

**AI Powered News Search App - Rishabh Karmakar,
TheSmartBridge**

SMARTBRIDGE

REMOTE SUMMER INTERNSHIP

PROJECT REPORT

ON

AI POWERED NEWS SEARCH APP

BY

RISHABH KARMAKAR

**AI Powered News Search App - Rishabh Karmakar,
TheSmartBridge**

TABLE OF CONTENTS

- 1. Introduction**
 - a. Overview**
 - b. Purpose**
- 2. Literature Survey**
 - a. Existing Problem**
 - b. Proposed Solution**
- 3. Theoretical Analysis**
 - a. Block Diagram**
 - b. Hardware/ Software Designing**
- 4. Experimental Investigations**
- 5. Flowchart**
- 6. Result**
- 7. Advantages and Disadvantages**
- 8. Applications**
- 9. Conclusion**
- 10. Future Scope**
- 11. Bibliography**
- 12. Appendix**
 - a. Overview**

AI Powered News Search App - Rishabh Karmakar, TheSmartBridge

AI Powered News Search App (Level-1)



By: Rishabh Karmakar (SB20200000570)

(rishabh.rk1705@gmail.com)

Category: IBM Cloud Application

Skills Required: IBM Cloud, IBM Watson, GitHub

Project Description: AI Powered News Search APP (Level-1 Project)

1. INTRODUCTION

In this project, we'll start with the basics and build our own news mining web application using Node-RED / Python Web App and the IBM Watson Discovery Service.

a. Overview

To do this, we'll:

1. Build a Server Side Application using Node-RED
2. Use the pre-built Watson Discovery News collection
3. Access the Watson Discovery Service through the Discovery API

Optionally, we can choose to:

1. Use a Slack interface to query the data
2. Push news alerts out to web notification
3. Deploy the app on IBM Cloud

b. Purpose

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.

This app allows the user to search for any news at any time. The Watson discovery service used has an inbuilt collection of news in various topics which brings news to the doorstep in a jiffy.

2. LITERATURE SURVEY

AI Powered News Search App - Rishabh Karmakar, TheSmartBridge

The following observations were made while surveying the problem statement.

a. Existing problem

Searching and surfing on the web is a very common occurrence and may comprise most of our internet usage. But using popular search engines like google, bing or duckduckgo, we get only the search results listed by the SEO, but we don't get any other metadata like sentiment, score, the url from where it is submitted. And if we want to implement that, we have to take explicit permission and apply for api or even have to pay huge amount of money to just get data.

Even after getting the data, we have to create and deploy a model to perform sentiment analysis and then its integration in any web framework like Flask, Django or Node.js makes the website loading very slow and unresponsive.

So to rectify the existing problems and get the kind of 'everything-in-a-place' dashboard, we propose the below solution.

b. Proposed solution

This code pattern shows you how to tap into massive data sets to mine insight. The app demonstrates two use cases using Watson Discovery News:

- Search: 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.
- Trending topics in the news: Identify popular topics over the past 24 hours. Topics can be general or specific to an industry or category.

We propose the use of IBM Discovery, IBM Watson and IBM Node-RED.

IBM Discovery:

With IBM Watson Discovery, you can ingest, normalize, enrich, and search your unstructured data (JSON, HTML, PDF, Word, and more) with speed and accuracy. It packages core Watson APIs such as Natural Language Understanding and Document Conversion along with UI tools

AI Powered News Search App - Rishabh Karmakar, TheSmartBridge

that enable you to easily upload, enrich, and index large collections of private or public data.

The following image shows a high-level view of all of the components that make up the Discovery pipeline.

IBM Watson Platform:



Infuse Watson into your apps and workflows to tap into organizational data and put AI to work across multiple departments – from finance, to customer care, to supply chain. With Watson, you can create better, more personalized experiences for customers, scale the expertise of your best people across the organization, and make smarter decisions based on deep insights from data.

IBM Watson™ Discovery News is included with Discovery. Watson Discovery News is an indexed dataset that is pre-enriched with the following cognitive insights: **Keyword Extraction, Entity Extraction, Semantic Role Extraction, Sentiment Analysis, Relation Extraction, and Category Classification.** (To learn more about enrichments, see [Adding enrichments](#).) The following additional metadata is also added: crawl date and publication date. Historical search is available for the past 60 days of news data. See a demo of what you can build with IBM Watson™ Discovery News [here](#).

Watson Discovery News is updated continuously with new articles, and is available in English, Spanish, German, Korean, French, and Japanese. Discovery News English is updated with approximately 300,000 new articles daily; Discovery News Spanish is updated with approximately

AI Powered News Search App - Rishabh Karmakar, TheSmartBridge

60,000 new articles daily; Discovery News German is updated with approximately 15,000 new articles daily; Discovery News Korean with 10,000 new articles daily; Discovery News French with 23,000 new articles daily; Discovery News Japanese is updated with approximately 17,000 new articles daily. The news sources vary by language, so the query results for each collection are not identical.

Use cases for Watson Discovery News:

1. News alerting - Create news alerts by taking advantage of the support for entities, keywords, categories, and sentiment analysis to watch for both news and how it is perceived.
2. Event detection - The subject/action/object semantic role extraction checks for terms/actions such as "acquisition", "election results", or "IPO".
3. Trending topics in the news - Identify popular topics and monitor increases and decreases in how frequently they (or related topics) are mentioned.

IBM Node-RED:

Node-RED is a flow-based development tool for visual programming developed originally by IBM for wiring together hardware devices, APIs and online services as part of the Internet of Things. Node-RED provides a web browser-based flow editor, which can be used to create JavaScript functions.

Slack Integration:

So, finally after all this is done, we can integrate it with a slackbot so that, we can get the information using singular commands.

So, using the integration of these technologies, we can create a dashboard which on searching the news can get it from IBM Discovery, use sentiment analysis, show charts and create a fully deployable dashboard/ website on the web using NODE-RED. Finally, integrating it with slack, we

AI Powered News Search App - Rishabh Karmakar, TheSmartBridge

can get the information anywhere which is much better than the existing methods.

3. THEORITICAL ANALYSIS

Below is the Theory, including the Block Diagram and Designing.

a. Block Diagram

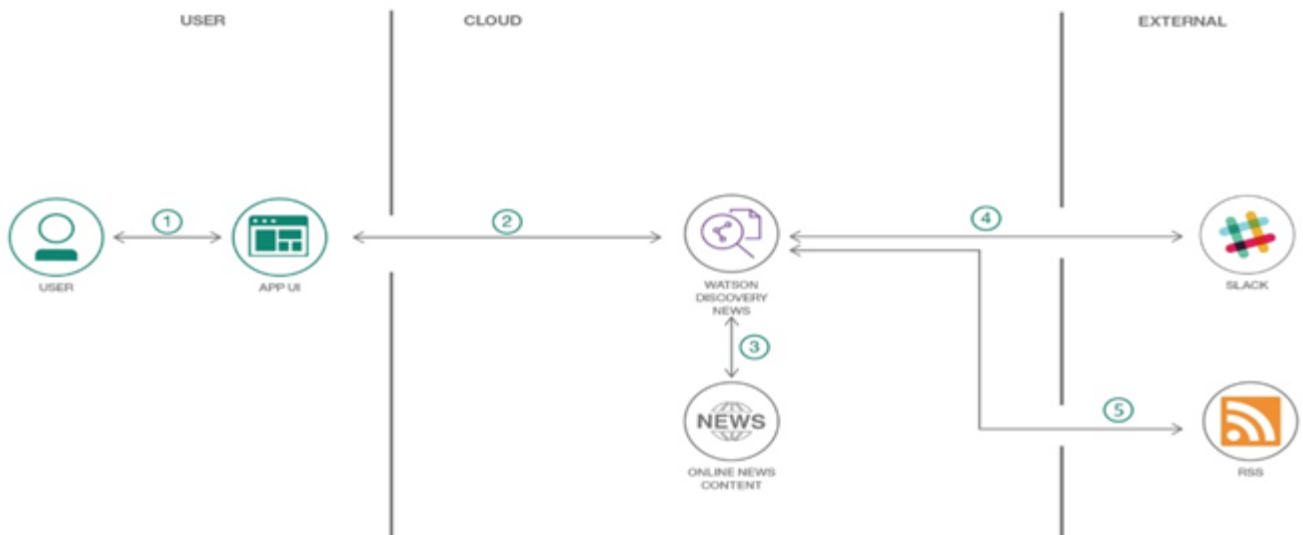
We'll:

1. Build a Server Side Application using Node-RED
2. Use the pre-built Watson Discovery News collection
3. Access the Watson Discovery Service through the Discovery API

Optionally, we choose to:

1. Use a Slack interface to query the data
2. Push news alerts out to web notification
3. Deploy the app on IBM Cloud

Flow



b. Hardware/ Software Designing

AI Powered News Search App - Rishabh Karmakar, TheSmartBridge

1. The user interacts with the app UI(Built with Node-RED or Cloud or Local) to request relevant news content.
2. The app sends user requests to Watson Discovery News.
3. The Watson Discovery Service is continually crawling the web to update its Discovery News collection.
4. The Watson Discovery Service responds to Slack search requests.

SOFTWARE:

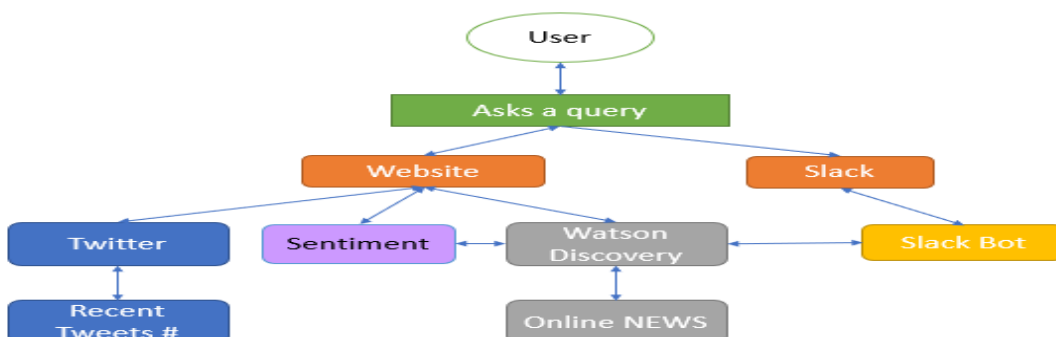
- Code is written in Node.js, with the server-side using the Express framework and the client using ReactJS.
- The pre-built Watson Discovery News collection was used.
- Access the Watson Discovery Service through the Discovery API.
- Use a Slack interface to query the data

4. EXPERIMENTAL INVESTIGATIONS

The following experimental investigations were carried out:

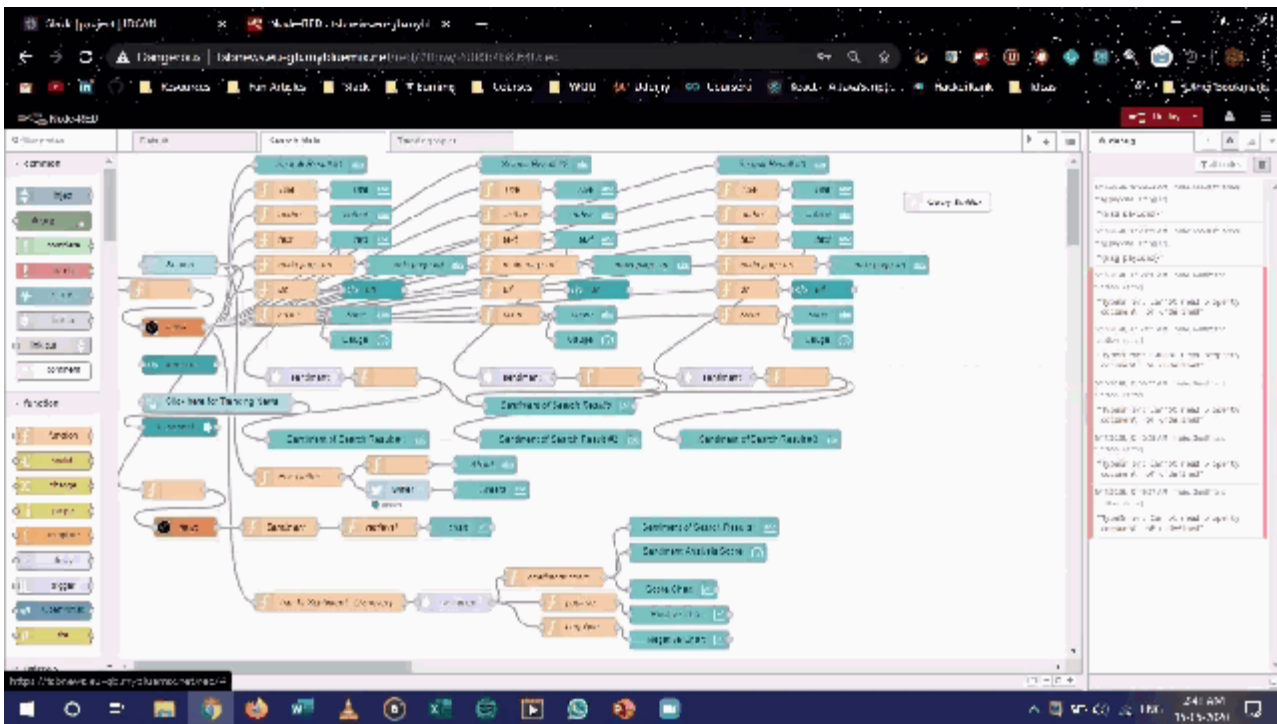
- Playing with the output of Discovery, to get different output styles
- Creating multiple outputs from discovery
- Manipulating the meta_data to get score, url, hostname, titles and so.
- Playing with msg.payload (debug) to get various outputs
- Whole sentiment analysis with taking the top 10 results and running algorithm on that
- Integrating it with twitter to get the recent tweets regarding the search term.

5. FLOWCHART



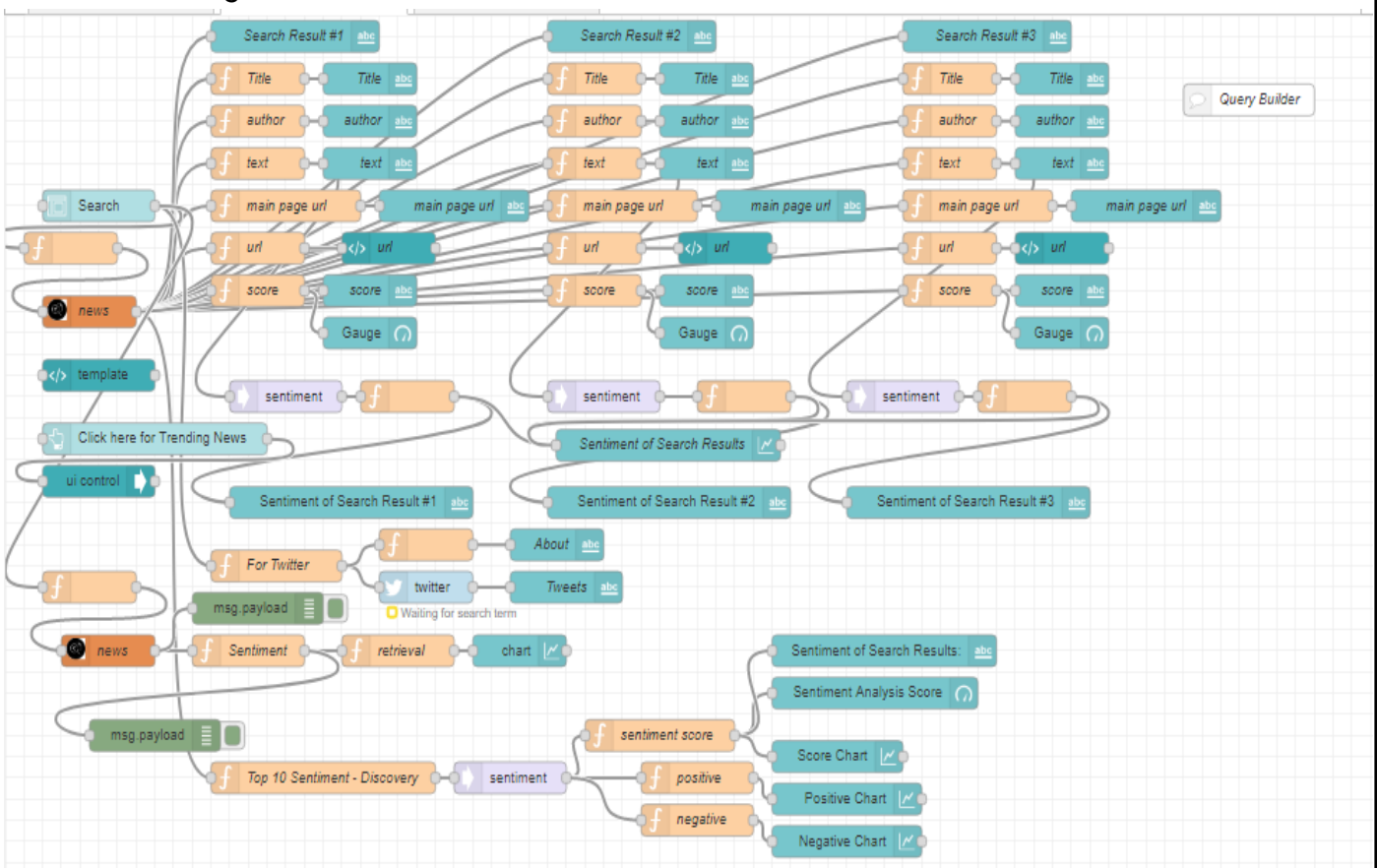
6. RESULT

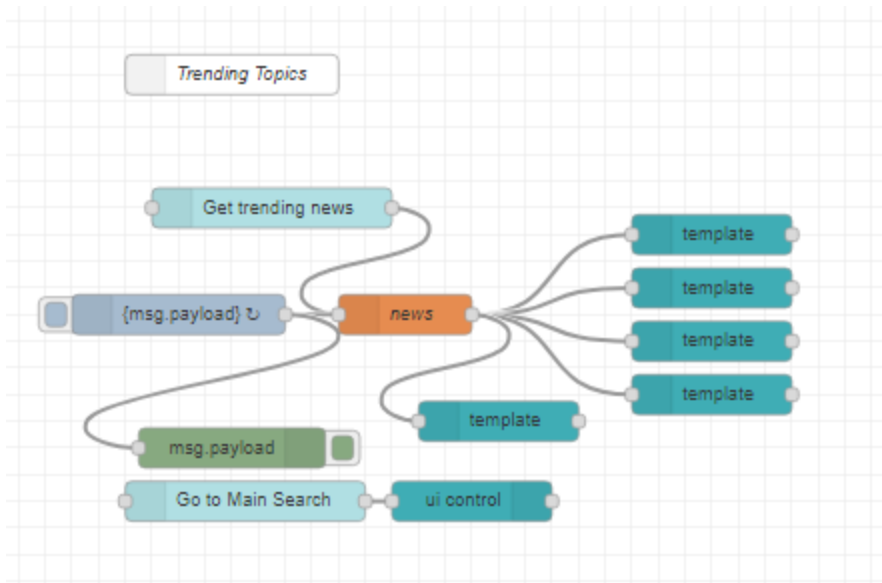
AI Powered News Search App - Rishabh Karmakar, TheSmartBridge



The NODE-RED Flow:

Main Search Page:

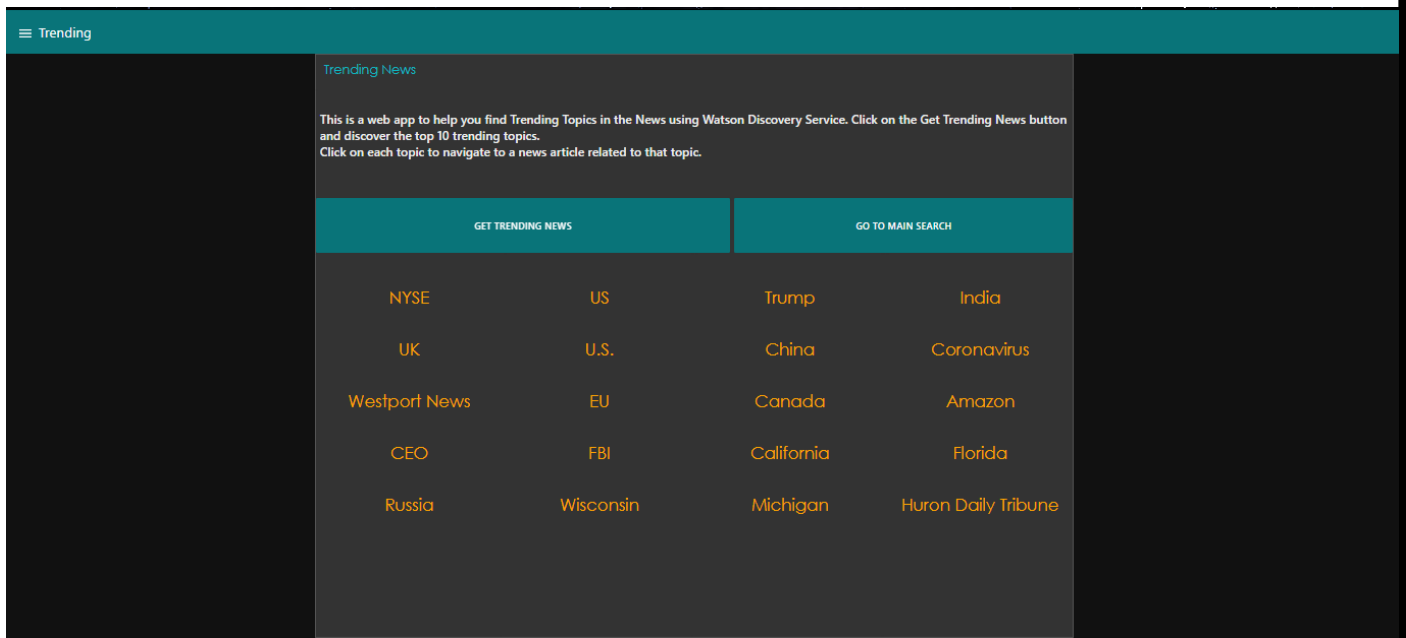




AI Powered News Search App - Rishabh Karmakar, TheSmartBridge



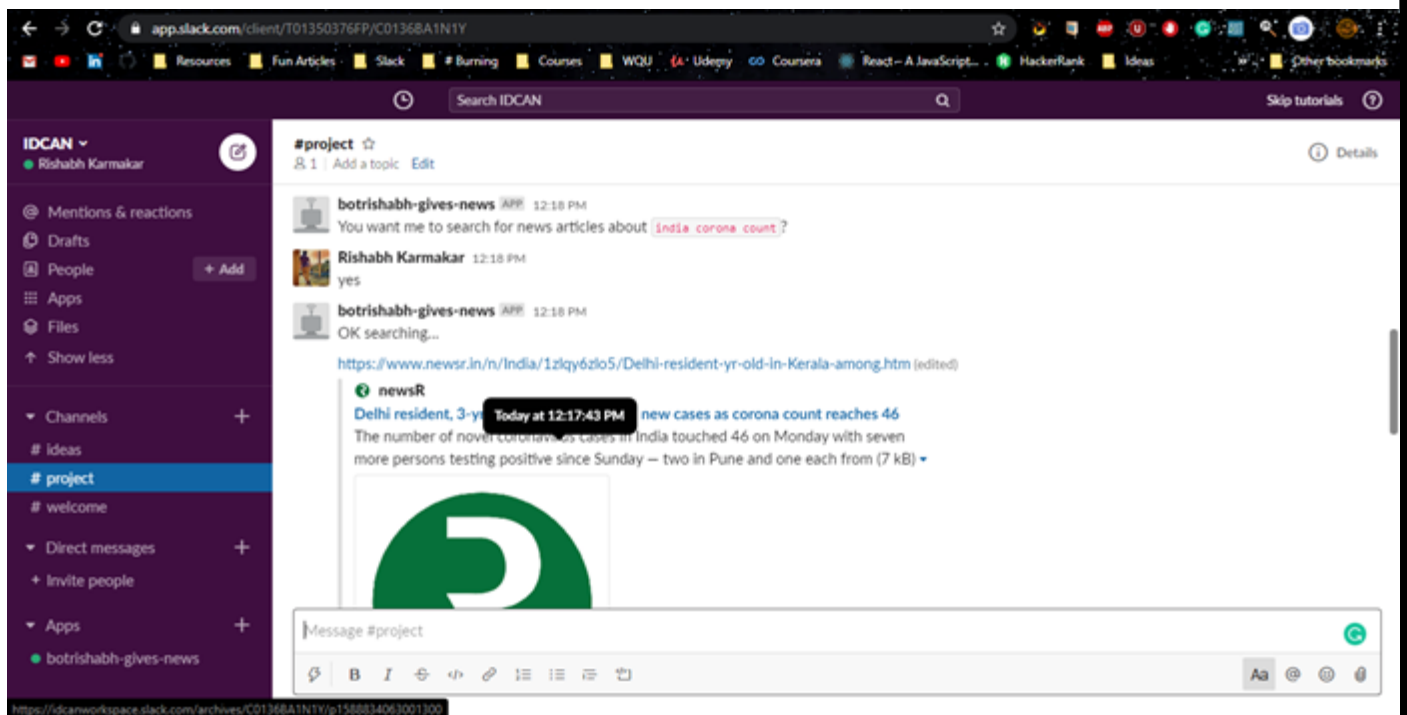
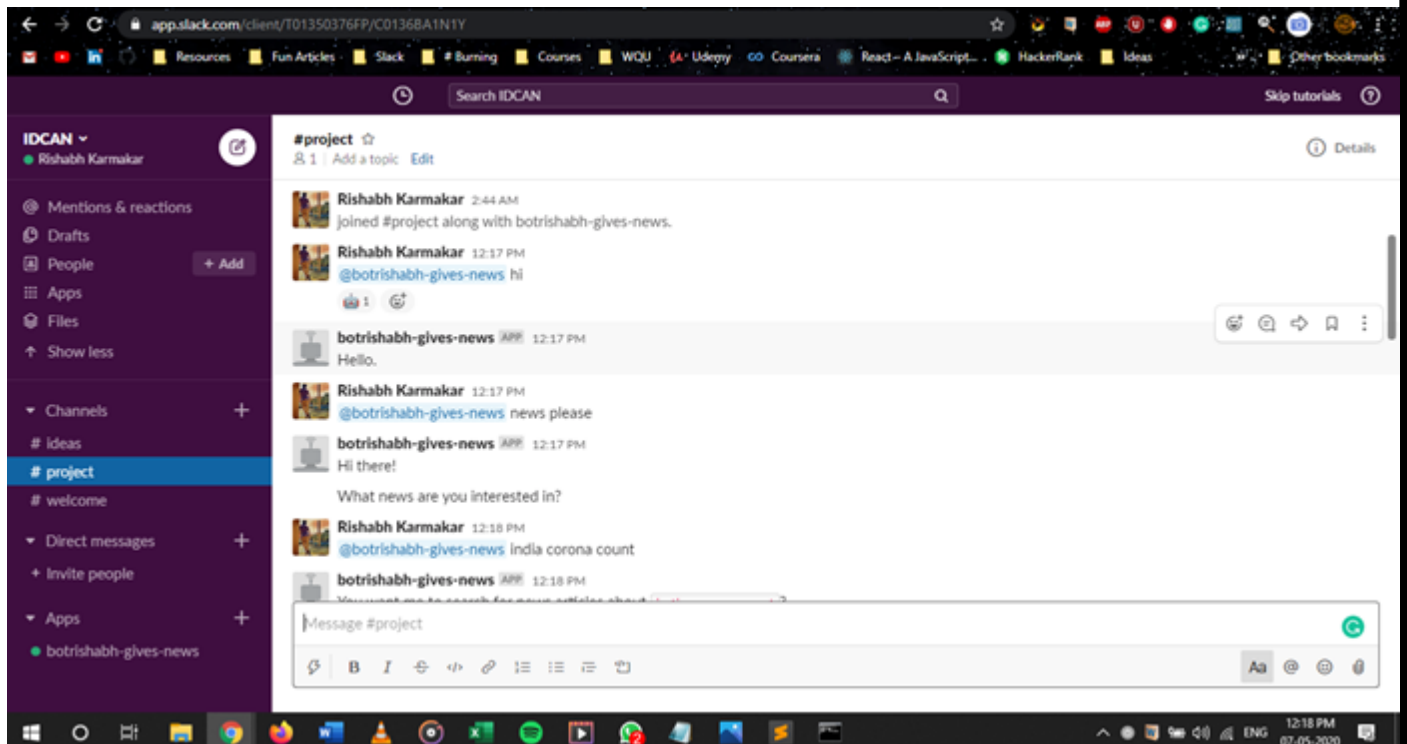
Trending News Page:



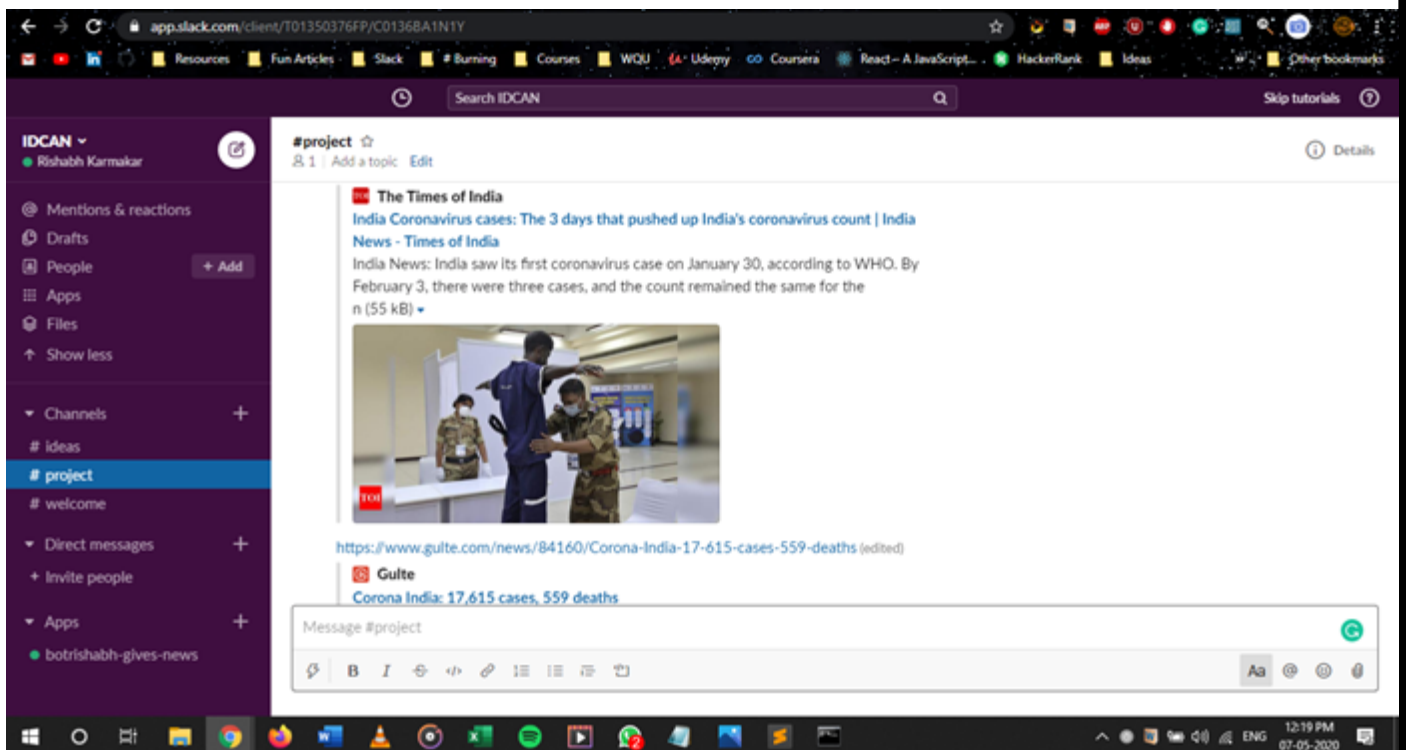
Website: <https://tsbnews.eu-gb.mybluemix.net/ui/>

AI Powered News Search App - Rishabh Karmakar, TheSmartBridge

Slack Integration:



AI Powered News Search App - Rishabh Karmakar, TheSmartBridge



7. ADVANTAGES & DISADVANTAGES

The following advantages and disadvantages are observed.

a. Advantages:

- Complete integration with Discovery, Twitter and Slack
- News coming with sentiment analysis done
- Singular commands needed for slack to get news
- Getting recent tweets real-time.
- Option to get trending news realtime

b. Disadvantages

- Currently only 3 search results, but that can be easily increased with more functions.
- Since it does store tweets and only displays recent ones, searching for non-popular keywords may find less occurent tweets.

AI Powered News Search App - Rishabh Karmakar, TheSmartBridge

8. APPLICATIONS

- A full fledged software or website with real-time applications
- Getting news articles fast and reliable with score and other metadata informations
- Sentiment analysis on the data
- Can be used as an everyday-app to browse and get news real-time, anytime in the day

9. CONCLUSION

So here, we created this application on the web with sentiment analysis, slack integration and embedded using Node-Red and Watson Discovery.

10. FUTURE SCOPE

The future scope may include:

- Adding Watson assistant to it.
- News search using audio and converting that using speech-to-text.
- Using language translator to view news in regional languages.
- Adding language translator with audio to enable search and results in regional language.

11. BIBLIOGRAPHY

<https://developer.ibm.com/tutorials/how-to-create-a-node-red-starter-application/>
<https://medium.com/segmentx/writing-a-simple-slack-bot-using-node-red-fast-6bb131af4ea1>
<http://nodeme.blogspot.com/2017/02/getting-started-with-node-red-and-iot.html>
<http://lucbors.blogspot.com/2018/09/how-to-use-node-red-to-interact-with.html>

12. APPENDIX

a. Source code

Source code can be found at:

<https://eu-gb.git.cloud.ibm.com/rishabh.rk1705/NodeREDTSmartB>

Youtube video: <https://www.youtube.com/watch?v=7wZutVgPWuo>