**PENTESTING RECONNAISSANCE TOOLS**

Table of Contents

[WIRESHARK 1](#_Toc176875742)

[WHOIS 4](#_Toc176875743)

[HOSTING 6](#_Toc176875744)

[IP2LOCATION 6](#_Toc176875745)

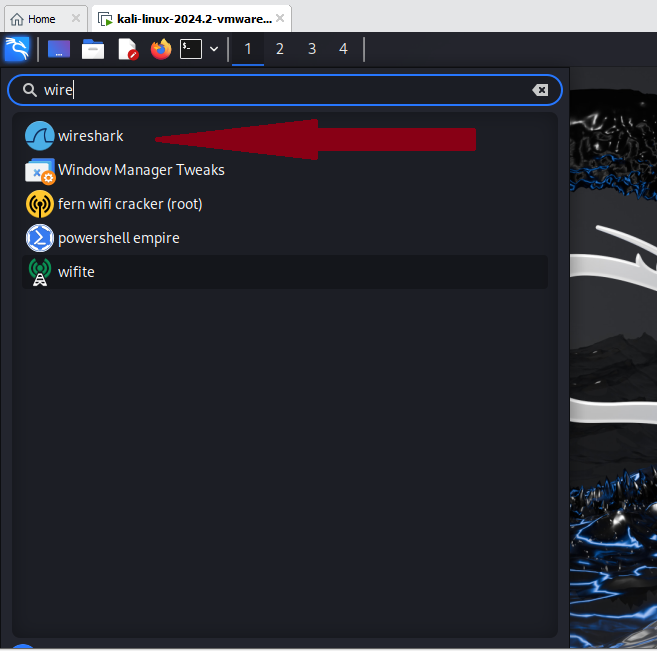
[GOOGLE MAPS 8](#_Toc176875746)

[WAYBACK MACHINE 10](#_Toc176875747)

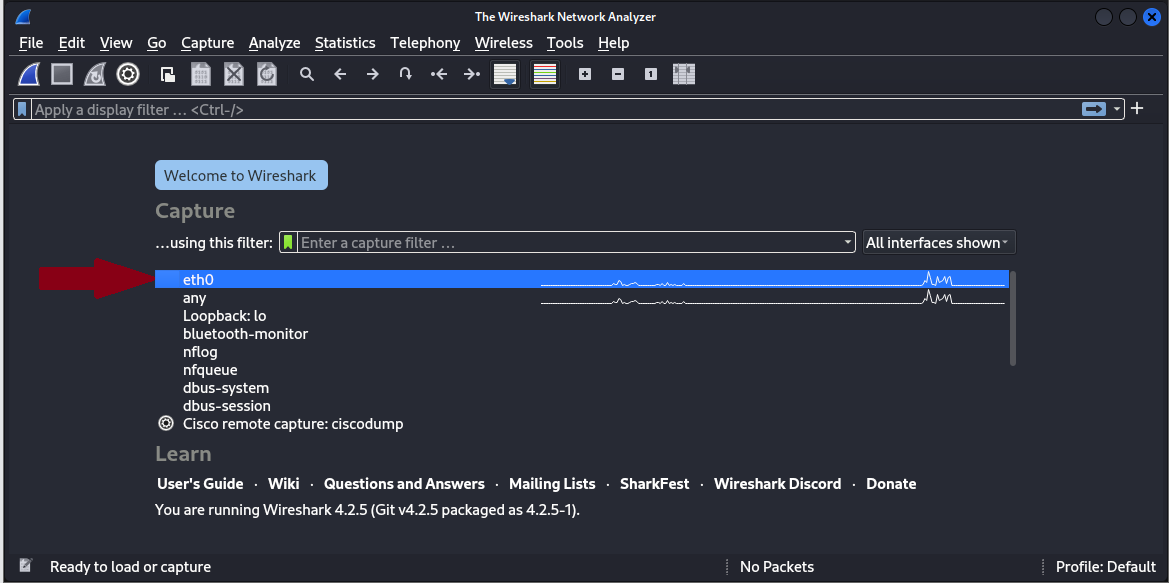
[GOOGLE DORKING 13](#_Toc176875748)

## WIRESHARK

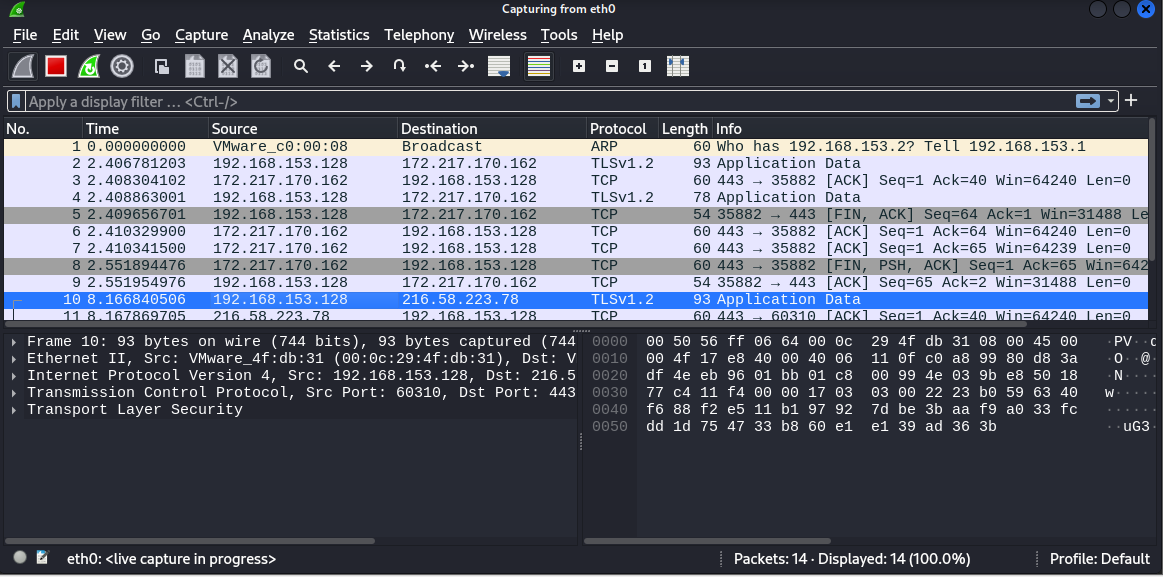
Wireshark is an inbuilt software readily available in Kali Linux Operating Systems (it does not need to be downloaded). Wireshark is an open source network analyzer that can capture and display real-time details of network traffic. It is particularly useful for troubleshooting network issues, analyzing network protocols and ensuring network security. Saved Captures from the Wireshark application use the file extension .pcapng



While using VMware, in order to record/scan activity from your computer, change the Network Adapter setting to **NAT** (Network Address Translation), but if you’re monitoring activity from outside your computer but on the network you are on, Use the Network Adapter setting **Bridged**.



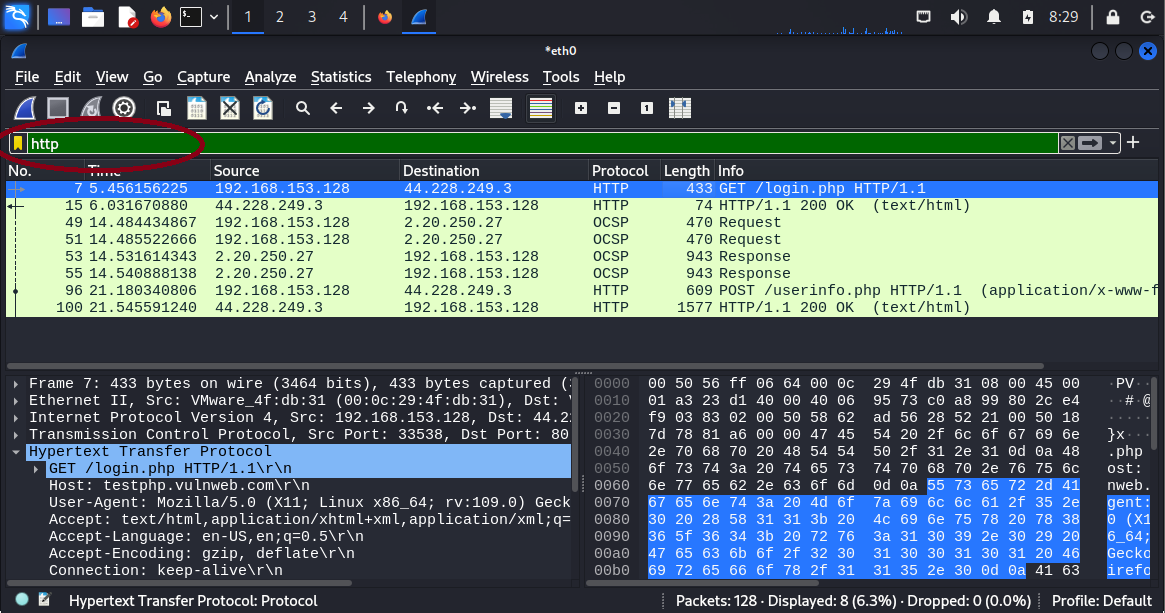
Upon opening Wireshark you will quickly notice “Eth0” is your host machine’s network card. When trying to hack into a Wi-Fi network you will need an external network adapter. (Alfa Card Network Adapter.)



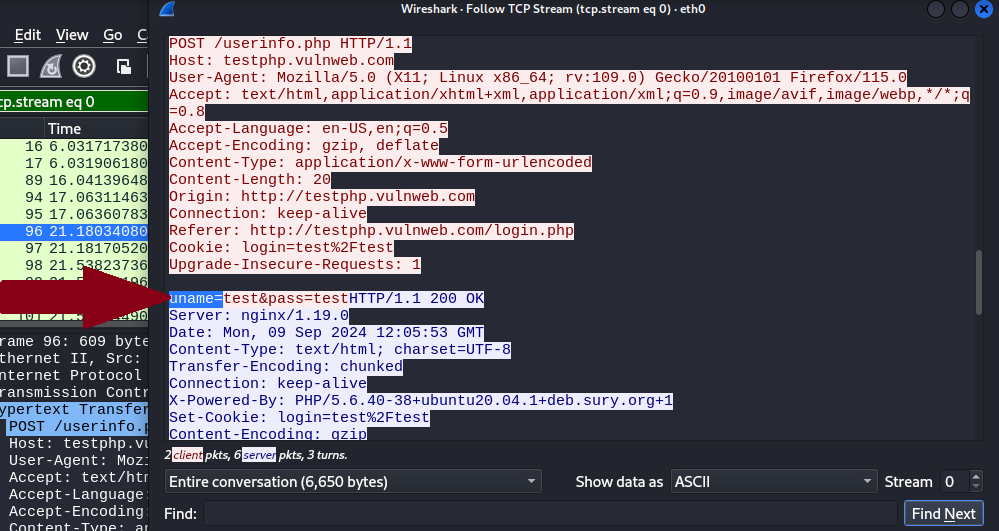
When you click into the “Eth0” adapter, you will be able to monitor the packets being transferred through the network you are currently connected to. For example you can try opening this webpage on your browser <testphp.vulnweb.com/login.php>

This webpage is used by learners to use vulnerability scanners and other tools. So for example you want to capture the information you key in specifically into this website, perform the following:

1. Open the website on your browser.
2. Go to Wireshark and make sure that it is not recording any data or reset and clear any old data, (this step reduces traffic information flowing through. Makes it easy to monitor the exact data we’re looking for rather than sifting through a lot of figures).
3. Go to back to the browser and type in the Username and Password for the page. Click login and observe what happened on Wireshark. Be sure to stop the recording.
4. Filter out the target protocols and observe for further information.
   1. The packets captured are still plenty which is why it is important to know who or what it is that you’re targeting beforehand. Important info may include: The Protocol for example a website using **HTTP/HTTPS** (Simply type it into the filter), the IP Address of the target machine (ip.addr==192.168.0.100).



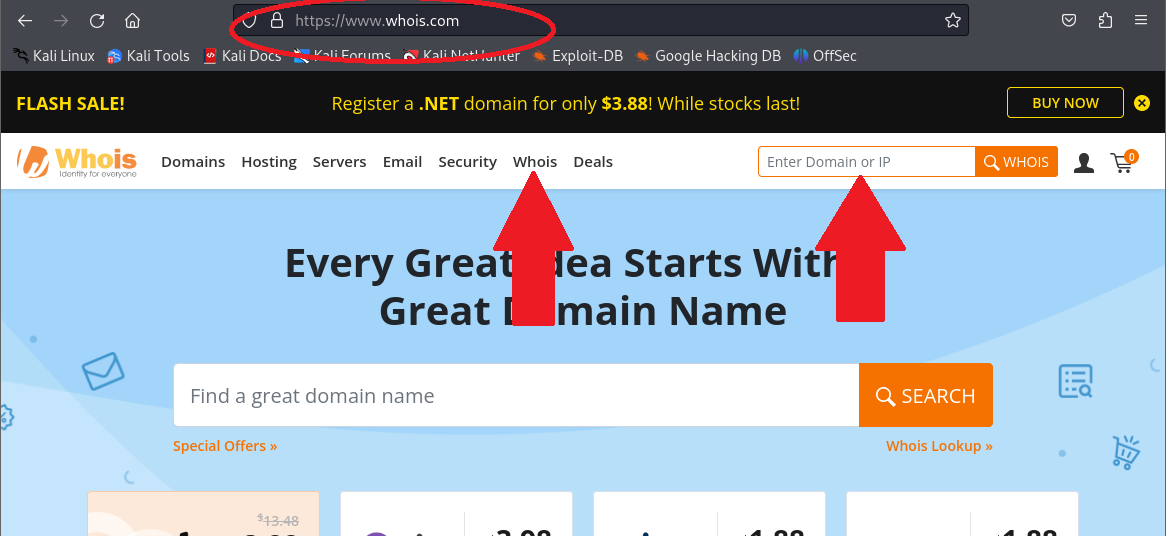
1. For this example, we know that the website we’re monitoring is HTTP, we know that there is a login page and we know that some information was sent through the network to enable the user to login. Now for us to know what exactly what transferred through that network, we need to right click on the exact packet => click on follow => then click on TCP Stream.
   1. **TAKE NOTE:** The information in red is the devices’ sent data packet, while the blue section is the response from the server.
   2. From the image below, the highlighted section will reveal the Username and Password that was input. These are the parameters of the data packet sent from the computer to the server.



## WHOIS

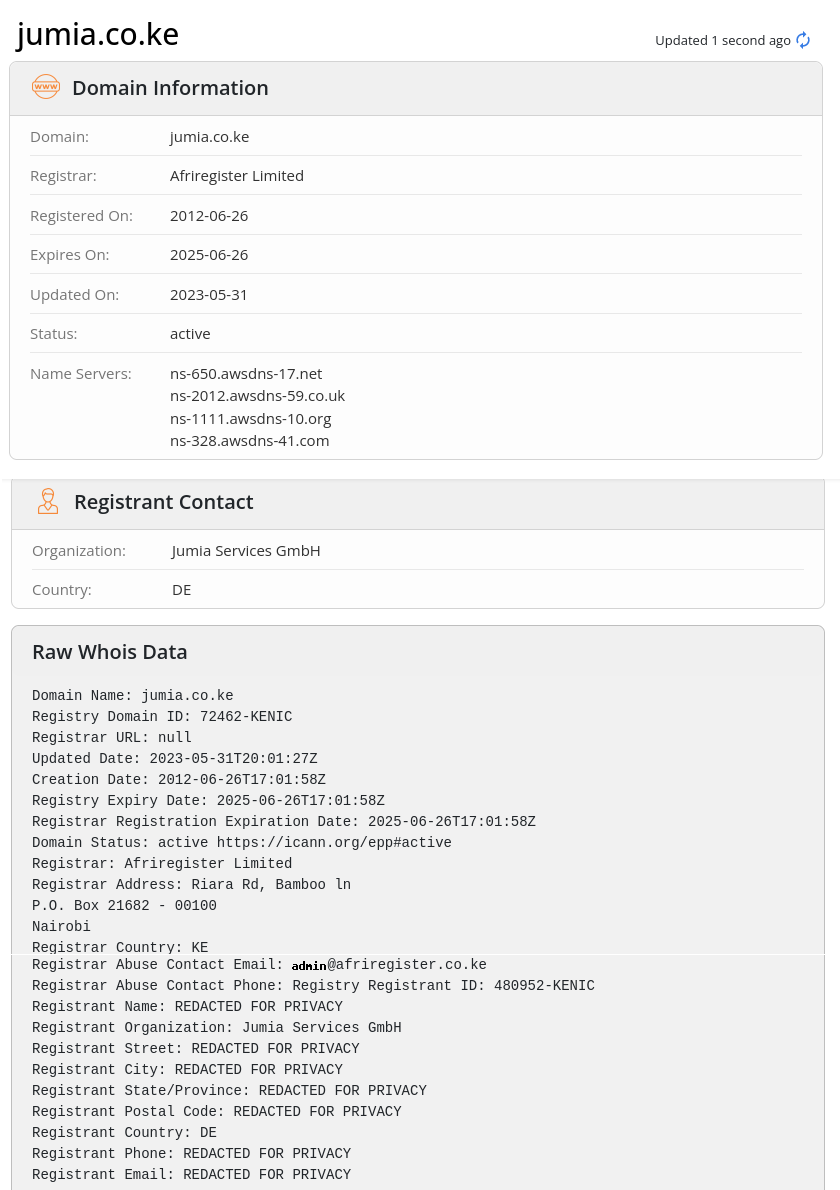
This is a website ([whois.com](https://whois.com/)) used to register domains and keeps track of your registered data when claiming a domain. It allows anyone to do a domain lookup (investigating for more information about a specific website). You’re able to find out, **who owns** a website, whether it’s a company or an individual details provided may include: Name, Email and Phone Number, Country, P.O Box, and Street Info (unless redacted for privacy purposes). You can also find out when the domain was **registered**, when the domain license **expires** and when it was **last updated.**

A similar website is [arin.net](https://arin.net/)



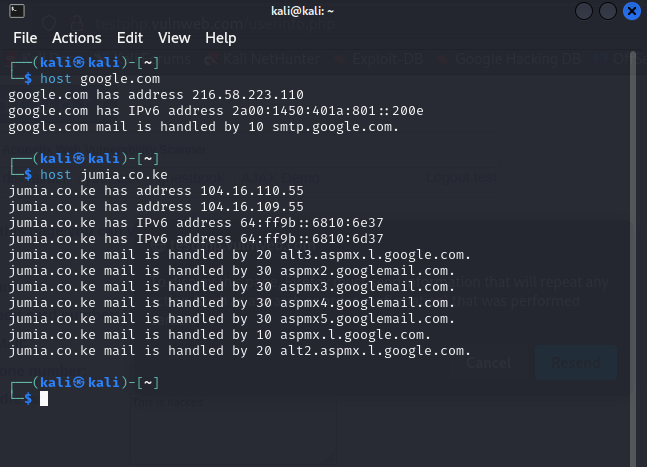
Use the Whois Search

From this point onwards, we will perform a full reconnaissance of Jumia through their website jumia.co.ke. See results below:



## HOSTING

To find out the IP Address of a website open the Linux Terminal => use the host key word e.g. host jumia.co.ke



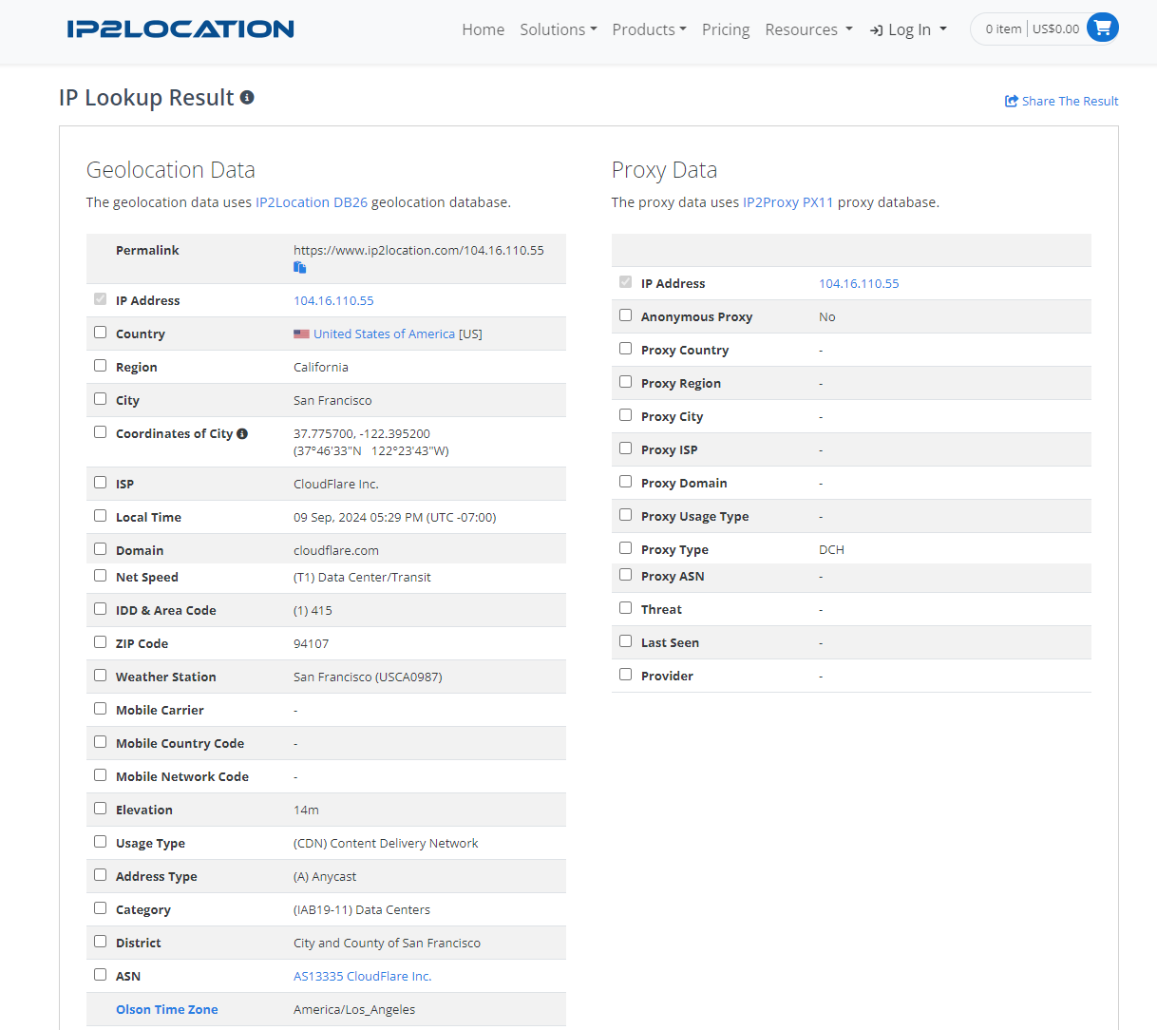
With these IP Addresses you can not only launch a Whois search, but you can now also find out the exact location of the servers using ip2location.

## IP2LOCATION

IP2Location is a non-intrusive website ([ip2location.com](https://ip2location.com)) with IP location lookup technology that retrieves geolocation information with no explicit permission required from users. In simple terms, if you have the IP Address of a hosted website, you can track their location from this website.

Why is this important? Having gathered the IP Address of the target website, it is important to confirm not only the location of the servers but also to perform a further survey of the layout of where the serve is.

Let’s continue with the example of Jumia’s website and find out more about the physical location of the company’s resources.



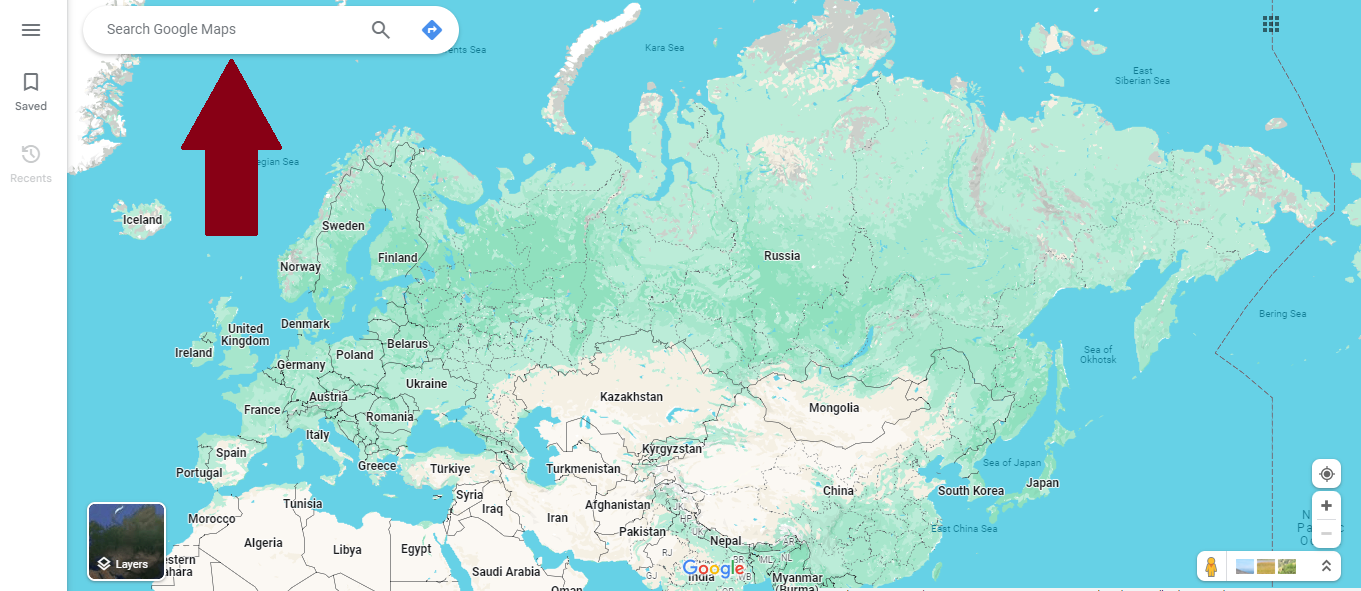
All of this is very important information but take note of the “Coordinates of the City”. Those are Longitude and Latitude showing the exact location of the servers. These can be copied into Google Maps so as to gain a layout of the environment.

## GOOGLE MAPS

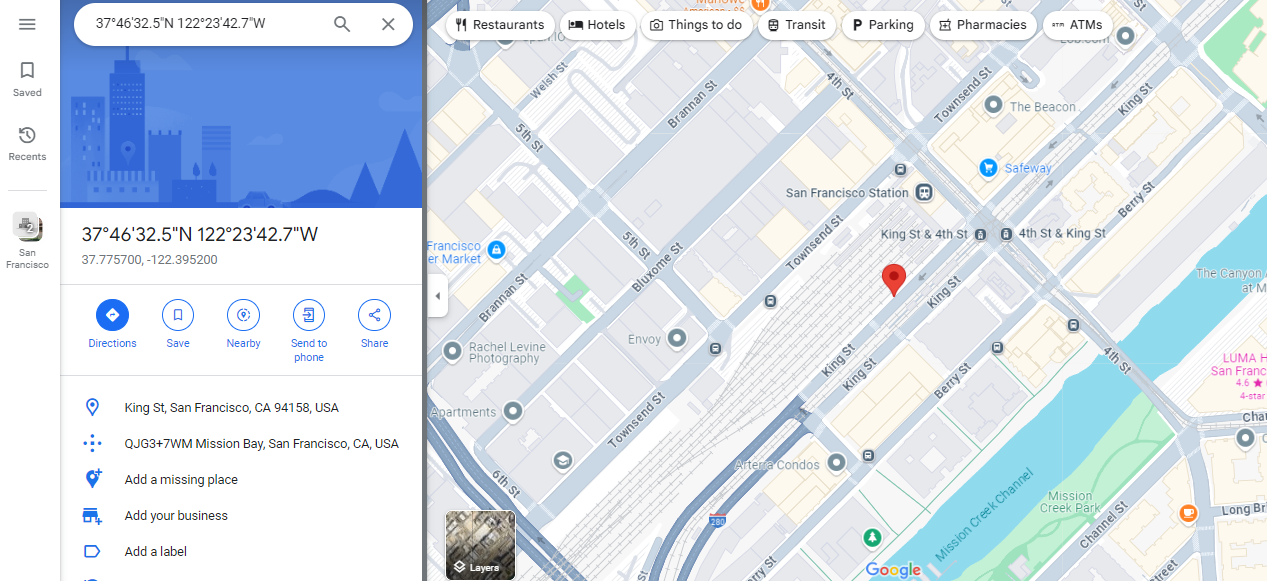
Google Maps is a widely used web mapping service developed by Google. It is valuable tool in reconnaissance because it provides, site assessment, route planning if you’re going to visit the place physically, has street view for casual inspection of buildings and provides a deeper understanding of the geographical layout and any local amenities and business that may be relevant to your mission. Visit [maps.google.com](https://maps.google.com) to explore the various tools.

Copy the coordinates given from the ip2location website. Take note that the coordinates are split into two types, either one of them will do.

Paste the coordinates to the search box on Google Maps.



If you are unfamiliar with the location it may be difficult to determine which exact building you’re targeting, but you can quickly survey the area around for CCTV cameras, police posts, crowded areas / tourist hotspots and so much more.



Above is the location of the Jumia Servers in San Francisco California.

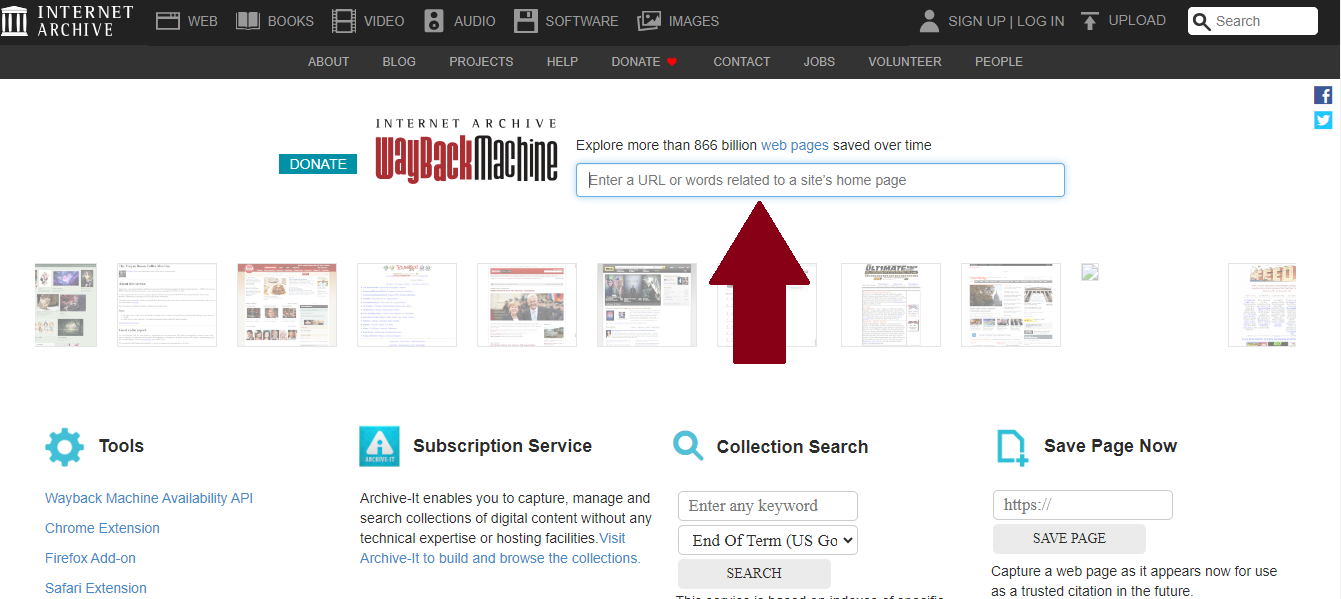


Using the street view tool, this is the supposed building that may contain the servers. One thing’s for sure… They’re in here somewhere.

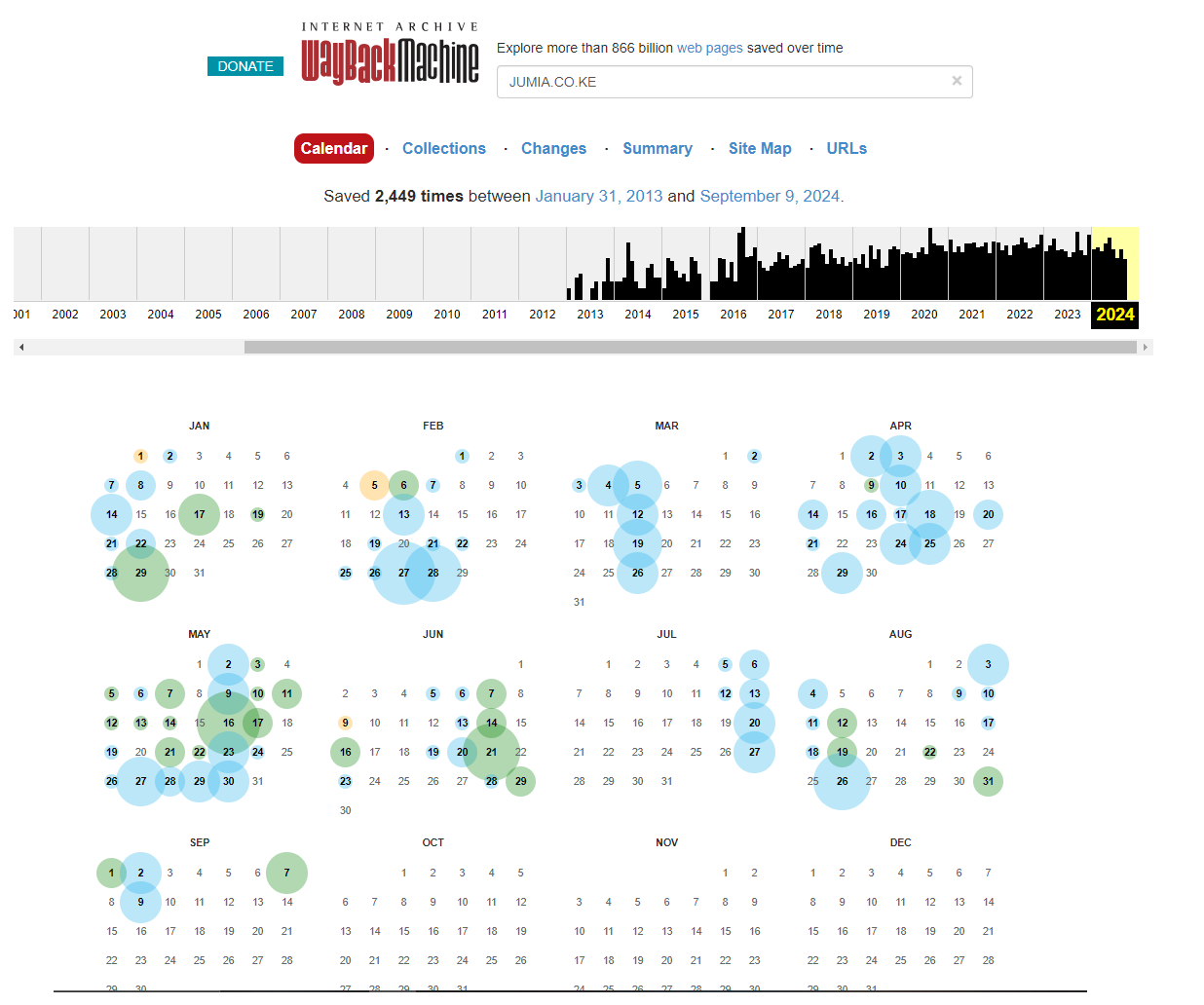
## WAYBACK MACHINE

The Wayback Machine is a digital archive of the Internet. It allows users to view archived versions of web pages dating back to the early days of the internet. Users can select different dates to view how a website looked at various points in time, allowing them to see changes and historical content.

As seen in the photo below, this is the [Wayback Machine (web.archive.org)](https://web.archive.org/)



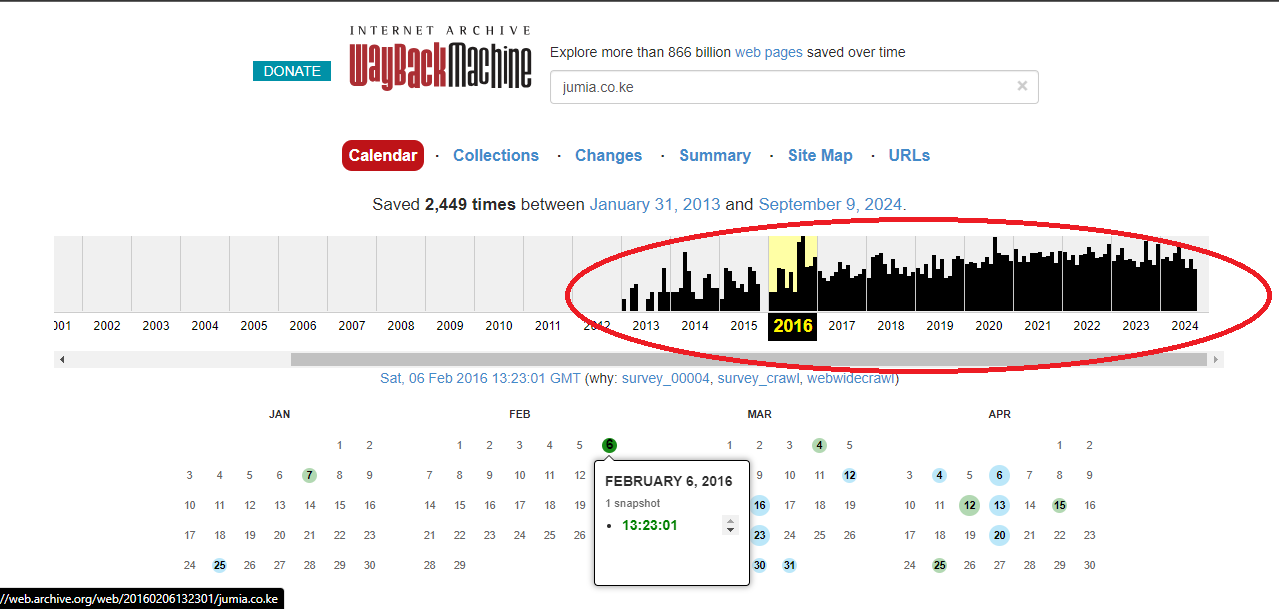
You can write down any webpage to try and explore the archived data (if any). For example we’re going to continue using the Jumia webpage and find out more about it



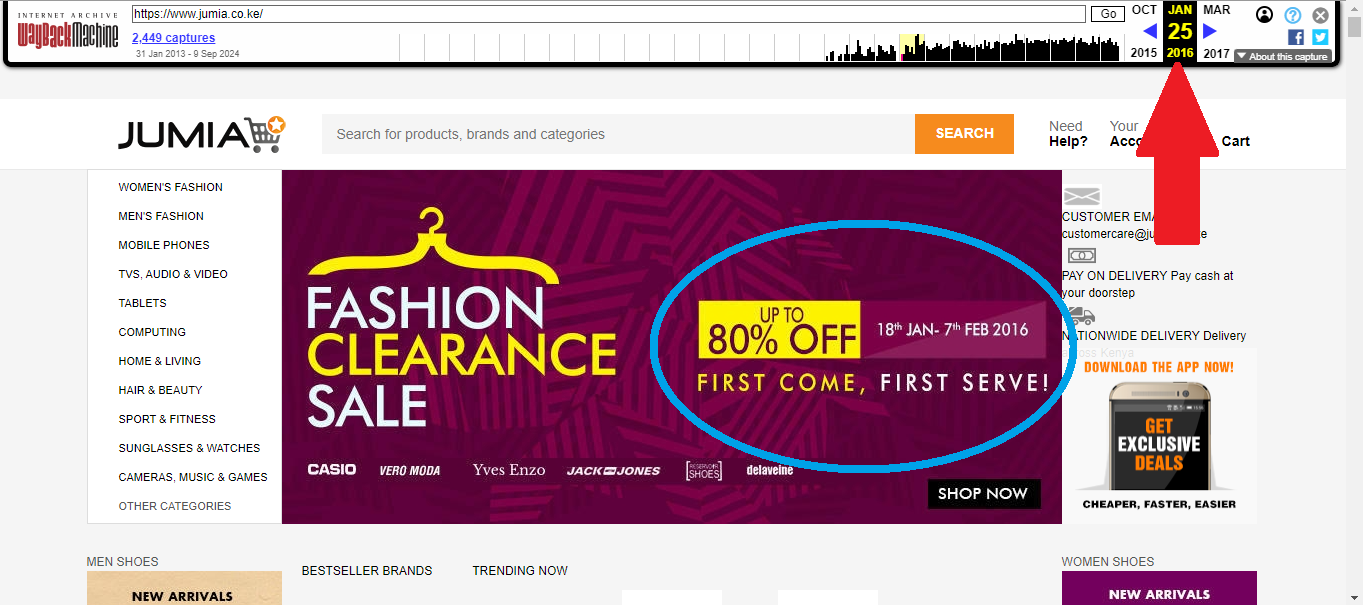
From this results page, we can take note of the following things:

* There were 2,449 captures / updates that have been archived in this website.
* The approximate date when the website was first launched / first updated (31st January 2013).
* The last update made to the website (9th September 2013).
* The frequency at which the website is being updated. From the screenshot above we can tell that there were many updates around April and May, probably in preparation for the “Back to School” discounts that guarantees many visitors to their sites.
* The various colours have meanings as well: Blue is a good result; the full page was archived without error, green means a redirect, orange represents a client error and red denotes a server error.

Take note of this highlighted timeline. This is the tool that will allow you to “travel back in time” specifically in relation to the website you are searching. You can click anywhere on the time line, in this example we clicked on 2016. Upon visiting that year, you will notice there are captures that were archived, for example this one on the 6th of February 2016.



Upon visiting the capture, you will get the website to open like in the screenshot below. You can confirm the exact date you have travelled to from the top right hand corner. The website itself will take time to load but usually you will get some kind of content to confirm that it is successful, such as what I have highlighted in blue (The discount offers at that point in time).



## GOOGLE DORKING

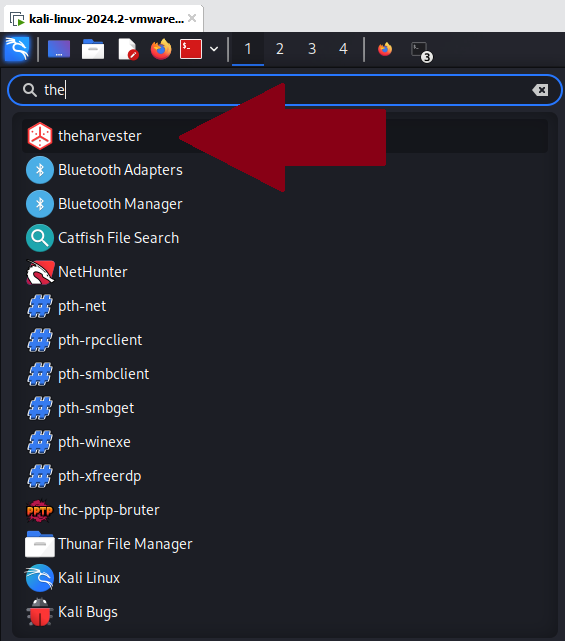
A Google Dork is a search query that looks for specific information on Google’s search engine. It uses advanced search filters that allow to retrieve more efficient results. It is a technique often used by cybersecurity professionals in order to find valuable information about a target. They can be used to find information that you didn’t even know existed.

Below is an example of Search Operators you can try.

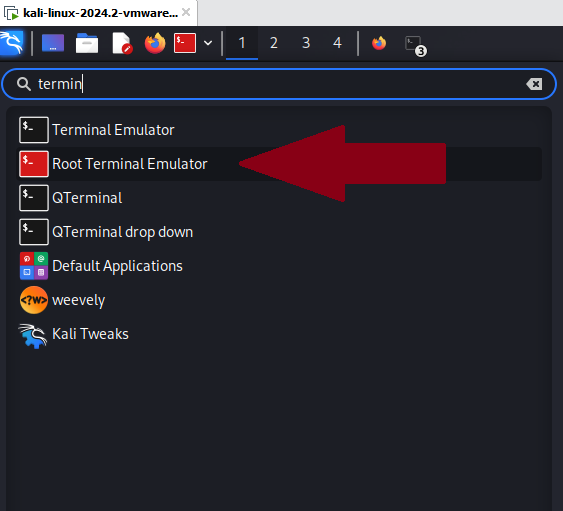
| **Operator** | **Description** | **Syntax** | **Example** |
| --- | --- | --- | --- |
| () | Group multiple terms or operators. Allows advanced expressions | (<term> or <operator>) | inurl:(html | php) |
| \* | Wildcard. Matches any word | <text> \* <text> | How to \* a computer |
| "" | The given keyword has to match exactly. *case-insensitive* | "<keywords>" | "google" |
| m..n / m...n | Search for a range of numbers. *n* should be greater than *m* | <number>..<number> | 1..100 |
| - | Documents that match the operator are excluded. *NOT*-Operator | -<operator> | -site:youtube.com |
| + | Include documents that match the operator | +<operator> | +site:youtube.com |
| | | Logical *OR-Operator*. Only one operator needs to match in order for the overall expression to match | <operator> | <operator> | "google" | "yahoo" |
| ~ | Search for synonyms of the given word. Not supported by Google | ~<word> | ~book |
| @ | Perform a search only on the given social media platform. Rather use **site** | @<socialmedia> | @instagram |
| after | Search for documents published / indexed after the given date | after:<yy(-mm-dd)> | after:2020-06-03 |
| allintitle | Same as **intitle** but allows multiple keywords seperated by a space | allintitle:<keywords> | allintitle:dog cat |
| allinurl | Same as **inurl** but allows multiple keywords seperated by a space | allinurl:<keywords> | allinurl:search com |
| allintext | Same as **intext** but allows multiple keywords seperated by a space | allintext:<keywords> | allintext:math science university |
| AROUND | Search for documents in which the first word is up to *n* words away from the second word and vice versa | <word1> AROUND(<n>) <word2> | google AROUND(10) good |
| author | Search for articles written by the given author if applicable | author:<name> | author:Max |
| before | Search for documents published / indexed before the given date | before:<yy(-mm-dd)> | before:2020-06-03 |
| cache | Search on the cached version of the given website. Uses Google's cache to do so | cache:<domain> | cache:google.com |
| contains | Search for documents that link to the given fileype. Not supported by Google | contains:<filetype> | contains:pdf |
| date | Search for documents published within the past *n* months. Not supported by Google | date:<number> | date:3 |
| define | Search for the definition of the given word | define:<word> | define:funny |
| ext | Search for a specific filetype | ext:<documenttype> | ext:pdf |
| filetype | Refer to **ext** | filetype:<documenttype> | filetype:pdf |
| inanchor | Search for the given keyword in a website's anchors | inanchor:<keyword> | inanchor:security |
| index of | Search for documents containing direct downloads | index of:<term> | index of:mp4 videos |
| info | Search for information about a website | info:<domain> | info:google.com |
| intext | Keyword needs to be in the text of the document | intext:<keyword> | intext:news |
| intitle | Keyword needs to be in the title of the document | intitle:<keyword> | intitle:money |
| inurl | Keyword needs to be in the URL of the document | inurl:<keyword> | inurl:sheet |
| link / links | Search for documents whose links contain the given keyword. Useful for finding documents that link to a specific website | link:<keyword> | link:google |
| location | Show documents based on the given location | location:<location> | location:USA |
| numrange | Refer to **m..n** | numrange:<number>-<number> | numrange:1-100 |
| OR | Refer to **|** | <operator> OR <operator> | "google" OR "yahoo" |
| phonebook | Search for related phone numbers associated with the given name | phonebook:<name> | phonebook:"william smith" |
| relate / related | Search for documents that are related to the given website | relate:<domain> | relate:google.com |
| safesearch | Exclude adult content such as pornographic videos | safesearch:<keyword> | safesearch:sex |
| source | Search on a specific news site. Rather use **site** | source:<news> | source:theguardian |
| site | Search on the given site. Given argument might also be just a TLD such as **com, net**, etc | site:<domain> | site:google.com |
| stock | Search for information about a market stock | stock:<stock> | stock:dax |
| weather | Search for information about the weather of the given location | weather:<location> | weather:Miami |

## the harvester

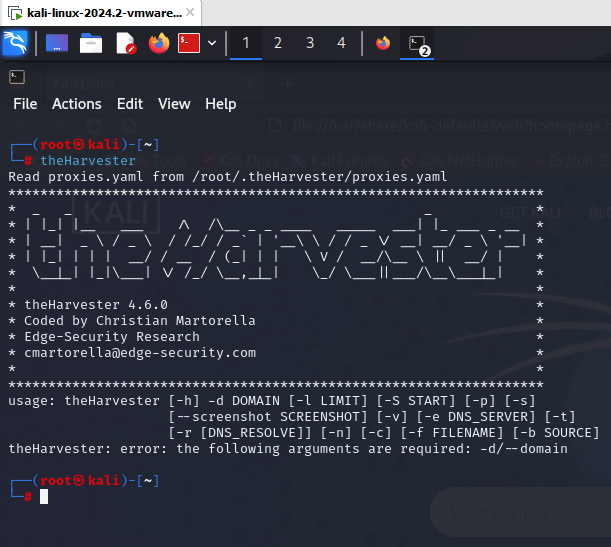
What is it?



T



T

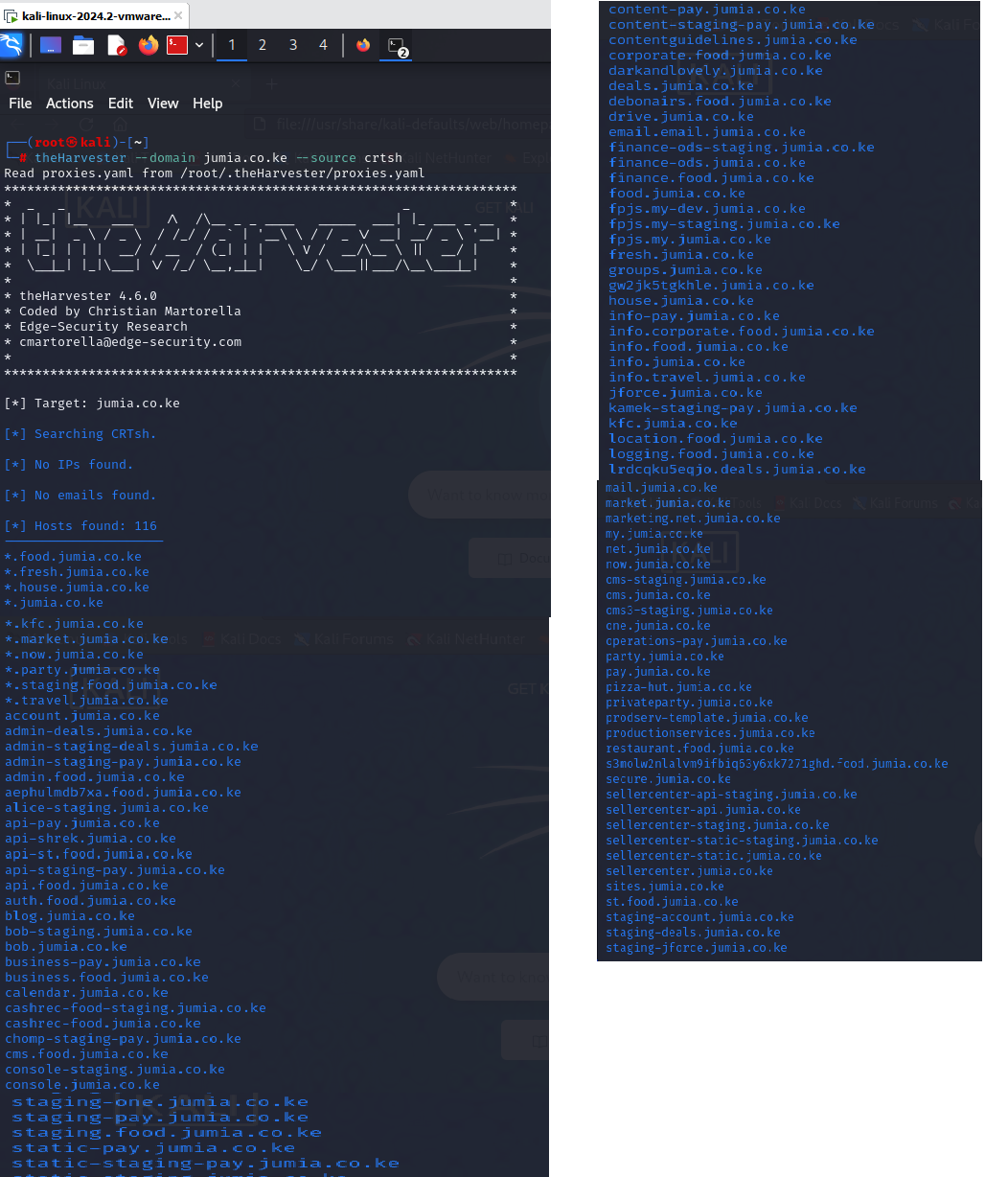


R

R



R



## BUILTWITH

This is a website [builtwith.com](http://www.builtwith.com)

Tells users what technology a website is created