Privacy Meta Search Engine

Mr.A.V.Santhosh babu Prof/CSE, S Sowbaranika, C. Premkumar, P. Kaviarasan,
Dept of CSE, Erode Sengunthar Engineering College,
Erode, Tamil Nadu

ABSTRACT

One of the most common online activities all the people in the world do is using search engines to surf through the internet. Additionally, consumers typically get good results and have a lot of faith in search engines' abilities. This paper's main goal is to offer a template for browsing the web and gathering pertinent data at the user's request. When a user submits a search query, using multiple search engines can occasionally produce different search results on the one hand, and a wealth of quality and quantity of results on the other. Additionally, by developing unique search profiles for each system search, the system gives users anonymity.

Keywords:*-Search Engine, Quality and Quantity,*

I. INTRODUCTION

A dynamic information source, the World Wide Web is always expanding. The exponential growth of the Web is outpacing the capabilities of individual search engines.

Meta-search engines are used as a solution to the problem of individual search engines' low information coverage rate on the web.

A meta-search engine is a sort of search tool that enables simultaneous access to many search engines by sending user queries to the underlying search engines on demand. The user is given a single, comprehensive result list when the results have been compiled.

According to industry experts, the quality of the pertinent search results obtained from a meta-search engine has improved dramatically.

We suggest a new model to improve these meta-search engines by precisely duplicating the search and optimizing the results, breaking down a result into multiple search results until the research search is exhausted, and using docker to create unique search profiles for each search, thereby improving the meta-search engines' performance and safeguarding users' privacy.

We are making a new model of the meta search engine for the better privacy and quality of the searches by hiding all the user data and their

searches from the search engine providers and making the meta search engine more secure and safe for surfing in them.

And also increasing the quality and the quantity of the search also without making any difference in them. A meta search engine is a tool that gathers information by combining the outcomes of many search engines. The results are ranked according to the meta search engine's specifications, after which they are combined into a single consolidated rank list. The gathering of results and the ranking process are key stages that are significant in comparing any two meta search engines.

It is an essential process that affects how well a meta search engine performs.The rank assigned by the underlying search engines was the main factor taken into account by the earlier created meta search engine models, and not many additional factors. The basic models lacked the ability to manage the dynamic nature of the web. This served as the primary inspiration for a revolutionary strategy that may be tailored to the dynamic nature of the web.

The suggested method employs the Best Worst Method (BWM) to weigh criteria and the Genetic Algorithm with Cluster Dependent Ordered Weighted Average (DOWA) as the fitness function to produce a list of documents that are ordered by relevance. Comparing this method to the current models, it is more precise.

II. META SEARCH ENGINE

A user submits a search using a user-friendly interface during a typical session utilising a metasearch engine. The user's query is subsequently forwarded by the Metasearch engine to several underlying search engines. A user presents a query to several information resources during the process of metasearching, which is also referred to as integrated searching, simultaneous searching, cross-database searching, parallel searching, broadcast searching, and federated searching.

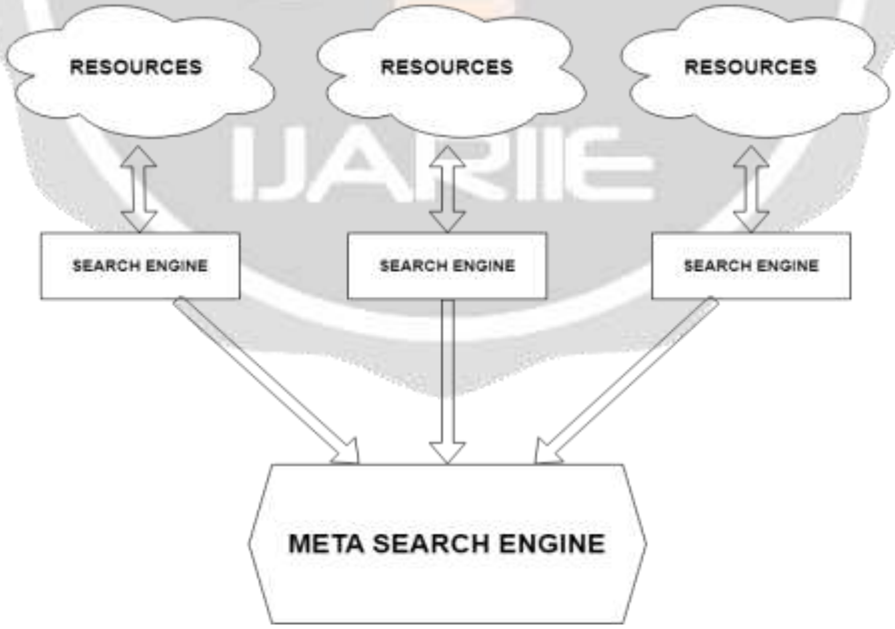
The resources may differ in various ways, including their location, the way they provide their information, the technology they use, the kinds of items they have, and more. Systems for metasearch rely on just-in-time processing. The system doesn't save any local indexes of its information landscape; it just accesses the various resources to get the information when it is needed.

2.1. THE META SEARCH ENGINE METHOD:

High-ranking web pages on numerous search engines are probably more relevant in terms of offering helpful information. However, each website is given a ranking score by each search engine, and these ratings are frequently different.

Because different criteria and scoring systems are prioritized by different search engines, a website may look highly rated on one search engine while being weakly placed on another. This is a concern since credible accounts are generated by metasearch engines, which rely substantially on the constancy of this data.

2.1.1. The Meta Search Engine Methodology



2.1.2. Docker:

Any Linux, Windows, or macOS computer may execute a virtual container that contains a program and all of its dependencies thanks to Docker.

This makes it possible for the program to run anywhere, including on-premises, in a public (see decentralized computing, distributed computing, and cloud computing) or private cloud. When running on Linux, Docker makes use of the kernel's resource isolation capabilities like cgroups and kernel namespaces as well as a union-capable file system like OverlayFS to enable containers to run within a single Linux instance without the need to launch and maintain virtual machines. The containers are run by a Linux virtual machine using Docker on macOS. To make the search profile different every time the docker helps in that.

.

2.1.3. The Searx Framework:

Searx is a free and open-source metasearch system and all the searx distributions was given the permission by the GNU. In order to accomplish this, Searx does not divulge users' IP addresses or search histories to the search engines from whom it receives results. The search engines do not serve tracking cookies, restricting the personalization of results based on user profiles. SearX searches are by default submitted using HTTP POST in order to prevent user-provided query keywords from appearing in web server logs. Searx organization was motivated by the Seeks idea, although not utilizing the peer-to-peer user-sourced results rating feature.

2.1.4. The Docker Compose:

A tool for creating and operating multi-container Docker applications is called Compose. The services of your application are configured using Compose using a YAML file. You then create and start each service from your setup with a single command.

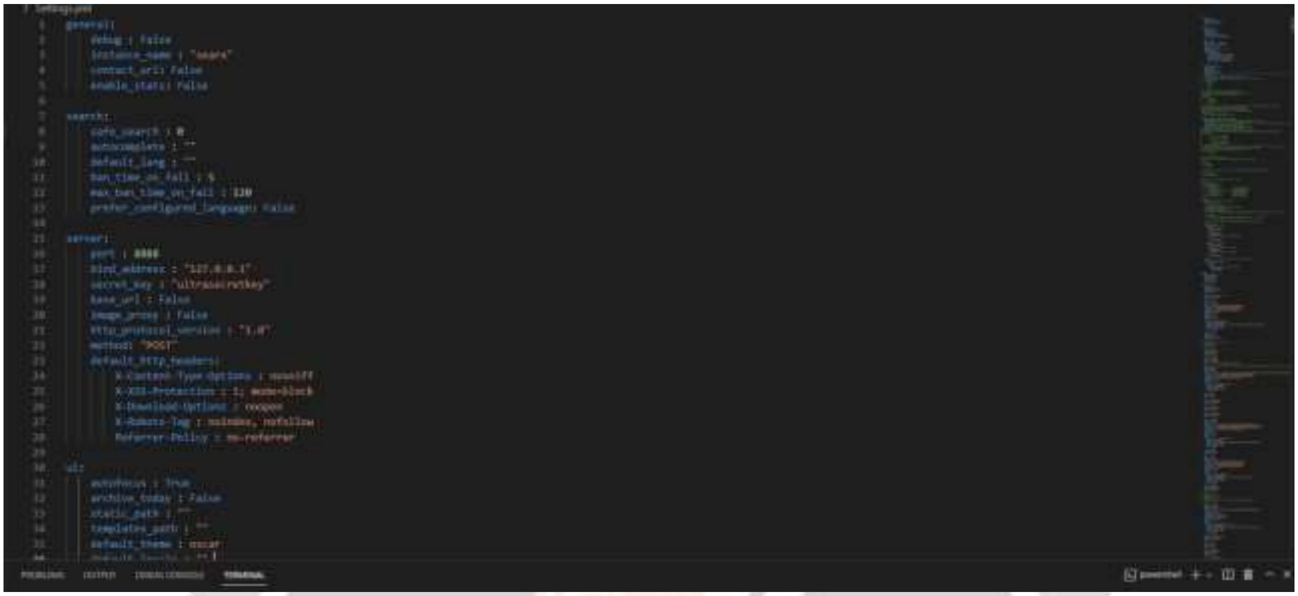
Compose functions in all contexts, including CI workflows, production, staging, development, and testing. Additionally, it offers instructions for controlling your application's entire lifecycle

III. SETTINGS

```
general:
    debug : False # Debug mode, only for development
    instance_name : "OurSearch"
    git_url: false
    git_branch: master
    issue_url: false
    docs_url: false
    public_instances: https://oursearch.in
    contact_url: False
    wiki_url: https://github.com/oursearch


search:
    safe_search : 2
    autocomplete : "duckduckgo"
    default_lang : "eng"
    ban_time_on_fail : 5
    max_ban_time_on_fail : 120
    prefer_configured_language: False

server:
    port : 8888
    bind_address : "localhost"
    secret_key : "Change this key"
    base_url : https://oursearch.in
    image_proxy : False
    default_locale : "google"
    default_theme :light
    default_http_headers:

ui:
    autofocus : True
    archive_today : False
    static_path : ""
    templates_path : ""
    default_theme : oscar
    default_locale : ""
    theme_args :
        oscar_style : negative

 outgoing:
    request_timeout : 2.0
    max_request_timeout: 10.0

checker:

    off_when_debug: True


   name : google
  engine : chrome
  shortcut : go
  base_url : 'https://google.com/'
  categories : general
  timeout : 2.0
  api_key : 'Enter Api key’'
  disabled : True
  language : en_US


      android: &test_android
        matrix:
          query: ['ios']
          lang: ['en', 'de', 'zh-CN']
        result_container:
          - not_empty
          - ['one_title_contains', 'google']
        test:
          - unique_results
engines:
  - name: google
    engine: google
    timeout: 4.0
    shortcut: goo
    disabled: True
```

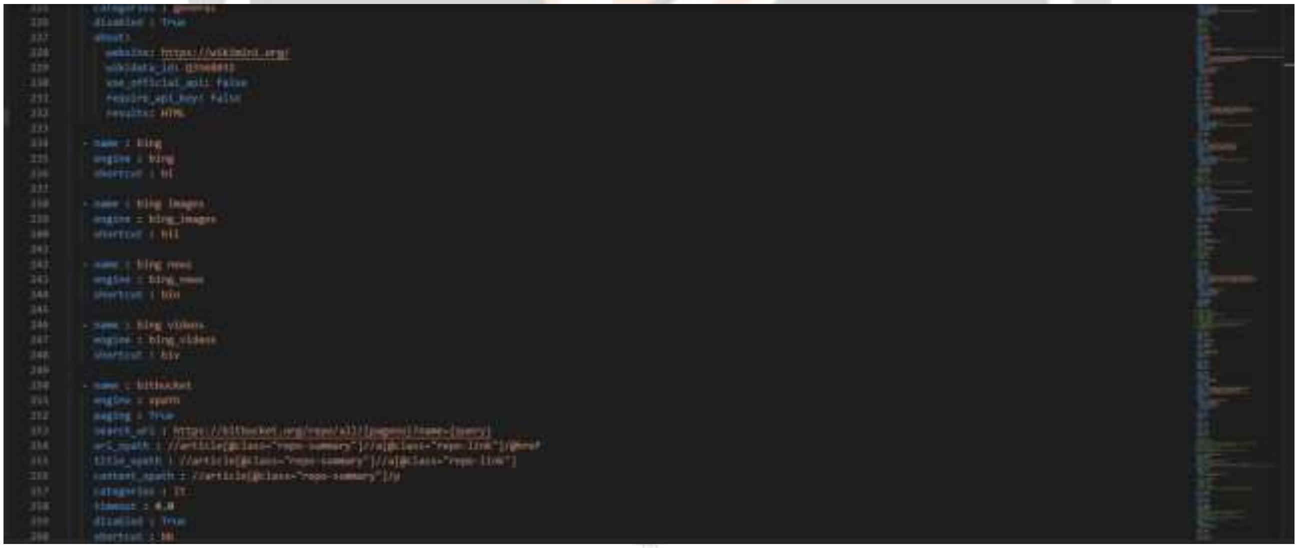
```
# Requires Tor
  - name : torconfig
    engine : ahmia
    categories : onions
    enable_http : True
    shortcut : a

  - name : arch  wiki
    engine : archlinux
    shortcut : al

  - name : archive is
    engine : xpath
    search_url : https://archive.is/search/?q={query}
    url_xpath : (//div[@class="TEXT-BLOCK"]/a)/@href
    title_xpath : false
    content_xpath : false
    categories : general
    timeout : 7.0
    disabled : True
    shortcut : ai
    soft_max_redirects: 1
    about:
      website: https://archive.is/
      wikidata_id: Q13515725
      official_api_documentation: false      use_official_api: false
      require_api_key: false
      results: HTML
```

IV. RESULT



```
general:
    debug : false
    instance_name : "searx"
    contact_url: false
    enable_stats: false

search:
    safe_search : 0
    autocomplete : ""
    default_lang : ""
    ban_time_on_fail : 5
    max_ban_time_on_fail : 120
    prefer_configured_language: false

server:
    port : 8888
    bind_address : "127.0.0.1"
    secret_key : "ultrasecretkey"
    base_url : False
    image_proxy : False
    http_protocol_version : "1.0"
    method: "POST"
    default_http_headers:
        X-Content-Type-Options : nosniff
        X-XSS-Protection : 1; mode=block
        X-Download-Options : noopen
        X-Robots-Tag : noindex, nofollow
        Referrer-Policy : no-referrer

ui:
    autofocus : True
    archive_today : False
    static_path : ""
    templates_path : ""
    default_theme : oscar
```

Figure 1: Network configuration



```
     categories : general
     disabled : True
     about:
       website: https://wikimini.org/
       wikidata_id: Q3568032
       use_official_api: false
       require_api_key: false
       results: HTML

   - name : bing
     engine : bing
     shortcut : bi

   - name : bing images
     engine : bing_images
     shortcut : bii

   - name : bing news
     engine : bing_news
     shortcut : bin

   - name : bing videos
     engine : bing_videos
     shortcut : biv

   - name : bitbucket
     engine : xpath
     paging : True
     search_url : https://bitbucket.org/repo/all/{pageno}?name={query}
     url_xpath : //article[@class="repo-summary"]//a[@class="repo-link"]/@href
     title_xpath : //article[@class="repo-summary"]//a[@class="repo-link"]
     content_xpath : //article[@class="repo-summary"]/p
     categories : it
     timeout : 4.0
     disabled : True
     shortcut : bb
```

Figure 2: Search Engine Configuration

V. CONCLUSION

After a thorough examination of the component's advantages, disadvantages, and limitations, it is safe to conclude that the product is extremely efficient. This program runs smoothly and complies with every search. There have been numerous instances of technical difficulties, content mistakes, and most significantly adjusted weights for features. So, it's possible that the aforementioned software will be enhanced in the near future to offer more dynamic weight adjustment and security.

With this search engine, the user's privacy is secure and they won't see any personally relevant advertisements. Also, the search engine is simply modifiable, and new search engines may be incorporated right away.

REFERENCES

[1]. "In-Depth Meta-Search Engine" by Belabed Mourad Dennai Abdeslam. “Published in the 2021 Tebessa in Algeria in the International Conference on Information Systems and Advanced Technologies”.

[2]. "MetaFusion: An effective meta search engine employing evolutionary algorithms," Daya Gupta and Devika Singh. Eighth International Conference on Modern Computing, Noida, India, 2016, published.

[3]. Rishi Soni, Shashi Kant Gupta, and Brijesh Kumar Chaurasia. 2011 International Conference on Computational Intelligence and Communication Networks, "Meta Search Engine," Gwalior, India

[4]. B. R. Kishore, V. R. Amogh, S. Vrushab, and C. H. Akshay "Using a Genetic Algorithm to Improve a Metan Search Engine's Coherence." In: Tirunelveli, India, 2022, 4th International Conference on Smart Systems and Inventive Technologies.

[5]. Juan Tang; Ya-Jun Du; Ke-Liang Wang 2017. 2007 International Conference on Machine Learning and Cybernetics, "Design and Development of a Bespoke Meta-Search Engine Utilizing FCA's Arabic-English Parallel Corpus Algorithm"

[6]. “Metasearch engine result optimization using the reformed methods of genetic algorithm”,Published in the 9th International Conference on Computer and Knowledge Engineering, Mashhad, Iran, Somayeh Adeli and Mohammad Pourmood Aghababa, 2019.