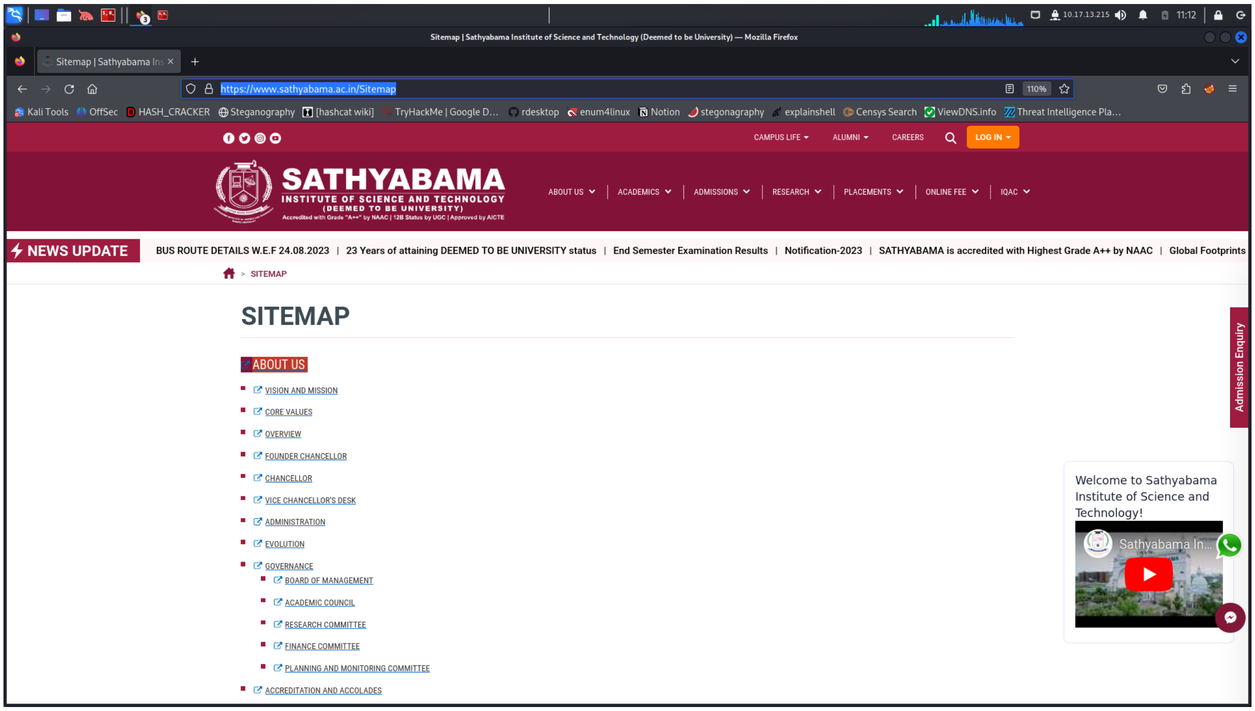
	wildcard, allowing **all "User-agents"**
Allow	Specify the directories or file(s) that the "Crawler" **can** index
Disallow	Specify the directories or file(s) that the "Crawler" **cannot** index
Sitemap	Provide a reference to where the sitemap is located (improves SEO as previously discussed, we'll come to sitemaps in the next task)

In this case:

1. Any "Crawler" can index the site

2. The "Crawler" is allowed to index the entire contents of the site

3. The "Sitemap" is located at http://mywebsite.com/sitemap.xml



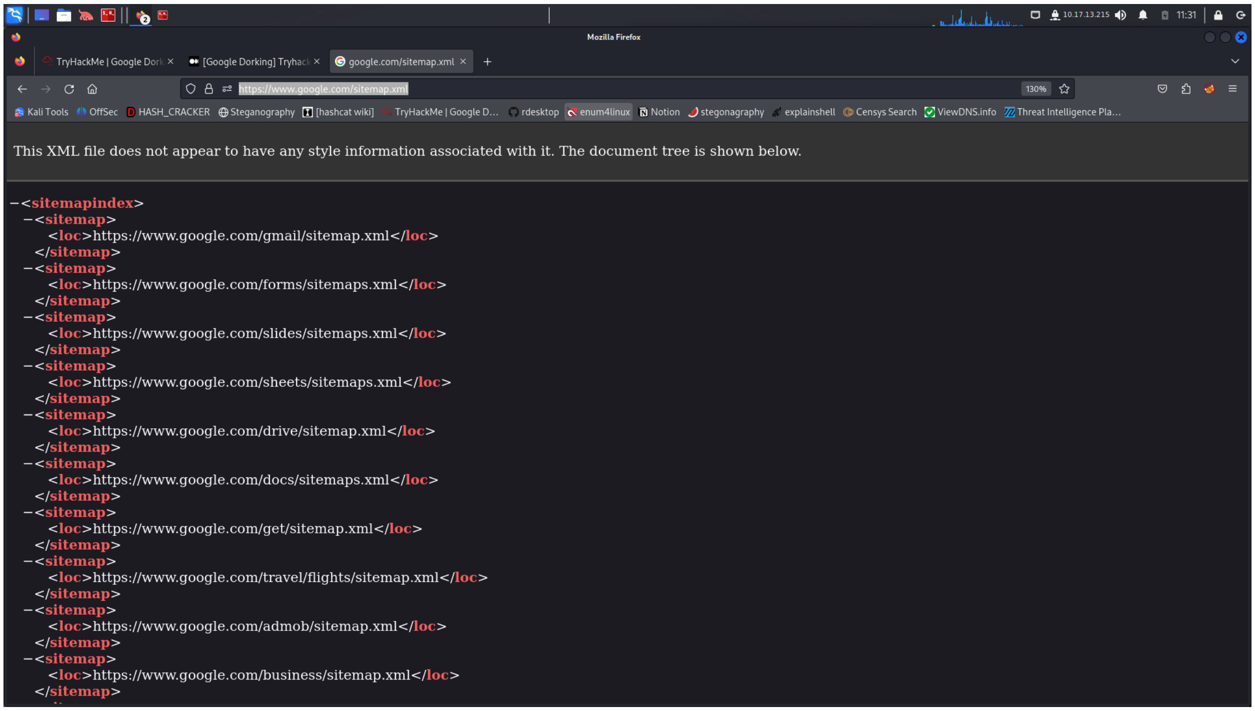
BUS ROUTE DETAILS W.E.F 24.08.2023 | 23 Years of attaining DEEMED TO BE UNIVERSITY status | End Semester Examination Results | Notification-2023 | SATHYABAMA is accredited with Highest Grade A++ by NAAC | Global Footprints

SITEMAP

ABOUT US

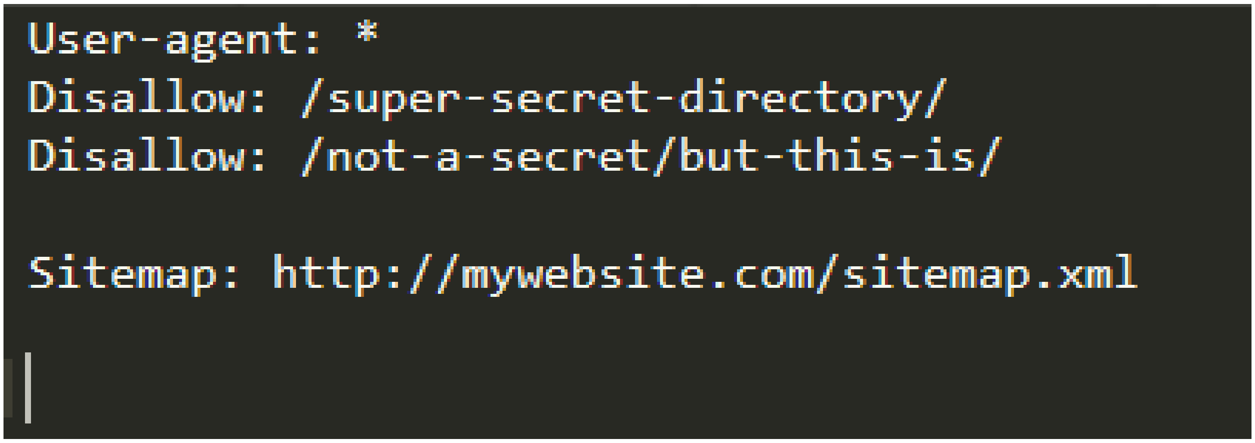
- VISION AND MISSION
- CORE VALUES
- OVERVIEW
- FOUNDER CHANCELLOR
- CHANCELLOR
- VICE CHANCELLOR'S DESK
- ADMINISTRATION
- EVOLUTION
- GOVERNANCE
  - BOARD OF MANAGEMENT
  - ACADEMIC COUNCIL
  - RESEARCH COMMITTEE
  - FINANCE COMMITTEE
  - PLANNING AND MONITORING COMMITTEE
- ACCREDITATION AND ACCOLADES

Welcome to Sathyabama Institute of Science and Technology!



This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
-<sitemapindex>
  -<sitemap>
     <loc>https://www.google.com/gmail/sitemap.xml</loc>
   </sitemap>
  -<sitemap>
     <loc>https://www.google.com/forms/sitemaps.xml</loc>
   </sitemap>
  -<sitemap>
     <loc>https://www.google.com/slides/sitemaps.xml</loc>
   </sitemap>
  -<sitemap>
     <loc>https://www.google.com/sheets/sitemaps.xml</loc>
   </sitemap>
  -<sitemap>
     <loc>https://www.google.com/drive/sitemap.xml</loc>
   </sitemap>
  -<sitemap>
     <loc>https://www.google.com/docs/sitemaps.xml</loc>
   </sitemap>
  -<sitemap>
     <loc>https://www.google.com/get/sitemap.xml</loc>
   </sitemap>
  -<sitemap>
     <loc>https://www.google.com/travel/flights/sitemap.xml</loc>
   </sitemap>
  -<sitemap>
     <loc>https://www.google.com/admob/sitemap.xml</loc>
   </sitemap>
  -<sitemap>
     <loc>https://www.google.com/business/sitemap.xml</loc>
   </sitemap>
```



```
User-agent: *
Disallow: /super-secret-directory/
Disallow: /not-a-secret/but-this-is/

Sitemap: http://mywebsite.com/sitemap.xml
```

In this case:

1. Any "Crawler" can index the site

2. The "Crawler" can index every other content that isn't contained within "/super-secret-directory/".

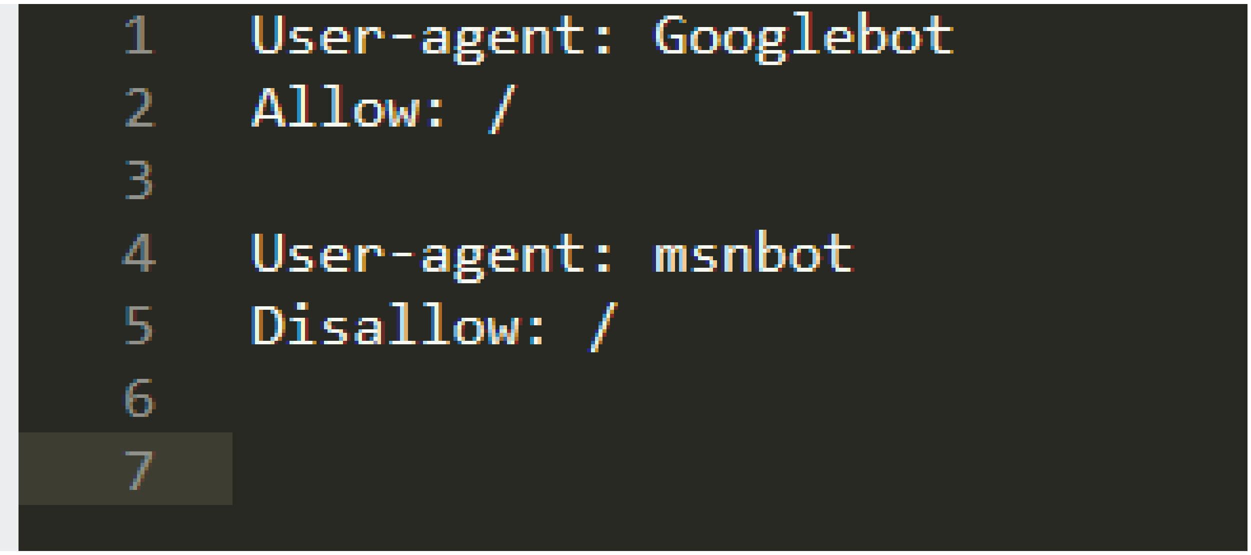
Crawlers
 also know the differences between sub-directories, directories and
files. Such as in the case of the second "Disallow:" ("/not-a-secret/but-this-is/")

The "Crawler" will index all the contents within "**/not-a-secret/**", but will not index anything contained within the sub-directory **"/but-this-is/"**.

3. The "Sitemap" is located at http://mywebsite.com/sitemap.xml

What if we Only Wanted Certain "Crawlers" to Index our Site?

We can stipulate so, such as in the picture below:



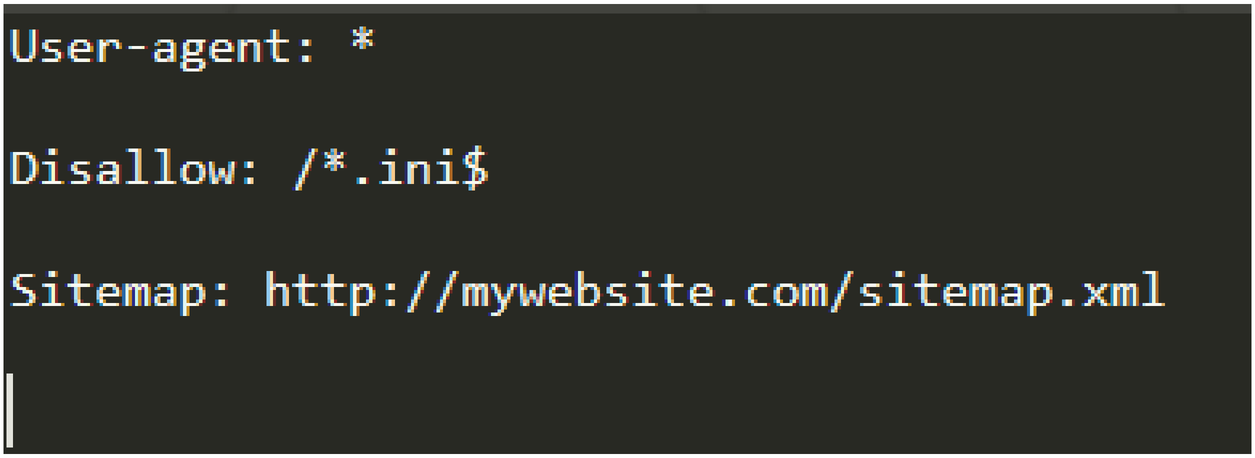
```
User-agent: Googlebot
Allow: /

User-agent: msnbot
Disallow: /

```

In this case:

1. The "Crawler" "Googlebot" is allowed to index the entire site ("Allow: /")
2. The "Crawler" "msnbot" is not allowed to index the site (Disallow: /")

How about Preventing Files From Being Indexed?

Whilst you can make manual entries for every file extension that you don't want to be indexed, you will have to provide the directory it is within, as well as the full filename. Imagine if you had a huge site! What a pain...Here's where we can use a bit of **regexing.**



```
User-agent: *

Disallow: /*.ini$

Sitemap: http://mywebsite.com/sitemap.xml
```

In this case:

1. Any "Crawler" can index the site

2. However, the "Crawler" cannot index **any** file that has the extension of **.ini** within any directory/sub-directory using ("$") of the site.

3. The "Sitemap" is located at http://mywebsite.com/sitemap.xml

Why would you want to hide a **.ini** file for example? Well, files like this contain sensitive configuration details. Can you think of any other file formats that might contain sensitive information?

1. **Where would "robots.txt" be located on the domain "ablog.com".**

ans: **ablog.com/robots.txt**

2. **If a website was to have a sitemap, where would that be located?**

ans: **sitemap.xml**

3. **How would we only allow "Bingbot" to index the website?**

ans: **User-Agent: Bingbot**

4. **How would we prevent a "Crawler" from indexing the directory "/dont-index-me/"?**

ans: **Disallow: /dont-index-me/**

5. **What is the extension of a Unix/Linux system configuration file that we might want to hide from "Crawlers"?**

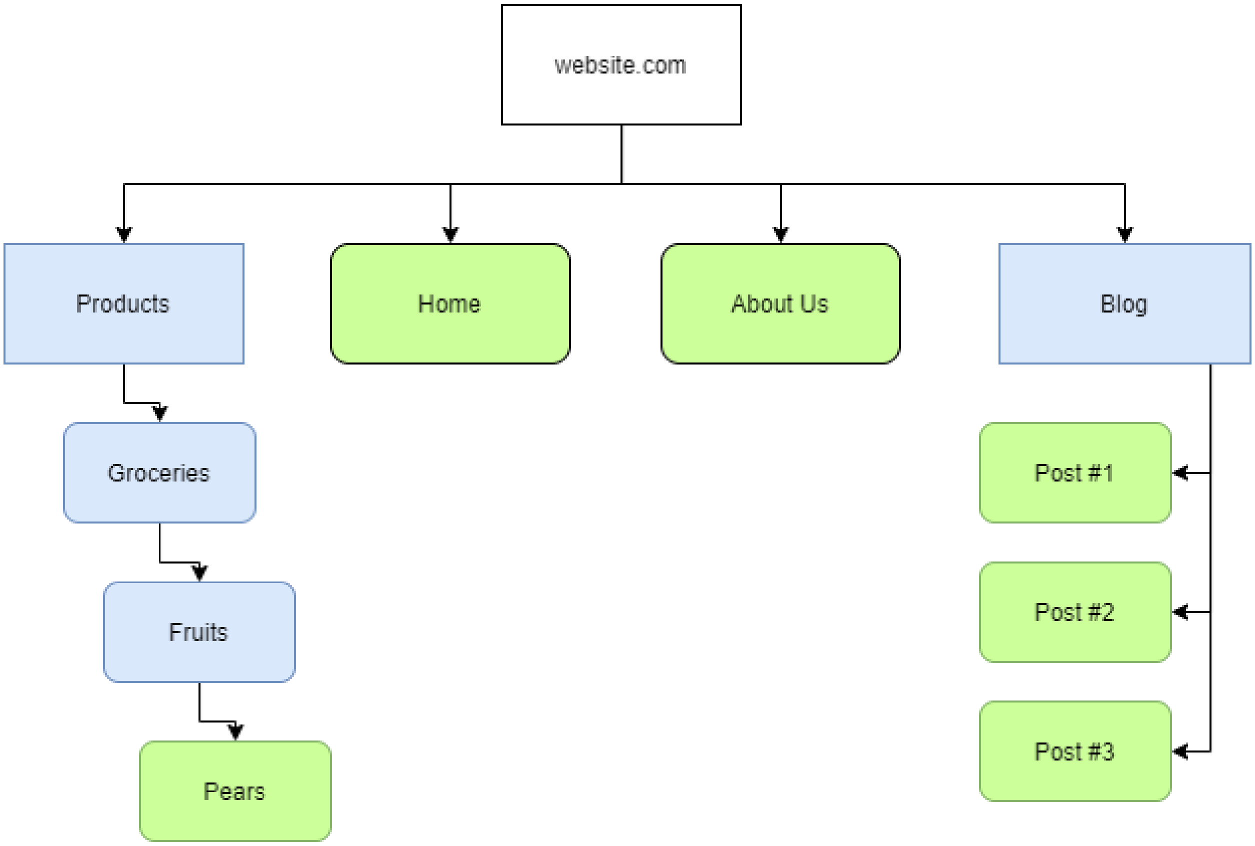
ans: .conf

Task 5

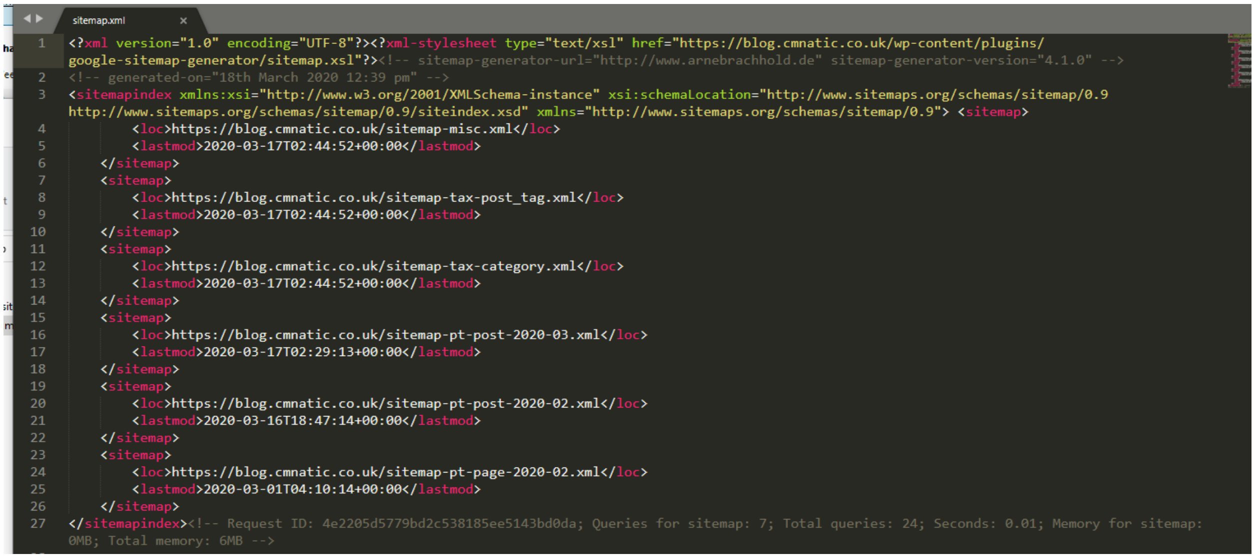
Sitemaps

Comparable to geographical maps in real life, “Sitemaps” are just that - but for websites!

“Sitemaps” are indicative resources that are helpful for crawlers, as they specify the necessary routes to find content on the domain. The below illustration is a good example of the structure of a website, and how it may look on a "Sitemap":



The blue rectangles represent the **route** to nested-content, similar to a directory I.e. “Products” for a store. Whereas, the green rounded-rectangles represent an actual page. However, this is for illustration purposes only - “Sitemaps” don't look like this in the real world. They look something much more similar to this:



```xml
<?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="https://blog.cmnatic.co.uk/wp-content/plugins/
google-sitemap-generator/sitemap.xsl"?><!-- sitemap-generator-url="http://www.arnebrachhold.de" sitemap-generator-version="4.1.0" -->
<!-- generated-on="18th March 2020 12:39 pm" -->
<sitemapindex xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/siteindex.xsd" xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"> <sitemap>
		<loc>https://blog.cmnatic.co.uk/sitemap-misc.xml</loc>
		<lastmod>2020-03-17T02:44:52+00:00</lastmod>
	</sitemap>
	<sitemap>
		<loc>https://blog.cmnatic.co.uk/sitemap-tax-post_tag.xml</loc>
		<lastmod>2020-03-17T02:44:52+00:00</lastmod>
	</sitemap>
	<sitemap>
		<loc>https://blog.cmnatic.co.uk/sitemap-tax-category.xml</loc>
		<lastmod>2020-03-17T02:44:52+00:00</lastmod>
	</sitemap>
	<sitemap>
		<loc>https://blog.cmnatic.co.uk/sitemap-pt-post-2020-03.xml</loc>
		<lastmod>2020-03-17T02:29:13+00:00</lastmod>
	</sitemap>
	<sitemap>
		<loc>https://blog.cmnatic.co.uk/sitemap-pt-post-2020-02.xml</loc>
		<lastmod>2020-03-16T18:47:14+00:00</lastmod>
	</sitemap>
	<sitemap>
		<loc>https://blog.cmnatic.co.uk/sitemap-pt-page-2020-02.xml</loc>
		<lastmod>2020-03-01T04:10:14+00:00</lastmod>
	</sitemap>
</sitemapindex><!-- Request ID: 4e2205d5779bd2c538185ee5143bd0da; Queries for sitemap: 7; Total queries: 24; Seconds: 0.01; Memory for sitemap:
0MB; Total memory: 6MB -->
```

“Sitemaps” are XML formatted. I won't explain the structure of this file-formatting as the room XXE created by falconfeast does a mighty fine job of this.

The presence of "Sitemaps" holds a fair amount of weight in influencing the "optimisation" and favorability of a website. As we discussed in the "Search Engine Optimisation" task, these maps make the traversal of content much easier for the crawler!

Why are "Sitemaps" so Favourable for Search Engines?

Search engines are lazy! Well, better yet - search engines have a lot of data to process. The efficiency of how this data is collected is paramount. Resources like "Sitemaps" are extremely helpful for "Crawlers" as the necessary routes to content are already provided! All the crawler has to do is scrape this content - rather than going through the process of manually finding and scraping. Think of it as using a wordlist to find files instead of randomly guessing their names!

The easier a website is to "Crawl", the more optimised it is for the "Search Engine”

1. **What is the typical file structure of a "Sitemap"?**

ans: **xml**

2. **What real life example can "Sitemaps" be compared to?**

ans: **map**

3. **Name the keyword for the path taken for content on a website**

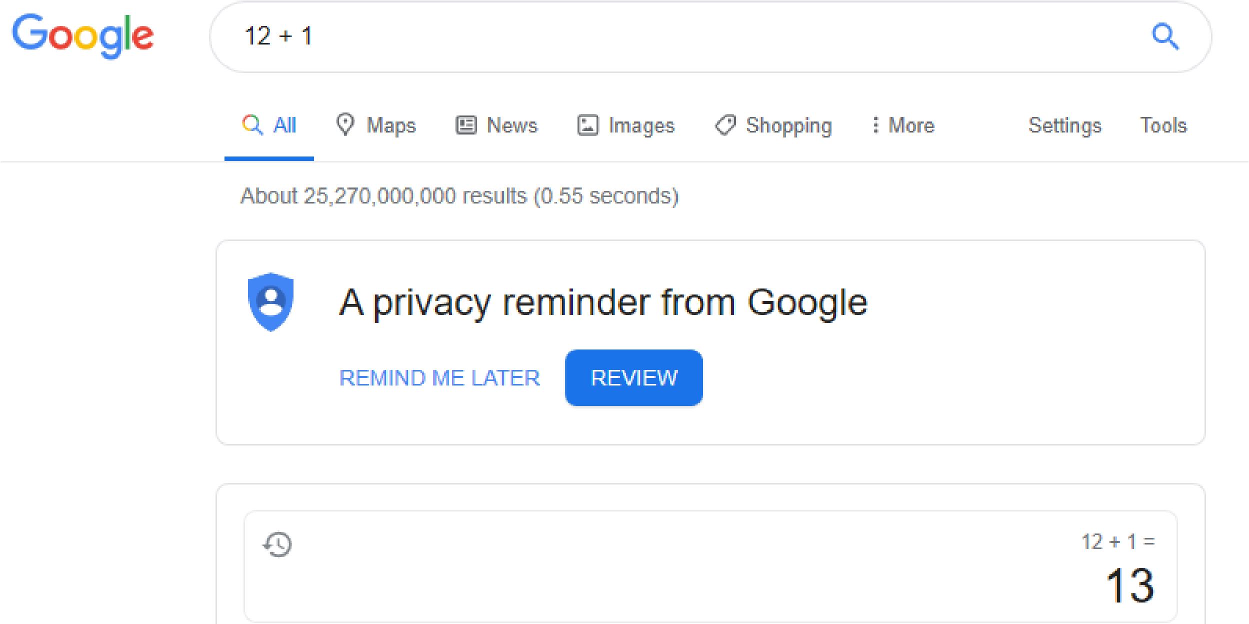
ans: **route**

Task 6

What is Google Dorking

As we have previously discussed, Google has a lot of websites crawled and indexed. Your average Joe uses Google to look up Cat pictures (I'm more of a Dog person myself...). Whilst Google will have many Cat pictures indexed ready to serve to Joe, this is a rather trivial use of the search engine in comparison to what it can be used for.

For
example, we can add operators such as that from programming languages to either increase or decrease our search results - or perform actions such as arithmetic!



12 + 1

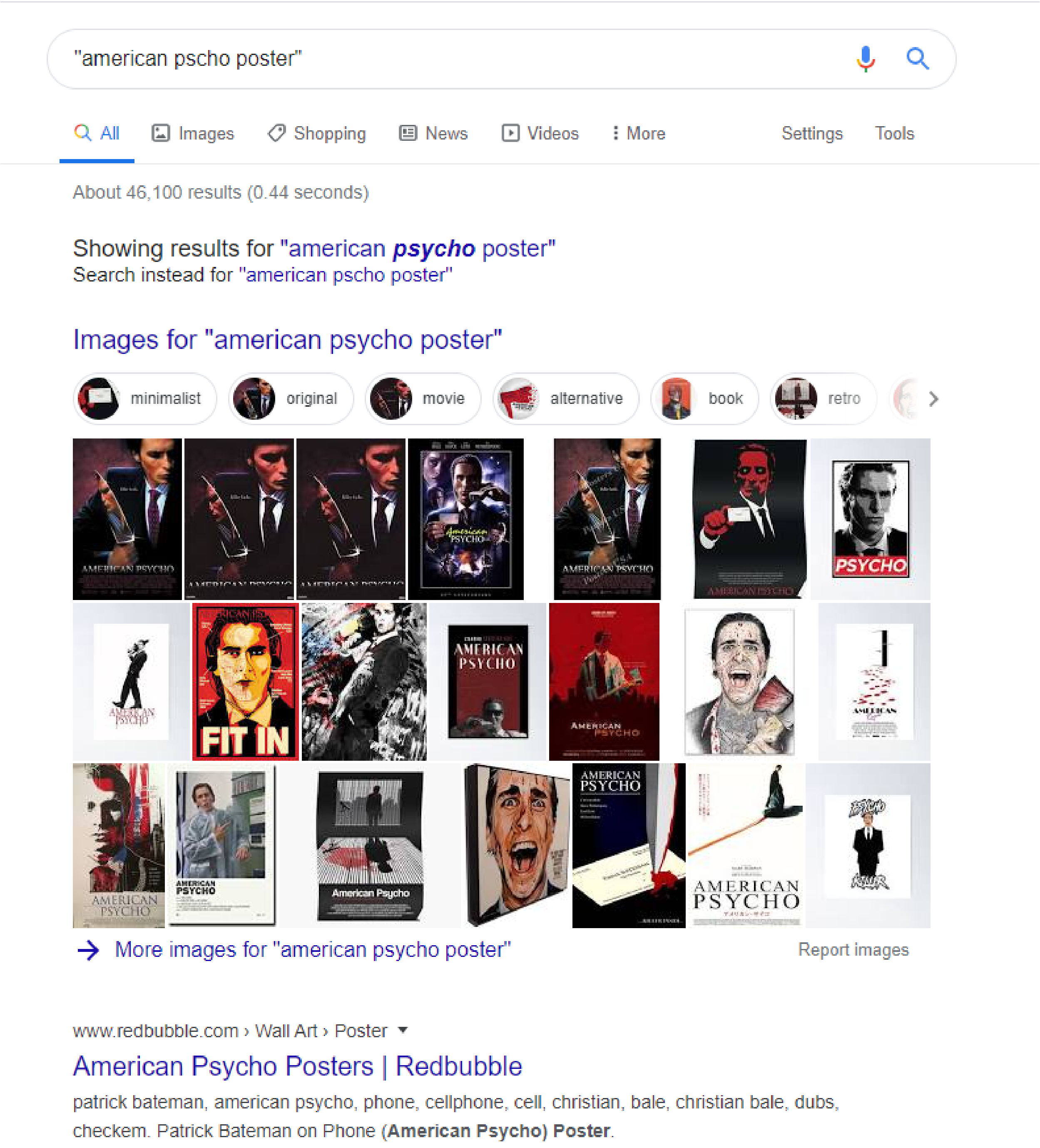
About 25,270,000,000 results (0.55 seconds)

A privacy reminder from Google

12 + 1 =

13

Say if we wanted to narrow down our search query, we can use quotation marks. Google will interpret everything in between these quotation marks as exact and only return the results of the exact phrase provided...Rather useful to filter through the rubbish that we don't need as we have done so below:



About 46,100 results (0.44 seconds)

Showing results for "american ***psycho*** poster"
Search instead for "american pscho poster"

Images for "american psycho poster"

More images for "american psycho poster"

www.redbubble.com › Wall Art › Poster

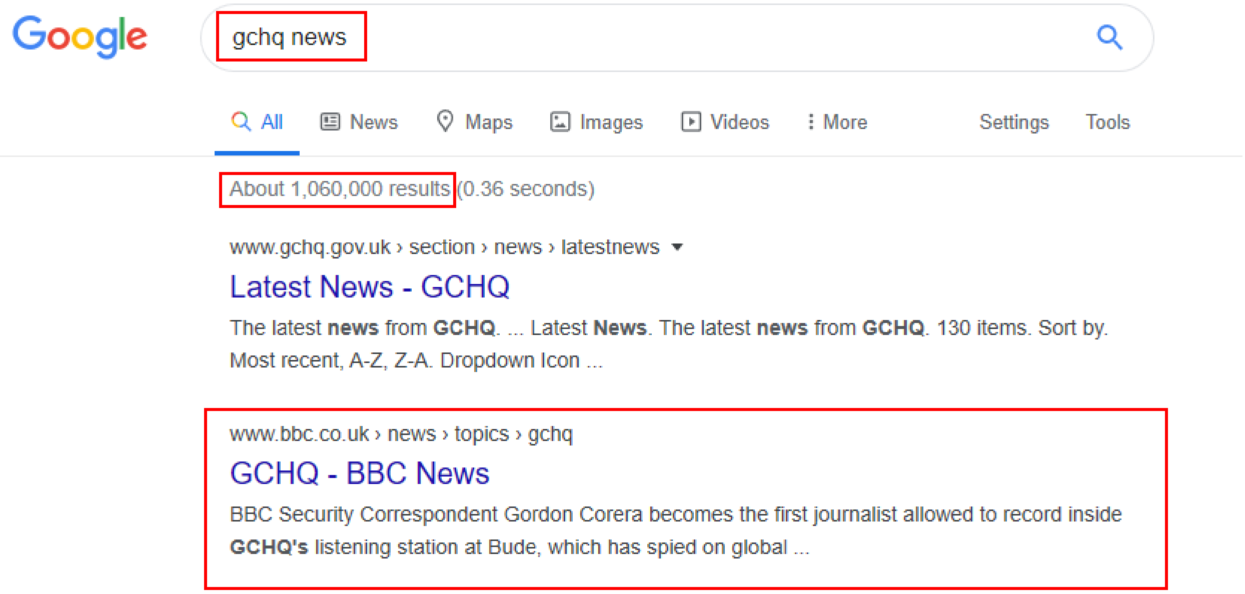
American Psycho Posters | Redbubble

patrick bateman, american psycho, phone, cellphone, cell, christian, bale, christian bale, dubs, checkem. Patrick Bateman on Phone (**American Psycho**) **Poster**.

Refining our Queries

We can use terms such as “**site**” (such as bbc.co.uk) and a query (such as "gchq news") to search the specified site for the keyword we have provided to filter out content that may be harder to find otherwise. For example, using the “site” and "query" of "bbc" and "gchq", we have modified the order of which Google returns the results.

In the screenshot below, searching for “gchq news” returns approximately 1,060,000 results from Google. The website that we want is ranked behind GCHQ's actual website:



gchq news

About 1,060,000 results (0.36 seconds)

www.gchq.gov.uk › section › news › latestnews

Latest News - GCHQ

The latest **news** from **GCHQ**. ... Latest **News**. The latest **news** from **GCHQ**. 130 items. Sort by. Most recent, A-Z, Z-A. Dropdown Icon ...

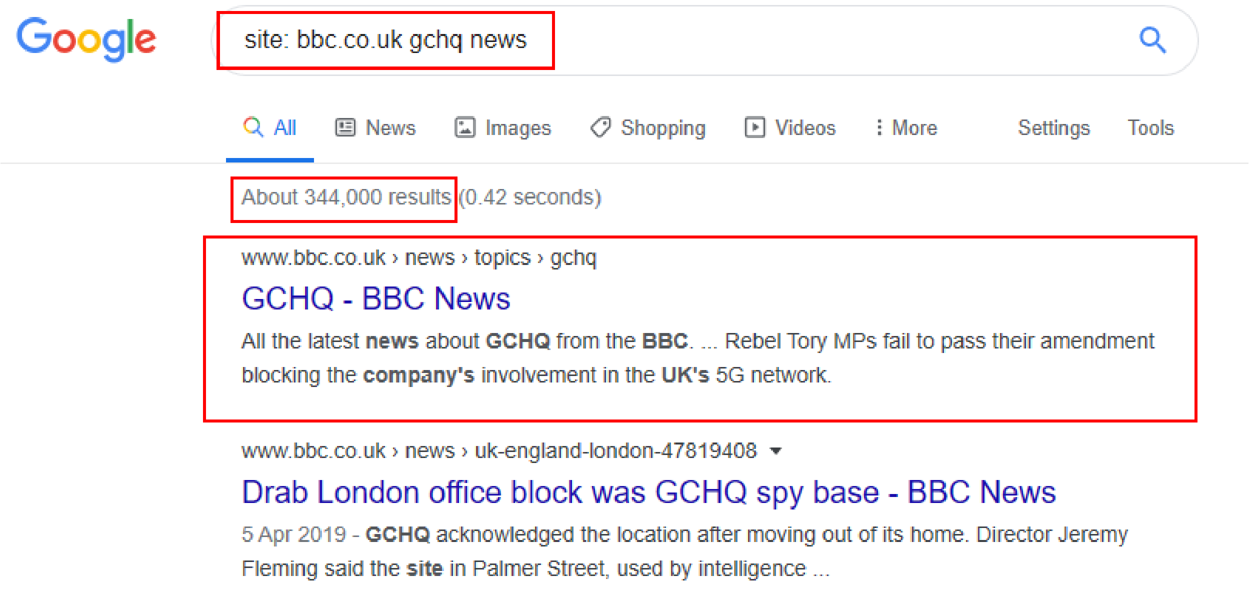
www.bbc.co.uk › news › topics › gchq

GCHQ - BBC News

BBC Security Correspondent Gordon Corera becomes the first journalist allowed to record inside **GCHQ's** listening station at Bude, which has spied on global ...

But we don't want that...We wanted “**bbc.co.uk**” first, so let's refine our search using the “**site**”

term. Notice how in the screenshot below, Google returns with much fewer results? Additionally, the page that we didn't want has disappeared, leaving the site that we did actually want!


site: bbc.co.uk gchq news

About 344,000 results (0.42 seconds)

www.bbc.co.uk › news › topics › gchq

GCHQ - BBC News

All the latest **news** about **GCHQ** from the **BBC**. ... Rebel Tory MPs fail to pass their amendment blocking the **company's** involvement in the **UK's** 5G network.

www.bbc.co.uk › news › uk-england-london-47819408

Drab London office block was GCHQ spy base - BBC News

5 Apr 2019 - **GCHQ** acknowledged the location after moving out of its home. Director Jeremy Fleming said the **site** in Palmer Street, used by intelligence ...

Of course, in this case, GCHQ is quite a topic of discussion - so there'll be a load of results regardless.

So What Makes "Google Dorking" so Appealing?

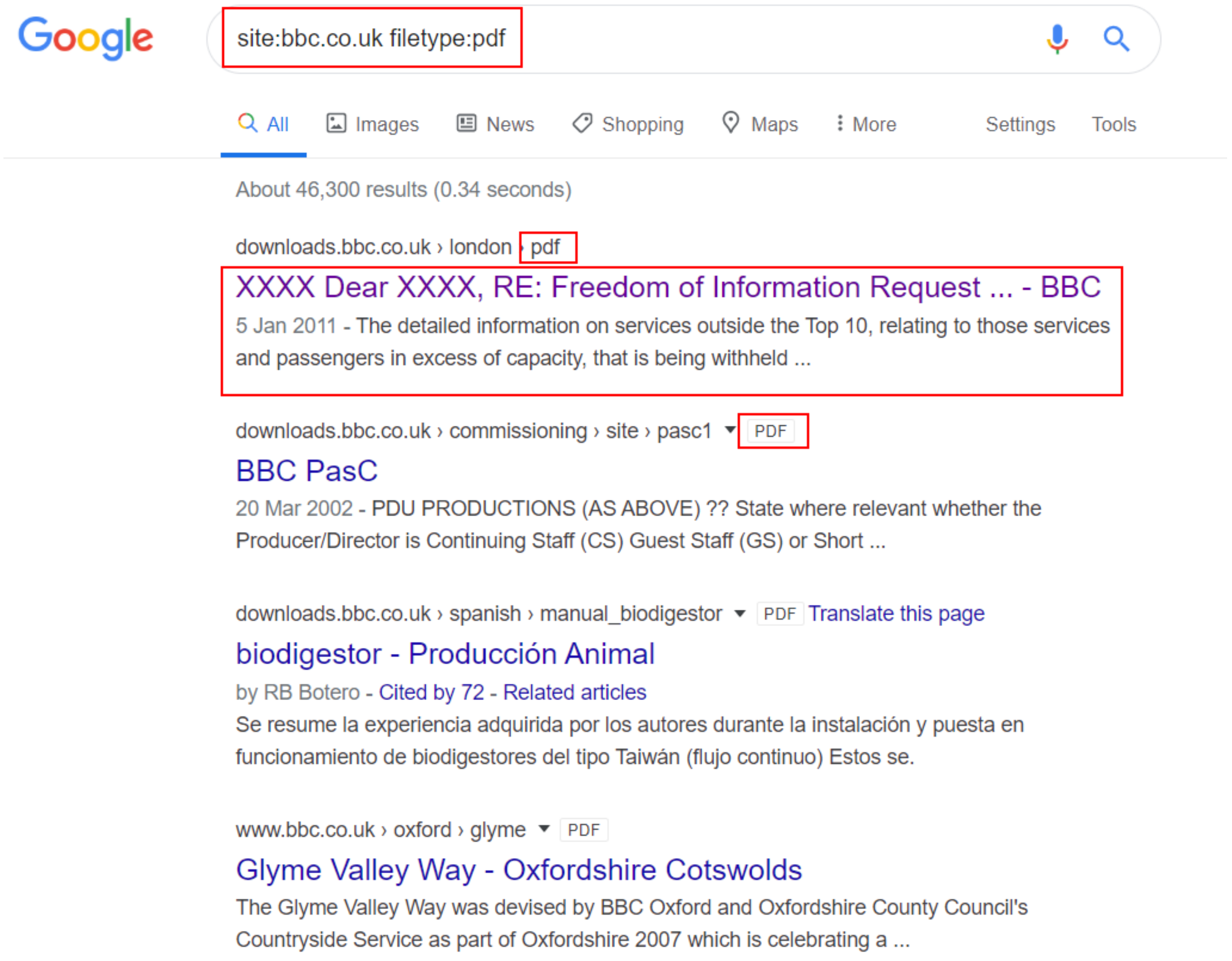
First of all - and the important part - it's legal! It's all indexed, publicly available information. However, what you do with this is where the question of legality comes in to play…

A few common terms we can search and combine include:

Term	Action
filetype:	Search for a file by its extension (e.g. PDF)
cache:	View Google's Cached version of a specified URL
intitle:	The specified phrase MUST appear in the title of the page

For example, let's say we wanted to use Google to search for all PDFs on bbc.co.uk:

`site:bbc.co.uk filetype:pdf`



site:bbc.co.uk filetype:pdf

About 46,300 results (0.34 seconds)

downloads.bbc.co.uk › london › pdf
XXXX Dear XXXX, RE: Freedom of Information Request ... - BBC
5 Jan 2011 - The detailed information on services outside the Top 10, relating to those services and passengers in excess of capacity, that is being withheld ...

downloads.bbc.co.uk › commissioning › site › pasc1 ▾ PDF
BBC PasC
20 Mar 2002 - PDU PRODUCTIONS (AS ABOVE) ?? State where relevant whether the Producer/Director is Continuing Staff (CS) Guest Staff (GS) or Short ...

downloads.bbc.co.uk › spanish › manual_biodigestor ▾ PDF Translate this page
biodigestor - Producción Animal
by RB Botero - Cited by 72 - Related articles
Se resume la experiencia adquirida por los autores durante la instalación y puesta en funcionamiento de biodigestores del tipo Taiwán (flujo continuo) Estos se.

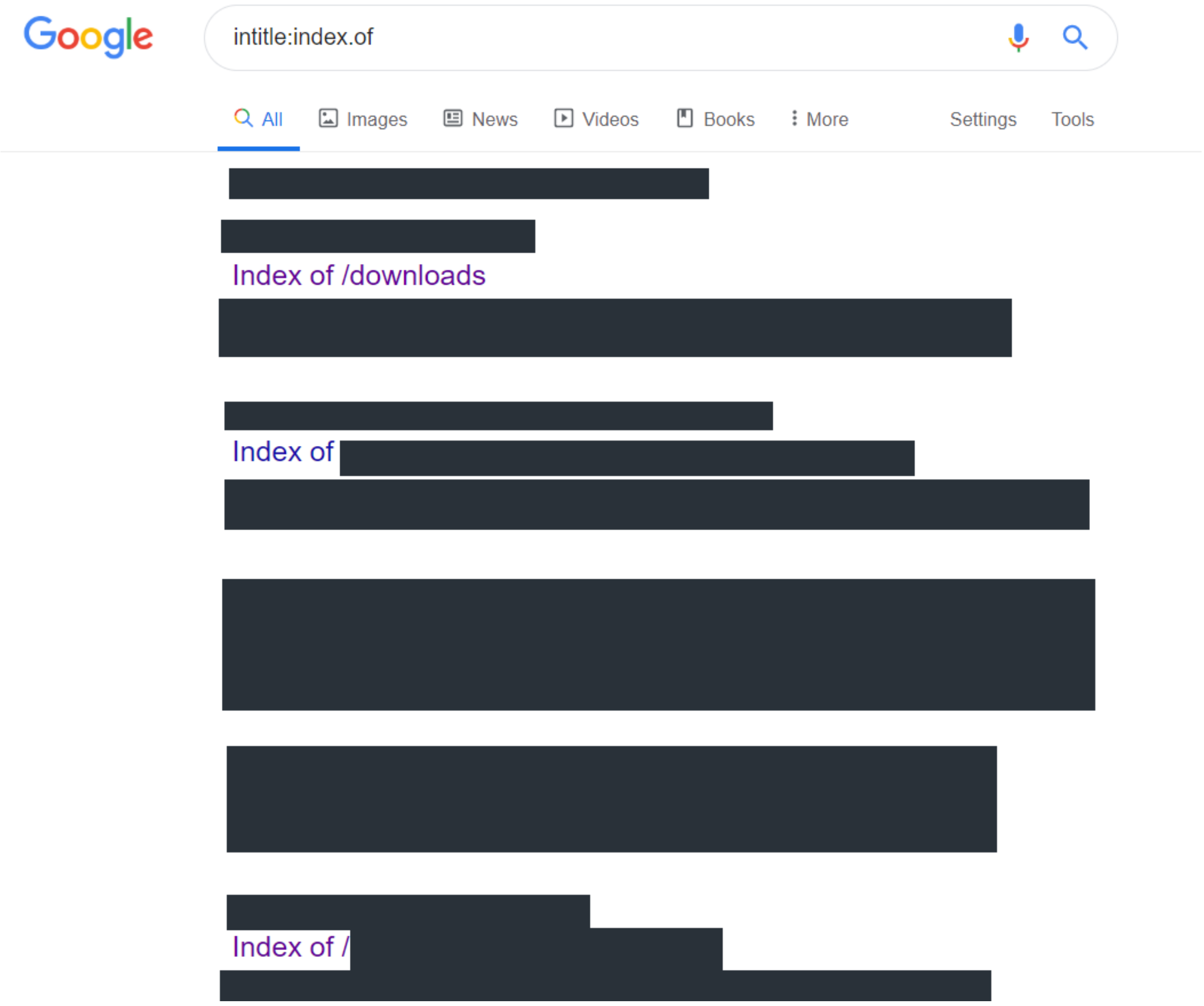
www.bbc.co.uk › oxford › glyme ▾ PDF
Glyme Valley Way - Oxfordshire Cotswolds
The Glyme Valley Way was devised by BBC Oxford and Oxfordshire County Council's Countryside Service as part of Oxfordshire 2007 which is celebrating a ...

Great, now we've refined our search for Google to query for all publically accessible PDFs on "**bbc.co.uk**" - You wouldn't have found files like this "Freedom of Information Request Act" file from a wordlist!

Here we used the extension **PDF**, but can you think of any other file formats of sensitive nature that **may** be publically accessible? (Often unintentionally!!) Again, what you do with any results that you find is where the legality comes into play - this is why "Google Dorking" is so great/dangerous.

Here is simple directory traversal.

I have blanked out a lot of the below to cover you, me, THM and the owners of the domains:



intitle:index.of

[redacted]

Index of /downloads

[redacted]

[redacted]

Index of [redacted]

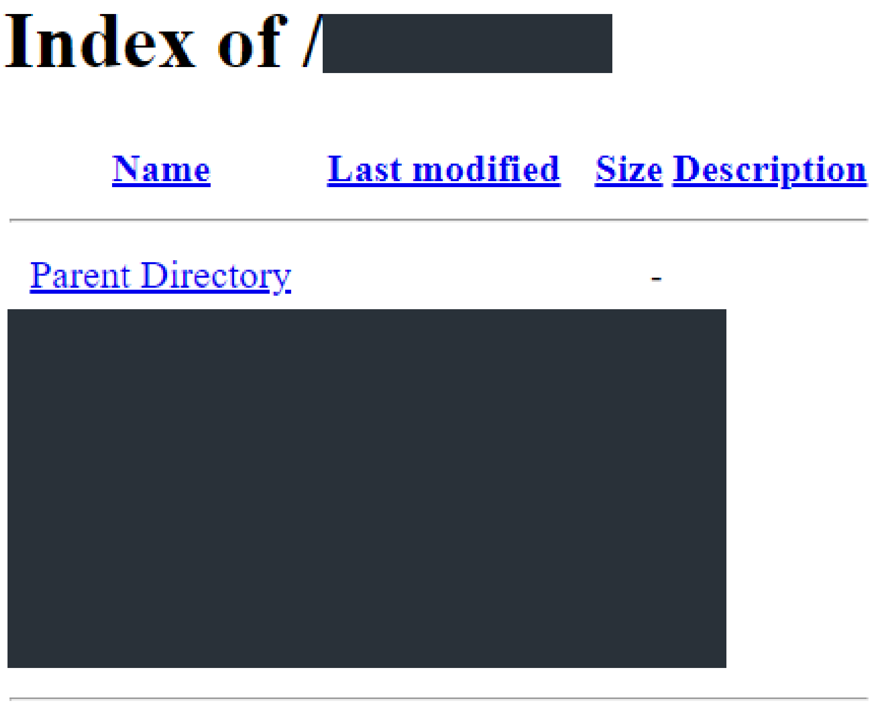
[redacted]

[redacted]

[redacted]

[redacted]

Index of / [redacted]



Index of /[redacted]

Name	Last modified	Size	Description
Parent Directory		-	

[redacted]

1. What would be the format used to query the site bbc.co.uk about flood defences

ans: site: bbc.co.uk flood defences

2. What term would you use to search by file type?

ans: filetype:

3. What term can we use to look for login pages?

ans: intitle: login

CHEET SHEET

Let's look at the most popular Google Dorks and what they do.

cache: this dork will show you the cached version of any website, e.g. cache:securitytrails.com

allintext: searches for specific text contained on any web page, e.g. allintext: hacking tools

allintitle: exactly the same as allintext, but will show pages that contain titles with X characters, e.g. allintitle:"Security Companies"

allinurl: it can be used to fetch results whose URL contains all the specified characters, e.g: allinurl:clientarea

filetype: used to search for any kind of file extensions, for example, if you want to search for pdf files you can use: email security filetype: pdf

inurl: this is exactly the same as allinurl, but it is only useful for one single keyword, e.g. inurl:admin

intitle: used to search for various keywords inside the title, for example, intitle:security tools will search for titles beginning with "security" but "tools" can be somewhere else in the page.

inanchor: this is useful when you need to search for an exact anchor text used on any links, e.g. inanchor:"cyber security"

intext: useful to locate pages that contain certain characters or strings inside their text, e.g. intext:"safe internet"

site: will show you the full list of all indexed URLs for the specified domain and subdomain, e.g. site:securitytrails.com

*: wildcard used to search pages that contain "anything" before your word, e.g. how to * a website, will return "how to…" design/create/hack, etc… "a website".

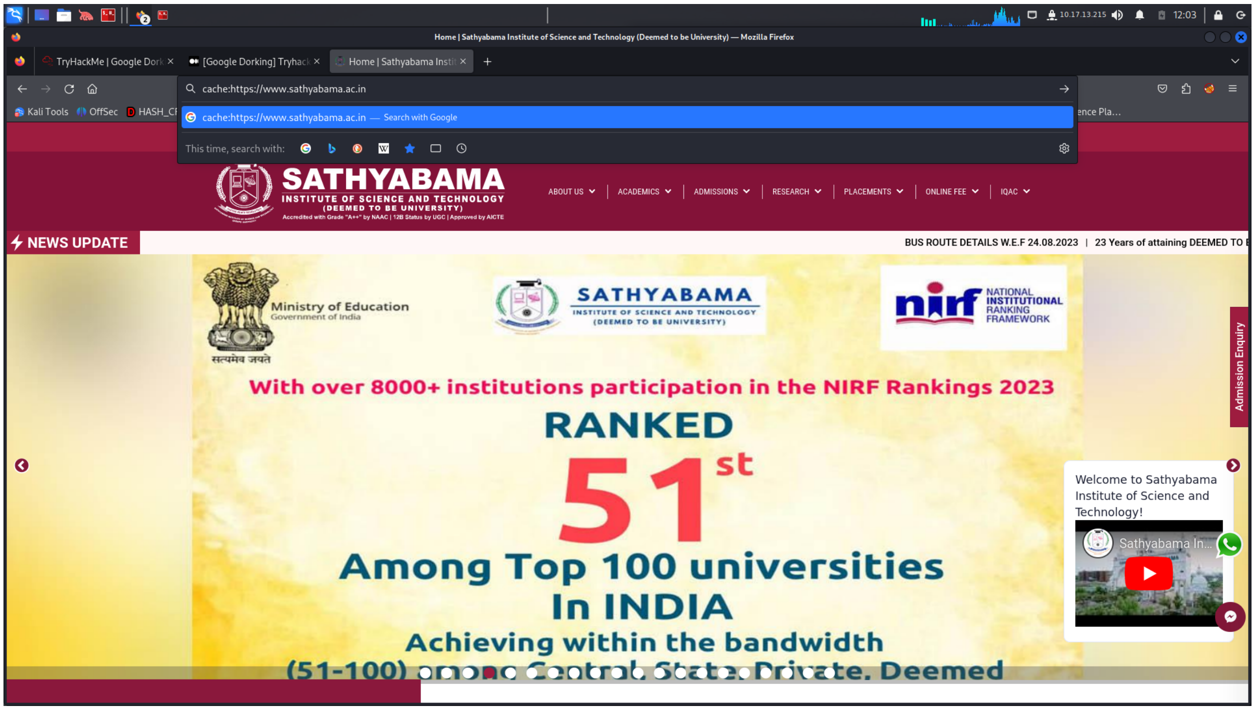
|: this is a logical operator, e.g. "security" "tips" will show all the sites which contain "security" or "tips," or both words.

+: used to concatenate words, useful to detect pages that use more than one specific key, e.g. security + trails

–: minus operator is used to avoiding showing results that contain certain words, e.g. security -trails will show pages that use "security" in their text, but not those that have the word "trails."

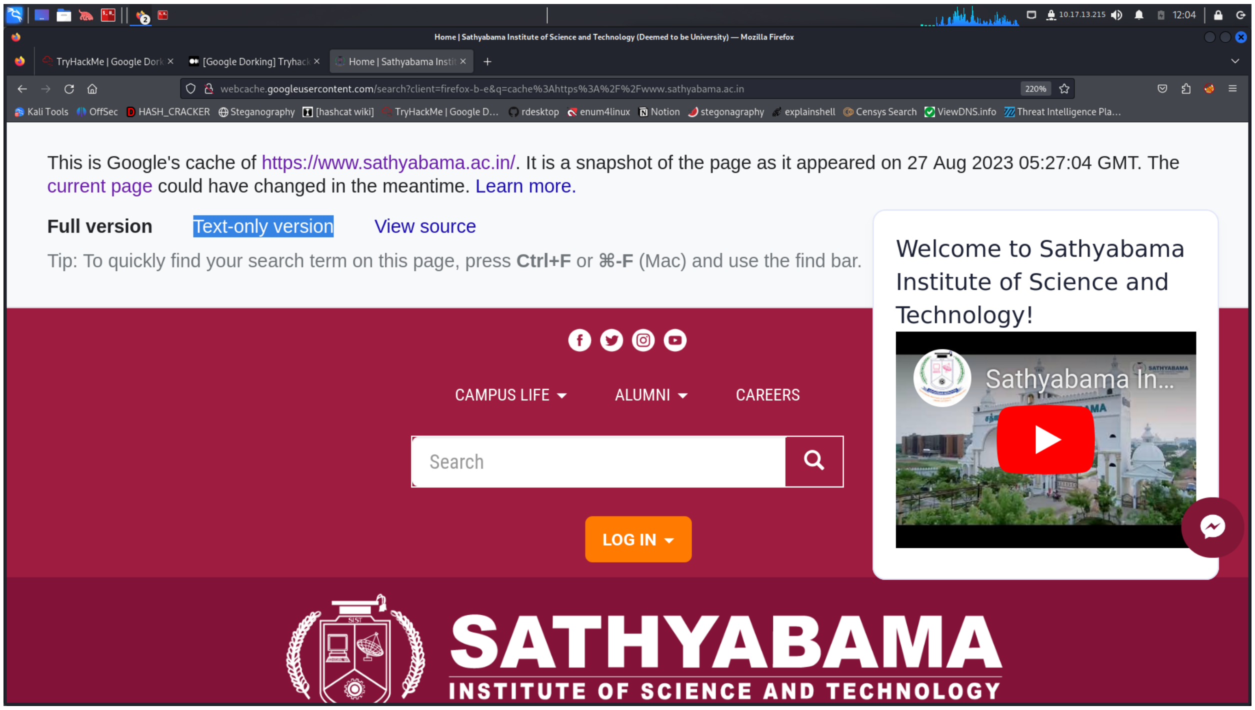
---

my searches


Home | Sathyabama Institute of Science and Technology (Deemed to be University) — Mozilla Firefox
TryHackMe | Google Dork
[Google Dorking] Tryhack
Home | Sathyabama Instit
Kali Tools
OffSec
HASH_CR
This time, search with:
SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited with Grade "A++" by NAAC | 12B Status by UGC | Approved by AICTE
ABOUT US
ACADEMICS
ADMISSIONS
RESEARCH
PLACEMENTS
ONLINE FEE
IQAC
NEWS UPDATE
BUS ROUTE DETAILS W.E.F 24.08.2023 | 23 Years of attaining DEEMED TO
Ministry of Education
Government of India
सत्यमेव जयते
NATIONAL INSTITUTIONAL RANKING FRAMEWORK
With over 8000+ institutions participation in the NIRF Rankings 2023
RANKED
51st
Among Top 100 universities
In INDIA
Achieving within the bandwidth
(51-100) among Central, State, Private, Deemed
Admission Enquiry
Welcome to Sathyabama Institute of Science and Technology!
Sathyabama In...

cache:https://www.sathyabama.ac.in

cache:https://www.sathyabama.ac.in — Search with Google

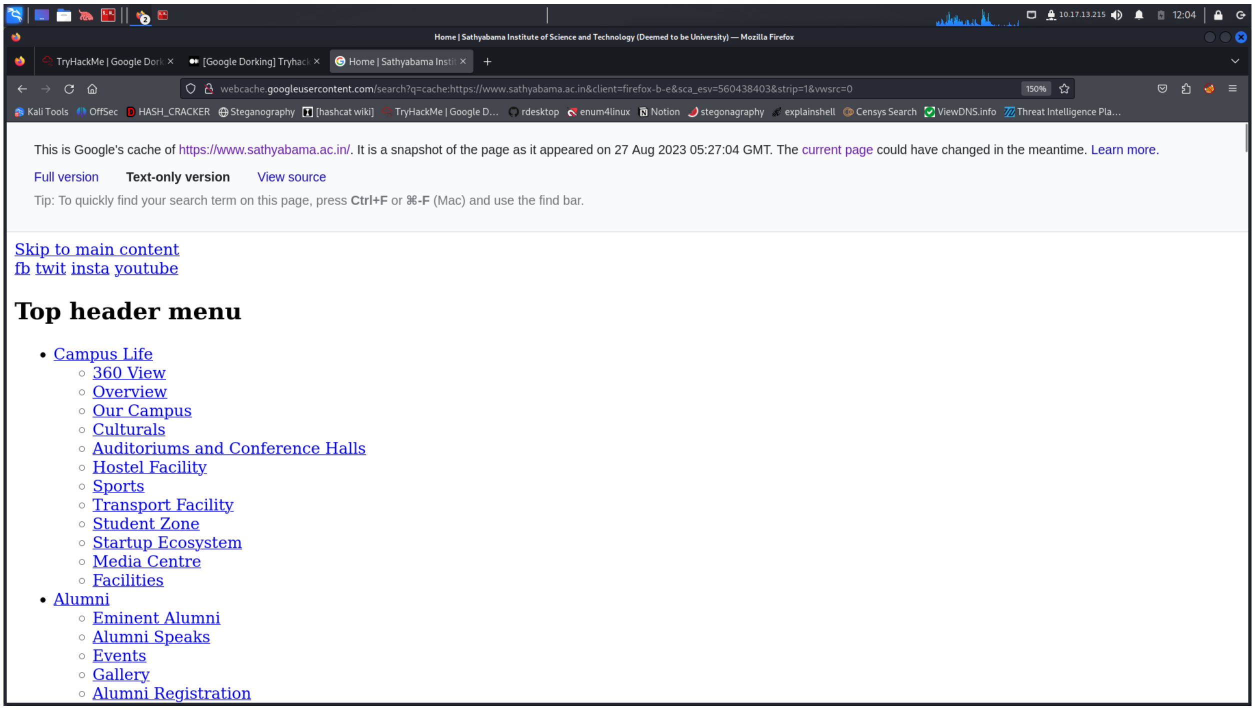


webcache.googleusercontent.com/search?client=firefox-b-e&q=cache%3Ahttps%3A%2F%2Fwww.sathyabama.ac.in

This is Google's cache of https://www.sathyabama.ac.in/. It is a snapshot of the page as it appeared on 27 Aug 2023 05:27:04 GMT. The current page could have changed in the meantime. Learn more.

Full version Text-only version View source

Tip: To quickly find your search term on this page, press **Ctrl+F** or **⌘-F** (Mac) and use the find bar.


10.17.13.215
12:04
Home | Sathyabama Institute of Science and Technology (Deemed to be University) — Mozilla Firefox
TryHackMe | Google Dork
[Google Dorking] Tryhack
Home | Sathyabama Instit
webcache.googleusercontent.com/search?q=cache:https://www.sathyabama.ac.in&client=firefox-b-e&sca_esv=560438403&strip=1&vwsrc=0
150%
Kali Tools
OffSec
HASH_CRACKER
Steganography
[hashcat wiki]
TryHackMe | Google D...
rdesktop
enum4linux
Notion
stegonagraphy
explainshell
Censys Search
ViewDNS.info
Threat Intelligence Pla...

This is Google's cache of https://www.sathyabama.ac.in/. It is a snapshot of the page as it appeared on 27 Aug 2023 05:27:04 GMT. The current page could have changed in the meantime. Learn more.

Full version **Text-only version** View source

Tip: To quickly find your search term on this page, press **Ctrl+F** or **⌘-F** (Mac) and use the find bar.

Skip to main content
fb twit insta youtube

Top header menu

- Campus Life
  - 360 View
  - Overview
  - Our Campus
  - Culturals
  - Auditoriums and Conference Halls
  - Hostel Facility
  - Sports
  - Transport Facility
  - Student Zone
  - Startup Ecosystem
  - Media Centre
  - Facilities
- Alumni
  - Eminent Alumni
  - Alumni Speaks
  - Events
  - Gallery
  - Alumni Registration