PROJECT REPORT

**AI Powered News Search App**

Client : Smartbridge

Project Manager : Devi Harika Kavala

**Table of Contents**

**1** **[Introduction:](#_Tocbpfoqlg9b7lk)****[3](#_Tocbpfoqlg9b7lk)**

1.1 [Overview: 3](#_Toctxeprh97i5cc)

1.2 [Purpose: 3](#_Tock0dj4v3r4sxx)

**2** **[Literature Survey:](#_Tocopq21xyl5ls6)****[3](#_Tocopq21xyl5ls6)**

2.1 [Existing Problem: 4](#_Tocxac17z0ebnbe)

2.2 [Proposed Solution: 4](#_Tocubtnmv86fo6n)

**3** **[Theory Analysis:](#_Toc0chtcvgk7vjg)****[4](#_Toc0chtcvgk7vjg)**

3.1 [Block diagram: 4](#_Toct03tuwgz43tr)

3.2 [Hardware/Software Designing: 5](#_Toco1myqbr27zs0)

*3.2.1* *[Technical Requirements :](#_Tocddw6tacmrroh)**[5](#_Tocddw6tacmrroh)*

*3.2.2* *[Software Requirements :](#_Toc92uld0vxduye)**[5](#_Toc92uld0vxduye)*

*3.2.3* *[Hardware Requirements:](#_Tocmcidagknklm7)**[5](#_Tocmcidagknklm7)*

**4** **[Experimental Investigation:](#_Tocyu42mjxhe63r)****[5](#_Tocyu42mjxhe63r)**

**5** **[FlowChart:](#_Tocx30ig3eg6ni4)****[7](#_Tocx30ig3eg6ni4)**

5.1 [Module-1: Explore IBM Cloud Platform: 7](#_Tocdpgo03wuldvm)

5.2 [Module-3: Explore IBM Watson Services: 8](#_Tocreu4n72zvd8v)

5.3 [Module-4 : Build Intelligent News Search App: 8](#_Tocwla9obt365c4)

5.4 [Module-5:Integrate Slack with Watson Discovery(locally). 9](#_Tocznddn4a0i7hm)

**6** **[Result:](#_Toccp5kmk5wsb6e)****[13](#_Toccp5kmk5wsb6e)**

6.1 [Output of UI: 13](#_Tocssbzao3yph6l)

6.2 [Slack Output: 14](#_Toceiz19xbwieqk)

**7** **[Advantages and Disadvantages:](#_Tocu3ls4x1eczpt)****[15](#_Tocu3ls4x1eczpt)**

7.1 [Advantages: 15](#_Toc29hzmpha574e)

7.2 [Disadvantages: 15](#_Tocuyt8brfp3hab)

**8** **[Applications:](#_Toctfxubap2s1yb)****[15](#_Toctfxubap2s1yb)**

**9** **[Conclusion:](#_Toctm823je24ljq)****[16](#_Toctm823je24ljq)**

**10** **[Future scope:](#_Tocfwm77nvxbh7m)****[16](#_Tocfwm77nvxbh7m)**

**11** **[Bibiliography:](#_Tocbc37mfyez1j7)****[17](#_Tocbc37mfyez1j7)**

**12** **[Appendex A:](#_Toc8we0p1sp45j4)****[17](#_Toc8we0p1sp45j4)**

# Introduction:

The web is home to massive amounts of data, with more being created every day. Organizations can harness this constant stream of information to gain understanding, plan strategies, and find opportunities. Enriched news data can help your application make dynamic connections across current events faster.

## Overview:

The basic idea of this app is to create a simple UI that interacts with the user and provides the latest news on the query given by the user.The UI uses Watson Discovery Service to return the news feeds along with the details of the author and the article it is being taken from.We also perform the sentimental analysis of the news feeds and integrate the app to slack so that we can directly interact with the slack bot to get the required news information.

## Purpose:

The main purpose of this app is to gain the required information related to news using the Watson Discovery Service. We can directly interact with the slack bot to get the required news information.

This creates a need for an enriched news mining web application to provide the user

with the latest news regarding any topic.

# Literature Survey:

The main idea behind creating this project is to minimize the problems that are

prevalent in many news mining applications like retrieving irrelevant data and providing

information that is not up to date.The web crawlers downloads pages that do not update their database from time to time, as a result it provides outdated information to

the users.The Watson Discovery Service is a solution to all these problems as it is

updated continuously with new articles from various sources. It also uses techniques like news alerting and event detection to provide the trending news in any field that is being searched.Thus, we can track recent historical trends and stay abreast of latest information about various products, events and many more.

## Existing Problem:

* The next problem is that the users do not specify the exact search words that are required, thereby resulting in poor queries.
* Many web mining applications do not provide results that are relevant to the subject being searched.Most of them use personalized searches that raise privacy concerns.
* The applications that are presently used are poor in performance.They are quite inefficient in returning the results, thus decreasing the precision and recall of the data that is being retrieved.
* These are few of the current problems that hinder the applications in providing better user experience and create better opportunities to make connections across current events in a faster way.

## Proposed Solution:

In the proposed solution, we are building a web mining application using Node-Red and

Watson Discovery Service.The user enters the query, then the Watson Discovery News

collection searches for the trending news across millions of articles and stories regarding that particular topic and returns the latest news.We build the server side application using Node Red by making use of the configuration nodes and embedding them in the flows.Then, we use the pre-built Watson Discovery News collection and access the Watson Discovery Service through the Discovery API to retrieve the result and also perform sentimental analysis on the data to gain an overview of the wider public opinion behind certain topics.We finally integrate the Watson discovery service with slack, and query the data using slack bot.

# Theory Analysis:

## Block diagram:

APP UI

USER

SLACK

WATSON DISCOVERY NEWS

ONLINE NEWS CONTENT

## Hardware/Software Designing:

### Technical Requirements :

* Use a Slack interface to query the data
* Push news alerts out to web notification
* Deploy the app on IBM Cloud
* Integrate Slack with Watson Discovery

### Software Requirements :

IBM Cloud

IBM Watson

Node-Red

GitHub

Slack

### Hardware Requirements:

Processor

Ethernet Connection (LAN) or Wireless adapter (WiFi)

Memory (RAM)

Screen for display.

# Experimental Investigation:

To create our own News Mining Web Application, we initially create an IBM Cloud

account and provide the cloudant details to create a cloud foundry app.We then create a

Watson Discovery Service and then choose the default Watson Discovery News

Collection to be used with our application.We can even create a custom collection by

providing our own dataset.And, then we launch the discovery service to make use of the

collection.Now, to create the UI, we use a flow based development tool called Node-Red

that is used to create various flows by including a wide range of nodes present in thee

palette.In this tool, we create flow using the discovery node to query the data and

provide the output.We also perform sentimental analysis on the retrieved information to

find out the latest trends.Later, we integrate our application with Slack wherein, we can

make conversation with the slack bot and also get information from it.The slack bot

listens to our messages and also responds to it by providing suitable replies.Thus, not

only making it interactive but also interesting.

**STEPS:**

The step-by-step planned process of this project is listed as follows-

1. Begin the documentation of the project kickoff, analyze the contract properly and understand the needs of the project.

2. Gather all the technical and software requirements and set up the environment.

3. Using IBM Cloud for creating a starter application using Node Red.

4. We create an app and deploy it.

5. The we create a simple web page.

6. The we utilize the servives of Watson Discovery to build

intelligent news search app.

7.Integrate the Watson Discovery with Slack.

**Project Deiverables :**

- A Server Side Application using Node-RED

- Pre-built Watson Discovery News collection

- The Watson Discovery Service through the Discovery API

-The Integration of Slack with Watson Discovery

**Project Team :**

Member : Devi Harika Kavala

**Project Schedule :**

Kickoff at April 21, 2020

Deadline at May 20, 2020

**PROJECT DESCRIPTION IN BRIEF:**

**Module-1:** Project Planning and Kickoff

**○ Step-1:** An initial step of documentation is done by noting down all the details about the project i.e., project name, a detailed summary, team members, client, deadline, software requirements, technical requirements, functional requirements, plan of the project, deliverables and so on.

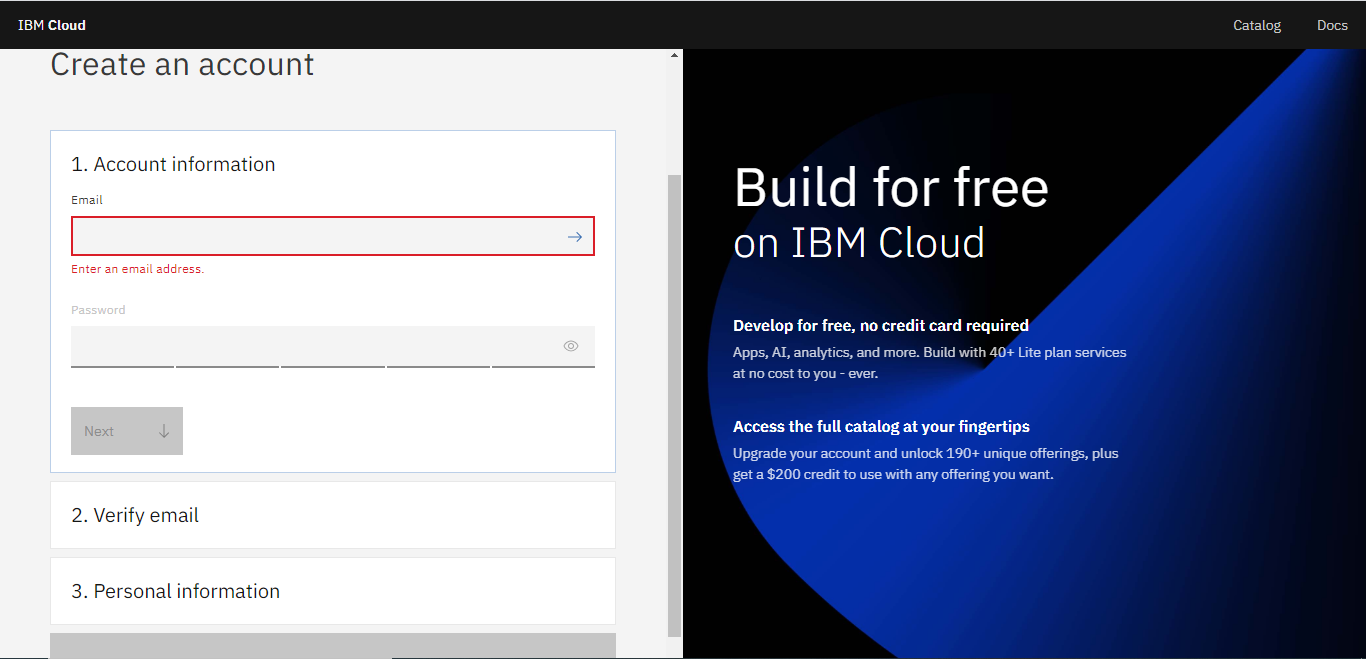
**○ Step-2:** The after listing out the whole needs and idealogies of the project, the enironment set-up that is required for the project is done. This project included Slack,IBM Cloud,IBM Watson,IBM Discovery,Zoho Writer for documentation and

Github.

# FlowChart:

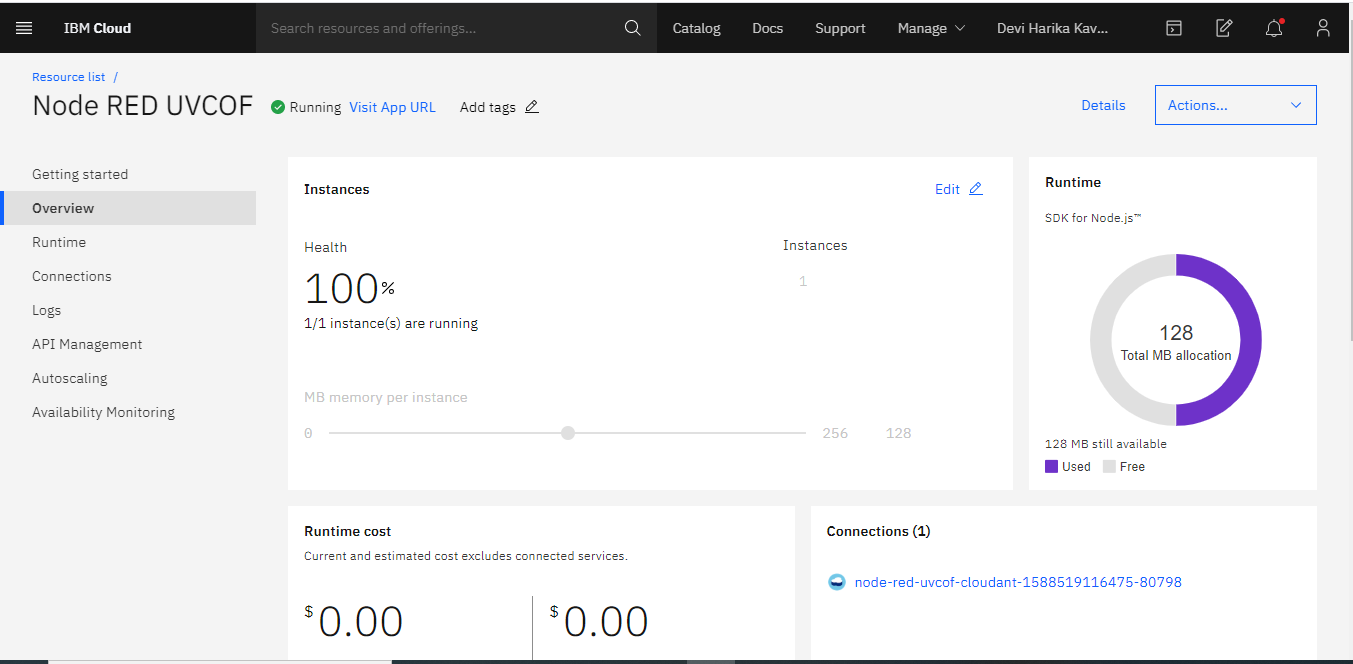
## Module-1: Explore IBM Cloud Platform:

**○ Step-1:** Creating an account with a lite plan in the IBM Cloud.



<https://cloud.ibm.com/registration>

**○ Step-2:** Creating a Node-RED starter application using one of the IBM free starter-kit software called Node-RED. Initially we created an app and deployed it through which a delivery pipeline was automatically created and a Cloudant Service was also assigned to it through which we could access Node-RED Flow Editor.



## Module-3: Explore IBM Watson Services:

**Step-1:**

* Exploring and getting to know the Watson products and services and how they are used to infuse Artificial Intelligence into our applications to make more accurate predictions.
* Getting to know how the Watsons works and how it is functioning as a leading AI platform in business.

<https://cloud.ibm.com/catalog#services>

**Step-2:** Getting familiar with Watson Discovery, its uses and functionalities.

**Step-3:** Now we explore and get familiar with the query concepts and how to make an app using the Watson Discovery News.

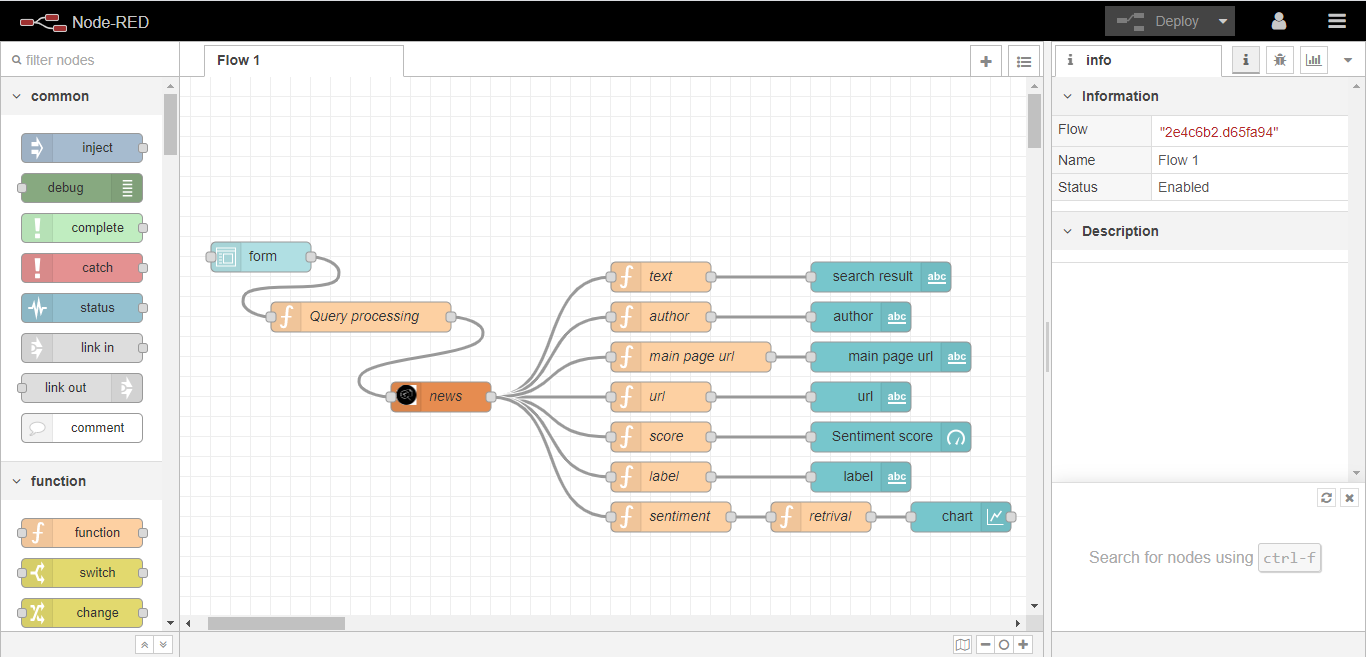
## Module-4 : Build Intelligent News Search App:

**Step-1:** Create Watson Discovery Service:

Go to the Discovery service from catalog and create a discovery service.

**Step-2:** Configure the service ad integrate it with the Node-RED starter application.

**Step-3:** Build a Node-RED user interface using the flow editor.



<https://node-red-uvcof.eu-gb.mybluemix.net/red/#flow/2e4c6b2.d65fa94>

**Step-4:**Access the news search app by simple copying the URL (i.e Node-RED user interface flow page URL) and using it followed by **/ui.**

<https://node-red-uvcof.eu-gb.mybluemix.net/ui/#!/0?socketid=wdpjVkX_A7LQbxD0AAAA>

## Module-5:Integrate Slack with Watson Discovery(locally).

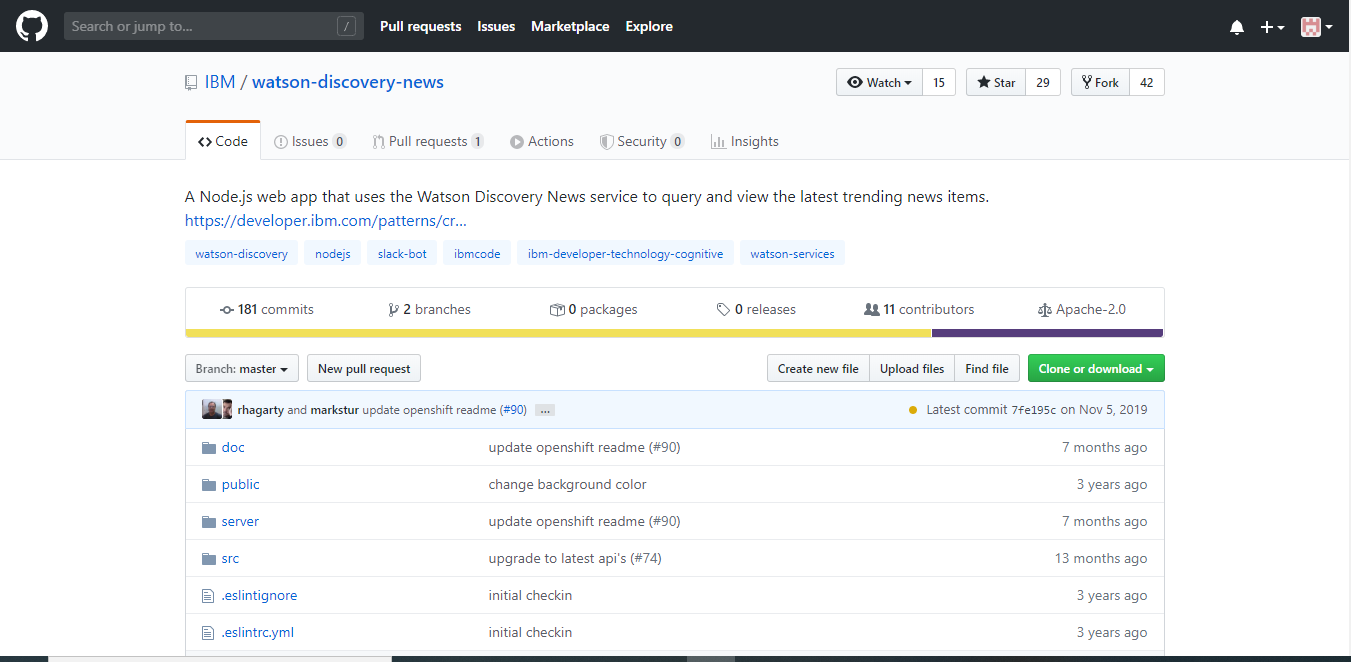
**Step1:**

Download Git for windows by using the link below.

<https://git-scm.com/download/win>

And clone the watson-discovery-news Github repository by executing the command.

<https://github.com/ibm/watson-discovery-news>



**Step2:**

From the home directory of your cloned local repo, create a .env file by copying it

from the sample version.

copy env.sample .env #Command for Windows user

**Step3:**

Copy all the service credentials listed on the homepage of yor Discovery service

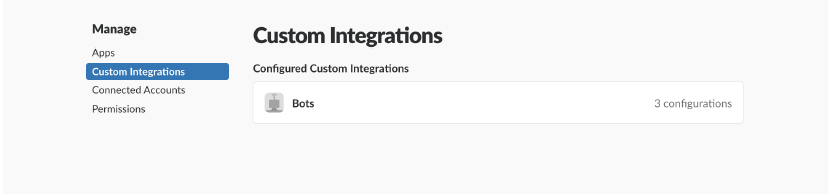
and paste it in the .env file.

**Step4:**

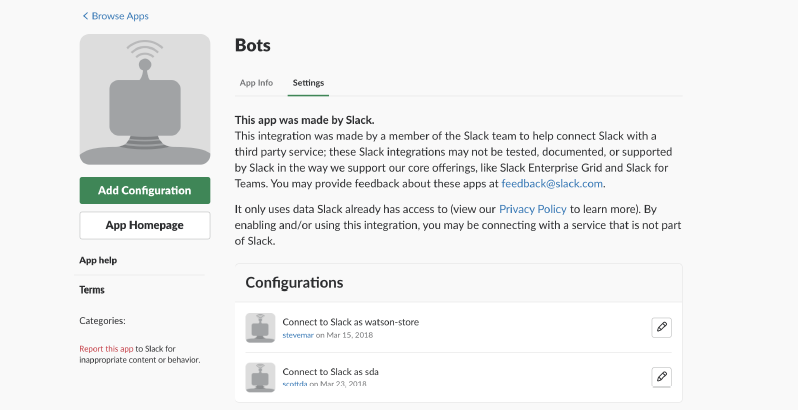
To integrate a new Slack Bot into your existing Slack team, navigate to

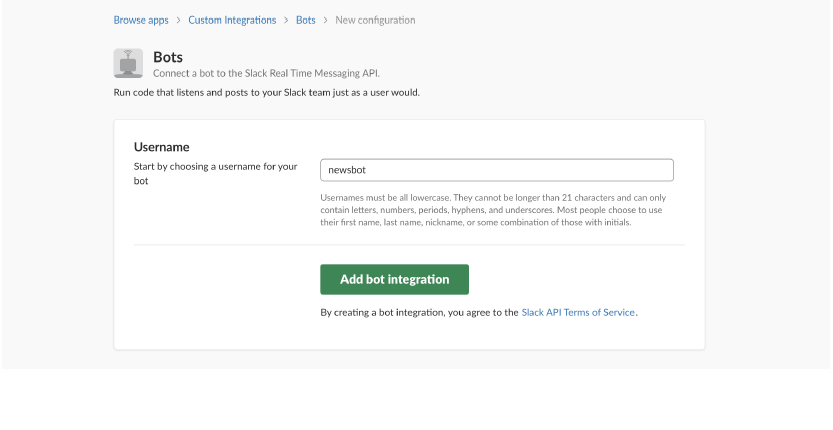
<https://smartbridgebasic.slack.com/apps/manage/custom-integrations>

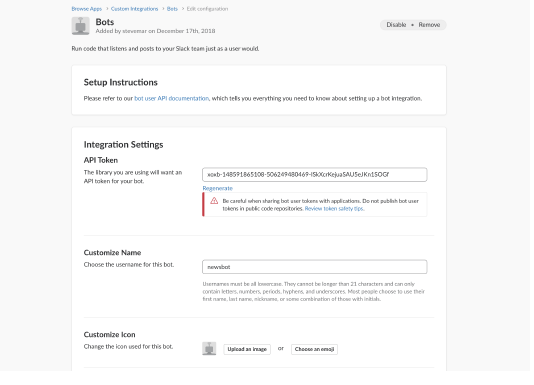
* From the Custom Integrations page, select the Bots option.



* To add a new bot, select the Add Configuration button.



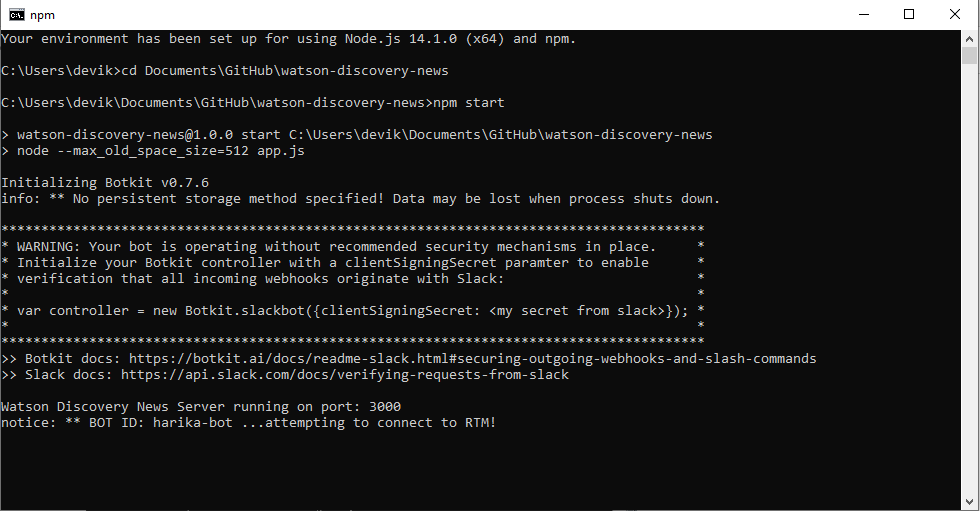
* Enter a username for the bot and click Add bot integration.
* Once created, save the API Token that is generated.



* Edit the .env file and enter the Slack Bot API Token saved in the previous step.
* Finally, run the application by executing the following commands.

npm install

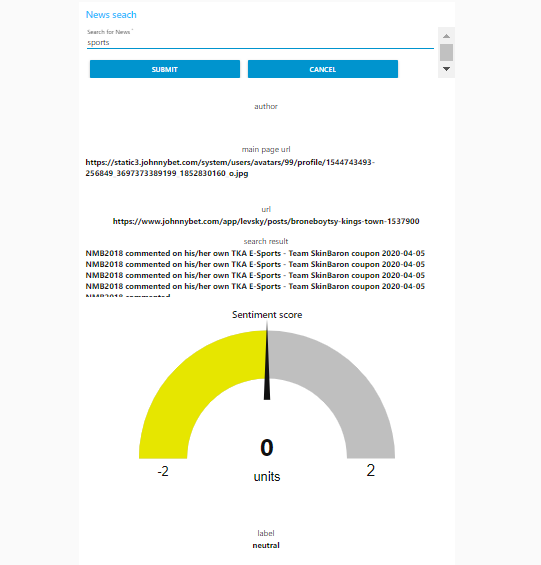
npm start

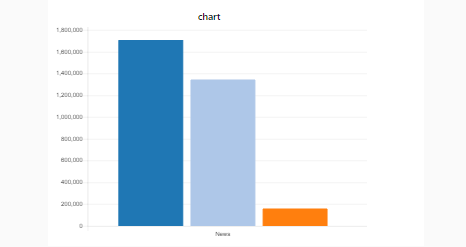


# Result:

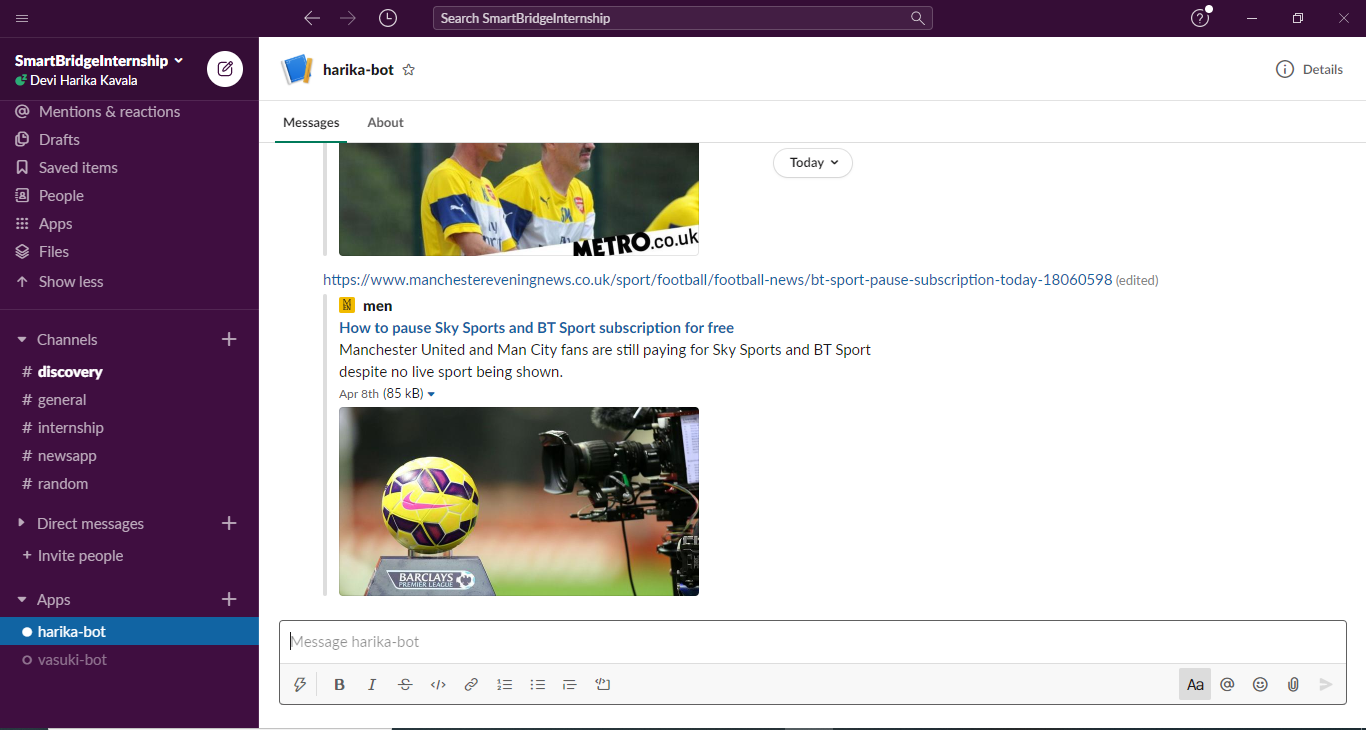
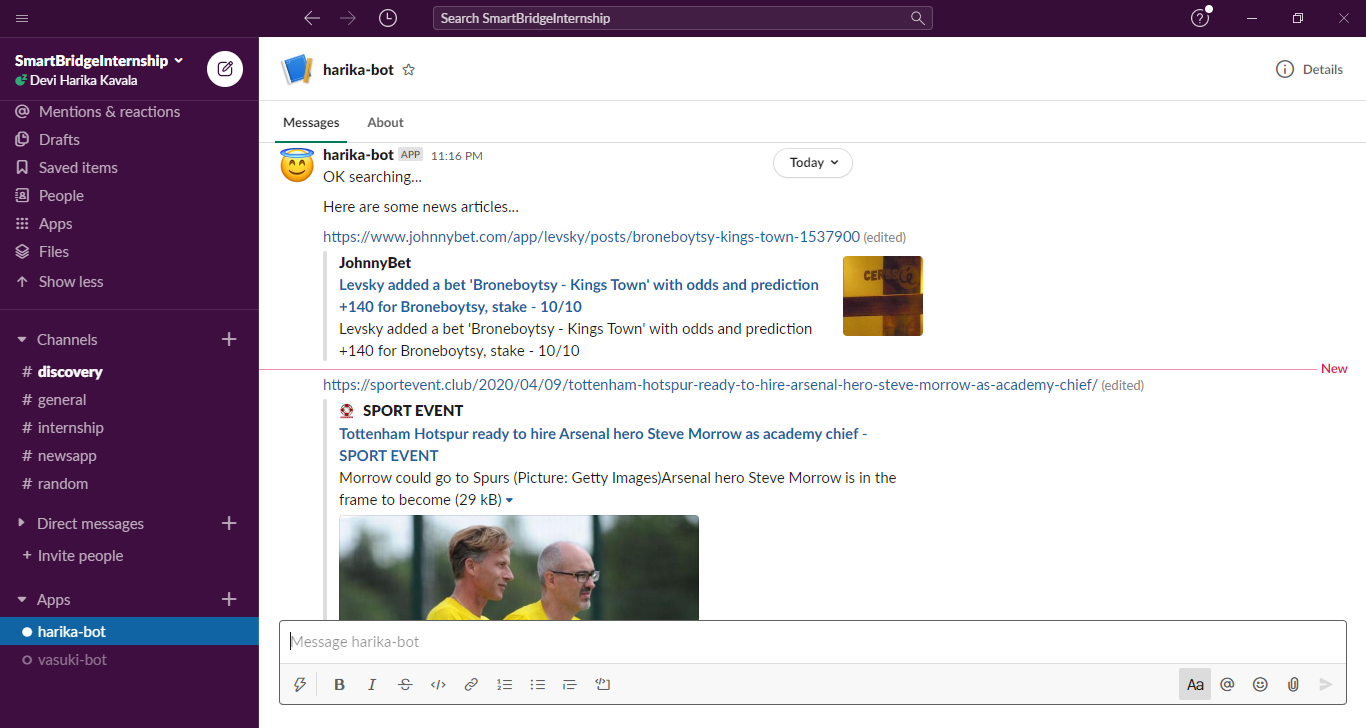
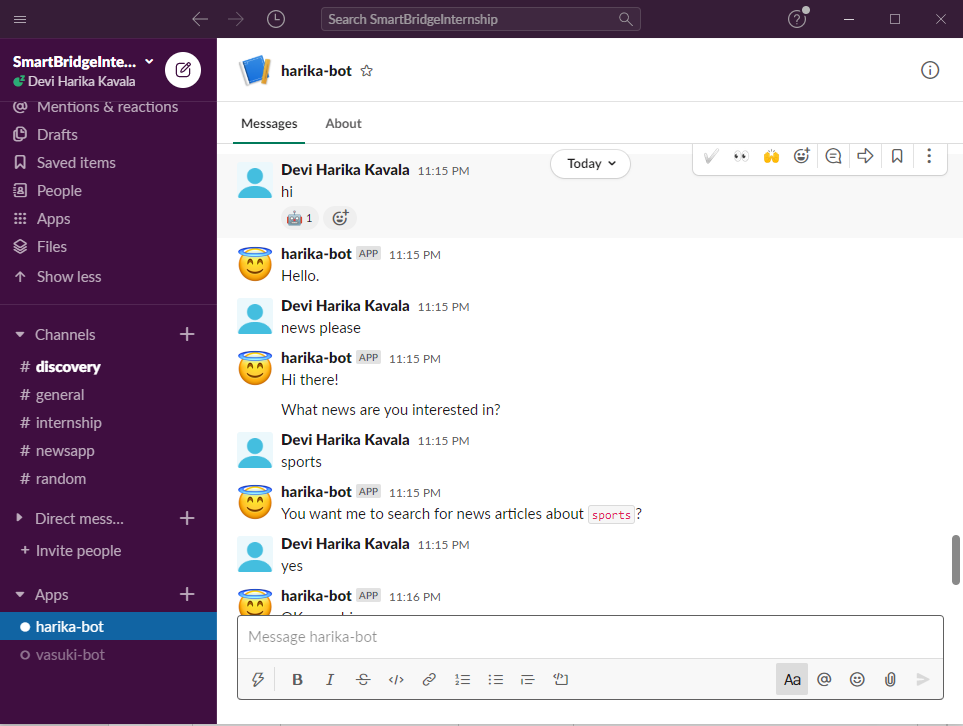
## Output of UI:

address:<https://node-red-uvcof.eu-gb.mybluemix.net/ui/#!/0?socketid=IDM8CiEGEz9HlPJvAAAB>





## Slack Output:



# Advantages and Disadvantages:

## Advantages:

* Highly accurate data is retrieved everytime the query is processed.
* Flexible access - The users can access from any part of the world.
* Always up-to-date - Since, Watson Discovery Service keeps updating the data from time to time, the user always gets the latest information.
* Integrating to slack is also highly efficient as anny request to the bot is sent to the server that starts the connection.
* With the availability of cloud, the storage space can be increased anytime.

## Disadvantages:

* Wifi and Network issues:if you happen to lose connection you will not be able to access the app.
* Not platform independent.
* The application does not process structured data directly, thus resulting in slower performance.

# Applications:

The main application of this project is to provide enriched news data about any particular topic.It not only helps to gain an understanding about it, but also proves helpful to make dynamic connections across the events in a faster way.The user can query for the most relevant new articles about a specific topic or subject. Because the news collection is pre-enriched with natural language processing, you can query not just on keywords or categories but also on concepts, sentiment, and relations to get richer search responses.We can also identify popular topics over the past 24 hours. Topics can be general or specific to an industry or category.

The Watson Discovery service also identifies important meta-information – like authors, publication dates, and relevant keywords in addition to sentimental analysis of the data.By discovering trends and patterns in sentiment with aggregate analysis, you’ll see new perspectives on how news unfolds across the globe.We can even create news alerts by taking advantage of the support for entities, keywords, categories.

# Conclusion:

This project gives you some basic working knowledge of the Watson Discovery Service and shows you how to use Discovery along with JavaScript and Node.js to build your own news mining web application.With open source toolkit Botkit, it’s easier than ever to create your own chatbot to respond to the user by giving them all the latest news.In this project, which demonstrates how to serve news articles with the Watson Discovery API, we use Botkit to create a chatbot that acts as a middleman to facilitate queries to the Watson Discovery Service.This project satisfies it's objective of creating a news mining application that provides the user with alll the latest trends and patterns.

# Future scope:

This application can further be improved by including multiple language querying.That is, the user can ask queries in multiple languages and the application can retrieve the data by interpreting the query.We can also include customer feedback option to make changes to the UI and make it much better.The application can even be integrated with Watson Assistant so that chatbots can access the Discovery service and reply to the user's questions.Even, the layout of the UI can be optimized to make it more simple and clean.Thus, making the UI more attractive.We can also develop a more Mobile optimized design so that the users can access the application even on their mobiles.The load time of the graphics can also be decreased by minifying the page and making the user experience far more pleasurable.

# Bibiliography:

**References:**

* <https://developer.ibm.com/technologies/artificial-intelligence/patterns/create-a-cognitive-news-search-app/>
* <https://www.ibm.com/watson/services/discovery-news/>
* <https://developer.ibm.com/tutorials/integrating-slack-with-watson-discovery-news/>
* <https://developer.ibm.com/articles/introduction-watson-discovery/>

# Appendex A:

**Source code:**

[{"id":"2e4c6b2.d65fa94","type":"tab","label":"Flow 1","disabled":false,"info":""},{"id":"f2f2649a.0d0d98","type":"debug","z":"2e4c6b2.d65fa94","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"payload","targetType":"msg","x":650,"y":240,"wires":[]},{"id":"a477336d.f7a54","type":"ui\_form","z":"2e4c6b2.d65fa94","name":"","label":"","group":"c3405adb.67a528","order":0,"width":0,"height":0,"options":[{"label":"Enter your Query","value":"INPUT","type":"text","required":true,"rows":null}],"formValue":{"INPUT":""},"payload":"","submit":"submit","cancel":"cancel","topic":"","x":140,"y":200,"wires":[["b359fbc.385a808"]]},{"id":"b359fbc.385a808","type":"function","z":"2e4c6b2.d65fa94","name":"","func":"msg={\n discoveryparams:\n {\n \"environment\_id\":\"system\",\n \"query\":msg.payload.input\n }\n}\nreturn msg;","outputs":1,"noerr":0,"x":260,"y":140,"wires":[["25950cf1.f14734"]]},{"id":"25950cf1.f14734","type":"watson-discovery-v1","z":"2e4c6b2.d65fa94","name":"NEWS","environmentname":"","environmentId":"system","collectionId":"news-en","configurationname":"","configurationId":"","language\_code":"en","collection\_name":"","count":"1","passages":true,"nlp\_query":true,"query":"","filter":"","aggregation":"","return":"","description":"","size":"LT","discovery-method":"query","service-endpoint":"https://api.eu-gb.discovery.watson.cloud.ibm.com/instances/858fde5a-b8c6-4829-8dbe-a4c787e312ad","x":390,"y":80,"wires":[["3ed2b49b.99d1dc"]]},{"id":"3ed2b49b.99d1dc","type":"function","z":"2e4c6b2.d65fa94","name":"","func":"msg.payload=msg.search\_results.results[0].text\nreturn msg;","outputs":1,"noerr":0,"x":540,"y":140,"wires":[["f2f2649a.0d0d98","d78a7b36.270538"]]},{"id":"d78a7b36.270538","type":"ui\_text","z":"2e4c6b2.d65fa94","group":"c3405adb.67a528","order":1,"width":"10","height":"4","name":"","label":"SEARCH RESULT","format":"{{msg.payload}}","layout":"col-center","x":730,"y":300,"wires":[]},{"id":"c3405adb.67a528","type":"ui\_group","z":"","name":"Watson Discovery News Search","tab":"f2df76ea.dc9138","order":1,"disp":true,"width":"14","collapse":false},{"id":"f2df76ea.dc9138","type":"ui\_tab","z":"","name":"IBM Watson Discovery","icon":"NEWS","disabled":false,"hidden":false}]