



## IBM Developer SKILLS NETWORK

# Exercise: Analyzing documents with Watson Discovery!

Objective for Exercise:

- How to create an IBM Cloud account.
- How to create a Watson Discovery service instance.
- How to utilize Watson Discovery to extract valuable information from documents.

When handling an enormous amount of text documents, Watson Discovery is a valuable tool to optimize your work flow by applying Natural Language Processing (NLP) algorithm to extract, valuable information, sentiment, concepts, semantic roles from your collection of documents.

NOTE: For learners working to get a certificate for this course (i.e. not auditing), please go through this exercise carefully and thoroughly it will be used for the graded assignment.

FYI: In order to complete this exercise you will be creating an IBM Cloud account and provisioning an instance for Watson Discovery service. A credit card is NOT required to sign up for IBM Cloud Lite account and there is no charge associated in creating a Lite plan instance of the Watson Discovery service.

## Step 1: Create an IBM Cloud Account

To access the resources and services that the IBM Cloud provides, you need an IBM Cloud account.

If you already have an IBM Cloud account, you can skip Tasks 1 and 2 and proceed with *Task 3: Login to your IBM Cloud account*.

## Step 2: Sign up for IBM Cloud

Go to: <https://cloud.ibm.com/registration> to create a free account on IBM Cloud.

Note: If you already have an IBM Cloud account, Login by going to <https://cloud.ibm.com/login> and skip ahead to Exercise 2.

Enter your company **Email** address and a strong **Password** and then click the **Next** button.

IBM Cloud

Already have an IBM Cloud account? [Log in](#)

Create an account

1. Account information

Email

you@yourcompany.com

Password

Next

↓

2. Verify email

3. Personal information

Create account

→

© Copyright IBM Corp. 2014, 2020. All rights reserved.

Build for free  
on IBM Cloud

Develop for free, no credit card required

Apps, AI, analytics, and more. Build with 40+ Lite plan services at no cost to you - ever.

Access the full catalog at your fingertips


Upgrade your account and unlock over 200 unique products, plus get a \$200 credit to use with any product you want. [Learn more.](#)

An email is sent to the address that you signed up with to confirm your email address. Check your email and copy and paste **Verification code**. Then click **Next**.

https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDDeveloperSkillsNetwork-AI0103EN-SkillsNetwork/labs/FinalAssignment-coursera/Exercise\_Analyzing\_documents\_with\_Watson\_Discovery.m...

2/14

IBM Cloud



Already have an IBM Cloud account? [Log in](#)

Create an account

1. Account information ✓Edit

.....

2. Verify email

We sent a 7-digit verification code to . This code is valid for 30 minutes. Want to create this account with a different email? [Edit email](#).

Verification code

Verification code is required.

Next

↓


Resend code

3. Personal information

Create account


Once your email is successfully verified, enter your **First name**, **Last name**, and **Country** and click **Next**.

### 1. Account information ✓

Edit 

.....

### 2. Verify email ✓

Edit 

Your email is verified.

### 3. Personal information

Lite plan services are always free, guaranteed. You'll be charged only if you use any billable plans.  
[Learn more](#)


First name


Last name


Rav

Ahuja

Country or region

Canada 

Next 

Create account 

© Copyright IBM Corp. 2014, 2020. All rights reserved.

After your personal information is accepted Click on **Create account** button.

1. Account information ✓

Edit

.....

2. Verify email ✓

Edit

Your email is verified.

3. Personal information ✓

Edit

Personal account

Rav Ahuja

Canada

Create account

→

IBM may use my contact data to keep me informed of products, services and offerings:

☐ by email

You can withdraw your marketing consent at any time by submitting an [opt-out request](#). Also you may unsubscribe from receiving marketing emails by clicking the unsubscribe link in each email.


More information on our processing can be found in the [IBM Privacy Statement](#). By submitting this form, I acknowledge that I have read and understand the IBM Privacy Statement.

I accept the product [Terms and Conditions](#) of this registration form.

© Copyright IBM Corp. 2014, 2020. All rights reserved.

It takes a few seconds to create and set up your account.

You will see this page which confirms your account creation and allows you to login.



Account setup in progress

Now proceed to login by using the button below:

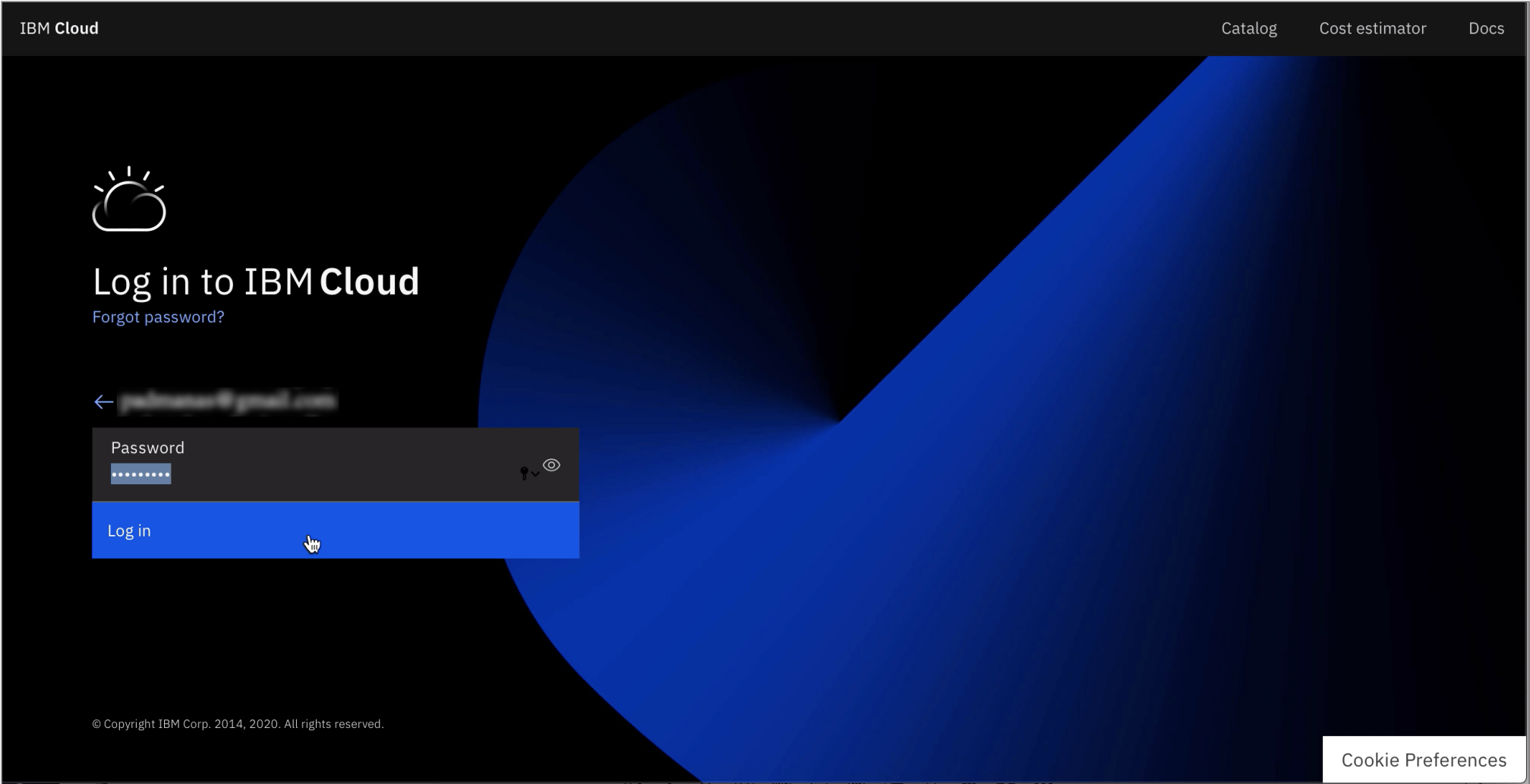
Login

Additionally, you should receive a welcome email with a login link at

If you can't log in and you don't receive an email within a few hours, [contact support](#).

### Step 3: Login to Your Account

The username (which is your email address) is already populated. Enter your password and login.



Read carefully about the IBMid Privacy and click on proceed to acknowledge and login.

About your IBMid Account Privacy

This notice provides information about accessing your IBMid user account (Account). If you have previously been presented with a version of this notice, please refer to “Changes since the previous version of this notice” below for information about the new updates.

Updates to the [IBM Privacy Statement](#) since this notice was originally published provide additional information about how your personal information is processed by IBM.

+ Changes since the previous version of this notice

+ What data does IBM collect?

+ Why IBM needs your data

+ How your data was obtained

+ How IBM uses your data

+ How IBM protects your data

+ How long we keep your data

Your rights

Our [Privacy Statement](#) provides more information about your personal data rights. It also provides contact information if you have questions or concerns regarding our handling of your personal data.

Acknowledgement

I acknowledge that I understand how IBM is using my Basic Personal Data and I am at least 16 years of age.

Proceed

Cancel Sign In

This document was last updated on 2020-03-05

## Step 4: Create a New Resource

Once you successfully login, you should see the dashboard click on the *Create resource* button. You can now proceed to the next step.

https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDDeveloperSkillsNetwork-AI0103EN-SkillsNetwork/labs/FinalAssignment-coursera/Exercise\_Analyzing\_documents\_with\_Watson\_Discovery.m...

6/14



**Dashboard**

[Upgrade account](#) [Create resource](#) +

**Resource summary** [View all](#)

Welcome!

Get started with IBM Cloud by creating services and other resources.

[Create a resource](#) →

Try spinning up an app using [starter kits](#)

**Planned maintenance** [View all](#)

Clear skies!  
You can view your scheduled maintenance events here.

**For you** ⓘ

Watson Studio provides a suite of tools and a collaborative environment for data scientists, developers and domain experts.

**News** [View all](#)

Introducing IBM Automation Document Processing

OpenShift Version 4.5 Now Available in Red Hat OpenShift on IBM Cloud

**Recent support cases** [View all](#)

FEEDBACK

## Step 5: Create a Watson Discovery Resource

On the [Catalog](#) page.

**Search for Discovery** in the search field or find it in the AI category of the catalog and then **click on that tile**.

A page will offer you to create the service. Change the name of the service or keep the default, if you wish. Select **DALLAS** as the Region

Catalog / Services /

**Discovery**

IBM • Date of last update: 11/21/2020 • [Docs](#) • [API docs](#)

[Create](#) [About](#)

Select a region

Select a region

Dallas

Select a pricing plan

Displayed prices do not include tax. Monthly prices shown are for country or region: [United States](#)

Plan	Features	Pricing
Lite	0 - 1,000 documents per month 200 news queries per month 1 custom model <a href="#">See documentation for plan details</a>	Free

The Lite plan is a great starter plan to trial features at no cost. With 1000 documents, 200 news queries, and 1 custom model, test-drive the service's foundational capabilities. When you upgrade to a paid plan, you'll keep any content you have ingested. See documentation for details.

Lite plan services are deleted after 30 days of inactivity.

[Create](#)

[Add to estimate](#)

[View terms](#)

**Summary**

**Discovery** **Free**

Region: Dallas

Plan: Lite

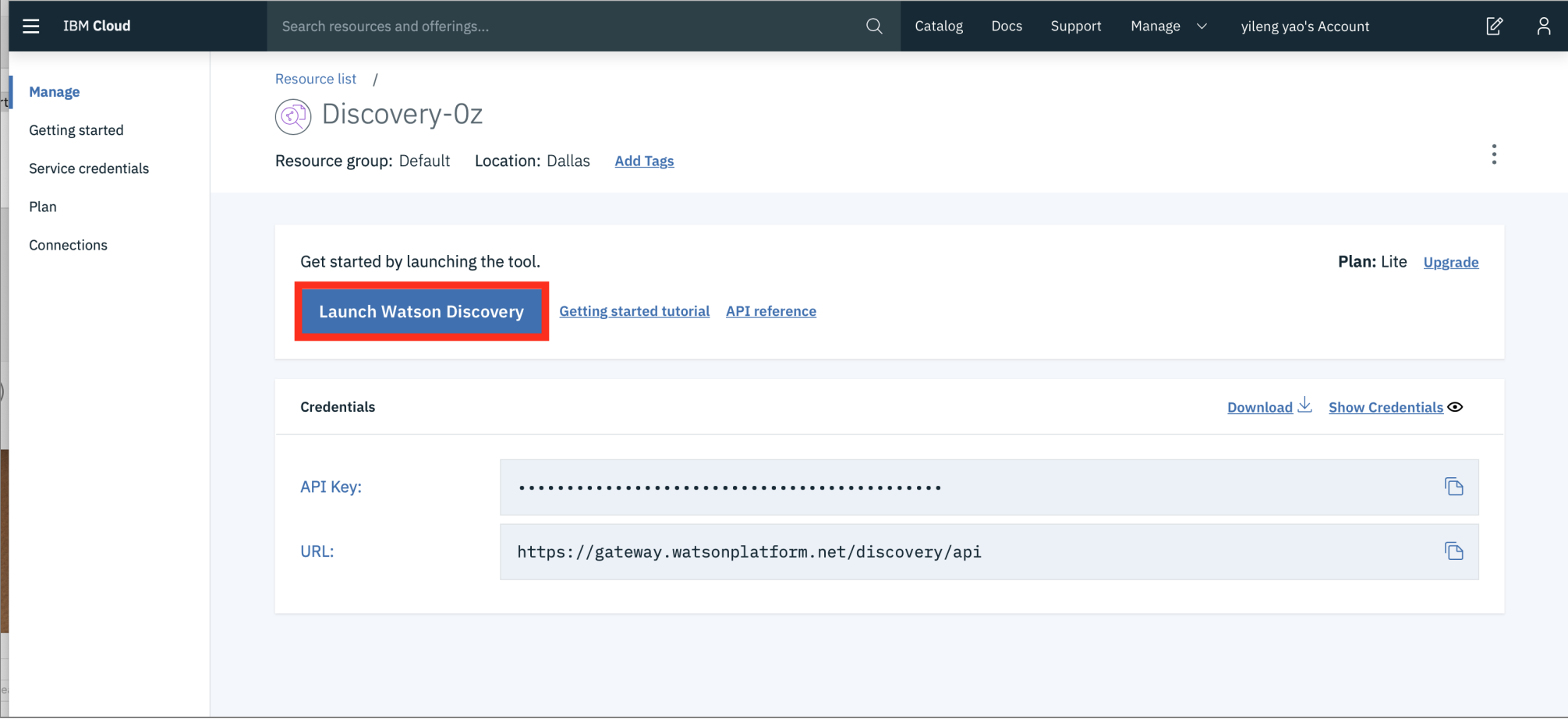
Service name: Discovery-yh

Resource group: Default

FEEDBACK

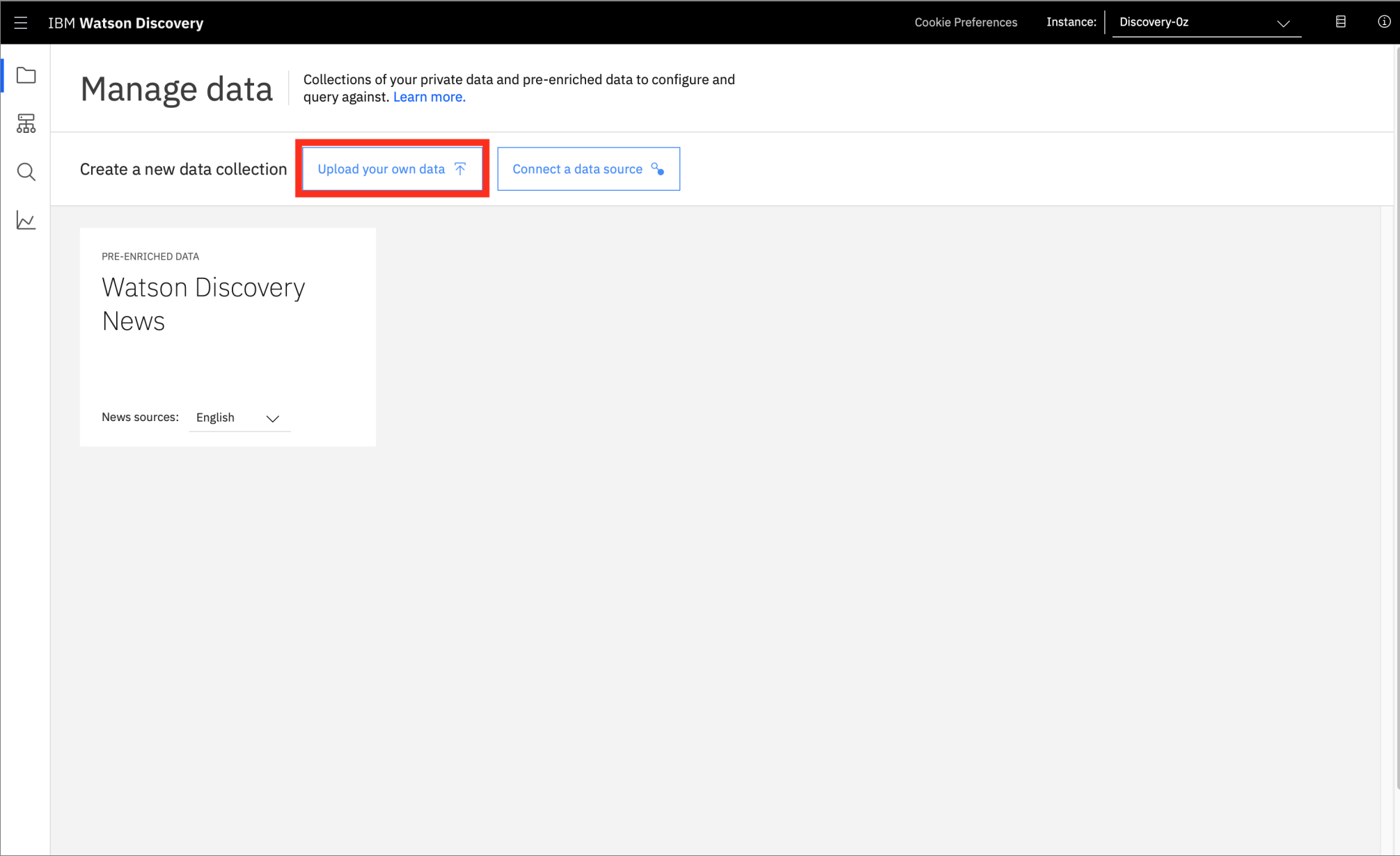
Read the details of the various plans offered to you, and then, **with the free Lite plan selected, click Create**.

Now, Launch your service by clicking "Launch Watson Discovery".



## Step 6: Setting up Your Collection

Your collection will be hosted in a Collection, click on "Upload your own data" to create your collection.



You will be prompted to upgrade to the Advance plan, but for this exercise you only need the Lite plan, so you could Set up with current plan.



# Set up Discovery for your private data.

Your current **Lite plan** is suitable for trial & experimentation.

## Building a production application?

The Advanced plan provides more storage, better performance with dedicated clusters and support for a production grade workload.

[Upgrade plan in IBM Cloud](#)

Cancel

Set up with current plan

Since we are creating a collection for our documents, you can give your Collection a name, and set the language of your documents to English.

## Manage data

Collections of your private data and pre-enriched data to configure and query against. [Learn more.](#)

Create a new data collection

[Upload your own data](#)

[Connect a data source](#)

PRE-ENRICHED DATA

Watson Discovery  
News

News sources: English

### Name your new collection

Collection name

Annual-Report

Select the language of your documents

English

[Advanced](#)

Cancel

Create

## Step 7: Uploading your Documents

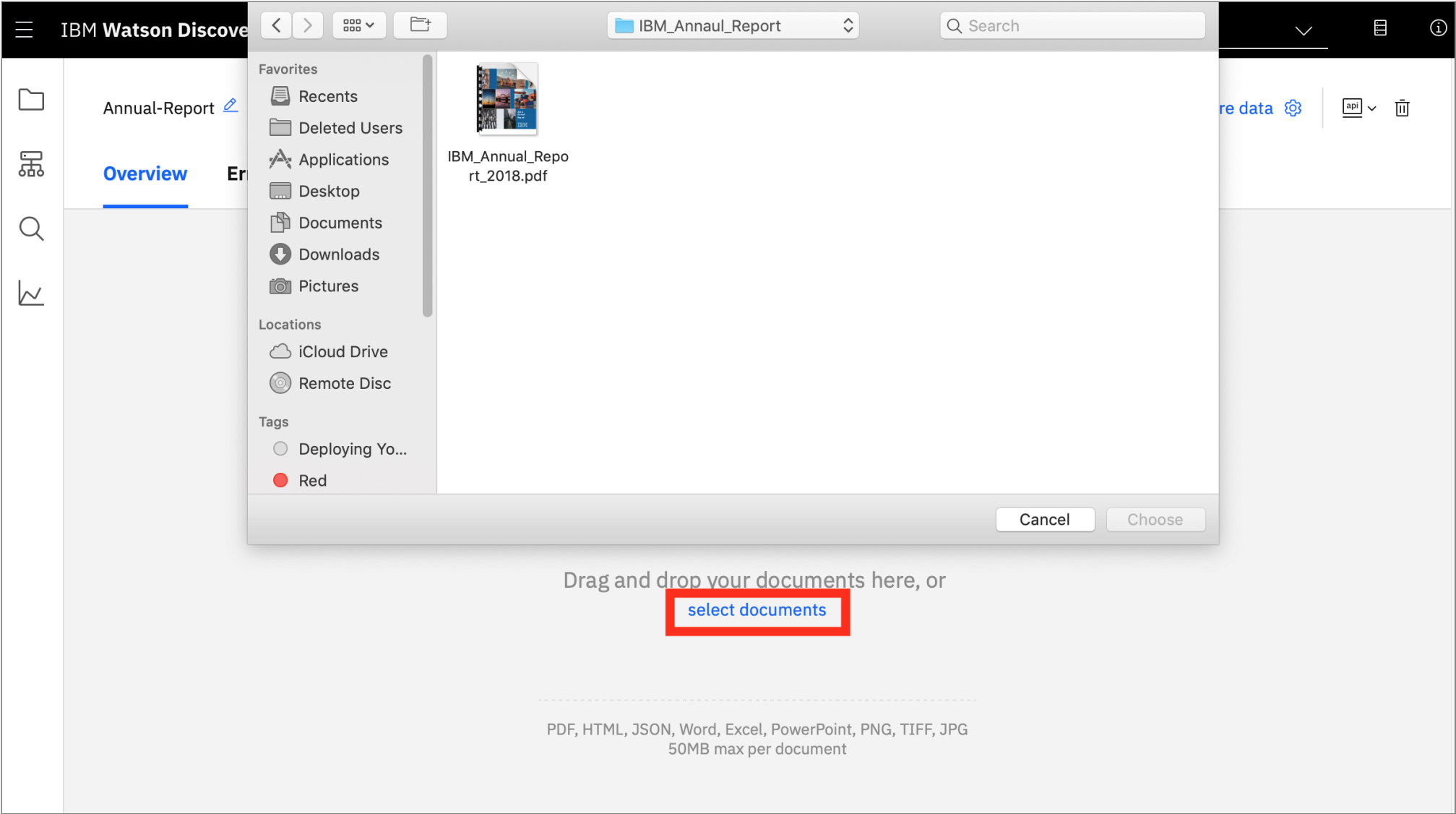
For the purpose of this lab, we will be analyzing IBM's 2018 Annual Report, you can download the report from the following link.

[https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDDeveloperSkillsNetwork-AI0103EN-SkillsNetwork/labs/FinalAssignment-coursera/data/IBM Annual Report 2018.pdf?utm\\_medium=Exinfluencer&utm\\_source=Exinfluencer&utm\\_content=000026UJ&utm\\_term=10006555&utm\\_id=NA-SkillsNetwork-Channel-SkillsNetworkCoursesIBMDDeveloperSkillsNetworkAI0103ENSkillsNetwork20648538-2021-01-01>IBM Annual Report\ 2018.pdf](https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDDeveloperSkillsNetwork-AI0103EN-SkillsNetwork/labs/FinalAssignment-coursera/data/IBM%20Annual%20Report%202018.pdf?utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=10006555&utm_id=NA-SkillsNetwork-Channel-SkillsNetworkCoursesIBMDDeveloperSkillsNetworkAI0103ENSkillsNetwork20648538-2021-01-01>IBM%20Annual%20Report%202018.pdf)

Upload the IBM 2018 Annual Report onto Watson Discovery.

Click on **select document** and choose the IBM 2018 Annual Report.

**Note:** You can upload documents in the form of PDF, HTML, JSON, WORD, Excel, PowerPoint, PNG, TIFF, JPG. Any of the above filetypes for up to 50 megabytes.



## Step 8: Analyzing your Document

After your document has finished processing, click on the "Schema" tag highlighted by the red box, and click on document view. As you scroll down you can see the sentiment level of each entity. In this case it is positive for IBM Watson, IBM Services, IBM Cloud, and AI. So this tool has improved my workflow by providing me the sentiment for IBM Watson, IBM Services, and IBM Cloud without someone reading through the whole document.

IBM Watson Discovery

Cookie PreferencesInstance: Discovery-0z

Annual-Report / View data schema

Build queries →

Collection viewDocument view

Documents

IBM\_Annual\_Report\_2018

enriched_text.entities.disambiguation.dbpedia_resource	http://dbpedia.org/resource/IBM
enriched_text.entities.count	12
enriched_text.entities.sentiment.score	0.48946
enriched_text.entities.sentiment.label	positive
enriched_text.entities.text	IBM Watson
enriched_text.entities.relevance	0.642748
enriched_text.entities.type	Company
enriched_text.entities.count	14
enriched_text.entities.sentiment.score	0.505348
enriched_text.entities.sentiment.label	positive
enriched_text.entities.text	IBM Services
enriched_text.entities.relevance	0.554695
enriched_text.entities.type	Company
enriched_text.entities.count	13
enriched_text.entities.sentiment.score	0.530233
enriched_text.entities.sentiment.label	positive
enriched_text.entities.text	IBM Cloud
enriched_text.entities.relevance	0.513838
enriched_text.entities.type	Company
enriched_text.entities.count	25
enriched_text.entities.sentiment.score	0.173435
enriched_text.entities.sentiment.label	positive
enriched_text.entities.text	AI

## Step 9: Querying your Document

You can also extract important information from your document by using keywords. Click on the Search tab, and enter your keywords in the Search for documents section. In this case we are searching for "Intellectual Property", Watson Discovery has returned the result for "Intellectual Property" in our document.

IBM Watson Discovery

Cookie Preferences

Instance: Discovery-0z

Annual-Report / Build queries

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

Intellectual Property

+ Include analysis of your results

+ Filter which documents you query

> More options

Summary

JSON

Query URL

https://gateway.watsonplatform.net/discovery/api/v1/envIRON

Passages

"IBM continues to expand its reach to new and existing clients through digital marketplaces, digital sales and local Business Partner resources. Research, Development and Intellectual Property IBM's research and development (R&D) operations differentiate the company from its competitors."

"percent in 2018 compared to the prior year. The year-to-year performance was driven by decline in intellectual property (IP) income (2 points) and a higher level of workforce rebalancing charges (1 point), offset by continued focus on efficiency resulting in lower spending (3 points)."

"IBM Worldwide Organizations The following worldwide organizations play key roles in IBM's delivery of value to its clients: • Global Markets • Research, Development and Intellectual Property Global Markets IBM has a global presence, operating in more than 175 countries with a broad-based geographic distribution of revenue. The company's Global"

Results

Showing 1 of 1 matching documents

IBM\_Annual\_Report\_2018

Sentiment

positive

Text

"...Percent Change Licensing of **intellectual property** including royalty-based fees \$ 723 \$1,193 (39.4)% Custom development income 275 252 9.3 Sales/other transfers of **intellectual property** 28 21 35.0 Total \$1,026 \$1,466 (30.0)% Licensing of **intellectual property** including royalty-based fees decreased 39.4 percent in 2018 compared to 2017...."

"...Percent Change Licensing of **intellectual property** including royalty-based fees \$1,193 \$1,390 (14.1)% Custom development income 252 214 17.5 Sales/other transfers of **intellectual property** 21 27 (24.2) Total \$1,466 \$1,631 (10.2)% Licensing of **intellectual property** including royalty-based fees decreased 14.1 percent in 2017 compared to 2016...."

"...If the nature of the company's promise is to provide a right to access the company's **intellectual property** throughout the license period (i.e., the license is symbolic **intellectual property**), such as a trademark license, income is recognized over time...."

## Step 10: Save a screen shot

NOTE: This step is Optional!

📷 From Steps 8, take and save a screenshot in .jpeg or .jpg format including the Watson Discovery sentiment analysis of the document. See a sample screenshot below.

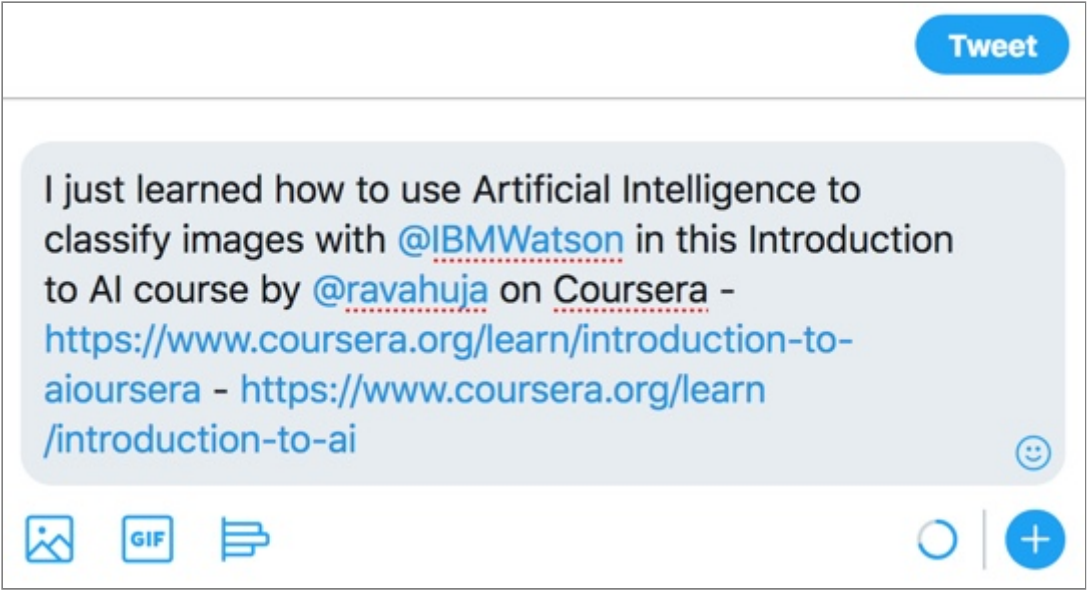
https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDDeveloperSkillsNetwork-AI0103EN-SkillsNetwork/labs/FinalAssignment-coursera/Exercise\_Analyzing\_documents\_with\_Watson\_Discovery...

12/14

enriched_text.entities.sentiment.label	positive
enriched_text.entities.text	IBM
enriched_text.entities.relevance	0.721457
enriched_text.entities.type	Company
enriched_text.entities.disambiguation.subtype.0	SoftwareLicense
enriched_text.entities.disambiguation.subtype.1	OperatingSystemDeveloper
enriched_text.entities.disambiguation.subtype.2	ProcessorManufacturer
enriched_text.entities.disambiguation.subtype.3	SoftwareDeveloper
enriched_text.entities.disambiguation.subtype.4	CompanyFounder
enriched_text.entities.disambiguation.subtype.5	ProgrammingLanguageDesigner
enriched_text.entities.disambiguation.subtype.6	ProgrammingLanguageDeveloper
enriched_text.entities.disambiguation.name	IBM
enriched_text.entities.disambiguation.dbpedia_resource	http://dbpedia.org/resource/IBM
enriched_text.entities.count	12
enriched_text.entities.sentiment.score	0.48946
enriched_text.entities.sentiment.label	positive
enriched_text.entities.text	IBM Watson
enriched_text.entities.relevance	0.642748
enriched_text.entities.type	Company

## Step 11: Share Your Results!

Follow us on Twitter and send us some of your funniest and most interesting results you found with IBM Watson Visual Recognition!



<a href=[https://twitter.com/intent/tweet?utm\\_medium=Exinfluencer&utm\\_source=Exinfluencer&utm\\_content=000026UJ&utm\\_term=10006555&utm\\_id=NA-SkillsNetwork-Channel-SkillsNetworkCoursesIBMDDeveloperSkillsNetworkAI0103ENSkillsNetwork20648538-2021-01-01](https://twitter.com/intent/tweet?utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=10006555&utm_id=NA-SkillsNetwork-Channel-SkillsNetworkCoursesIBMDDeveloperSkillsNetworkAI0103ENSkillsNetwork20648538-2021-01-01)>Click here to share the above Tweet.

## Author(s)

[Rav Ahuja](#)

## Changelog

Date	Version	Changed by	Change Description
2020-08-27	2.0	Anamika	Migrated Lab to Markdown and added to course repo in GitLab
2020-01-06	3.0	Shubham	Updated lab with new IBM Cloud screenshots
2021-06-08	3.1	Anamika	Updated lab with new IBM Cloud screenshots

© IBM Corporation 2020. All rights reserved.