

# I. Introduction

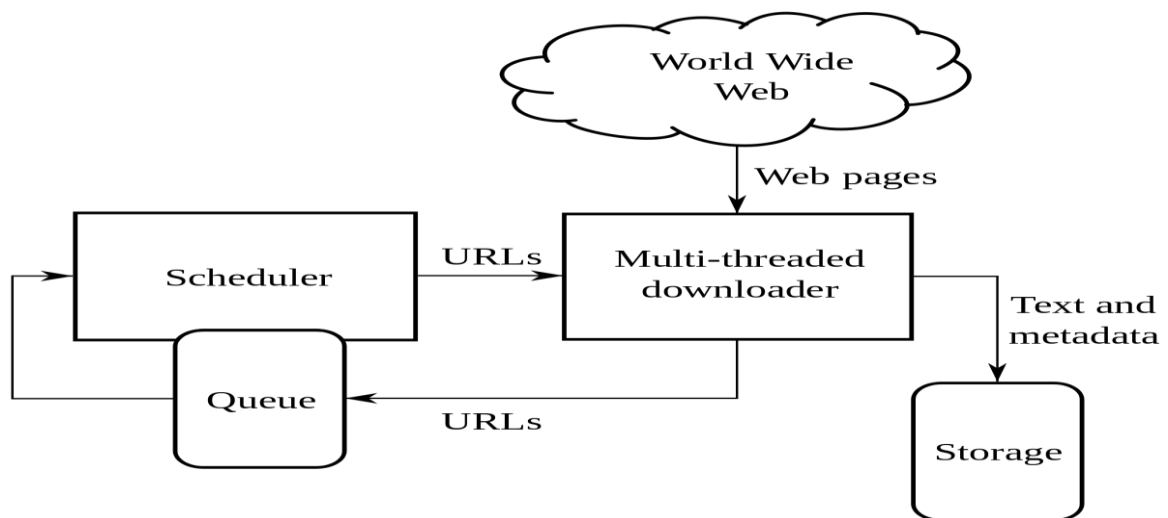
## 1.1 Basic Information

In day to day life internet is becoming more and more popular. That helps us to get Any kind of information in very short, That all is managed by Web Applications(Websites) where we can visit and find or read suitable information that we requires. Now as we know there are thousands of web Application are there on internet, they Are providing same information's and how can we reach very perfect information which we want to see, without knowing URL. Now Search Engine comes in picture, that helps us to find suitable information and best information from the multiple Web Applications with the help of keywords. Askme is Also a web Application based Search Engine that helps to find best and suitable Information across WWW(World Wide Web), with multiple features ex best search, manages web log, page marks, visitor management, search Images, search news, Search Shopping, Search Books, Search Documents.

Also Askme have subproject like Amail (Askme mail), Askme news, Askme Sports, Askme Healthcare,(Askme Pay , Askme Ture and travels, AMart (Askmart). In more details about subproject are explained below.

## 1.2 Proposed System

### i.Web Crawler



A web crawler (also known as a web spider or web robot) is a program or automated script which browses the World Wide Web in a methodical, automated manner.

This process is called Web crawling or spidering.

Many legitimate sites, in particular search engines, use spidering as a means of providing up-to-date data.

Web crawlers are mainly used to create a copy of all the visited pages for later processing by a search engine, that will index the downloaded pages to provide fast searches.

Crawlers can also be used for automating maintenance tasks on a Web site, such as checking links or validating HTML code.

Also, crawlers can be used to gather specific types of information from Web pages, such as harvesting e-mail addresses (usually for spam).

## ii. MapReduce

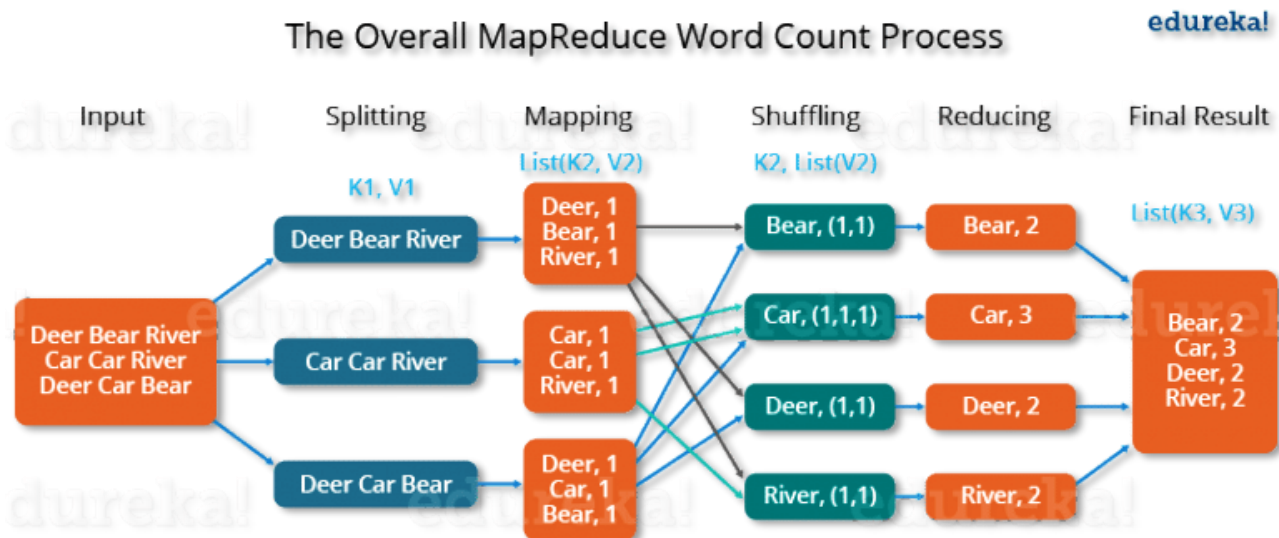


Fig.02: MapReduce – Example

MapReduce is a programming paradigm that enables massive scalability across hundreds or thousands of servers in a Hadoop cluster. As the processing component, MapReduce is the heart of Apache Hadoop. The term "MapReduce" refers to two separate and distinct tasks that Hadoop programs perform. The first is the map job, which takes a set of data and converts it into another set of data, where individual elements are broken down into tuples (key/value pairs).

The reduce job takes the output from a map as input and combines those data tuples into a smaller set of tuples. As the sequence of the name MapReduce implies, the reduce job is always performed after the map job.

### 1.3 Abstract

This Project is based on Search Engine That data Comes from WWW. And Stored as Physical data in database. Ask Email is helps for Emailing user. Ask Healthcare is for Giving information About Health Tips, online Appointment System etc. this project is have feature of comparing data provide simple Information about search.

### 1.4 Advantages

- **Time Savings**

A search engine saves you time in two ways: by eliminating the need to find information manually, and by performing searches at high speeds. Without a search engine, you would have to look at sites one by one and pore over the contents of each carefully – a tedious prospect. A search site automatically compares your criteria to billions of Web pages and gives you results in a fraction of a second. You can perform dozens of searches in the course of a few minutes, altering the criteria as you narrow down results.

- **Relevance**

When a search engine scans a website, it scores the content for relevance to particular search words. For example, the site, "Joe's Pizza" scores higher for the search term "pizza restaurant" than the site "Acme Pizza Dough Distributors." Although search companies keep their ranking formulas secret, they take into consideration factors such as repetition of related words and links from other sites. A search engine sorts its results page by relevance to your criteria, with the score in descending order. You see the highest-scoring results at the top; as you move down the list, websites become less relevant.

- **Free Access**

Some search engines, such as [LexisNexis](#), specialize in legal or other specialized, scholarly information; these sites charge a fee to use their services. Google, Bing and Yahoo pay for their operations through advertising; searches are free to the user, without restrictions for the information you seek, the time spent on the site or the number of searches you perform. Although this benefits all users, it is a particular advantage for students, job seekers and others of limited means.

- **Comprehensive**

Search engines scan the entire Web and keep comprehensive data on every page they catalog. Because search companies hold so much data, they help you find obscure sites about which you would not otherwise know. Search results are more likely to give you too much information rather than too little.

- **Advanced Search**

In addition to keywords, search engines let you use advanced search options to refine your results. These options help make your searches more flexible and sophisticated. For example, to exclude results containing a certain word, type a minus sign before the word. To look for an exact phrase, surround it with quotation marks. When you want to search only a specific site or group of sites, type "site:" without quotes followed by the site's Web address. You can, for

example, search for mentions of beer on [Facebook](#) by typing, "beer site:www.facebook.com."  
To search all college and university sites for nuclear physics, use, "nuclear physics site:.edu."

- **Search Sports Data**

Search Live Sports Data (Score, upcoming Matches) And Sports information and Sportsman Information.

- **Search Health Related Data**

Search Health Related Information, Symptom's information, Medicine Information etc.

## 1.5 Modules Used In Project

1. **User Authentication:** User Sign in, Sign up, forgot password.
2. **Manage Websites details :** Add Website, Delete Web Site, block Website;
3. **Manage Spiders :** Find Public URL from website and it's details(Title, keyword, Description etc....).
4. **Manage Search Results :** What-what data have to show on search.
5. **Manage webpage ranking:** assign the position of web page according to keywords.
6. **Manage keyword:** Add keyword, manage pages according to that keyword.
7. **Manage vide data:** show video file when user search for video.
8. **Manage Image data:** show Image file when user search for Image(jpg,gif, png. Etc....).
9. **Manage Audio data:** show Audio file when user search for Audio.

### 2: ASports (AskMe Sports).

10. **Show sports score cards:** Show All Current Live Scores.
11. **Manage live score:** Update Score Time by time Automatically.
12. **Manage Events:** List All The current events with Date and information and upcoming events.
13. **Manage Schedule of sports:** Show current and Upcoming schedule information.
14. **Manage sports information:** Information of sports ex About etc.
15. **Manage Sports Player Information:** Information of Players According to their Sports ex-Football, cricket etc.....
16. **Manage Sports related News:** manage news of Sports.

### 3: AMail (AskMe Mail)

17. **Manage Email:** Manage Email (ID, Passwords, Emails, Security).
18. **Manage Inbox:** Manage Incoming Emails of user.
19. **Manage sent Emails:** Manage outgoing Emails of user.
20. **Manage draft:** Manage pending email of user.
21. **Manage Spam Emails:** Manage spam email of user.

- 22. **Manage Blocked Emails:** Manage Blocked email of user.
- 23. **Manage Settings of Emails:** manage Accounts Change Email id, Name, Password, Permissions etc.

#### 4: Askme news

- 24. **Manage Breaking news:** manage breaking news of current days according to category.
- 25. **Manage news data:** Fetch and show news data from different news websites.
- 26. **Manage and Rank News Websites:** Rank News Websites according to there view Comments
- 27. **Manage news by keyword:** Save store news keywords.
- 28. **Manage Comment, Like and Share:** Save, store, show, delete, update comments and like and share,
- 29. **Manage news by Time:** Show and update news time by time.
- 30. **Manage News category:** Management of News category.

#### 5: Askme Healthcare

- 31. **Show health tips:** Find and show best health tips from different web.
- 32. **Manage disease information:** show disease information.
- 33. Show keywords(healthcare) related question on web.
- 34. **Manage hospitals data:** store and show hospitals data.
- 35. **Manage doctors:** Save doctors data, (Name, Specialist, Address, fees etc.) and show to user.
- 36. **Request Appointment:** help to request Appointment.

## 1.6 Limitations

Limitations of the Website are only the ones that have been imposed by the existing technology.

## II. Related Work

### 2.1 Major Features

#### I. Technology-Related

##### 1. Google

Google Search Engine is the best search engine in the world and it is also one of most popular products from Google. Almost 70 percent of the Search Engine market has been acquired by Google. The tech giant is always evolving and looking to improve the search engine algorithm to provide best results to the end-user. Although Google appears to be the biggest search engine, as of 2015 YouTube is now more popular than Google (on desktop computers).



##### 2. Bing

Bing is Microsoft's answer to Google and it was launched in 2009. Bing is the default search engine in Microsoft's web browser. At Bing, they are always striving to make it a better search engine but it's got a long way to go to give Google competition. Microsoft's search engine provides different services including image, web and video search along with maps. Bing introduced Places (Google's equivalent is Google My Business), this is a great platform for business to submit their details to optimise their search results



##### 3. Yahoo

Yahoo & Bing compete more with each other than with Google. A recent report on netmarketshare.com tells us that Yahoo have a market share of 7.68 percent. Although a leader as a free email provider, this is declining significantly though with their recent acknowledgement that User Details & Passwords were hacked last year.



##### 4. Baidu

Baidu is the most used search engine in China and was founded in Jan, 2000 by Chinese Entrepreneur, Eric Xu. This web search is made to deliver results for website, audio files and images. It provides some other services including maps, news, cloud storage and much more.



### 5. AOL

Aol.com is also among the top search engines. These are the guys that used to send out CD's which you'd load onto your PC to install their browser and modem software. Once the pre-eminent player they now have a market share of 0.59 percent. Verizon Communication bought AOL for \$4.4 billion. It was started back in 1983 as Control Video Corporation. It was named America Online in 1991 and in 2009 as AOL Inc. AOL is a global mass media company which is based in New York. The company also provides advertising services as AOL Advertising, AOL mail and AOL Platform.

# Aol.

### 6. Ask.com

Founded in 1995, Ask.com, previously known as Ask Jeeves. Their key concept was to have search results based on a simple question + answer web format. It is a question & answer community where you can get the answers for your question and it integrates a large amount of archive data to answer your question. Because of this dependency on archived and active user contributions the results will not be as current as those you get in Google, Bing and Yahoo. They've tried to counter where their resources don't have the answer they take help from a third-party search engine. Interestingly they don't name who this is.



### 7. Excite

Excite is not widely known but is one that still gets into the top 10. Excite is an online service portal that provides internet services like email, search engine, news, instant messaging and weather updates. This also provides us with the latest trends, topics and search for phrases such as: What can President Trump actually do?



### 8. DuckDuckGo

DuckDuckGo is a popular search engine known for protecting the privacy of the users. Unlike Ask.com they are quite open about who they use to generate search results; they've partnered with Yahoo, Bing and Yummly. It was founded back in 2008 by Gabriel Weinberg in California and its revenue come from Yahoo-Bing search alliance network and Affiliates.



## i. Domain-Related

### o Google.com

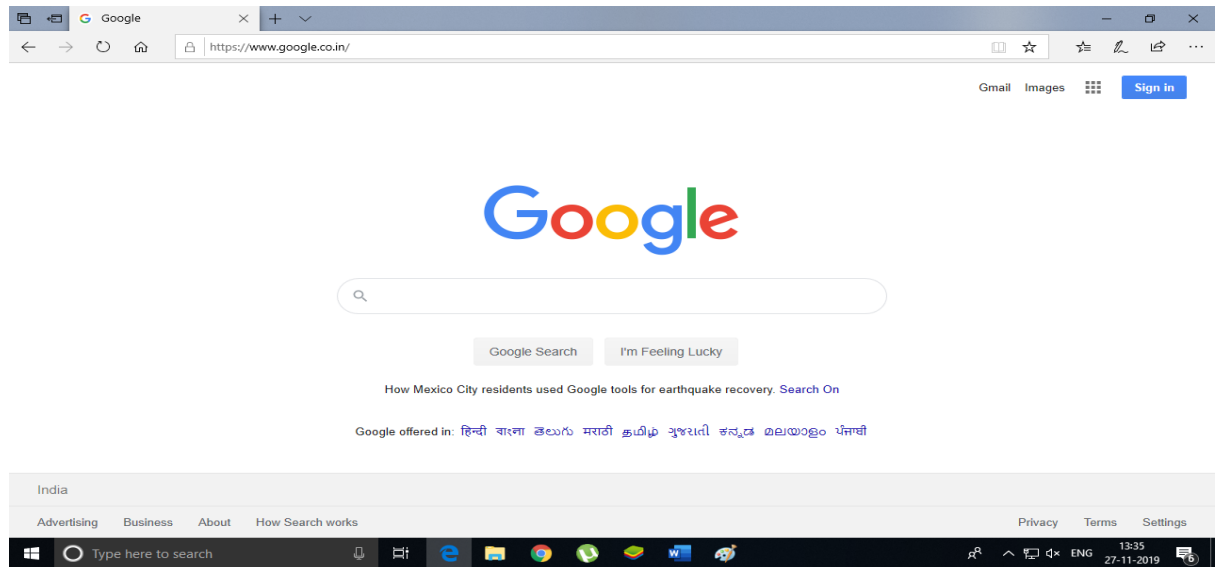
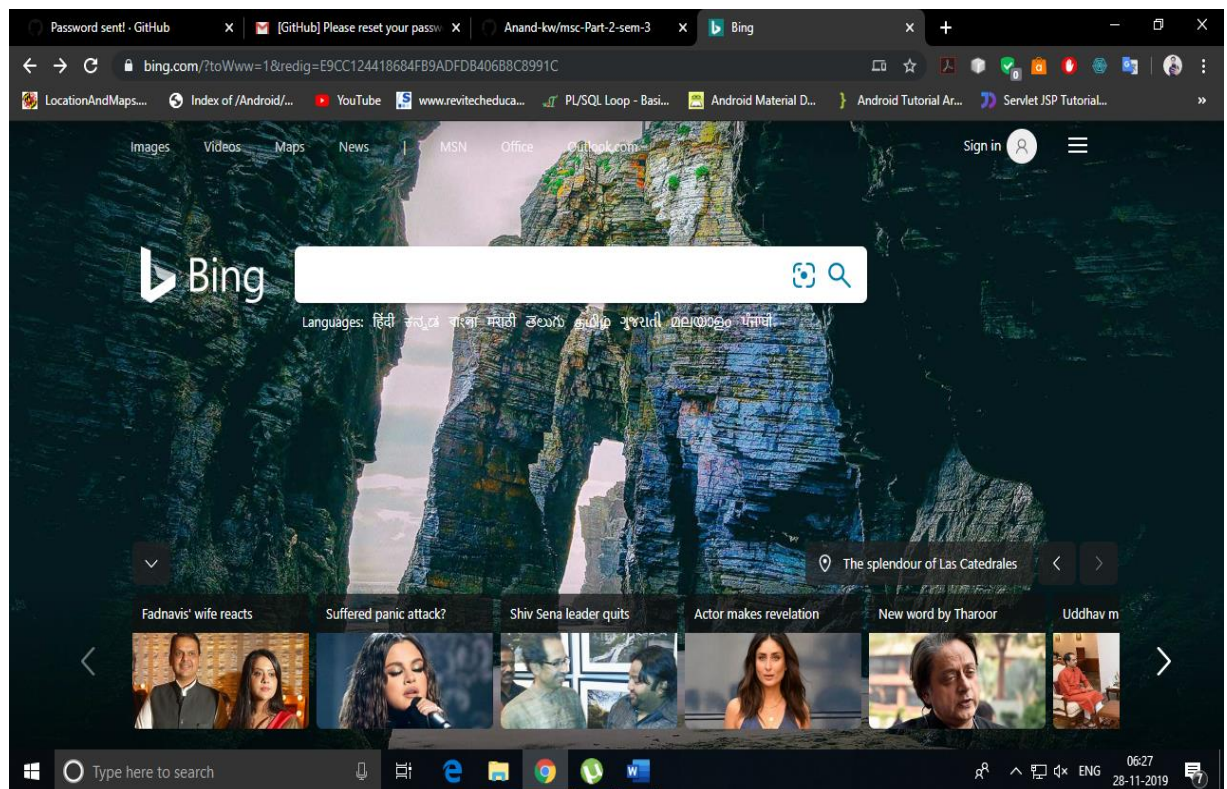


Fig.06: <https://www.google.co.in/>

### o Being.com





### III. Objective

This project is based on Algorithms, Cloud Computing and Business Intelligence Technologies. the main objective of this project is to Provide best search Results from WWW on the basis of users Input. **Amail** provides Email system, Ex. Email id, send Email, Receive Email etc.. **Askme sports** helps to find live scores of cricket football, hockey etc. and it also show sports related news. **Askme news** helps to find latest news, filter news. **Askme healthcare** helps to read health tips, nearest gyms, Medical , hospital clinic etc... and also user can do quick Appointment.+

### IV. Methodology

#### 4.1 Software Requirements

This project as title “Askme” and its Sub Projects is comes under the Algorithms, Cloud Computing and Business Intelligence Relational Database Management System (RDBMS). This application is developed with the help of HTML5,CSS3,JavaScript, Angular.js,Node.js, Bootstrap, SASS, LESS, Ajax, Java.1.8 and MySQL. This application can also be run on the Any operating System, All browser.

This project is developed using the tools, which are most suited for development of the Application Package. These tools are as follows: -

1. HTML5, CSS3,JavaScript, Angular.js,Node.js, Bootstrap, SASS, LESS, Ajax, Jsp (For front end)
2. Java.1.8 and MySQL (For Server Side, Database Storage as Back end).

#### HARDWARE SOFTWARE REQUIREMENTS

##### HARDWARE:

**Processor:** Any

**Hard Disk Space:** 20 GB (min.)

**Ram Memory:** 32 MB (64 MB recommended)

**Monitor:** Any

##### SOFTWARE:

**Operating System:** Windows, Linux, Android. Mac OS.

**Database Server:** Xampp server/phpMyAdmin)

**Front end:** NetBeans 8.2, Notepad++, Atom

**Server:** GlassFish 4.1.1 and Apache 8.0.27.0

**SCSS Compiler:** Prepros 6.3.0

## Bibliography

### Future Scope of Project

This Project is a search engine project named as AskMe, This helps user to find better result for his/her query, It is based on simple search so in the future this project will attract user to use this Application for their solutions,

Initial level data size can be low, so it won't be helpful but in future when data will get increased is it will become useful. New search algorithms can be create for making application unique in Market.

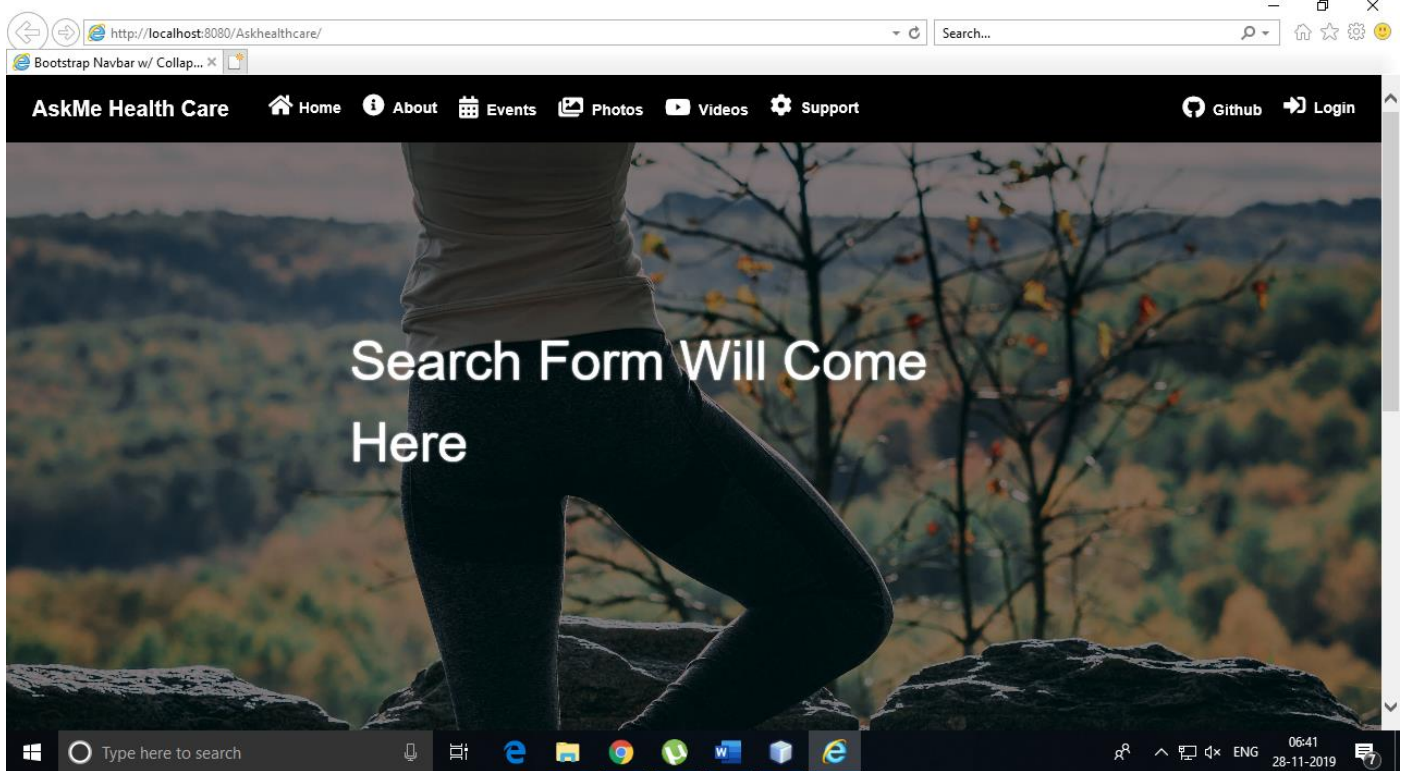
In future more module can be Added example security feature, more sub project can also be added Askme Pay , Askme Ture and travels, AMart(Askmart) etc.

### Reference

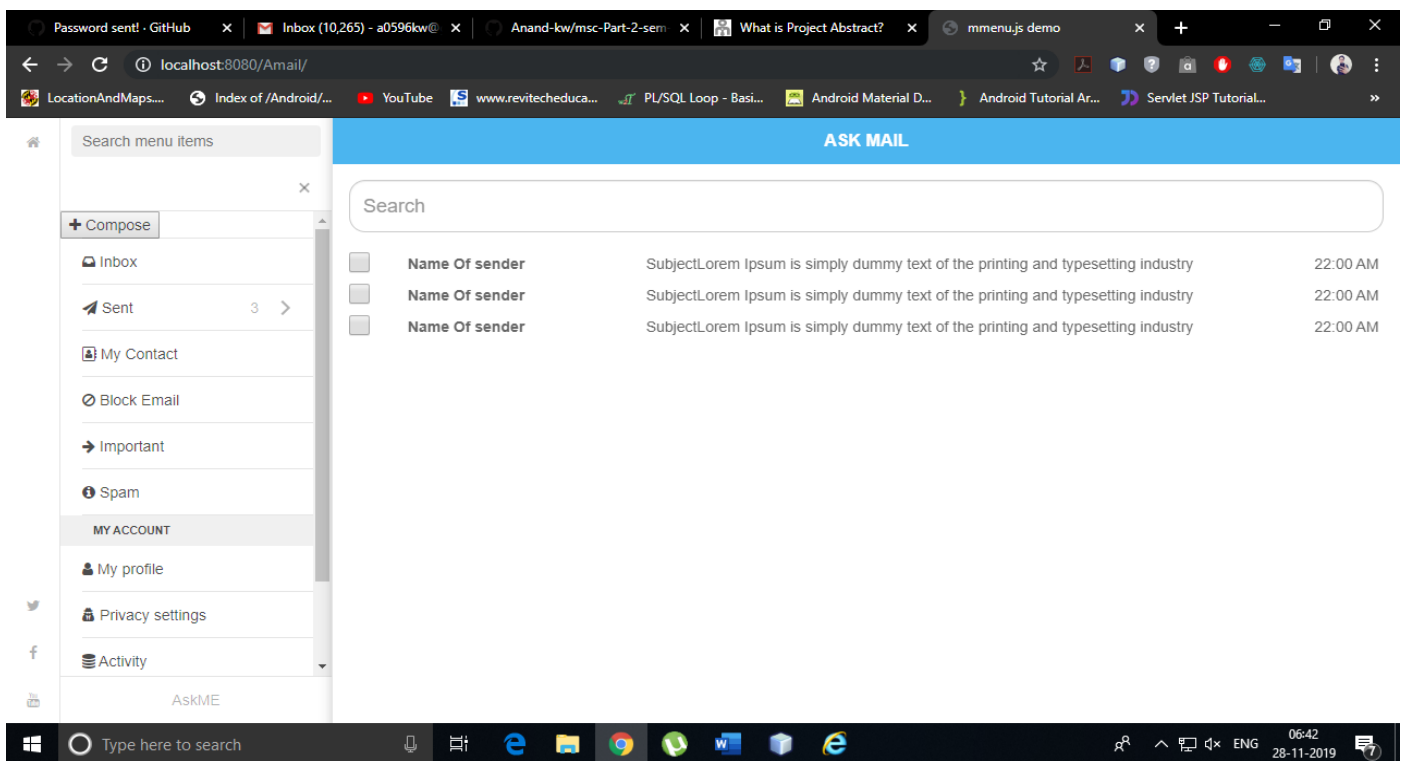
- <https://www.google.co.in> (Search Engine)
- <https://www.bing.com> (Search Engine)
- <https://www.ask.com>(Search Engine)
- <https://www.mayoclinic.org> (Health care)
- <https://www.healthline.com> (Health care)
- <https://www.cricbuzz.com> (Sports)
- <https://www.espn.in> (sports)
- <https://mail.google.com/mail/u/0/#inbox> (Mail)
- <https://mail.rediff.com/cgi-bin/login.cgi> (Mail)
- <https://login.yahoo.com>(Mail)
- [https://outlook.live.com/owa/\(Mail\)](https://outlook.live.com/owa/(Mail))

## Screen Shorts of Projects

### ○ AskHeathCare



### ○ Ask Mail



## ○ SearchWorld(AskMe)

The screenshot displays the SearchWorld.com website interface. The browser's address bar shows the URL `localhost:8080/SearchWorld.com/`. The website has a dark header with the logo "SearchWorld.com" and a navigation menu with links: Home, Person, Country, Books, City, Quiz, Download, Page 2, and Sign In. Below the header, there is a search bar containing the text "Af". To the right of the search bar are two buttons: "Search For Image" and "Comapre".

The search results are displayed in a list format. The first result is for "Afghanistan (AF)". It includes the flag of Afghanistan, the country name "Afghanistan (AF)", and details: Capital: Kabul, Currency: AFN, and Language: fa-AF,ps,uz-AF,tk. The second result is for "Central African Republic (CF)". It includes the flag of Central African Republic, the country name "Central African Republic (CF)", and details: Capital: Bangui, Currency: XAF, and Language: fr-CF,sg,ln,kg.

Below the search results, there is a detailed view for "AFGHANISTAN". It features a large banner image of a stadium at night with the word "AFGHANISTAN" overlaid. To the left of the banner is the flag of Afghanistan. Below the flag is a button labeled "Add Image". To the right of the flag is a large gray rectangular area. Below the gray area is a dropdown menu labeled "Select Language".



SearchWorld.com

Home Person Country Books City Quiz Download Page 2 Login

Select Language  
Powered by Google Translate

Afghanistan, officially the Islamic Republic of Afghanistan, is a landlocked country located within South Asia.

### > Population of Afghanistan


World level Rank:	39
Country (or dependency):	Afghanistan
Population(2018):	36,373,176
Yearly Change:	2.37 %
Net Change:	843095
Density(P/Km²):	56
Land Area(Km²):	652860

Type here to search

SEARCHWORLD

Google Translated to: Hindi Show original Options

# अफगानिस्तान



छवि जोड़ें

Hindi  
Powered by Google Translate

अफगानिस्तान, आधिकारिक तौर पर अफगानिस्तान के इस्लामी गणराज्य, एक landlocked दक्षिण एशिया के भीतर स्थित देश है।

SearchWorld.com

username Password Login

Search world is your world  
it is 100% free for you

Register with us!!!

First name  
Last name  
Email  
Password  
☐ Male ☐ Female  
☐ Student ☐ Teacher ☐ Other  
Register

Home Person Country Books City Quiz Download Page 2 Sign In

SearchWorld.com

Search For Image Comapre

in VS rd

Argentina(Country)  
Bosnia and Herzegovina(Country)  
Burkina Faso(Country)  
Bahrain(Country)  
Benin(Country)  
Saint Barthélemy(Country)  
Bonaire, Sint Eustatius and Saba(Country)  
Cocos [Keeling] Islands(Country)  
China(Country)

Cameroon(Country)  
Micronesia(Country)  
Faroe Islands(Country)  
Croatia(Country)  
Comoros(Country)  
Morocco(Country)  
Montenegro(Country)  
Romania(Country)  
Aurora(City of U.S.A.)