Querying Watson Discovery News

Overview

In this lab, you will familiarize yourself with the Watson Discovery service by using the tooling to query Discovery News,   an existing dataset of English-language news sources that is updated continuously, with approximately 300,000 new articles and blogs added daily.

Prerequisites

You need the following accounts and software:

• An IBM Bluemix account

• An Internet Explorer, Firefox, or Chrome web browser

Section 1. Create your instance of Discovery Service

The Discovery service enriches/indexes the documents into collections and provides the options to query the collections. In this section, you'll create an instance of the Discovery Service

1. In a web browser navigate to https://bluemix.net/ and login to Bluemix

2. Click on **Catalog**
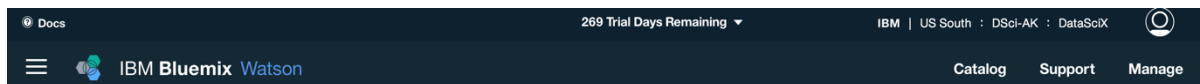
3. On the navigation menu on the left, select **Watson->Discovery**
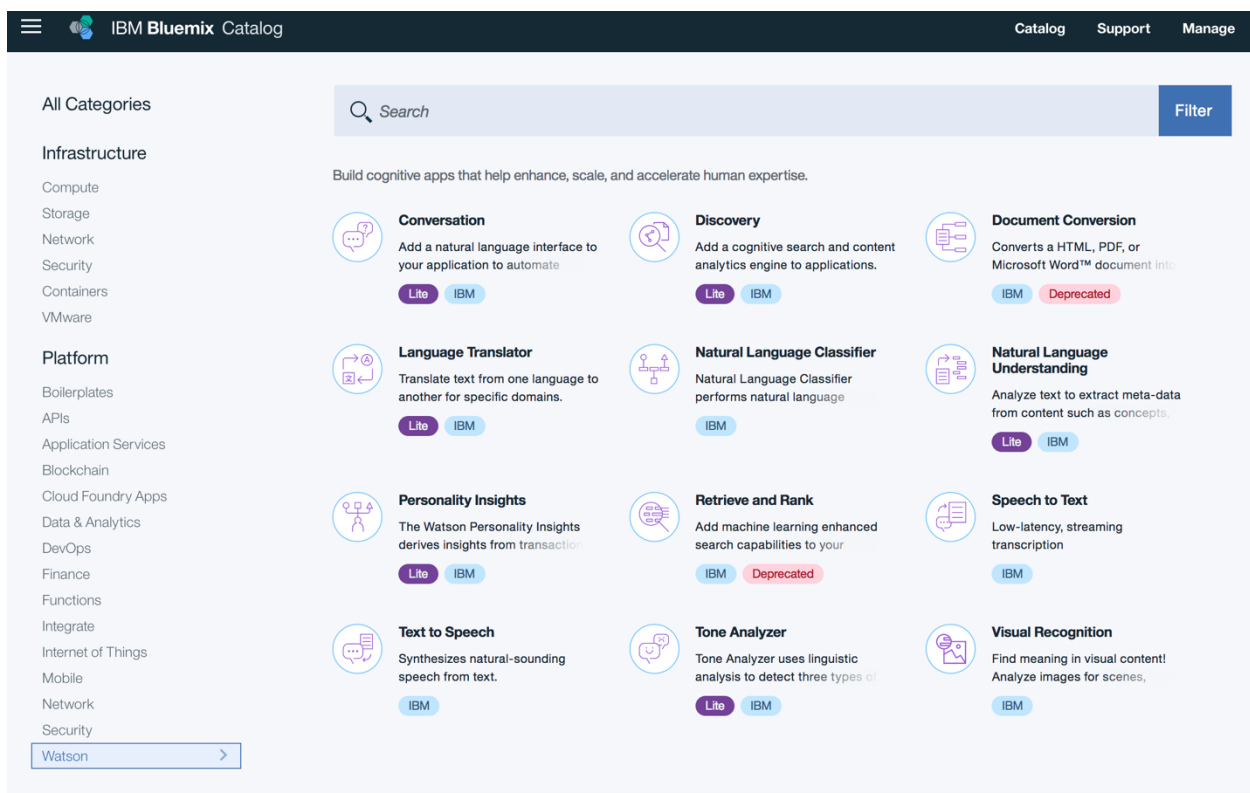
4. Enter service name, region to deploy, organization and space. Select lite plan in pricing plans

5. Click **Create**. The service is created in the respective region and space. The Apps and Service Dashboard is displayed after service creation is completed. The newly created service is displayed under "All Services".

Section 2. Launch the Discovery News Service Tool

In this lab section, you will query the Discovery News Service from the instance created above.

1. In your browser go to the Bluemix Dashboard.

2. Click on the Service name, which we created in section 1 under All Services section

3. Click **"Launch Tool"**, in Discovery Service Manage Section.

4. Click the Build Queries icon on the side navigation.
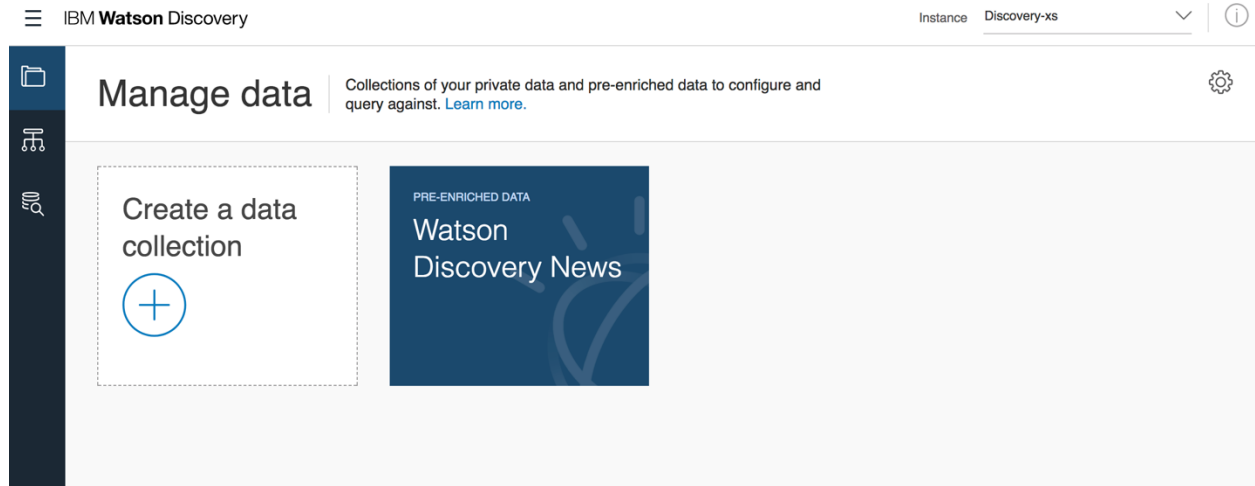
5. You should see the page "Build queries" page. Choose "Watson Discovery News" in the drop down. Click **Get started**.
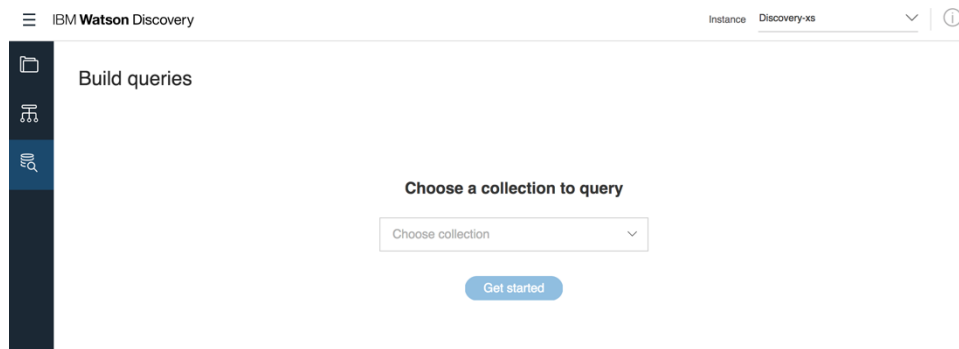
6. A page will load allowing you to build queries and see the results. Expand the components on the build side to practice with the queries at the end of this lab. Let's walk through the first one in the next section.

Build queries

Collection: Watson Discovery News

Summary   JSON

Build a query using one or more of these components. Learn more.

Use a sample query

⊕ Search for documents

⊕ Include analysis of your results

⊕ Filter which documents you query

> More options

Your results will appear here when you run a query.

Run query   Close

Instance   Discovery-xs
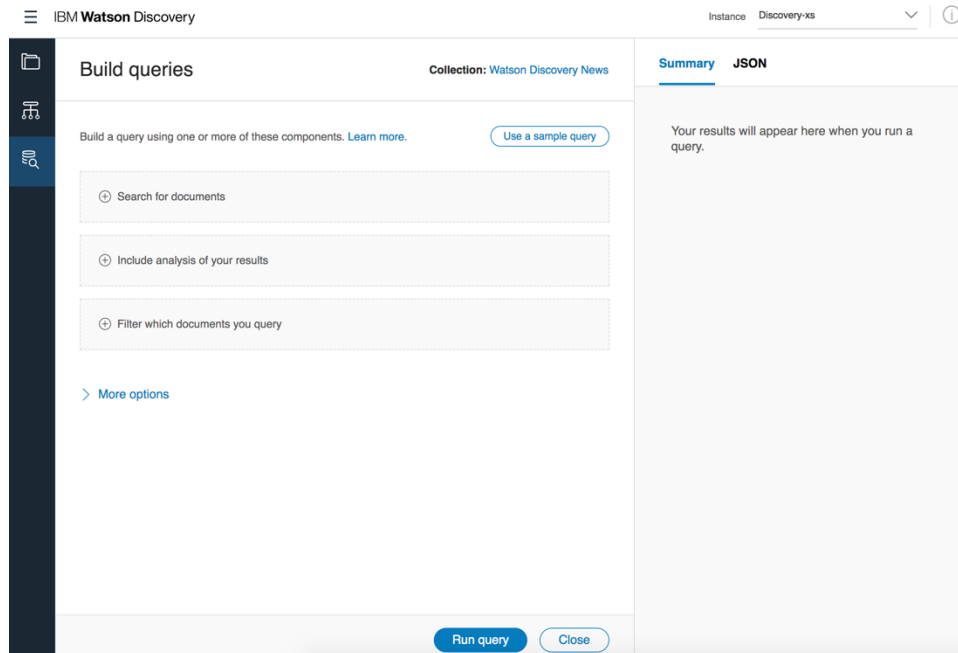
Section 3. Query the Discovery News Service

In this section, we will do a basic search for documents with "IBM in finance". The following shows the format for the queries.

Search:

IBM in finance

No aggregation.

No filter.

Resulting query string:

https://gateway.watsonplatform.net/discovery/api/v1/environments/system/collections/news/query?version=2017-08-01&count=&offset=&aggregation=&filter=&passages=true&highlight=true&return=&natural_language_query=IBM%20in%20finance

Now let's go step by step to build this query.

1. Expand the "Search for documents" component and type in "IBM in finance".

2. Leave analysis (aggregation) field blank. Leave filter field blank.

3. Click **Run query**. Your results will appear in the Summary and JSON section on the right. Click **JSON** to see the resulting query string at the top.

Build queries

Collection: Watson Discovery News

Build a query using one or more of these components. Learn more.

[Use a sample query]

Search for documents

Use natural language | Use the Discovery Query Language

IBM in finance

⊕ Include analysis of your results

⊕ Filter which documents you query

〉 More options

[Run query] [Close]

Query URL https://gateway.watsonplatform.net/discovery/api/v

Results

Showing 10 of 3366099 matching documents

⌄ **IBM, ZF and UBS Partner to Develop Blockchain Car eWallet · Crypt0 News**

Sentiment	neutral
Keywords	**IBM** blockchain technology, **IBM** Cloud
Entities	**IBM**
Categories	/**finance**/bank, /**finance**/personal finance/financial planning/asset and portfolio management
Text	"...Sep 13 **IBM**, ZF and UBS Partner to Develop Blockchain Car eWallet The platform is built on **IBM**..."
	"... blockchain technology, which is provided through the **IBM** Cloud and driven by Hyperledger Fabric 1.0, a blockchain … more ...."
Title	**IBM**, ZF and UBS Partner to Develop Blockchain Car eWallet · Crypt0 News
Url	http://crypt0.net/post/**ibm**-zf-and-ubs-partner-to-develop-blockchain-car-ewallet/

〉 Banking giants join UBS-backed IBM blockchain

Build queries

Collection: Watson Discovery News

Build a query using one or more of these components. Learn more.

[Use a sample query]

Search for documents

Use natural language | Use the Discovery Query Language

IBM in finance

⊕ Include analysis of your results

⊕ Filter which documents you query

〉 More options

[Run query] [Close]

Query URL https://gateway.watsonplatform.net/discovery/api/v

```
⌄ {
    "matching_results": 3366099,
    "passages": [],
    "results": [
      ⌄ {
          "id": "-
sMaJuwsnDUeIjZyQWNp1NE20mZA_v6b0jr_9I_McBD5NKc9U8hC
EKdBMB6gWBwW",
          "score": 4.031971,
        ▸ "enriched_title": {…},
          "crawl_date": "2017-09-13T11:38:49Z",
          "url": "http://crypt0.net/post/ibm-zf-and-
ubs-partner-to-develop-blockchain-car-ewallet/",
          "host": "crypt0.net",
          "text": "Sep 13 IBM, ZF and UBS Partner to
Develop Blockchain Car eWallet The platform is
built on IBM blockchain technology, which is
provided through the IBM Cloud and driven by
Hyperledger Fabric 1.0, a blockchain … more .",
          "country": "US",
          "source_type": "mainstream",
          "language": "en",
          "publication_date": "2017-09-13T00:00:00Z",
        ⌄ "enriched_text": {
          ▸ "entities": […],
          ▸ "sentiment": {…},
          ▸ "semantic_roles": […],
          ▸ "concepts": […],
          ▸ "categories": […],
          ▸ "relations": […],
          ▸ "keywords": […]
          },
        ▸ "extracted_metadata": {…},
          "title": "IBM, ZF and UBS Partner to Develop
Blockchain Car eWallet · Crypt0 News",
        ▸ "highlight": {…}
        },
      ⌄ {
          "id":
"vRcuejUvpgpnhsY4lTMhSrre7s0ftt0g3fLCTaMeuAHaXZlUyD
ynm8qHhcmiLMs0",
```

Using the same process, build the queries that follow. Remember to expand the component applicable to the query.

Query: 50 most mentioned companies in the tech industry

No search query.

Aggregation:

nested(enriched_text.entities).filter(enriched_text.entities.type::Company).term(enriched_text.entities.text,count:50)

Filter:

enriched_text.categories.label::/technology and computing

Resulting query string:
https://gateway.watsonplatform.net/discovery/api/v1/environments/system/collections/news/query?version=2017-08-01&count=&offset=&aggregation=nested%28enriched_text.entities%29.filter%28enriched_text.entities.type%3A%3ACompany%29.term%28enriched_text.entities.text%2Ccount%3A50%29&filter=enriched_text.categories.label%3A%3A%2Ftechnology%20and%20computing&passages=true&highlight=true&return=&query=

Query: Use sentiment analysis to see if a given host site (Gizmodo) has been publishing positive, negative, or neutral articles

Search (Use natural language):

Host::gizmodo.com

Aggregation:

term(enriched_text.sentiment.document.label)

No filter query.

Resulting query string:

https://gateway.watsonplatform.net/discovery/api/v1/environments/system/collections/news/query?version=2017-08-01&count=&offset=&aggregation=term%28enriched_text.sentiment.document.label%29&filter=&passages=true&highlight=true&return=&natural_language_query=Host%3A%3Agizmodo.com

Query: Remove the junk

Search (Use natural language):

Host::gizmodo.com, title:!deadspin

Aggregation:

term(enriched_text.sentiment.document.label)

No filter query.

Resulting query string:

https://gateway.watsonplatform.net/discovery/api/v1/environments/system/collections/news/query?version=2017-08-01&count=&offset=&aggregation=term%28enriched_text.sentiment.document.label%29&filter=&passages=true&highlight=true&return=&natural_language_query=Host%3A%3Agizmodo.com%2C%20title%3A%21deadspin

Query: Find the top 10 posts on Facebook Live with Oculus/VR

Search:

Facebook Live

Aggregation:

filter(enriched_title.entities.type::Company).top_hits(10)

Filter:

Oculus

Resulting query string:

https://gateway.watsonplatform.net/discovery/api/v1/environments/system/collections/news/query?version=2017-08-01&count=&offset=&aggregation=filter%28enriched_title.entities.type%3A%3ACompany%29.top_hits%2810%29&filter=Oculus&passages=true&highlight=true&return=&natural_language_query=Facebook%20Live