

Report

Topic : AI Powered News Search App

Project_ID : SPS_PRO_67

INTRODUCTION

Ongoing exploration of the ways that IBM Watson® services can make sense of large datasets led us to a demonstration of exploring a live corpus of English language news stories from around the world. Watson Discovery News is an indexed dataset with news articles from the past 60 days — approximately 300,000 English articles daily. The dataset is pre-enriched with the following cognitive insights: Keyword Extraction, Entity Extraction, Semantic Role Extraction, Sentiment Analysis, Relation Extraction, and Category Classification. IBM Watson Discovery makes it possible to rapidly build cognitive, cloud-based exploration applications that unlock actionable insights hidden in unstructured data — including your own proprietary data, as well as public and third-party data. With Discovery, it only takes a few steps to prepare your unstructured data, create a query that will pinpoint the information you need, and then integrate those insights into your new application or existing solution.

In order to process all of this unstructured information quickly, the ability to understand language at scale is required.

LITERATURE SURVEY

Existing Problem

Traditional enterprise search and search engines don't provide employees and customers with exact answers. They can't understand the nuances of phrases and acronyms in industry and accurately search through complex documents in a timely manner.

Proposed solution

Watson Discovery solves these challenges. It is enterprise search that delivers specific answers to queries while also serving up the entire document and supporting links, allowing employees and customers to make informed decisions with confidence. The Proposed solution is used to broaden people's typical field of vision for news, to encourage curiosity, and to facilitate research within the ever expanding landscape of news. We wanted to invigorate the familiar routine of browsing a daily news feed while demonstrating the power of a Watson service, Watson Discovery News. One of the strengths of this service is its ability to analyze unstructured text to extract meta-data from content. This is made possible with embedded natural language processing capabilities. The service provided data such as: concepts, entities, sentiments, semantic roles, authors, publication dates, and relevant keywords with associated confidence scores. Visualizing news stories using this information help users find stories they might not have in a typical day of browsing the news. Also it include sentiment analysis of searched data.

REQUIREMENT ANALYSIS

Watson Discovery: A cognitive search and content analytics engine for applications to identify patterns, trends, and actionable insights.

Node-RED: Node-RED is built on Node.js, taking full advantage of its event-driven, non-blocking model. This makes it ideal to run at the edge of the network on low-cost hardware such as the Raspberry Pi as well as in the cloud.

Slack: Slack is a cloud-based set of team collaboration tools and services with chat bot integration

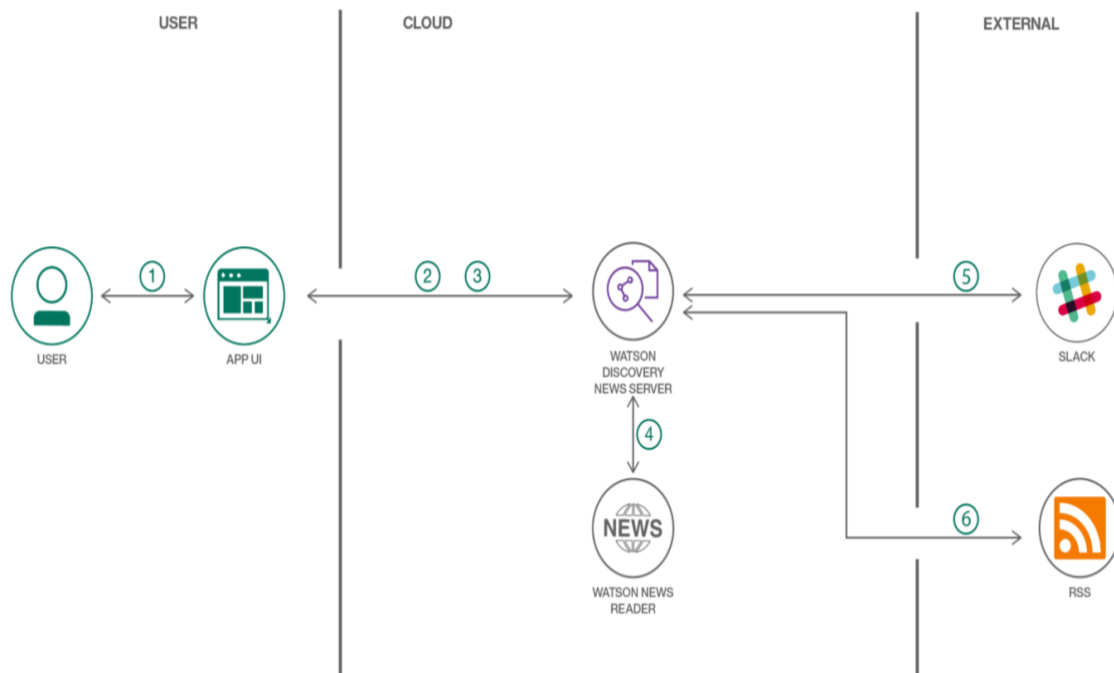
RSS: RSS (Rich Site Summary) is a format for delivering regularly changing web content in our case it will be trending topics

EXPERIMENTAL INVESTIGATIONS

This code pattern shows you how to tap into massive data sets to mine insight. You'll build a news mining web application with the Watson Discovery service using the Watson Node.js SDK. 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. Topics can be general, or specific to an industry or category.

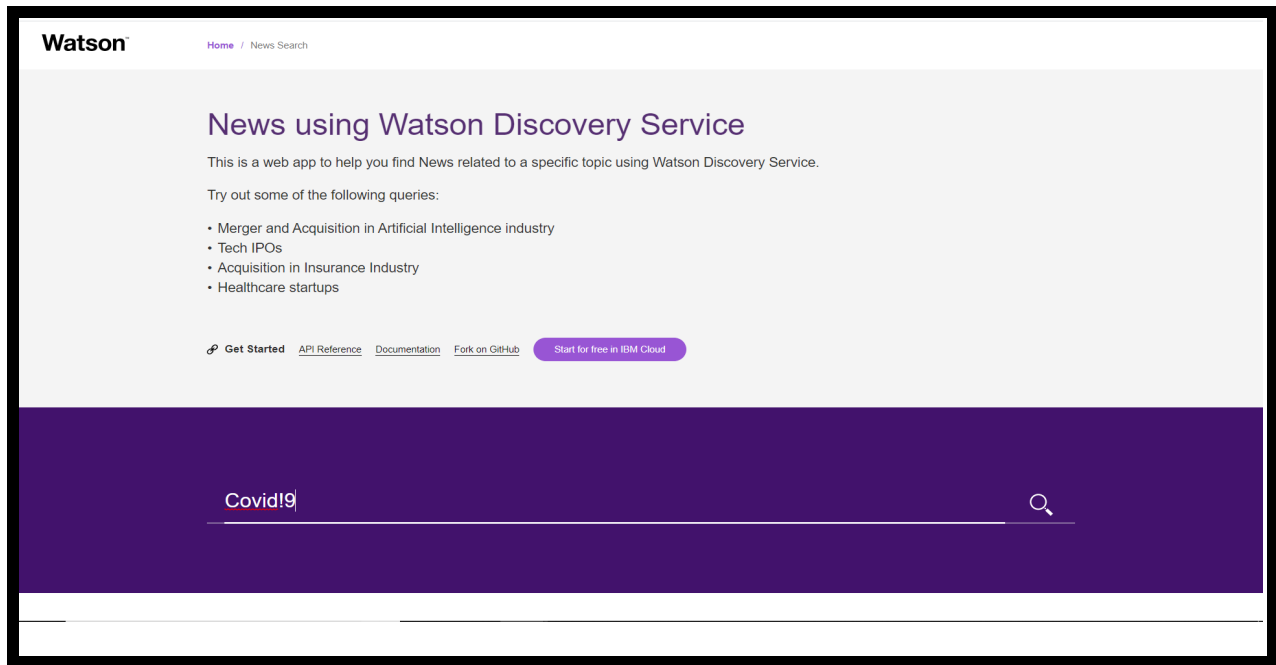
FLOWCHART



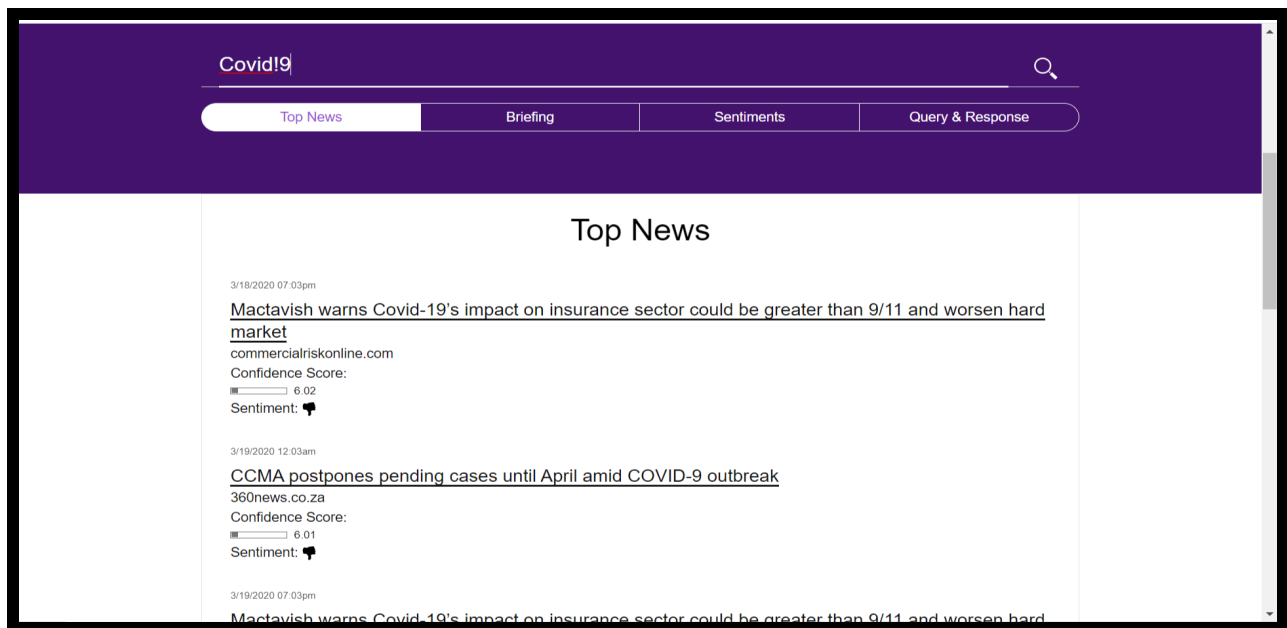
- A. The user interacts with the Watson Discovery News Server via the app UI.(Built with Node-Red or Cloud or Local) to request relevant news content.
- B. User input is processed and routed to the Watson Discovery News Server.
- C. The Watson Discovery News Server sends user requests to the Watson Discovery Service.
- D. The Watson Discovery Service queries the Watson News Collection.
- E. The Watson Discovery Service responds to Slack search requests.

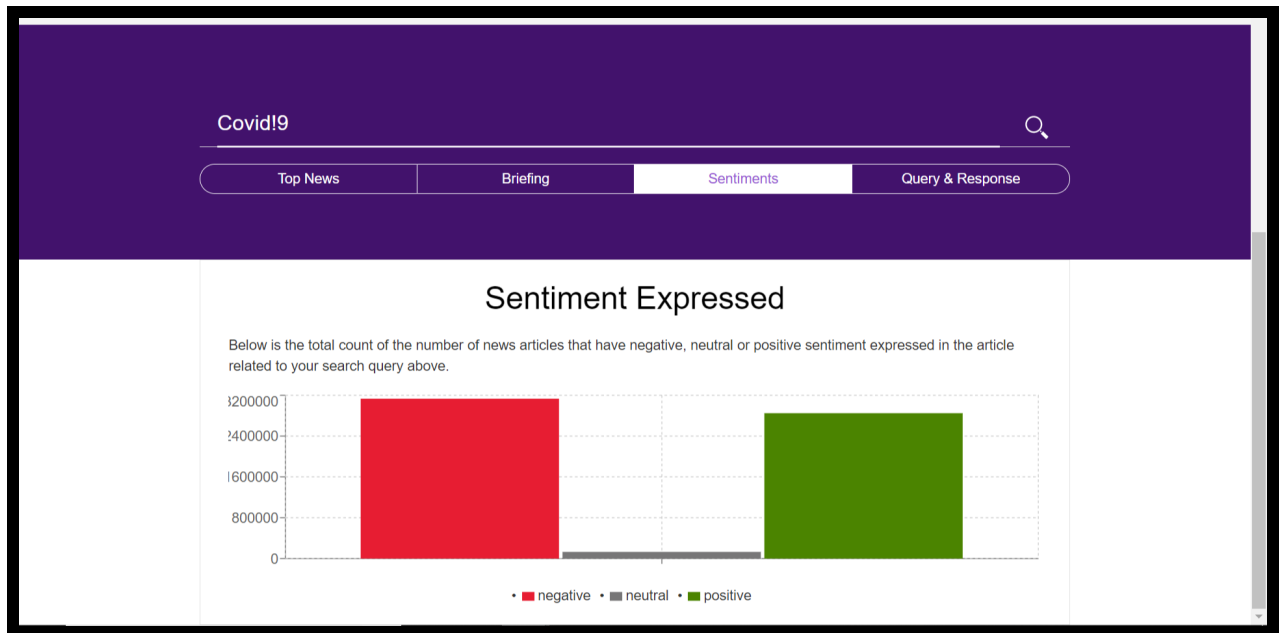
RESULTS

Input :



Output:





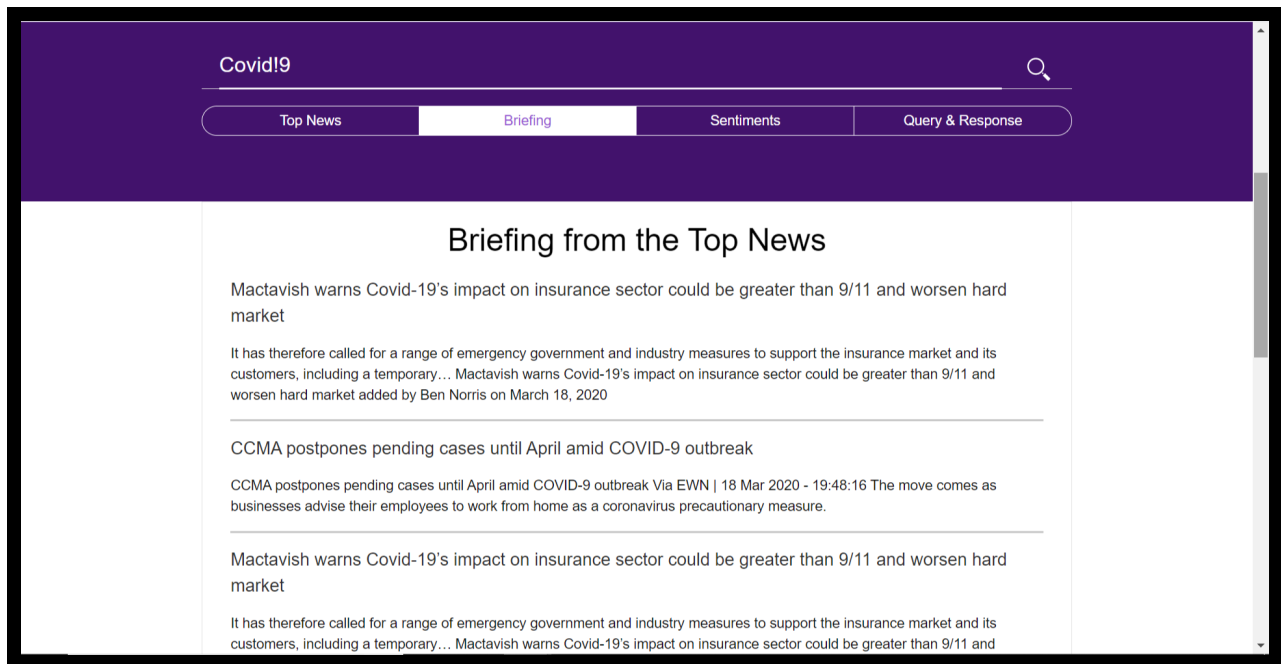
Covid19

Top News Briefing Sentiments Query & Response

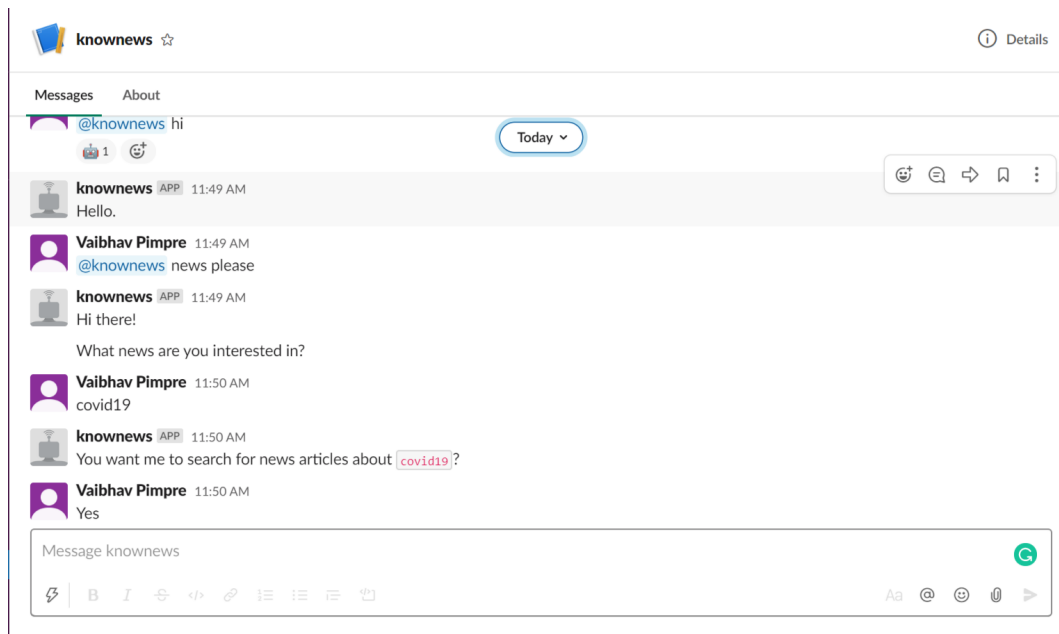
Query to and Response from the Discovery Service

Query Response


```
{
  "count": 10,
  "return": "title,text,url,host,crawl_date,result_metadata.score,id,enriched_text.entities.text,enriched_text.senti",
  "aggregation": "term(enriched_text.sentence.document.label)",
  "natural_language_query": "Covid19"
}
```




Slack Input -



Slack Output -

 knownews ☆

Messages About

 knownews APP 11:50 AM Today ▾

OK searching...

<https://newsday.co.tt/2020/05/07/survey-trinis-split-on-covid19-being-man-made/> (edited)

Trinidad and Tobago Newsday

Survey: Trinis split on covid19 being man-made - Trinidad and Tobago Newsday

TRINIDADIANs who participated in a covid19 consumer outlook study were split on whether covid19 is a man-made virus and whether it could be contracted from pets. The study, done by consultancy firm Market Facts and Opinion (MFO) and released Tuesday, was done via an online platform over a two-week period (April 8-22) and a total ... Continue reading Survey: Trinis split on covid19 being man-made



May 8th (92 kB) ▾

CHART 3.1: ATTITUDES TOWARD THE PRESENCE OF COVID-19



Statement	Percentage	Mean	Std. Dev.
I am taking extra care to avoid bringing COVID-19 home to my family	97%	2.87	0.174
I am worried about my family experiencing financial hardship due to COVID-19	90%	2.79	0.208
The economic impact of COVID-19 will be greater than the outbreak itself	70%	2.49	0.383

Message knownews

 knownews ☆

Messages About

 knownews APP Today ▾

<https://cyberwarzone.com/20-covid19-sites-03-05-2020/> (edited)

Cyberwarzone

20 COVID19 sites [03-05-2020] | Cyberwarzone

covid-entraide.fr, softsimulation.com and covid-kosova.xyz are just some of the COVID19 sites that we have listed down today. The COVID19 sites have been listed in such a way, that you can utilize the...

May 3rd (7 kB) ▾



<http://www.andamansheekha.com/84793/>

Message knownews

ADVANTAGES & DISADVANTAGES

- 1) Provide your organization with insights extracted from exploding volumes of data.
- 2) Integrate growing volumes of unstructured data to reveal previously undiscovered trends.
- 3) Rapidly scale the learning curve and overcome three obstacles organizations face when deploying cognitive solutions: identify and ingest data sources; enrich data so it's searchable; and build queries.

APPLICATIONS

- 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.
- Event detection - The subject/action/object semantic role extraction checks for terms/actions such as "acquisition", "election results", or "IPO".
- Trending topics in the news - Identify popular topics and monitor increases and decreases in how frequently they (or related topics) are mentioned.

CONCLUSION

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.

Source Code

```
[
{
  "id": "8b51f984.04d3e8",
  "type": "tab",
  "label": "Discovery",
  "disabled": false,
  "info": ""
},
{
  "id": "7e57dcd.d171d24",
  "type": "ui_tab",
  "z": 0,
  "name": "Home",
  "icon": "dashboard",
  "disabled": false,
  "hidden": false
},
{
  "id": "6e027282.6ca6cc",
  "type": "ui_group",
  "z": 0,
  "name": "News Search",
```

```

    "tab": "7e57dcd.d171d24",
    "order": 1,
    "disp": true,
    "width": "6",
    "collapse": false
  },
  {
    "id": "f2f2649a.0d0d98",
    "type": "debug",
    "z": "8b51f984.04d3e8",
    "name": "",
    "active": true,
    "tosidebar": true,
    "console": false,
    "tostatus": false,
    "complete": "true",
    "targetType": "full",
    "x": 650,
    "y": 20,
    "wires": []
  },
  {
    "id": "b200797b.0630d8",
    "type": "function",
    "z": "8b51f984.04d3e8",
    "name": "Query Processing",
    "func": "msg={\ndiscoveryparams:\n  {\n    \"environment_id\": \"system\", \n    \"query\": \"msg.payload.input\" \n  }\n}\nreturn msg",
    "outputs": 1,
    "noerr": 0,
    "x": 130,
    "y": 160,
    "wires": [
      [
        "ab7001d9.b8528"
      ]
    ]
  }
]

```

```

},
{
  "id": "6743ab39.297ef4",
  "type": "function",
  "z": "8b51f984.04d3e8",
  "name": "Receive News",
  "func": "msg.payload = msg.search_results.results[0].text\nreturn msg;",
  "outputs": 1,
  "noerr": 0,
  "x": 480,
  "y": 60,
  "wires": [
    [
      "f2f2649a.0d0d98",
      "a4edd67.465d528"
    ]
  ]
},
{
  "id": "ab7001d9.b8528",
  "type": "watson-discovery-v1",
  "z": "8b51f984.04d3e8",
  "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": "term(enriched_text.sentiment.document.label,count:10)",
  "return": ""
}

```

```
"description": "",
"size": "LT",
"discovery-method": "query",
"service-endpoint":
"https://api.us-south.discovery.watson.cloud.ibm.com/instances/2cd12cba-a32d-470c-
a957-6b0d5910f233",
"x": 290,
"y": 160,
"wires": [
  [
    "6743ab39.297ef4",
    "fda1884f.9b7758",
    "db21954f.d0b968",
    "40122435.8c932c",
    "8e015d55.5b201",
    "8d56e6d9.ac3e08",
    "baab3ef8.f7e0f",
    "fae86c43.69f4e"
  ]
],
},
{
  "id": "24965aad.2a1386",
  "type": "ui_form",
  "z": "8b51f984.04d3e8",
  "name": "",
  "label": "Form",
  "group": "6e027282.6ca6cc",
  "order": 0,
  "width": 0,
  "height": 0,
  "options": [
    {
      "label": "Search For News",
      "value": "Input",
      "type": "text",
      "required": true,
```

```
        "rows": null
      }
    ],
    "formValue": {
      "Input": ""
    },
    "payload": "",
    "submit": "submit",
    "cancel": "cancel",
    "topic": "",
    "x": 130,
    "y": 40,
    "wires": [
      [
        "b200797b.0630d8"
      ]
    ]
  },
  {
    "id": "a4edd67.465d528",
    "type": "ui_text",
    "z": "8b51f984.04d3e8",
    "group": "6e027282.6ca6cc",
    "order": 1,
    "width": 0,
    "height": 0,
    "name": "",
    "label": "Search Result",
    "format": "{{msg.payload}}",
    "layout": "col-center",
    "x": 680,
    "y": 80,
    "wires": []
  },
  {
    "id": "fda1884f.9b7758",
    "type": "function",
```

```
"z": "8b51f984.04d3e8",
"name": "Author",
"func": "msg.payload=msg.search_results.results[0].author\nreturn msg;",
"outputs": 1,
"noerr": 0,
"x": 450,
"y": 120,
"wires": [
  [
    "2d28f521.7d19aa"
  ]
]
},
{
  "id": "db21954f.d0b968",
  "type": "function",
  "z": "8b51f984.04d3e8",
  "name": "Main Page Url",
  "func": "msg.payload=msg.search_results.results[0].main_image_url\nreturn msg;",
  "outputs": 1,
  "noerr": 0,
  "x": 480,
  "y": 180,
  "wires": [
    [
      "c62cac73.cfb43"
    ]
  ]
},
{
  "id": "40122435.8c932c",
  "type": "function",
  "z": "8b51f984.04d3e8",
  "name": "url",
  "func": "msg.payload=msg.search_results.results[0].url\nreturn msg;\n",
  "outputs": 1,
  "noerr": 0,
```



```

    "x": 450,
    "y": 240,
    "wires": [
      [
        "10a900d6.1b6f5f"
      ]
    ]
  },
  {
    "id": "8e015d55.5b201",
    "type": "function",
    "z": "8b51f984.04d3e8",
    "name": "score",
    "func":
    "msg.payload=msg.search_results.results[0].enriched_text.sentiment.document.score\n
return msg;",
    "outputs": 1,
    "noerr": 0,
    "x": 450,
    "y": 300,
    "wires": [
      [
        "3e5ac325.6ee08c"
      ]
    ]
  },
  {
    "id": "8d56e6d9.ac3e08",
    "type": "function",
    "z": "8b51f984.04d3e8",
    "name": "label",
    "func":
    "msg.payload=msg.search_results.results[0].enriched_text.sentiment.document.label\nr
eturn msg;",
    "outputs": 1,
    "noerr": 0,
    "x": 450,

```

```

    "y": 360,
    "wires": [
      [
        "8fcdeae3.4ab5e8"
      ]
    ]
  },
  {
    "id": "baab3ef8.f7e0f",
    "type": "function",
    "z": "8b51f984.04d3e8",
    "name": "sentiment",
    "func": "var
val=msg.search_results.aggregations[0].results[0].matching_results;\nval.push(msg.se
arch_results.aggregations[0].results[1].matching_results);\nval.push(msg.search_result
s.aggregations[0].results[2].matching_results);\nmsg.payload=val;\nreturn msg;",
    "outputs": 1,
    "noerr": 0,
    "x": 460,
    "y": 420,
    "wires": [
      [
        "5bcd3fe4.b97b3"
      ]
    ]
  },
  {
    "id": "5bcd3fe4.b97b3",
    "type": "function",
    "z": "8b51f984.04d3e8",
    "name": "retrival",
    "func": "/*msg.payload=\n { \n    \"series\": \"News\", \n
\\data\":msg.payload[0],msg.payload[1],msg.payload[2],    \n
\\labels\": \"Positive\", \"Negative\", \"Neutral\" \n  } \nreturn msg*/\nmsg.payload=\n[ \n
{ \n    \"series\": [ \"News\" ], \n
\\data\": [msg.payload[0],msg.payload[1],msg.payload[2]], \n
\\labels\": [ \"Positive\", \"Negative\", \"Neutral\" ] \n  } \n] \nreturn msg;",

```

```
"outputs": 1,
"noerr": 0,
"x": 590,
"y": 460,
"wires": [
  [
    "2a64e7f0.4944e8"
  ]
],
},
{
  "id": "2d28f521.7d19aa",
  "type": "ui_text",
  "z": "8b51f984.04d3e8",
  "group": "6e027282.6ca6cc",
  "order": 0,
  "width": 0,
  "height": 0,
  "name": "",
  "label": "author",
  "format": "{{msg.payload}}",
  "layout": "row-spread",
  "x": 650,
  "y": 120,
  "wires": []
},
{
  "id": "c62cac73.cfb43",
  "type": "ui_text",
  "z": "8b51f984.04d3e8",
  "group": "6e027282.6ca6cc",
  "order": 0,
  "width": 0,
  "height": 0,
  "name": "",
  "label": "main page url",
  "format": "{{msg.payload}}",
```

```
"layout": "row-spread",
"x": 680,
"y": 180,
"wires": []
},
{
  "id": "10a900d6.1b6f5f",
  "type": "ui_text",
  "z": "8b51f984.04d3e8",
  "group": "6e027282.6ca6cc",
  "order": 0,
  "width": 0,
  "height": 0,
  "name": "",
  "label": "url",
  "format": "{{msg.payload}}",
  "layout": "row-spread",
  "x": 650,
  "y": 240,
  "wires": []
},
{
  "id": "3e5ac325.6ee08c",
  "type": "ui_gauge",
  "z": "8b51f984.04d3e8",
  "name": "",
  "group": "6e027282.6ca6cc",
  "order": 0,
  "width": 0,
  "height": 0,
  "gtype": "gage",
  "title": "Sentiment Score",
  "label": "units",
  "format": "{{value}}",
  "min": 0,
  "max": 10,
  "colors": [
```

```
        "#00b500",
        "#e6e600",
        "#ca3838"
    ],
    "seg1": "",
    "seg2": "",
    "x": 690,
    "y": 320,
    "wires": []
},
{
    "id": "8fcdeae3.4ab5e8",
    "type": "ui_text",
    "z": "8b51f984.04d3e8",
    "group": "6e027282.6ca6cc",
    "order": 0,
    "width": 0,
    "height": 0,
    "name": "",
    "label": "label",
    "format": "{{msg.payload}}",
    "layout": "row-spread",
    "x": 660,
    "y": 380,
    "wires": []
},
{
    "id": "2a64e7f0.4944e8",
    "type": "ui_chart",
    "z": "8b51f984.04d3e8",
    "name": "",
    "group": "6e027282.6ca6cc",
    "order": 0,
    "width": 0,
    "height": 0,
    "label": "chart",
    "chartType": "bar",
```

```
"legend": "false",
"xformat": "HH:mm:ss",
"interpolate": "linear",
"nodata": "",
"dot": false,
"ymin": "",
"ymax": "",
"removeOlder": 1,
"removeOlderPoints": "",
"removeOlderUnit": "3600",
"cutout": 0,
"useOneColor": false,
"useUTC": false,
"colors": [
  "#1f77b4",
  "#aec7e8",
  "#ff7f0e",
  "#2ca02c",
  "#98df8a",
  "#d62728",
  "#ff9896",
  "#9467bd",
  "#c5b0d5"
],
"useOldStyle": false,
"outputs": 1,
"x": 690,
"y": 500,
"wires": [
  []
]
},
{
  "id": "fae86c43.69f4e",
  "type": "debug",
  "z": "8b51f984.04d3e8",
  "name": "",
```

```
"active": true,  
"tosidebar": true,  
"console": false,  
"tostatus": false,  
"complete": "payload",  
"targetType": "msg",  
"x": 450,  
"y": 20,  
"wires": []  
}  
]
```

App URL

<https://ANewsApp.eu-gb.mybluemix.net>