

Players Research and Information Searching Model.

Table of Contents:

Description:	Page Numbers:
1) Abstract.	2
2) Introduction.	3
3) Literature Review.	5
4) Comparative Analysis.	6
5) Design Details.	7
6) Summary.	8
7) References.	9

Abstract:

Improvised on a previously existing static database which stores the information of all the different versions of football players in an online game (FIFA Mobile). This already existing static database does not show the user all the parameters at once, hence to compare between the statistics of players is almost impossible.

Implemented the storage of different accounts of different users making the database comparatively more secure and personalized. Implemented a better way of visualizing and comparing the differences between different versions of players. Included and visualized personalized and detailed information of each player in the database. The strength of this project is that the algorithm that is used is faster and local and visualization is not only much more appealing but also more functional and user friendly.

Introduction:

What is a search engine?

A search engine is an online tool that searches for results in its database based on the search query (keyword) submitted by the internet user. The results are usually websites that semantically match with the search query.

Search engines find the results in their database, sort them and make an ordered list of these results based on the search algorithm. This list is generally called the search engine results page (SERP).

There are many search engines on the market, while the most widely used is Google. Many website browsers such as Chrome, Firefox, Safari or Edge usually come with a default search engine set as a home page or starting page.

How search engines work?

There may be some differences in how the search engines work but the fundamentals remain the same. Each of them has to do the following tasks:

1. Crawling

Search engines have their own crawlers, small bots that scan websites on the world-wide-web. These little bots scan all sections, folders, subpages, content, everything they can find on the website. Crawling is based on finding hypertext links that refer to other websites. By parsing these links, the bots are able to recursively find new sources to crawl.

2. Indexing

Once the bots crawl the data, it's time for indexing. The index is basically an online library of websites. Your website has to be indexed in order to be displayed in the search engine results page. Keep in mind that indexing is a constant process. Crawlers come back to each website to detect new data.

3. Creating results

Search engines create the results once the user submits a search query. It's a process of checking the query against all website records in the index. Based on the algorithm, the search engine picks the best results and creates an ordered list.

What is a search engine algorithm?

A search engine algorithm is a unique formula that determines how the websites are sorted in the search engine results page. It's a trademark of each search engine, therefore, it's kept secret.

The algorithm is a mixture of various ranking factors. You will find many articles dealing with the "real" Google ranking factors. The truth is that even when you know what the factors are, you don't know their exact weights.

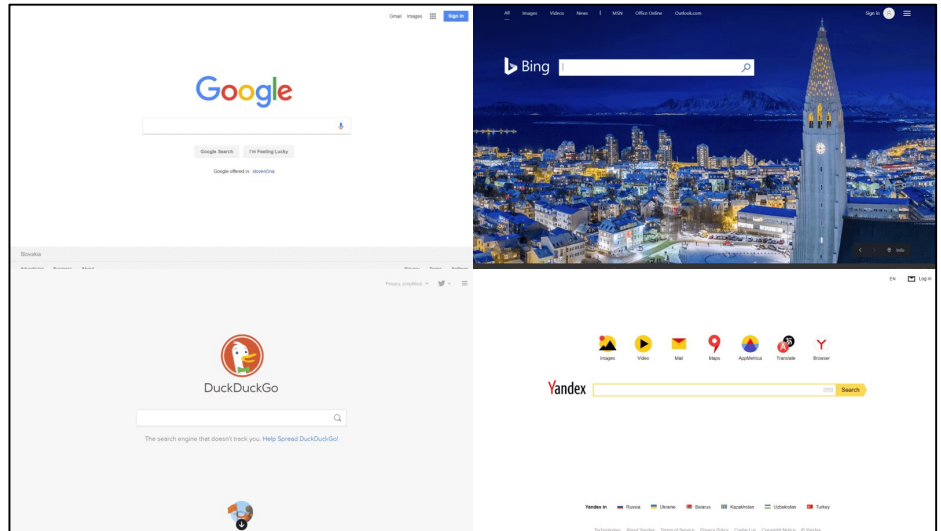
The most important ranking factor of all search engines should be the *relevancy*. The main goal of search engines is to find what the internet user is looking for.

When it comes to Google, the major ranking factors are:

- 1) Website/content relevancy
- 2) Strength and relevancy of backlinks

Some of the other very important ranking factors are:

- 1) Mobile optimization
- 2) Content structure and optimization
- 3) Usability
- 4) Accessibility
- 5) Page speed
- 6) Social signals



7) Overall domain authority

The most popular search engines

In terms of worldwide popularity, Google has been #1 for years. This is the list of top 10 most popular search engines:

1. Google

Google is the giant in the industry and has probably the most sophisticated algorithm. It includes machine learning, AI and RankBrain, another algorithm that can tweak the weights of ranking factors according to user's behaviour and quality of previous results. More than 70% internet users worldwide do their searches on Google since 1998.

2. Baidu

Baidu is the first search engine launched in China in 2000. It's like Chinese Google. Baidu cooperates with companies such as Microsoft, Qualcomm, Intel or Daimler on various AI projects. Similarly to Google, they offer a lot of other solutions such as cloud services, maps, social networking, image and video search and many others.

3. Bing

Microsoft launched their search engine in 2009 as a new project after earlier search engines MSN Search and Windows Live Search. The main goal was to develop a competitor for Google. From the global point of view, it's not really there but in the US, Bing is the 2nd most popular tool for the internet searches.

4. Yahoo!

Originally, it was one of the most widely used email providers and search engines. The company grew significantly in the 1990s but after 2000, they somehow started lacking the innovation and lost their value. In 2017, Yahoo! was acquired by Verizon Communications.

5. Yandex

Yandex Search is the major Russian search engine. According to Wikipedia, Yandex generates more than 50% of all searches in Russia. Though the algorithm is not as sophisticated as Google, it constantly gets better by integrating AI and machine learning that analyze searches and learn from them.

6. Ask

Ask (formerly Ask Jeeves) was launched in 1996. It was designed to answer questions submitted to the search form. Thanks to the Ask toolbar, this search engine was able to compete with big players such as Bing, Yahoo! and Google. Unfortunately, the toolbar was many times installed as an unwanted browser feature.

7. DuckDuckGo

DuckDuckGo is a bit different search engine. They protect the users' privacy by not tracking any information. DuckDuckGo doesn't show personalized results based on your previous searches. Likewise, advertisers can't follow the behavior of the users. On the other hand, you can launch ads via Bing since DuckDuckGo is their search partner together with Yahoo.

8. Naver

Naver is the Google of South Korea. This search engine covers around 75% of searches in the country. It was launched in 1999 and in 2000 it was able to pull out various types of results that match the entered keywords. The results included websites, images, blogs, restaurants, shops, etc. Google launched this feature 5 years later.

9. AOL

In the 1990's, AOL was one of the biggest crawler-based search engines. As a company, AOL offers a lot of other services: email service, instant messenger, video content, yellow pages, city guides. The AOL Search is nowadays used by not more than 0.5% of the internet users.

10. Dogpile

Dogpile is a metasearch engine, so it creates the search engine results page by doing simultaneous search requests for the same query in other search engines, namely: Google, Yahoo!, Yandex and others. Dogpile was launched in 1995.

Literature Review:

Search engine symbolizes an extremely powerful and valuable tool for fetching any sort of information from Internet. There has been numerous researches carried on search engines techniques, the major ones are syntactic and semantic. Referring to the Syntactic web, the results obtained are purely as per the keyword match. That is the query outputs numerous web pages against the keyword match that may not even be relevant or meaningful. Whereas, unlike the syntactic web, the semantic web is a revised or upgraded version of the web which produces quiet meaningful and specific output as it has the potential to comprehend the query effectively. Few examples of Semantic based search engines include Kosmix, Hakia, Cognition, Swoogle and Lexxe. Whereas syntactic based search engines are Google, Yahoo, Ask. The work performs a comparison amidst the performance of semantic and syntactic based search engine and evaluates them by employing certain queries.

Our search engine is a syntactic search engine as it searches purely based on the name of the player.

With the rapid development of World Wide Web, search engines have become the main tool for people to get network information. However, the search results are widely criticized due to the poor accuracy and redundancy disadvantages. After the advent of semantic Web, new search engine with the ability of understanding queries and documents has attracted more and more attentions. This paper starts from the traditional search engine, and firstly introduces its classification, popular technology, advantages and disadvantages, thus leads to the semantic search engine model. Then we research the semantic search technology in depth, which can be divided into enhanced semantic search based on traditional search, knowledge semantic search based on ontology inference and other semantic search types.

Semantic based search engine are able to provide more relevant information because they understand the meaning of the term and relationship between the term and web pages but keyword based search engine not able to understand these thing. This paper evaluates the performance of Semantic web based search engine and Keyword based search engine. We selected three Semantic web based search engine (Bing, DuckDuckGo and Lexxe) and two Keyword based search engine (Google and Yahoo) to compare their search performance.

Since the use of internet has incredibly increased, it becomes an important source of knowledge about anything for everyone. Therefore, the role of search engine as an effective approach to find information is critical for internet's users. The study of search engine users' behaviour has attracted considerable research attention. These studies are helpful in developing more effective search engine and are useful in three points of view for users at the personal level, for search engine developers. These kinds of studies can be done through analysing the log file of search engine wherein the interactions between search engine and the users are captured.

These days, the growth of web has led it to a big source of information. Web search engine plays an important role of searching desired information from this enormous web. However, search engine provides the same result independently to the user while actually each user has different preference. We benefit from proxy servers, which are widely used in a community network to reduce bandwidth needs. Proxy servers are, actually, providing the user preference within its access log that contains accessed URLs. Instead of web crawler, we will use this logs, which is always updated as users browse the web through this proxy. This would be the base of our customized web search. As the proxy log only covers URL list, we still need to crawl the information contained in an URL. When the crawling method has completed, document vector is created to make those data to be more machine friendly. Eventually, searching process is carried out by utilizing the vector space model.

Investigation of information retrieval from distributed databases. The most popular example of such systems are the regular web search engines. In depth search engines systems and have made an evaluation of different search engines based on popular queries searched. Exploring the structure of the search engines and explain its functions behind the scenes, how they can get the results for a specific query. The key idea behind search engines is the ranking algorithms used to bring the appropriate results. This research has chosen Google, Yahoo, and Bing to do evaluation experiments, since they are top search engines used nowadays.

Previous work in this area was done by Fifarenderz.com which scans the FIFA Mobile app in its entirety to get newly created / updated players (versions) from which we have taken reference from and made the UI and scraping algorithm local and faster.

Comparative Analysis:

There are multiple cons in this previously existing technology.

- 1) Scanning the app in its entirety takes a long time especially when the outcome (tag to be found) is unknown.
- 2) The servers are online hence accessing data without an internet connectivity is not feasible.
- 3) The visualization of data in an in-depth format is very hard and not user friendly.

Hence modifications are instantiated in the program as to surpass these cons.

Algorithm:

The first con is surpassed as the tag to be found is already known and hence that makes our algorithm a read-once algorithm making it quicker than the already existing algorithm.

Step 1: Data.getsetdata() function is called, parameter – maximum number of pages.

Step 2: get() function is called , parameter – maximum number of pages.

Step 2.1: Creating a for loop with 1 loop variable from 1 to the maximum number of pages applicable up to step.

Step 2.2: Creating a link with page as loop variable.

Step 2.3: Calling getdata() function, parameter – link of the web site(page wise).

Step 2.3.1: Creating a connection to the web site(parameter).

Step 2.3.2: We read and select data from the web page for a particular tag (span).

Step 2.3.3: Converting the selected data into text.

Step 2.3.4: Using a loop segregating the read data into respective columns(array).

Step 2.4: Calling getadddata() function, parameter – link of the web site(page wise).

Step 2.4.1: Creating a connection to the web site(parameter).

Step 2.4.2: We read and select data from the web page for a particular tag (span).

Step 2.4.3: Converting the selected data into text.

Step 2.4.4: Checking if player is a goalkeeper.

Step 2.4.4.T: Using a loop segregating the read data into respective columns(array) like goalkeeper positioning, jumping, reflexes, etc.

Step 2.4.4.F: Using a loop segregating the read data into respective columns(array) like sprint speed, acceleration, shot power, etc.

Step 3: Calling the set() function.

Step 3.1: Creating a connection to the MySQL backend (port: 3306).

Step 3.2: Setting the username and password.

Step 3.3: Checking if there is a “ ‘ ” in the data and if it exists replacing it with a space(“ ”).

Step 3.4: Creating and executing insert statement for a part of the data read from function getdata() to table players.

Step 3.5: Checking if player is a goalkeeper.

Step 3.5.T: Creating and executing insert statement for the rest of the data read from function getdata() and from function getadddata() to table statkeepers and exstatkeepers respectively.

Step 3.5.F: Creating and executing insert statement for the rest of the data read from function getdata() and from function getadddata() to table statplayers and exstatplayers respectively.

Local Database:

The database for this program is local and not online hence users can access and use this application without the requirement of an internet connectivity.

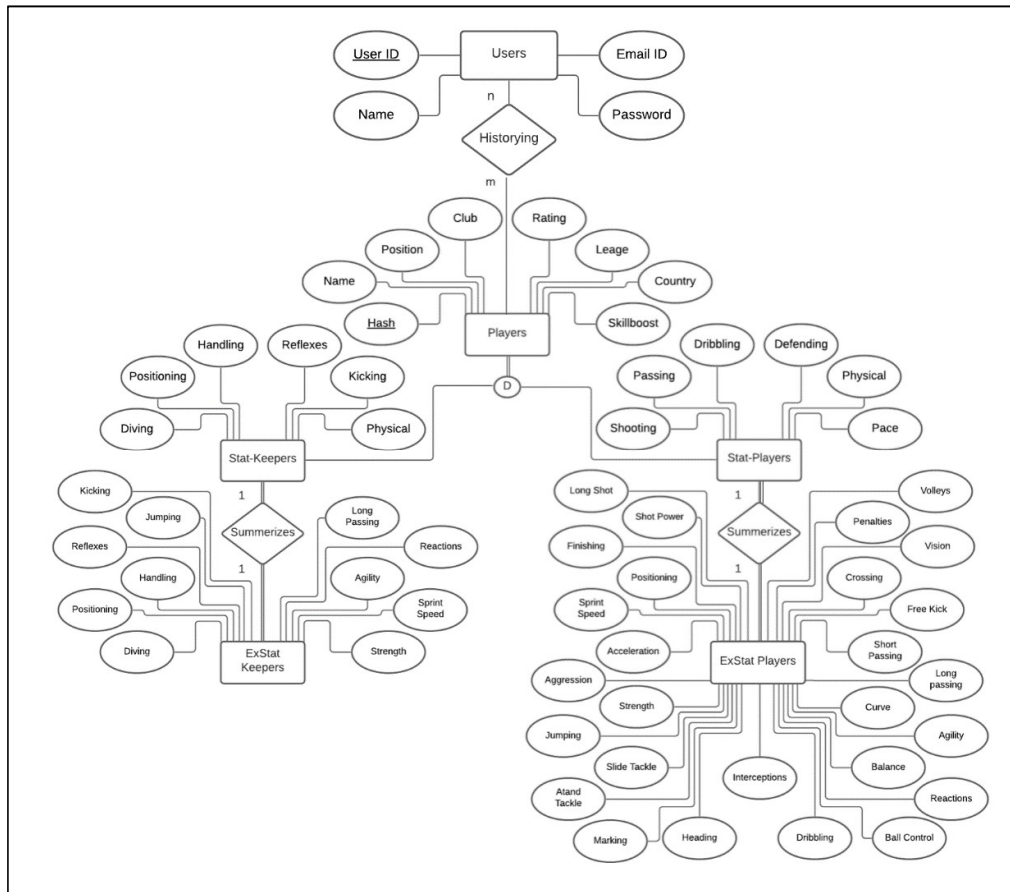
User Interface and In-Depth Data Visualization:

The User Interface is much simpler and much more functional as compared to that of Fifarenderz.com. With the help of having multiple user access this application secures the users search data in a better and a more personalized format. The program gives an In-Depth view of the data of a particular player in a static page, which is a page that is not affiliated to the program, hence multiple such pages could be opened and could be used to compare different versions of different players.

Design Details:

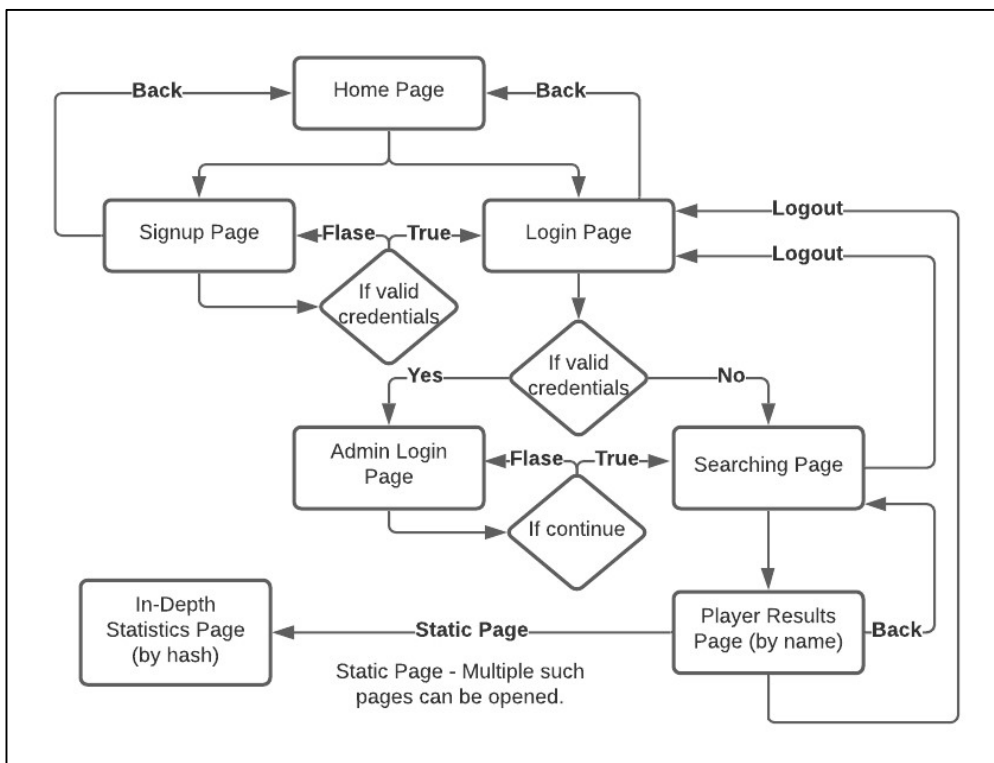
The EER diagram:

This is the backend structure of the application, the database,



The Flowchart:

This is the frontend structure of the application the user interface (UI).



Summary:

Conclusion:

Improvise on a previously existing static database which stores the information of all the different versions of football players in an online game (FIFA Mobile). This already existing static database did not show the user all the parameters at once, making comparison not viable. Created a data table which stores information of different accounts of different users which made the database comparatively more secure and personalized. Implemented a better way of visualizing and comparing the differences between different versions of players.

Future Scope:

The future scope of this application is that this application could be used for upcoming and even already existing versions of FIFA. This application could help us compare not only players from the same versions of FIFA but different too which is absolutely unheard of until now.

Reference:

David Hawking Web search engines. Part 1 CSIRO ICT Centre, IEEE, Australia 2006

A.C. Santha Sheela; C. Jayakumar Comparative Study of Syntactic Search Engine and Semantic Search Engine: A Survey, IEEE Chennai, India 2019

Farzaneh Shoeleh; Mohammad Sadegh Zahedi; Mojgan Farhoodi Search engine pictures: Empirical analysis of a web search engine query log, IEEE, Tehran, Iran 2017

Duygu Tümer; Mohammad Ahmed Shah; Yiltan Bitirim An Empirical Evaluation on Semantic Search Performance of Keyword-Based and Semantic Search Engines: Google, Yahoo, Msn and Hakia, IEEE, Venice, Italy, 2009

Cai Bo; Li Yang-Mei Design and Development of Semantic-Based Search Engine Model, IEEE, Changsha, China, 2014

Peiquan Jin; Jianlong Lian; Xujian Zhao; Shouhong Wan TISE: A Temporal Search Engine for Web Contents, IEEE, Shanghai, China, 2008

Bimo Sunarfri Hantono; Guntur Dharma Putra Generating customized web search result through community driven search engine, IEEE, Yogyakarta, Indonesia, 2013

Sarah Hamed Alahmadi Information Retrieval of Distributed Databases A Case Study: Search Engines Systems, IEEE, Riyadh, Saudi Arabia, 2018

Javed Ahmad Khan; Deepak Sangroha; Masroor Ahmad; Md. Tanzillur Rahman A performance evaluation of semantic based search engines and keyword based search engines, IEEE, Greater Noida, India, 2014

Wang Liangshen; Hou Jie; Xie Zaiyu; Wang Xiaochen; Qu Caiyue; Li Hui Problems and solutions of web search engines, IEEE, XianNing, China, 2011

Acknowledgement:

I would like to express my special thanks of gratitude to my mentor Ms. Jalpa Mehta and Ms. Nida Jawre, our principal Mr. Bhavesh Patel and the faculty of Shah and Anchor Kutchhi Engineering College for this wonderful project on the topic Search Engine which also helped me in doing a lot of Research. It had been a great journey in analysing, understanding and realizing the implementation of a real time project in JAVA and SQL. It helped me grow a lot not only academically but also in a broader and knowledgeable aspect. I would like to express my very great appreciation to my group members for their valuable and constructive execution and suggestions during the planning and development of this research work and in making of the project. The team building, managing and co-operative skills were also been imparted in us. Hence making us more professional and patient in work.