

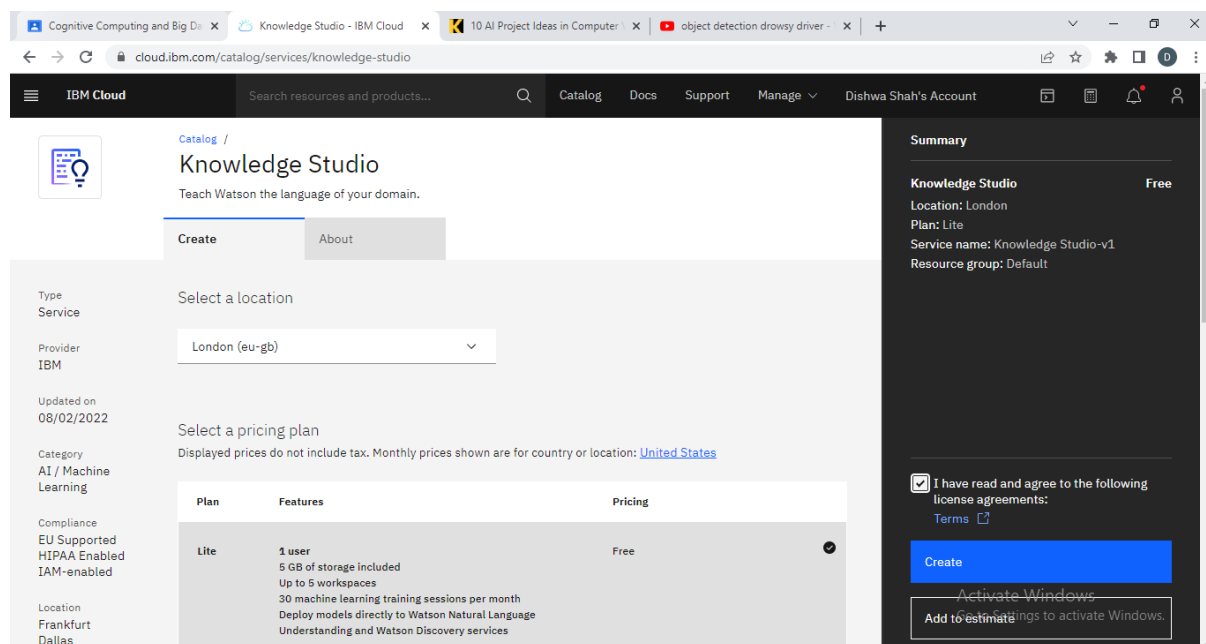
Institute of Computer Technology B. Tech Computer Science and Engineering Sub: Cognitive Computing (2CSE70E23)

PRACTICAL 8: WATSON KNOWLEDGE STUDIO

Aman got a task to extract information from news portals. However, unable to find a specific outcome. So for that, he needs to train a custom ML model to extract meaningful insights from news data. So he will use Watson Knowledge studio.

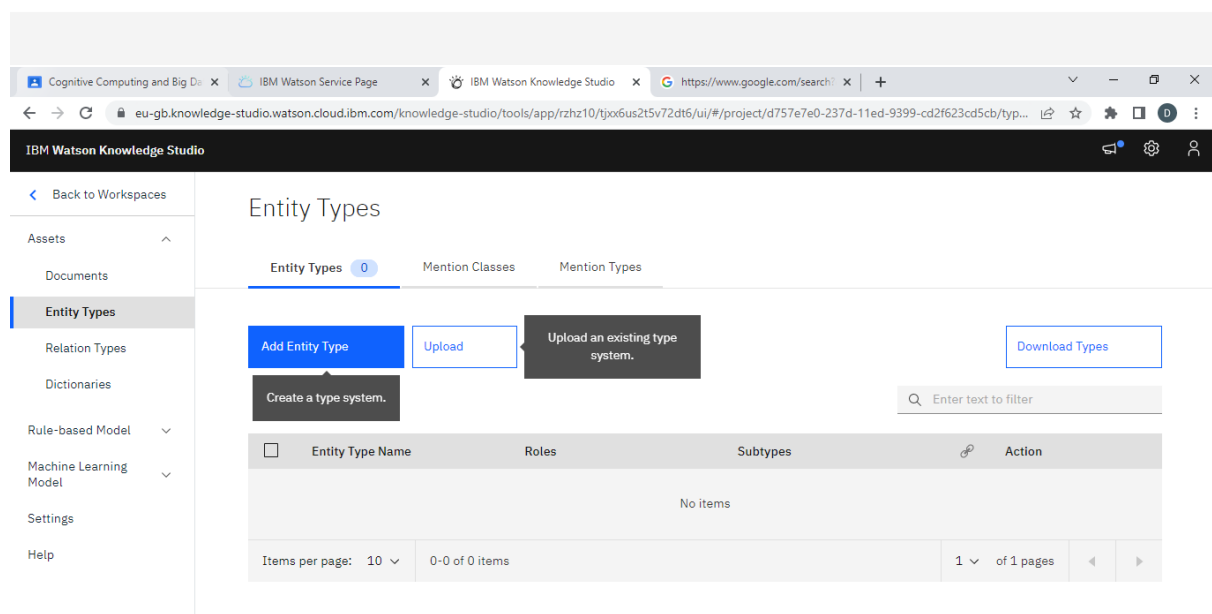
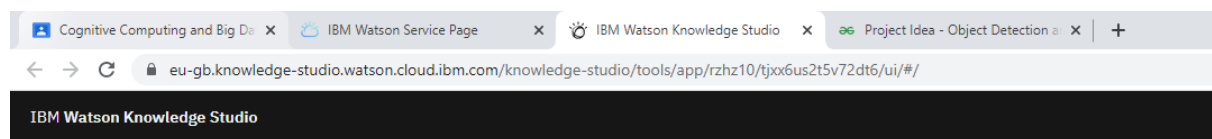
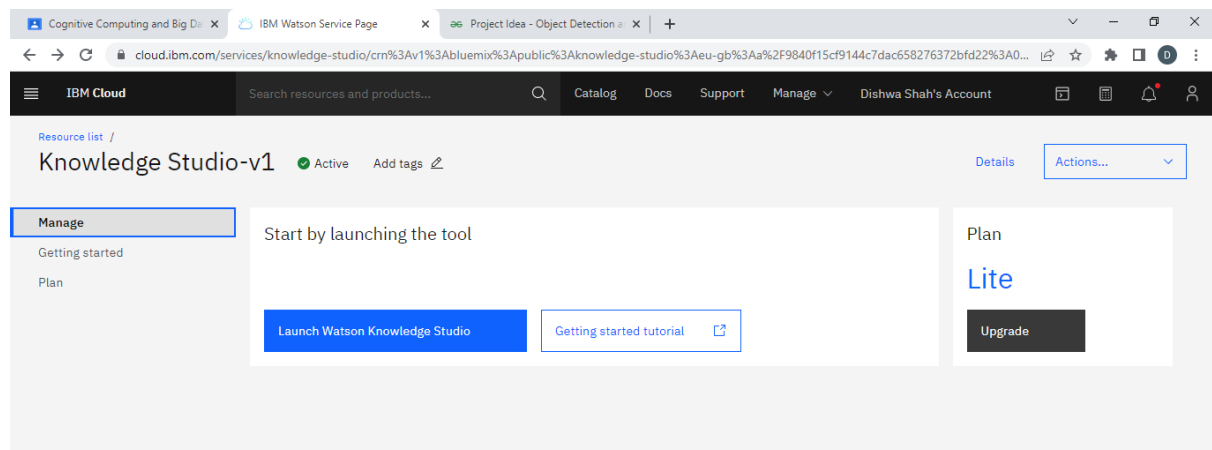
Use Watson Knowledge Studio to create a machine learning model that understands the linguistic nuances, meaning, and relationships specific to a particular industry or domain. Watson Knowledge Studio provides functionalities for annotating unstructured domain literature and uses those annotations to create a custom machine-learning model that understands the language of the domain.

Open Knowledge Studio in IBM Cloud in the same region as Watson Discovery.



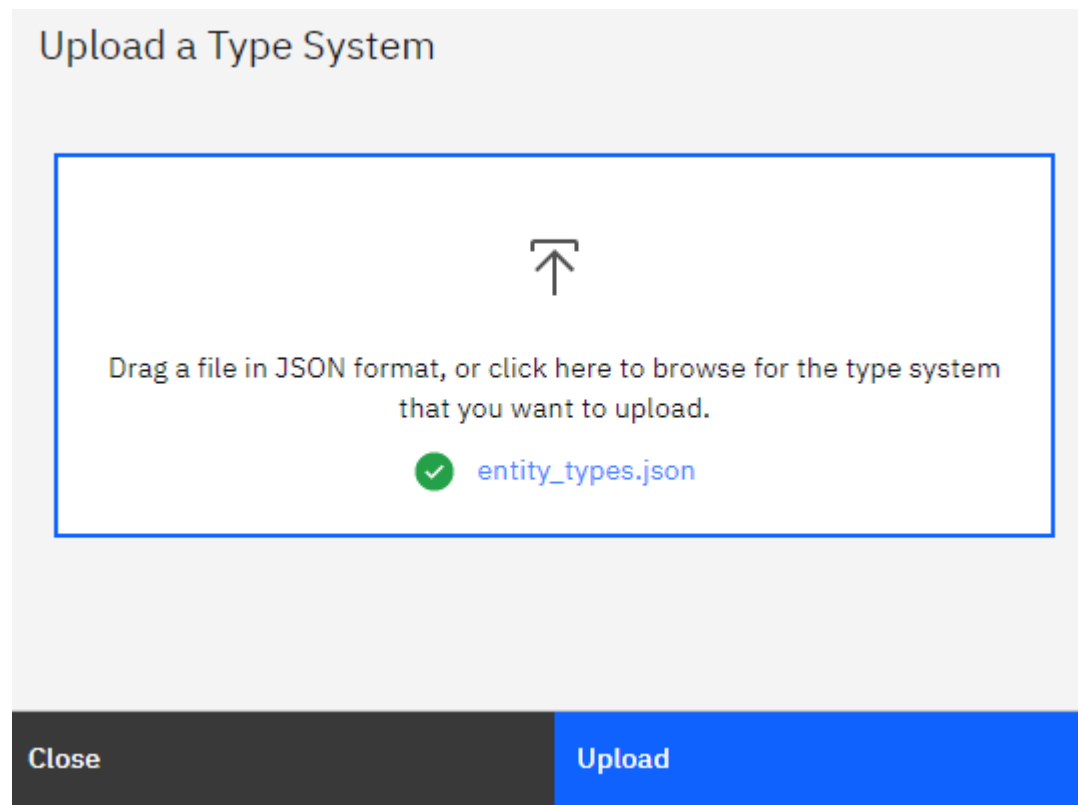
We can't just use NLP to detect entities since we have a lot of different varieties of entities which make it hard to detect them.

Knowledge Studio is used to create a completely customized model for us according to our needs. Then, we will use API of Watson Discovery so that it can work o]as our project.



The left portion showing assets and models is to customize your project.

Our file has entities and relations.



IBM Watson Knowledge Studio

Entity Types

Entity Types 52 | Mention Classes | Mention Types

Add Entity Type | Upload | Download Types

Enter text to filter

<input type="checkbox"/>	Entity Type Name	Roles	Subtypes	Action
<input type="checkbox"/>	ORDINAL	ORDINAL ANIMAL Show all...		Edit Delete
<input type="checkbox"/>	MONEY	MONEY AWARD	OTHER UNSPECIFIED	Edit Delete

Items per page: 10 | 1-10 of 52 items | 1 of 6 pages | [Go to Settings to activate Windows](#)

Entity Types

Entity Types **52** Mention Classes Mention Types

[Add Entity Type](#) [Upload](#) [Download Types](#)

Q Enter text to filter

<input type="checkbox"/>	Entity Type Name	Roles	Subtypes	Action
<input type="checkbox"/>	FACILITY	FACILITY		Edit Delete
<input type="checkbox"/>	WEB	WEB ORGANIZATION		Edit Delete
<input type="checkbox"/>	EVENT_BUSINESS	EVENT_BUSINESS		Edit Delete
<input type="checkbox"/>	EVENT_MEETING	EVENT_MEETING		Edit Delete

Items per page: 10 21-30 of 52 items 3 of 6 pages [Go to Settings to activate Windows.](#)

This is the type of entity (customized) that will not be present by default. This is why we add entities according to our needs.

You can also add more entities by clicking on Add Entity Type.

Cognitive Computing and Big D... IBM Watson Service Page IBM Watson Knowledge Studio <https://www.google.com/search>

eu-gb.knowledge-studio.watson.cloud.ibm.com/knowledge-studio/tools/app/rzhz10/tjxx6us2t5v72dt6/ui/#/project/d757e7e0-237d-11ed-9399-cd2f623cd5cb/typ...

IBM Watson Knowledge Studio

[Back to Workspaces](#)

Assets

Documents

Entity Types

Relation Types

Dictionaries

Rule-based Model

Machine Learning Model

Settings

Help

Relation Types **2177**

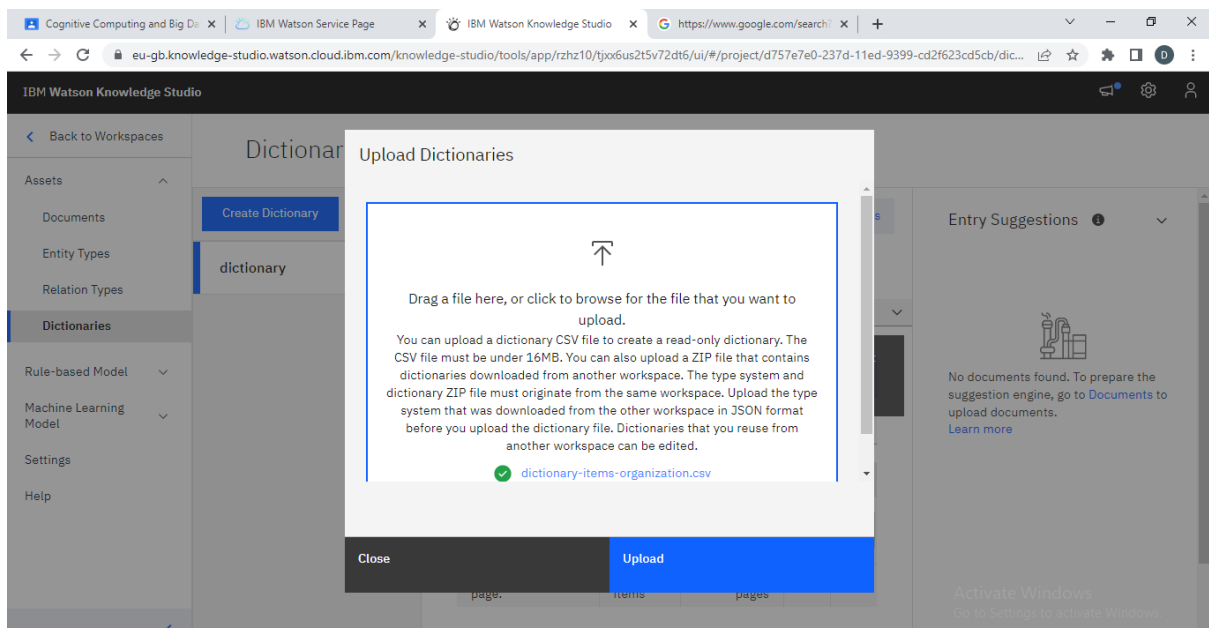
[Add Relation Type](#)

Q Enter text to filter

Relation Type	First Entity Type / Role	Second Entity Type / Role	Action
partOfMany	LAW	LAW	Edit Delete
siblingOf	ANIMAL	ANIMAL	Edit Delete
productOf	PRODUCT	ORGANIZATION	Edit Delete
partOfMany	EVENT_DEMONSTRATION	EVENT , EVENT_DEMONSTRATION	Edit Delete
spouseOf	ANIMAL	ANIMAL	Edit Delete

Items per page: 10 1-10 of 97 items 1 of 10 pages [Go to Settings to activate Windows.](#)

These are the relations between entities, which we uploaded through JSON file along with the entities.



Create a new dictionary and upload the dictionary file in it.

Create Dictionary

Preview only dictionary-items-... 24

Name: dict

Cancel Save

dict
Language: English | 24 entries

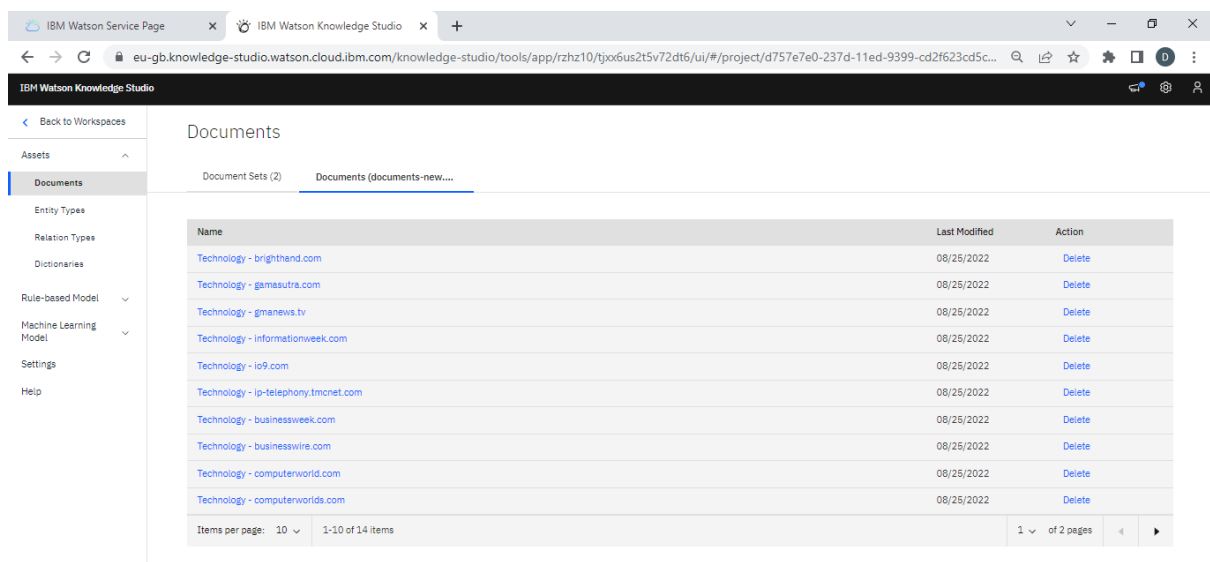
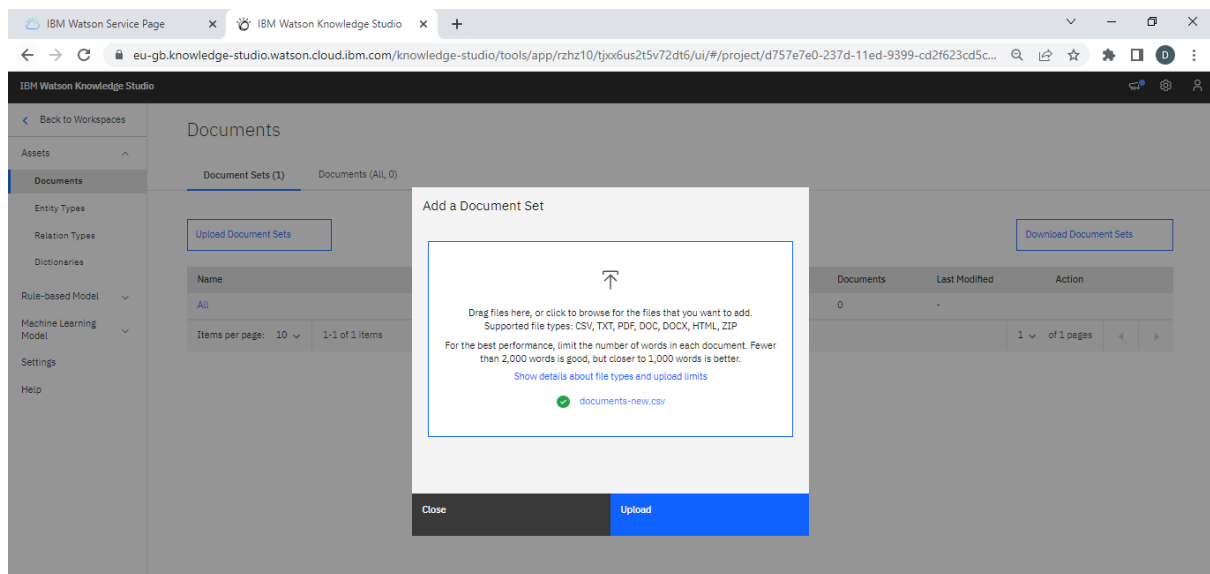
Entity type: None Rule class: None

Add Entry Delete Entries Upload Download

Enter text to filter

Lemma	Surface Forms	Part of Speech	Action
<input checked="" type="checkbox"/> IBM, International Business Machines	<input checked="" type="radio"/> IBM, International Business Machines <input type="radio"/> Enter text	Noun	Cancel Save
<input type="checkbox"/> Beverly Hills Hotel	Beverly Hills Hotel	Noun	Edit Delete
<input type="checkbox"/> British Academy	British Academy, BAFTA/LA	Noun	Edit Delete
<input type="checkbox"/> Golden Globes	Golden Globes	Noun	Edit Delete
<input type="checkbox"/> Writers Guild	Writers Guild, WGA	Noun	Edit Delete
<input type="checkbox"/> USA Today	USA Today	Noun	Edit Delete
<input type="checkbox"/> Paramount	Paramount	Noun	Edit Delete

Items per page: 100 1-24 of 24 items 1 of 1 pages



For entity type ORGANIZATION

IBM Watson Knowledge Studio

Back to Workspaces

Assets

- Documents
- Entity Types
- Relation Types
- Dictionaries**
- Rule-based Model
- Machine Learning Model
- Settings
- Help

Create Dictionary

dict 25

Language: English | 25 entries

Entity type: ORGANIZATION

Rule class: None

Add Entry Upload Download

Enter text to filter

Lemma	Surface Forms	Part of Speech	Action
<input type="checkbox"/>	IBM, International Business Machines	Noun	Edit Delete
<input type="checkbox"/>	Beverly Hills Hotel	Noun	Edit Delete
<input type="checkbox"/>	British Academy	Noun	Edit Delete
<input type="checkbox"/>	Golden Globes	Noun	Edit Delete
<input type="checkbox"/>	Writers Guild	Noun	Edit Delete
<input type="checkbox"/>	USA Today	Noun	Edit Delete
<input type="checkbox"/>	Paramount	Noun	Edit Delete

Items per page: 100 1-25 of 25 items 1 of 1 pages

Entry Suggestions

Add All Exclude All

- test
- s
- bandwidth
- like
- noted
- analyst
- blog
- post
- back
- industry
- nikkel
- than
- wednesday
- virtualization
- research
- digital
- may
- been

IBM Watson Knowledge Studio

Back to Workspaces

Assets

- Documents
- Entity Types
- Relation Types
- Dictionaries
- Rule-based Model
- Machine Learning Model
- Pre-annotation**
- Annotations
- Performance
- Versions
- Settings
- Help

Run Pre-annotators

Order Settings 11

Order	Pre-annotator	Status
1	Rule-based Model	Not available
2	Dictionaries	Available
3	Machine Learning Model	Not available
4	Natural Language Understanding	Not available

Pre-annotation: labelling

These are the 4 ways to do it.

We can only see dictionaries since we only added dictionaries.

Run pre-annotators:

Run Pre-annotators

Close Next

Select pre-annotators

Select the pre-annotators that you want to use.

Pre-annotator
<input type="checkbox"/> Dictionaries

Run Pre-annotators

Back

Run


Select document sets

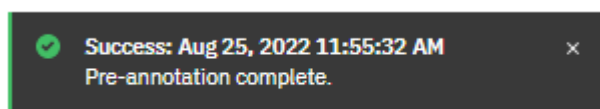
Check if you want to remove previous pre-annotation results from documents before running the pre-annotators. If not checked, all previous annotations are preserved.
 * Annotations made by humans outside of the pre-annotation process remain even if you check the wipe option.

☐ Wipe previous pre-annotation results

Select the document sets or annotation sets that you want to pre-annotate.

Document set	Documents	Pre-annotated documents	Human annotated documents
<input checked="" type="checkbox"/> documents-new.csv	14	0	0

 Running pre-annotators...Time elapsed: 10 seconds



and make the necessary changes to enable the pre-annotator

Now, the organizations' names in the documents will be highlighted since the pre-annotation has been completed.

The screenshot shows the IBM Watson Knowledge Studio interface. The left sidebar contains a navigation menu with options like Assets, Documents, Entity Types, Relation Types, Dictionaries, Rule-based Model, Machine Learning Model, Pre-annotation, Annotations (selected), Performance, Versions, Settings, and Help. The main content area is titled 'Annotations' and has two tabs: 'Ground Truth' and 'Annotation Tasks'. The 'Ground Truth' tab is active, displaying instructions on how to create ground truth by annotating documents. Below the instructions is a table with the following data:

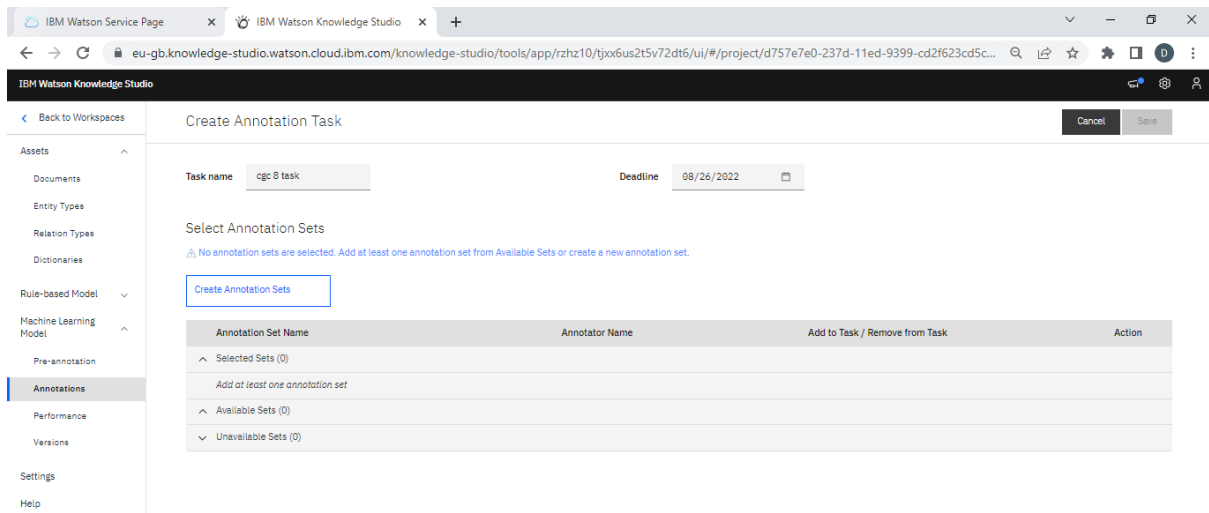
Document Set	Last Modified	Documents (Annotated/Total)	Annotations	Action
All	-	5 / 14	0	Annotate
documents-new.csv	08/25/2022	5 / 14	0	Annotate

We can see that only 5/14 documents have been annotated because only those had organization entities.

Here, we are not collaborating, so we will perform all the tasks that are usually divided amongst people.

Add task.

Set the task name and the due date.



Create annotation set.

Overlapping will be reduced if you want to divide the work between the team members. If we keep 100% overlap, all members will get all the files. Here, we will keep 100% since we are working alone.

Create Annotation Sets

Base set All ▼

Overlap 100 % of 14 documents

New Sets [+ Add another set and human annotator](#)

To view inter-annotator agreement scores, assign at least two human annotators and ensure that a percentage of documents overlap between the sets.

Annotator Dishwa Shah ▼

Set name SET_1

Annotator Dishwa Shah ▼ 🗑️

Set name SET_2

Cancel **Generate**

We created two sets to see how this works in collaboration.

IBM Watson Knowledge Studio

Back to Workspaces

Create Annotation Task

Task name: cgc 8 task

Deadline: 08/26/2022

Select Annotation Sets

Create Annotation Sets

Annotation Set Name	Annotator Name	Add to Task / Remove from Task	Action
Selected Sets (2)			
SET_1	Dishwa Shah	Remove from task	...
SET_2	Dishwa Shah	Remove from task	...
Available Sets (0)			
Unavailable Sets (0)			

Click on Save.

IBM Watson Knowledge Studio

Back to Workspaces

Annotations

Ground Truth

Annotation Tasks

Add Task

Refresh

cgc 8 task

Deadline: 08/26/2022

Progress: 0%

Deadline: 08/26/2022

Annotations added to annotation sets are not considered ground truth until the annotation sets are submitted and accepted.
When an annotation set is accepted, documents that are annotated in only one annotation set immediately become ground truth.
Documents that are annotated in two or more annotation sets become overlapping documents that will become ground truth after conflicts are resolved.

[Refresh](#) [Calculate Inter-Annotator Agreement](#) [Check Overlapping Documents for Conflicts](#) [Apply Type System Updates](#)

Annotation Set Name	Annotator Name	Status	Action
SET_2	Dishwa Shah	IN PROGRESS	Annotate
SET_1	Dishwa Shah	IN PROGRESS	Annotate

Activate Windows

Click on Annotate of SET_1.

Select Document

Annotation Set: SET_1
Annotator: dishwashah19@gnu.ac.in
Task: cgc 8 task

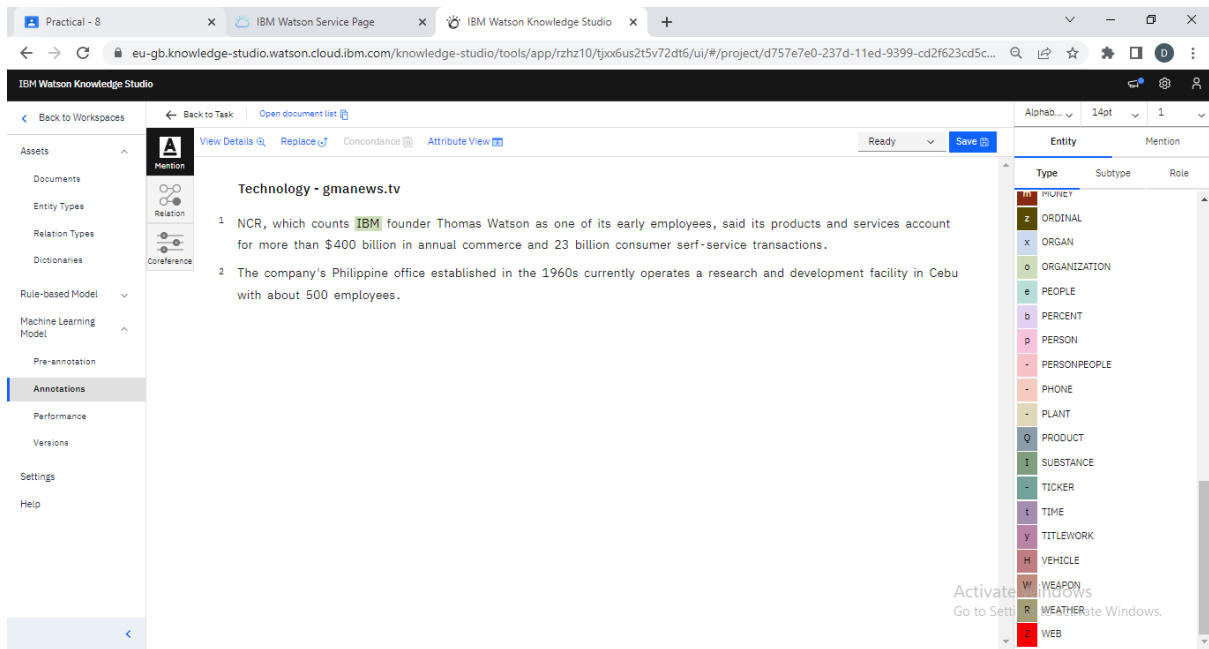
[Submit All Documents](#) Showing 1-14 of 14

Document Name	Status	Last Modified
Technology - brighthand.com	Ready	Aug 25, 2022 11:55:22 AM
Technology - gamasutra.com	Ready	Aug 25, 2022 11:55:22 AM
Technology - gmanews.tv	Ready	Aug 25, 2022 11:55:22 AM
Technology - informationweek.com	Ready	Aug 25, 2022 11:55:22 AM
Technology - io9.com	Ready	Aug 25, 2022 11:55:22 AM
Technology - ip-telephony.tmcnet.com	Ready	Aug 25, 2022 11:55:22 AM

Items per page: 50 1-14 of 14 items 1 of 1 pages

Click and open the third document.

Now we can see the content of the document and on the right side, we can see the entities we added.



Not all the organizations will be highlighted, only the ones mentioned in our dictionary will (according to the colors on the right side).

List of the documents where the annotations have been done:

BrightHand.com

Informationweek.com

Computerworld.com

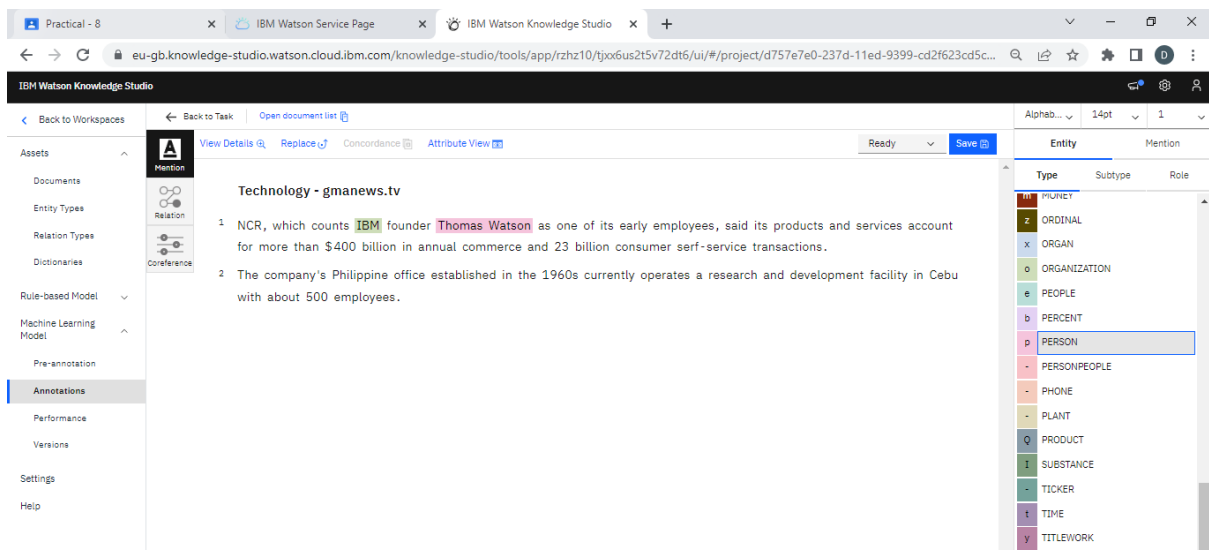
Computerworlds.com

Foxnews.com

Gmanews.tv

If you want to add any annotations, select that word and click on the entity on the right side.

This is called manual annotations.



The screenshot shows the IBM Watson Knowledge Studio interface. The left sidebar contains navigation options: Assets, Documents, Entity Types, Relation Types, Dictionaries, Rule-based Model, Machine Learning Model, Pre-annotation, Annotations (selected), and Performance. The main area displays a document titled "Technology - gmanews.tv" with two paragraphs. The first paragraph mentions "NCR, which counts IBM founder Thomas Watson as one of its early employees, said its products and services account for more than \$400 billion in annual commerce and 23 billion consumer self-service transactions." The second paragraph mentions "The company's Philippine office established in the 1960s currently operates a research and development facility in Cebu with about 500 employees." On the right, a panel shows a list of relation types: LOCATION, MEASURE, MONEY, ORDINAL, ORGAN, ORGANIZATION, PEOPLE, PERCENT, and PERSON. The "Mention" tab is selected.

Use CTRL Z to undo any changes.

Save the work when you are ready. These changes will be applied to all the documents.

Now, select the relations tab. You will see the relations and on the right side you will see all the relation types that we added.

The screenshot shows the IBM Watson Knowledge Studio interface with the "Relation" tab selected. The main area displays the same document titled "Technology - gmanews.tv" with two paragraphs. The first paragraph is annotated with a relation type "founderOf" connecting "IBM" and "Thomas Watson". The second paragraph is annotated with a relation type "PERSON" connecting "Thomas Watson" and "as one of its early employees". The left sidebar shows the "Relation" tab selected. The right panel shows a list of relation types: LOCATION, MEASURE, MONEY, ORDINAL, ORGAN, ORGANIZATION, PEOPLE, PERCENT, and PERSON.

Click on person and then click on organization and then select the relation type. Click on Save once you are ready.

The screenshot shows the IBM Watson Knowledge Studio interface with the "Relation" tab selected. The main area displays a document titled "There are two petaflop systems in the U.S., IBM's Roadrunner at Los Alamos National Laboratory, which passed the petaflop barrier last May, and Cray Inc.'s XT Jaguar at the Oak Ridge National Laboratory." The document is annotated with a relation type "affiliatedWith" connecting "IBM's Roadrunner at Los Alamos National Laboratory" and "Cray Inc.'s XT Jaguar at the Oak Ridge National Laboratory". The left sidebar shows the "Relation" tab selected. The right panel shows a list of relation types: LOCATION, MEASURE, MONEY, ORDINAL, ORGAN, ORGANIZATION, PEOPLE, PERCENT, and PERSON.

Submit all documents.

IBM Watson Knowledge Studio

Annotation Set: SET_1
Annotator: dishwa.shah19@gnu.ac.in
Task: cgc 8 task

[Submit All Documents](#) ←

Showing 1-14 of 14

Document Name	Status	Last Modified
Technology - io9.com	Ready	Aug 25, 2022 12:24:21 PM
Technology - ip-telephony.tmcnet.com	Ready	Aug 25, 2022 12:24:21 PM
Technology - businessweek.com	Ready	Aug 25, 2022 12:24:21 PM
Technology - businesswire.com	Ready	Aug 25, 2022 12:24:21 PM
Technology - computerworlds.com	Ready	Aug 25, 2022 12:24:21 PM
Technology - content.usatoday.com	Ready	Aug 25, 2022 12:24:21 PM
Technology - e.nikkei.com	Ready	Aug 25, 2022 12:24:21 PM
Technology - focusaiwan.tw	Ready	Aug 25, 2022 12:24:21 PM
Technology - foxnews.com	Ready	Aug 25, 2022 12:24:21 PM
Technology - gmanews.tv	In Progress	Aug 25, 2022 12:35:30 PM
Technology - computerworld.com	In Progress	Aug 25, 2022 12:43:04 PM

Items per page: 50 1-14 of 14 items 1 of 1 pages

IBM Watson Knowledge Studio

cgic 8 task

Deadline: 08/26/2022

Annotations added to annotation sets are not considered ground truth until the annotation sets are submitted and accepted.
When an annotation set is accepted, documents that are annotated in only one annotation set immediately become ground truth.
Documents that are annotated in two or more annotation sets become overlapping documents that will become ground truth after conflicts are resolved.

[Refresh](#) [Calculate Inter-Annotator Agreement](#) [Check Overlapping Documents for Conflicts](#) [Apply Type System Updates](#)

[Accept](#) [Reject](#)

Annotation Set Name	Annotator Name	Status	Action
SET_2	Dishwa Shah	IN PROGRESS	Annotate
<input type="checkbox"/> SET_1	Dishwa Shah	SUBMITTED	View

Now, we will annotate SET_2.

Technology - gmanews.tv

- 1 NCR, which counts IBM founder Thomas Watson as one of its early employees, said its products and services account for more than \$400 billion in annual commerce and 23 billion consumer self-service transactions.
- 2 The company's Philippine office established in the 1960s currently operates a research and development facility in Cebu with about 500 employees.

Add the same entity as in SET_1, but give a different relation.

Technology - gmanews.tv

- 1

employedBy

OR... PERSON

NCR, which counts IBM founder Thomas Watson as one of its early employees, said its products and services account for more than \$400 billion in annual commerce and 23 billion consumer self-service transactions.
- 2 The company's Philippine office established in the 1960s currently operates a research and development facility in Cebu with about 500 employees.

Save the changes and submit all documents.

cgic 8 task

Deadline: 08/26/2022

Annotations added to annotation sets are not considered ground truth until the annotation sets are submitted and accepted.
When an annotation set is accepted, documents that are annotated in only one annotation set immediately become ground truth.
Documents that are annotated in two or more annotation sets become overlapping documents that will become ground truth after conflicts are resolved.

Refresh Calculate Inter-Annotator Agreement Check Overlapping Documents for Conflicts Apply Type System Updates

Accept Reject

Annotation Set Name	Annotator Name	Status	Action
<input type="checkbox"/> SET_2	Dishwa Shah	SUBMITTED	View
<input type="checkbox"/> SET_1	Dishwa Shah	SUBMITTED	View

You can either Accept or Reject after reviewing. Here, we will accept.

We have a conflict.

[←](#) | cgc 8 task[Edit](#)

Deadline: 08/26/2022

Annotations added to annotation sets are not considered ground truth until the annotation sets are submitted and accepted.
When an annotation set is accepted, documents that are annotated in only one annotation set immediately become ground truth.
Documents that are annotated in two or more annotation sets become overlapping documents that will become ground truth after conflicts are resolved.

[Refresh](#) [Calculate Inter-Annotator Agreement](#) [Check Overlapping Documents for Conflicts](#) [Apply Type System Updates](#)

Annotation Set Name	Annotator Name	Status	Action
SET_2	Dishwa Shah	IN CONFLICT	View
SET_1	Dishwa Shah	IN CONFLICT	View

We can now check for overlapping documents for conflicts.

[← Back to cgc 8 task Task](#) | [Overlapping Documents in Task cgc 8 task](#)

These documents are included in two or more annotation sets. Click Check for Conflicts to open a document and check whether different human annotators applied different annotations to it. Resolve any inconsistencies that you find. If you know that no conflicts exist, click Accept to add the document to the ground truth.

Document Name	Annotation Set Name	Action
Technology - gmanews.tv	SET_2, SET_1	Check for Conflicts Accept
Technology - samasutra.com	SET_2, SET_1	Check for Conflicts Accept

Select the two sets of human annotation conflicts that you want to resolve.

- ☒ SET_2(dishwashah19@gnu.ac.in)
- ☒ SET_1(dishwashah19@gnu.ac.in)
- ☐ Current Ground Truth

[Cancel](#)[Check for Conflicts](#)

Now, click on the relation conflicts to resolve them.

Practical - 8 | IBM Watson Service Page | IBM Watson Knowledge Studio

eu-gb.knowledge-studio.watson.cloud.ibm.com/knowledge-studio/tools/app/rzhz10/tjxx6us2t5v72dt6/ui/#/project/d757e7e0-237d-11ed-9399-cd2f623cd5c

IBM Watson Knowledge Studio

Back to Workspaces | Back to List of Overlapping Documents

Assets

- Documents
- Entity Types
- Relation Types
- Dictionaries

Rule-based Model

- Machine Learning Model
- Pre-annotation
- Annotations
- Performance
- Versions
- Settings
- Help

Conflicts 3 remaining

Human Annotators

Dishwa Shah

Unresolved Resolved

Mention conflicts 1

Relation conflicts 2

Coreference chain conflicts 0

1

Reject (0) Accept (0)

Dishwa ... founderOf

Dishwa Sh... employedBy

OR... PERSON

NCR, which counts IBM founder Thomas Watson as one of its early employees, said its products

MONEY

and services account for more than \$400 billion in annual commerce and 23 billion consumer self-service transactions.

Select all annotations

2

Reject (1) Accept (1)

The company's Philippine office established in the 1960s currently operates a research and development facility in Cebu with about 500 employees.

Dishwa ... founderOf

Dishwa Sh... employedBy

OR... PERSON

NCR, which counts IBM founder Thomas Watson as one of its early employees, said its products

MONEY

and services account for more than \$400 billion in annual commerce and 23 billion consumer self-service transactions.

Select all annotations

Accept the other one.

Save your work. Click on promote to ground truth if you want.

Resolve all such conflicts.

Back to cgc 8 task | Overlapping Documents in Task cgc 8 task

These documents are included in two or more annotation sets. Click Check for Conflicts to open a document and check whether different human annotators applied different annotations to it. Resolve any inconsistencies that you find. If you know that no conflicts exist, click Accept to add the document to the ground truth.

Document Name	Annotation Set Name	Action
No overlapping documents exist in this task.		

← cgC 8 task

Edit

Deadline: 08/26/2022

Annotations added to annotation sets are not considered ground truth until the annotation sets are submitted and accepted.
 When an annotation set is accepted, documents that are annotated in only one annotation set immediately become ground truth.
 Documents that are annotated in two or more annotation sets become overlapping documents that will become ground truth after conflicts are resolved.

[Refresh](#) [Calculate Inter-Annotator Agreement](#) [Check Overlapping Documents for Conflicts](#) [Apply Type System Updates](#)

Annotation Set Name	Annotator Name	Status	Action
SET_2	Dishwa Shah	COMPLETED	View
SET_1	Dishwa Shah	COMPLETED	View

The task is now completed.

Now, in the Performance tab, select Train and Evaluate.

The screenshot shows the IBM Watson Knowledge Studio interface. The left sidebar contains a navigation menu with options: Back to Workspaces, Assets (Documents, Entity Types, Relation Types, Dictionaries), Rule-based Model, Machine Learning Model, Pre-annotation, Annotations, Performance (selected), Versions, Settings, and Help. The main area is titled 'Performance' and displays information for project 'cgC 8' (Language of documents: English). It shows the number of documents per set: 0 Training Set, 0 Test Set, and 0 Blind Set. There are buttons for 'View Ground Truth' for both Training and Test sets, and 'View Decoding Results' for the Test set. A 'Train and evaluate' button is present with a message: 'You have not trained the machine learning model. Click here to train and evaluate it.' Below this is a 'Document set evaluation' section with a 'Model over time' chart. The chart shows a score of 0.00 and a message: 'There is no evaluation result.' To the right of the chart, there are metrics for Mention, Relation, and Coreference, each with Precision and Recall values of '--'. A legend indicates a 'Low performing range' with a grey bar. The bottom right corner has the text 'Active Go to Se'.

The screenshot shows the 'Training / Test / Blind Sets' page in IBM Watson Knowledge Studio. The left sidebar contains a navigation menu with 'Performance' selected. The main area is titled 'Select Training/ Test/ Blind Sets' and features two buttons: 'Train' and 'Train & Evaluate'. Below these is a table of document sets:

Document Set	Task Status
<input checked="" type="checkbox"/> All	
<input type="checkbox"/> documents-new.csv	
<input type="checkbox"/> SET_1	COMPLETED
<input type="checkbox"/> SET_2	COMPLETED

To the right, a section titled 'Create new sets by splitting the selected document sets' shows a 'Ratio' input field set to '70'. Below it, the recommended splits are listed:

- Training Set (70% Recommended): 70
- Test Set (23% Recommended): 23
- Blind Set (7% Recommended): 7

Click on Train and evaluate.

The screenshot shows the 'Performance' page in IBM Watson Knowledge Studio. The left sidebar shows 'Performance' selected. The main area displays the 'cgic 8' model with a language of 'English'. A donut chart shows the distribution of documents: 9 Training Set, 3 Test Set, and 2 Blind Set. A 'Train and evaluate' button is present with the message: 'You have not trained the machine learning model. Click here to train and evaluate it.' Below this, the 'Document set evaluation' section shows a 'Model over time' graph with the text 'There is no evaluation result.' To the right, a table shows performance metrics for 'Mention', 'Relation', and 'Coreference', all with 'Precision: --' and 'Recall: --'. A 'Low performing range' indicator is shown at the bottom right.

Create version and deploy model.

The screenshot shows the IBM Watson Knowledge Studio interface. The left sidebar contains a navigation menu with options: Back to Workspaces, Assets, Documents, Entity Types, Relation Types, Dictionaries, Rule-based Model, Machine Learning Model (selected), Pre-annotation, Annotations, Performance, Versions (highlighted), Settings, and Help. The main content area is titled 'Versions' and 'Machine Learning Model'. It contains two boxes: 'Machine learning models, once trained, can be used to annotate new documents prior to human annotation.' with a 'Go to Pre-annotation page' button, and 'Export the current version of your model to use in other Watson applications such as Watson Explorer.' with an 'Export current model' button. Below these is a 'Version History and Deployment' table.

Version	Base	Creation Date	Entity Scores	Relation Scores	Description	Action
1.1		Current Version	1.00 (1.00 / 1.00)	N/A		Create Version
1.0		08/25/2022	1.00 (1.00 / 1.00)	N/A	cgc 8	Promote Delete Deploy

Click on Deploy. We will deploy it to Discovery.

The 'Deploy Model v1.0' dialog box is shown. It has a title bar 'Deploy Model v1.0'. Below the title, it says 'Select a service to deploy to.' and 'All services require a subscription. [Learn more](#)'. There are two radio button options: 'Natural Language Understanding' and 'Discovery'. The 'Discovery' option is selected. At the bottom, there are 'Cancel' and 'Next' buttons.

Region London. Service Watson Discovery. Click on Deploy.

The 'Deploy Model v1.0' dialog box is shown with deployment details. It has a title bar 'Deploy Model v1.0'. Below the title, it says 'Deploying to Discovery' and 'You must have a subscription to the IBM Watson™ Discovery service, and know the names of your IBM Cloud space and service instance. [Learn more](#)'. Below this is 'IBM Cloud Information' and 'IBM Cloud is the IBM cloud platform. Click [here](#) to open IBM Cloud and create an account or look up details for an existing service.' There are three dropdown menus: 'Region' (set to 'London'), 'Resource group' (set to 'Default'), and 'Service name' (set to 'Watson Discovery-m3'). At the bottom, there are 'Cancel' and 'Deploy' buttons.

You will get a model id (f2f3dfa5-39e4-46e9-a59e-b9a8cd6e2868). Save that model id so that you can use that later in Watson discovery under Enrichments option.

Deployment Started.

Deploying to Discovery

It might take a few minutes for publishing and deployment to complete, and for this model to be available to your applications.

You can view your deployed models, withdraw a model from deployment, or deploy a newer version.

Model ID: f2f3dfa5-39e4-46e9-a59e-b9a8cd6e2868

You can [view documentation](#) to learn how to implement the deployed model into your application.

OK

The screenshot shows the IBM Watson Knowledge Studio interface. On the left is a sidebar with navigation options: Assets, Documents, Entry Types, Relation Types, Dictionaries, Rule-based Model, Machine Learning Model, Pre-annotation, Annotations, Performance, Versions (selected), Settings, and Help. The main area is titled 'Versions' and 'Machine Learning Model'. It contains two boxes: 'Machine learning models, once trained, can be used to annotate new documents prior to human annotation.' with a 'Go to Pre-annotation page' button, and 'Export the current version of your model to use in other Watson applications such as Watson Explorer.' with an 'Export current model' button. Below these is a 'Version History and Deployment' table.

Version	Base	Creation Date	Entity Scores	Relation Scores	Description	Action						
1.1		Current Version	1.00 (1.00 / 1.00)	N/A		Create Version						
1.0		08/25/2022	1.00 (1.00 / 1.00)	N/A	cgic 8	Promote Delete Deploy						
<div>Deployed Models (1)</div> <table border="1"> <thead> <tr> <th>Model ID</th> <th>Service ID</th> <th>Action</th> </tr> </thead> <tbody> <tr> <td>f2f3dfa5-39e4-46e9-a59e-b9a8cd6e2868</td> <td>4c803f0c-9591-4f7b-bede-d57c17937873</td> <td>Undeploy Status</td> </tr> </tbody> </table>							Model ID	Service ID	Action	f2f3dfa5-39e4-46e9-a59e-b9a8cd6e2868	4c803f0c-9591-4f7b-bede-d57c17937873	Undeploy Status
Model ID	Service ID	Action										
f2f3dfa5-39e4-46e9-a59e-b9a8cd6e2868	4c803f0c-9591-4f7b-bede-d57c17937873	Undeploy Status										

Go to Watson Discovery and create a new collection by Uploading your own data:

The screenshot shows the IBM Watson Discovery 'Manage data' page. The header includes the IBM Watson Service Page, IBM Watson Discovery - Manage, IBM Watson Knowledge Studio, and a New Tab. The URL is eu-gb.discovery.watson.cloud.ibm.com/regions/eu-gb/services/crn%3Av1%3Abluemix%3Apublic%3Adiscovery%3Aeu-gb%3Aa%2F9840f15cf. The page title is 'IBM Watson Discovery'. The main content area is titled 'Manage data' and includes a description: 'Collections of your private data and pre-enriched data to configure and query against. [Learn more.](#)'. Below this is a 'Create a new data collection' section with two buttons: 'Upload your own data' (highlighted with a red arrow) and 'Connect a data source'. The page shows two data collections: 'PRE-ENRICHED DATA' named 'Watson Discovery News' with 'News sources: English' and 'PRIVATE DATA' named 'cgic 7' with 'Indexed on 8/18/2022'.

Give your collection a name and Create:

×

Name your new collection

Collection name

cgcb8

Select the language of your documents

English

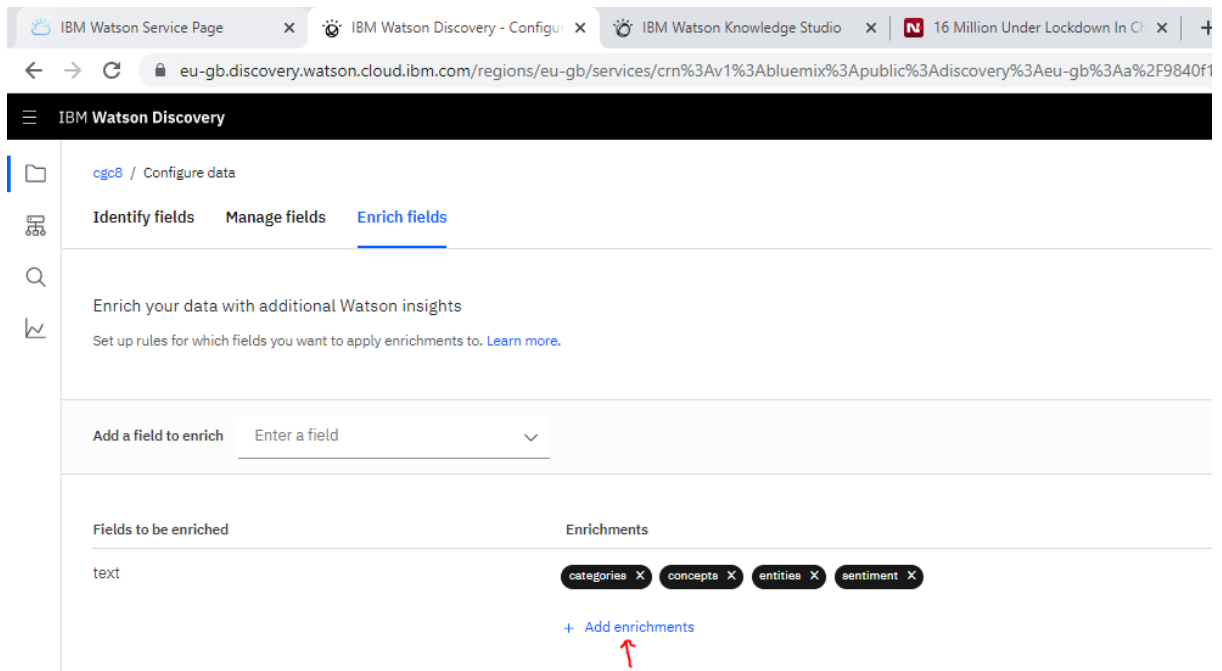
Cancel

Create

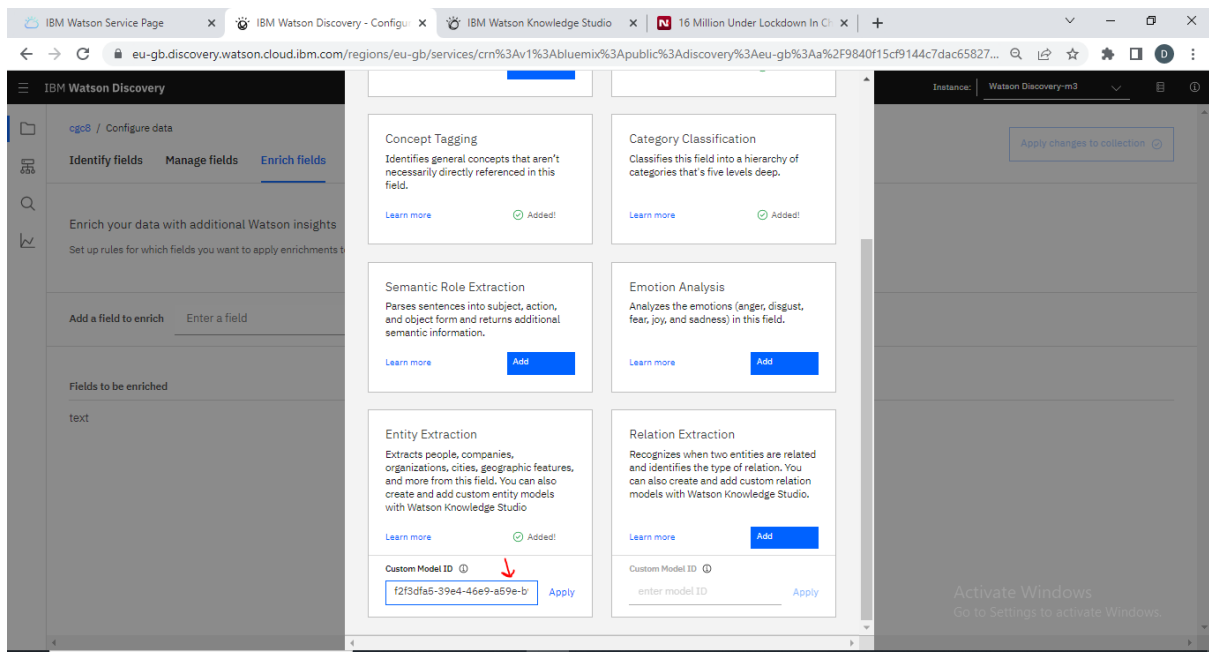
Now, upload your news article. Click on Add enrichments.

The screenshot shows the IBM Watson Discovery console for a collection named 'cgcb8'. The interface includes a sidebar with navigation icons and a main content area. The main area shows the collection overview with 1 document and 0 documents failed. Below this, there are sections for 'Identified 1 field from your data' (text), 'Added 4 enrichments to your data' (Entity Extraction, Sentiment Analysis, Concept Tagging, and Category Classification), and 'Now you're ready to query!' with several query options. A red arrow points to the 'Add enrichments' link at the bottom of the enrichment section.

Again, click on Add enrichments:



Enter your model id from Knowledge Studio and Apply:



Now, click on Apply changes to the collection. It will again ask you to upload your data. Upload it. Now, our model is deployed and the changes in entities will be updated:

The screenshot shows the IBM Watson Discovery Overview page for a project named 'cg08'. It indicates that 1 document has been uploaded and 0 documents have failed. The page shows that 4 enrichments have been added to the data:

- Entity Extraction:** 24 hours (1), Beijing (1), Chengdu (1), China (1), Hainan (1)
- Sentiment Analysis:** 0% positive, 0% neutral, 100% negative
- Concept Tagging:** City (1), Hui people (1), Infection (1), Qinghai (1), Qinghai Lake (1)
- Category Classification:** health and fitness → disease → cold and flu

On the right, there are suggestions for queries, such as 'Top people related to /health and fitness/disease/cold and flu' and 'Documents about 24 hours as a Quantity with a very negative sentiment'.

You can now see the entities:

The screenshot shows the IBM Watson Discovery Collection view for the 'cg08' project. The 'enriched_text.entities' field is selected, showing a list of fields and their example values:

Fields	Example values
enriched_text.entities.text	IBM
enriched_text.entities.type	Company
enriched_text.entities.relevance	0.916082
enriched_text.entities.count	12
enriched_text.entities.sentiment.score	0.666472

Below the table, there are sample queries using this enrichment, such as 'Entities of type Company which have negative sentiment'.