**1. INTRODUCTION**

**1.1 Overview**

**1.1 Project Summary:**

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.

In this project, we built a NEWS mining web application starting with the basics, using Node-RED / Python Web App and the IBM Watson Discovery Service.

To do this we start by building a Server Side Application using Node-RED then use the pre-built Watson Discovery News collection and access the Watson Discovery Service through the Discovery API. Optionally, we can choose to use a Slack interface to query the data, push news alerts out to web notification and deploy the app on IBM Cloud

**1.2 Purpose:**

The purpose of the report is to improve news fetching methodologies. Today, businesses want to know what buyers say about their brand and how they feel about their products. However, with all of the “noise” filling our email, social and other communication channels, listening to customers has become a difficult task. In this guide to sentiment analysis, you’ll learn how a bot uses artificial intelligence based approach can provide customer insight on a massive scale and ensure that you don’t miss a single conversation.

**2. LITERATURE SURVEY**

**2.1 Existing problem**

1) Too much information - It goes without saying that the distributed content production of individuals is an amazing and powerful phenomenon, but relying on search engines to index, sort and rank externally-produced content that isn't standardized for format doesn't work for the newsreader. There's still too much information and the organization is distorted by advertising dollars.

2) News is time-sensitive information – There should be constant organizing externally-produced content and figure out that new content was produced and, then organize all of the content that comes in many different structural formats and it has to passively (algorithmically) distinguish between news information (time-sensitive) and historical information that isn't time-sensitive.

3) News is geographically sensitive information - If you live in Dallas, TX do you want to have to sort through local content from Boston, MA or New York, NY or vice versa to find local news? No, you want to see time-sensitive information about what's happening in your community and NOW! While GeoRSS is out there, it's still nowhere near being adopted as a widely-used content standard.

4) Distilling quality news is difficult as link-backs aren't necessarily reflective of quality or demand. One has to filter the glut of information hitting the internet 24/7 while distilling the time-sensitive and geographically-sensitive content a newsreader wants.

**2.2 Proposed solution**

IBM has upgraded its Watson AI platform to boost the ability to understand business lingo that could be used to analyze the English language text and conversations.

The natural language processing capability represents the first commercialization of IBM Research’s Project Debater, billed as an AI system capable of debating humans. The tool is designed to help develop arguments that could ultimately be used to improve machine understanding and help analyze complex human speech. The integration of Project Debater technology with Watson includes a sentiment analysis and “summarization” tools along with the ability to cluster data by topic.

The sentiment analyzer being released this week can, for example, detect “sentiment shifters,” so-called because a negative modifier can reverse a speaker’s meaning using a phase like “hardly helpful.” The analysis tool would also enable the Watson API to identify common idioms like “calling it a day.”

The briefing tool extracts text data from different sources to provide a summary of speech and text on a specific topic. The clustering feature allows users to organize data and generate topics. IBM’s [Watson Discovery](https://www.ibm.com/cloud/watson-discovery?p1=Search&p4=p50290480283&p5=b&cm_mmc=Search_Google-_-1S_1S-_-WW_NA-_-%2Bwatson%20%2Bdiscovery_b&cm_mmca7=71700000060917569&cm_mmca8=kwd-302050444300&cm_mmca9=EAIaIQobChMInov-0NuS6AIVgbbICh1Unw65EAAYASAAEgJZ0vD_BwE&cm_mmca10=405891754210&cm_mmca11=b&gclid=EAIaIQobChMInov-0NuS6AIVgbbICh1Unw65EAAYASAAEgJZ0vD_BwE&gclsrc=aw.ds) enterprise search platform, the data organizer is aimed at subject experts in specific industries like health care, insurance and manufacturing.

**3. THEORITICAL ANALYSIS**

**3.1 Block diagram**

**SLACK**

**WATSON**

**DISCOVERY**

**NEWS**

**APP UI**

**USER**

**RSS**

**ONLINE**

**NEWS CONTENT**

**3.2 Hardware / Software designing**

IBM Cloud

Node Red

IBM Watson Discovery

Node.js

GitHub

Zoho Writer

Slack

Html

RSS

**4. EXPERIMENTAL INVESTIGATIONS**

This project used an experiment to examine differences in credibility assigned to news read in paper form and the same news read on a web site. Subsequent to exposure to six identical news stories, randomly assigned newspaper treatment and Internet treatment groups rated the credibility of the stories they read, using standard credibility measures. Generally, news stories appearing on a web site were evaluated as less credible than their identical counterparts on paper.

In the new digital era, brands and governments are targets for criticism and comment. They are challenged with how to manage large volumes of data and controversial news and how to address followers and skeptics. The beauty of the Cognitical solution is that it can deliver immediate insight with great precision and a remarkable level of detail.

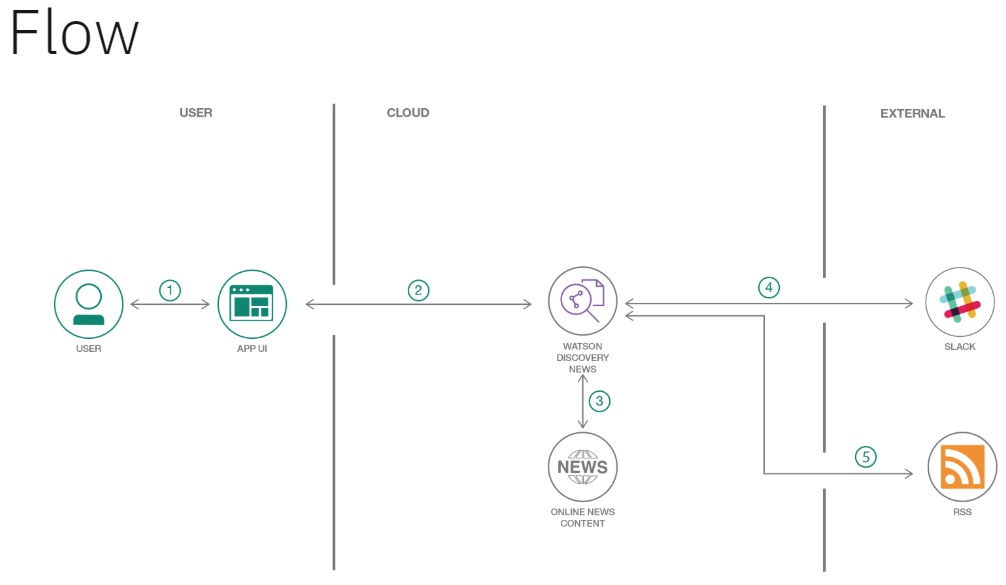
Cognitical can offer qualitative data about people’s opinions and feelings. The solution processes huge amounts of structured and unstructured data. The process is automated, so it’s fast. To compile a similar sentiment analysis report manually would take approximately a week, and by that time, the data would be old. Additionally, beyond just “listening,” Cognitical can understand sarcasm, satire and emotion.

The solutions can help brands or politicians improve how they engage, convert, and retain customers and followers on the digital front.

This project includes a bot it is the same as a regular app: it can access the same range of APIs and do all of the magical things that a Slack App can do. A bot is a type of Slack App designed to interact with users via conversation. But when you build a bot for your Slack App, you're giving that app a face, a name, and a personality and encouraging users to talk to it.

The bot can send DMs, it can be mentioned by users, it can post messages or upload files, and it can be invited to channels - or kicked out. Since bot is capable of doing everything that a Slack App can do, we're going to limit our focus to a common use-case for bots.

**5. FLOWCHART**



**6. RESULT**

The final application has

A user-interface is used to request the extraction of required news content.

The app generates the latest news regarding request generated by user using the Watson Discovery News.

It generates the URL it used to extract the news from.

It displays the author who had submitted the news report.

It provides with sentimental analysis that separate positive or negative opinion from good or bad news.

**7. ADVANTAGES & DISADVANTAGES**

**ADVANTAGES**

By querying and manipulating enriched data, you can build a more insightful search interface. This code pattern provides a Node.js app built on the Watson Discovery Service that does just that. The pattern demonstrates how you can use individual out-of-the-box UI components to extract and visualize the enriched data provided by the Discovery analytics engine.

The main benefit of using the Watson Discovery Service is its powerful analytics engine that provides cognitive enrichments and insights into your data. The app in this code pattern provides examples of how to showcase these enrichments through the use of filters, lists, and graphs. The key enrichments are:

Entities: People, companies, organizations, cities, and more

Categories: Classification of the data into a hierarchy of categories up to 5 levels deep

Concepts: Identified general concepts that aren’t necessarily referenced in the data

Keywords: Important topics typically used to index or search the data

Sentiment: The overall positive or negative sentiment of each document

The app uses standard search UI components such as filter lists, tag clouds, and sentiment graphs, but also more complex Discovery options such as the passages and highlight features. With these two features, the app identifies the most relevant snippets in your data based on your query and is more likely to return the data that you’re searching for.

**DISADVANTAGES**

Like other APIs and integrations, bot users are free. Unlike regular users, the actions they can perform are somewhat limited. For workspaces on the Free plan, each bot user counts as a separate integration.

Application Programming Interface (API) is Not Stable

One of the key problems that most of the developers encounter is the Application Programming Interface (API) keeps on changing at frequent intervals and does not remain stable.

At times, a new API appears having a number of backwards-incompatible changes. As a result the developers are forced to make changes in the accessible code bases to match the compatibility with the latest version of the Node.js API.

CHATBOT DISADVANTAGES:

• Complex Interface – Chatbots are often seen to be complicated and require a lot of time to understand user’s requirement. It is also the poor processing which is not able to filter results in time that can annoy people.

• Inability to Understand – Due to fixed programs, chatbots can be stuck if an unsaved query is presented in front of them. This can lead to customer dissatisfaction and result in loss. It is also the multiple messaging that can be taxing for users and deteriorate the overall experience on the website.

• Time-Consuming – Chatbots are installed with the motive to speed-up the response and improve customer interaction. However, due to limited data-availability and time required for self-updating, this process appears more time-taking and expensive. Therefore, in place of attending several customers at a time, chatbots appear confused about how to communicate with people.

• Increased Installation Cost – Chatbots are useful programs that help you save a lot of manpower by ensuring the all-time availability and serving to several clients at once. But unlike humans, every chatbot needs to be programmed differently for a new business which increases the initial installation cost. This also increases the time needed to prepare for the program and plan everything effectively. Considering the last-minute changes that can always happen, this is a risky investment as updating the program will invite added costs to it.

• Zero decision-making – Chatbots are known for being infamous because of their inability to make decisions. A similar situation has landed big companies like Microsoft etc. in trouble when their chatbot went on making a racist rant. Therefore, it is critical to ensure proper programming of your chatbot to prevent any such incident which can hamper your brand.

• Poor Memory – Chatbots are not able to memorize the past conversation which forces the user to type the same thing again & again. This can be cumbersome for the customer and annoy them because of the effort required. Thus, it is important to be careful while designing chatbots and make sure that the program is able to comprehend user queries and respond accordingly. Thus, chatbots are a thing of the future which is yet to uncover its potential but with its rising popularity and craze among companies, they are bound to stay here for long. Machine learning has changed the way companies were communicating with their customers. With new platforms to build various types of chatbots being introduced, it is of great excitement to witness the growth of a new domain in technology while surpassing the previous threshold.

Limitations of automated sentiment analysis

Sentiment analysis tools can identify and analyse many pieces of text automatically and quickly.

But computer programs have problems recognizing things like sarcasm and irony, negations, jokes, and exaggerations - the sorts of things a person would have little trouble identifying. And failing to recognize these can skew the results.

'Disappointed' may be classified as a negative word for the purposes of sentiment analysis, but within the phrase “I wasn't disappointed", it should be classified as positive.

We would find it easy to recognize as sarcasm the statement "I'm really loving the enormous pool at my hotel!", if this statement is accompanied by a photo of a tiny swimming pool; whereas an automated sentiment analysis tool probably would not, and would most likely classify it as an example of positive sentiment.

With short sentences and pieces of text, for example like those you find on Twitter especially, and sometimes on Facebook, there might not be enough context for a reliable sentiment analysis. However, in general, Twitter has a reputation for being a good source of information for sentiment analysis, and with the new increased word count for tweets it's likely it will become even more useful.

So, automated sentiment analysis tools do a really great job of analysing text for opinion and attitude, but they're not perfect.

When you're using a tool like Typely to analyse your text to see if it conveys the sentiment you want for your readers/audience, combine the results it gives you with your human judgement to identify anything the tool may not be able to easily determine.

Typely highlights phrases in your text by positive and negative sentiment, making it super easy for you to see where your document is either expressing exactly the sentiments you want it to, or where you may need to make some changes.

**8. APPLICATIONS**

10 Interesting Applications of Chatbots

A Chatbot is a program that can have a conversation with a person using rules and Artificial Intelligence (AI) in a way that mimics human-like conversations and interactions. Chatbots have become popular in the past few years as businesses discover innovative ways to put them to use. Having a Chatbot today has numerous benefits for businesses – they make life easier for customers, are available 24/7, save time (no more long waits to talk to a service rep) and they are easy to use.

These benefits have led to increased adoption of Chatbots by both businesses and consumers. This is great news for businesses and entrepreneurs as they can develop Chatbots to augment or improve the efficiency of any aspect of their business operations, improve customer engagement and enhance User Experience. The following examples highlight interesting applications of Chatbots in businesses across various industries –

#1. Content delivery

Media Publishers have realized that chatbots are a powerful way to engage with their audiences and monitor engagement to gain valuable insights on reader interests. Chat with the CNN and Wall Street Journal Chatbots on Facebook Messenger and receive the latest news directly in Messenger, without having to visit their websites.

#2. Order Food

Various fast food giants like KFC and Pizza Hut have invested in Chatbots that enable customers to place their orders through conversations. Taco Bell went a step further to improve the conversational experience by giving their Chatbot named Taco Bell some personality. It cracks jokes, uses emojis, answers trivia questions, and will even add a cup of water to an order if the customer mentions being hungover.

#3. Book Flights

Icelandair’s chatbot gives their customers the ability to search for and book flights in a text-based conversational manner. Instead of drop-down menus, customers enter the information themselves. These features give customers more control over how the flight is booked and it also keeps the entire conversation in one thread so that the purchase information can be reviewed and called up with ease.

#4. Companionship

Russian technology company Endurance developed its companion chatbot for Senior People and Patients with Alzheimer’s Disease. The primary function of the chatbot is to be a virtual companion – To speak with senior people on general topics like the weather, nature, hobbies, movies, music, news, etc. The chatbot asks questions, reacts to the answers, is able to speak on various topics, and share interesting news and facts from Google

#5. Market Research

The way market researchers get in touch with respondents is changing as many respondents may prefer to be contacted on WhatsApp or the Facebook messenger. Instead of having a dedicated survey app, many research firms are developing Chatbots to have personalized, engaging conversational surveys with respondents to improve the experience and increase completion rates.

#6. Transportation

Uber in partnership with Facebook has enabled users to sign up for Uber and request a ride, without having to leave Messenger or download the Uber app. Ride status updates and ride receipts are delivered to a private conversation between the customers and Uber on Messenger, making it easy to track Uber rides and payment history.

#7. Health Care

Chatbots have also made their way into health care by easing the burden on medical professionals by facilitating faster medical diagnosis, answering health-related questions, booking appointments and lots more. A Chatbot like Super Izzy can track menstrual cycles, dates and fertile windows. The bot also asks menstruation related questions and learns to give information about menstrual health and sexual issues.

#8. E-commerce

The e-commerce industry is also improving shopping experience with Chatbots. Customers can now search and shop more conveniently with the help of chatbots. eBay has invested in Chatbot technology with their Shop Bot, a virtual shopping assistant that helps customers find the items they want at the desired price range.

#9. Hospitality

Chatbots have been put to use in the hospitality industry in various ways. Radisson Blu Edwardian hotels use a virtual host called Edward to help guests with a variety of things from; online check-in, to handling customer inquiries, letting them know their rooms are ready and more.

#10. Legal

Chatbots have also shown their practicality in the Legal industry as well, they are a great way of optimizing work processes to save both lawyers and their clients time and money. Montreal-based law firm EXEO launched a smart virtual assistant called IVA to help future immigrants. IVA (Immigration Virtual Assistant) makes the process of immigration much easier and less stressful.

Applications and Benefits

Sentiment analysis has many applications and benefits to your business and organization. It can be used to give your business valuable insights into how people feel about your product brand or service.

When applied to social media channels, it can be used to identify spikes in sentiment, thereby allowing you to identify potential product advocates or social media influencers.

It can be used to identify when potential negative threads are emerging online regarding your business, thereby allowing you to be proactive in dealing with it more quickly.

**9. CONCLUSION**

This application uses individual out-of-the-box UI components to extract and visualize the enriched data provided by the Discovery analytics engine. The code pattern includes Watson Discovery series. It gets the customer sentiment insights from product reviews.

**10. FUTURE SCOPE**

Improving result relevance with the tooling:

When you provide a Discovery instance with training data, the service uses machine-learning Watson techniques to find signals in your content and questions. As you add more training data, the service instance becomes more accurate and sophisticated in the ordering of results it returns.

To train Watson, you must provide the following:

Example queries that are representative of the queries your users enter

Ratings that indicate which results for each query are relevant and not relevant

To use Continuous Relevancy Training:

Once trained, Continuous Relevancy Training is used to influence the results of a natural\_language\_query when using an environment-level query.

Continuous Relevancy Training can be used at query time by running a multi-collection natural\_language\_query across all collections in your environment

Making more complex bots

In the steps above, we made a lot of assumptions of simplicity. For example, we expected that users would respond with a very specific spelling, we assumed a test environment where there were no other conversations happening, and so on.

For a real bot in production, some of these assumptions would break the behavior of the bot. So let's cover some situations that you should address for your own bots - think of these as best practices rather than specific instructions to follow.

Tracking conversations

In our example bot, we've used a mention as the triggering point for a specific conversation, but you'll notice that your bot will still respond if you skip some of the steps - for example if you type Who's there?, your bot will respond to this message with A bot user, even if you didn't mention the bot or start at the beginning of the conversation.

A solution to this might involve tracking the beginning of a conversation, the participants involved, and the progress through the flow. For example, when the user first mentions the bot, a database entry is created that identifies that user and the open workflow with them.

As the user progresses through the flow, the database records this, and the user is unable to repeat earlier steps in the conversation (unless of course that is a desired behavior). Once the workflow is completed, the database entry is also marked as complete, and the bot waits for another mention before starting anew.

Threaded messages

Be aware that a user might choose to reply to your bot's messages in a thread rather than at the channel-level. Your bot will still receive message events for these threaded replies, but you will have to add some extra logic to ensure that your bot responds to the user in the relevant location.

Check out the Threading Messages docs for more information on how to spot the difference between messages and threaded messages.

Variations in phrasing

Because your bot will be interacting with humans, it's unlikely that you can expect consistent spelling and phrasing across messages from different people that might be trying to invoke the same thing. For example, our simple example bot used the phrase tell me a joke to trigger the start of the workflow, but at a very basic level a user might also try typing what's a good joke? or make me laugh.

Your bot can get more complex by broadening its understanding of natural language queries to capture a wider range of potential trigger phrases. Alternatively you can be more prescriptive about the exact phrasing to use, and provide user education to train correct usage.

Integrating with other services

The real magic of a bot comes when it is connected with external services, providing a seamless conversational interface for them from within Slack.

There's a huge range of possibilities for what your bot could do, so it might help to start with a great resource like the Botkit Community.

`

Usage monitoring

You can monitor and track usage of your Discovery instance and use this data to help you understand and improve your applications. The Events API can be used to create log entries that are associated with specific natural language queries and actions. For example, you can record which documents in a results set were "clicked" by a user, and when that click occurred.

**11. BIBLIOGRAPHY**

<https://www.youtube.com/watch?v=LOCkV-mENq8&feature=youtu.be>

<https://github.com/>

<https://www.youtube.com/watch?v=w3jLJU7DT5E&feature=youtu.be>

<https://www.youtube.com/watch?v=7YUTc4Cigc8&feature=youtu.be>

<https://www.ibm.com/cloud/get-started>

<https://developer.ibm.com/tutorials/how-to-create-a-node-red-starter-application/>

<https://www.w3schools.com/howto/howto_make_a_website.asp>

<https://www.youtube.com/watch?v=W3iPbFTAAds&feature=youtu.be>

<https://www.youtube.com/watch?v=W3iPbFTAAds&feature=youtu.be>

<https://discovery-news-demo.ng.bluemix.net/>

APPENDIX

A. Source code