

Contents

1. Prerequisitos	2
2. Conceptos utilizados	2
2.1 Métodos de extracción de datos	2
2.2 Configuration	2
2.3 Breve descripción de los componentes de la aplicación	3
2.3.1 Componente de extractor de datos personales:	3
2.3.2 Componente regex	3
2.3.3 Componente score:	3
2.3.4 Componente Viewer:	3
3. Despliegue de aplicación	4
3.1 Desplegar una aplicación de Java Liberty en IBM Cloud	4
3.1.1 Desplegar utilizando "Deploy to IBM Cloud"	4
3.1.2 Implementar mediante pasos manuales	4
4. Desplegar el modelo de Watson Knowledge Studio	7
4.1 Importar artefactos	7
4.1.1 Type Systems	7
4.1.2 Documentos.....	7
4.2 Create Workspace	7
4.3 Import type system	8
4.4 Importar documentos	10
4.5 Crear y asignar conjuntos de anotaciones	12
4.6 Modelo de Machine Learning	13
5. Desplegar un modelo de WKS en Watson Natural Language Understanding.....	23
6. Verify that configuration parameters are correct	25
7. Analyze Results	27
8. Consuming the output by other applications	29

1. Prerequisitos

- Cuenta en IBM Cloud: Si no tiene una cuenta en IBM Cloud, puedes crear una cuenta [aquí](#).
- Si opta por implementar la aplicación Liberty manualmente, entonces:
 - Cloud Foundry cli debe estar instalado. Si no está instalado, haga clic [aquí](#) para obtener instrucciones de instalación.
 - Maven debe estar instalado. Si no está instalado, puede descargar Maven desde [aquí](#). Puede consultar las instrucciones de instalación [aquí](#).
- Cuenta de Watson Knowledge Studio: el usuario debe tener una cuenta WKS. Si no tiene una cuenta, puede crear una cuenta gratuita [aquí](#). Tome nota de la URL de inicio de sesión, ya que es única para cada ID de inicio de sesión
- Conocimiento básico de modelos de construcción en WKS: el usuario debe poseer conocimientos básicos de modelo de construcción en WKS para poder construir un modelo personalizado. Consulta la documentación de inicio [aquí](#).

2. Conceptos utilizados

2.1 Métodos de extracción de datos

Tenemos que definir qué datos personales (por ejemplo, nombre, ID de correo electrónico) queremos extraer. Esto se hace de dos maneras en este patrón de código.

A) Uso del modelo personalizado creado con Watson Knowledge Studio (WKS) y

B) Uso de expresiones regulares. En las secciones siguientes se explican los detalles de cómo se utilizan.

2.2 Configuration

Utilizamos la configuración para extraer datos personales. Los datos personales se clasifican en diferentes categorías. A cada categoría se le asigna un peso. También se especificamos qué datos personales pertenecen a qué categoría.

Una configuración es como se muestra abajo

```
Categories: Very_High,High,Medium,Low
Very_High_Weight: 50
High_Weight: 40
Medium_Weight: 20
```

```

Low_Weight: 10
Very_High_PIIIs: MobileNumber,EmailId
High_PIIIs: Person,DOB
Medium_PIIIs: Name,DOJ
Low_PIIIs: Company
regex_params: DOB,DOJ
DOB_regex: (0[1-9]|[12][0-9]|3[01])[-
/.](Jan|Feb|Mar|Apr|May|Jun|Jul|Aug|Sep|Oct|Nov|Dec)[- /.](19|20)\d\d
DOJ_regex: (0[1-9]|[12][0-9]|3[01])[-
/.](Jan|Feb|Mar|Apr|May|Jun|Jul|Aug|Sep|Oct|Nov|Dec)[- /.]\d\d

```

Si desea cambiar la configuración, a continuación, siga la siguiente plantilla:

```

Categories: <new set of categories which are comma separated>. e.g. Categories:
MyCategory1,MyCategory2,MyCategory3
<category_name>_Weight: Weightage for each category. e.g. MyCategory1_Weight: 40
<category>_PIIs: Personal data (Entity types). e.g. MyCategory1_PIIIs: EmailId, Employee
Id
regex_params: Entity types which have to be extracted using regular expressions. e.g.
regex_params:
Date
<regex_param>_regex: Regular expression using which an entity needs to be extracted
from text e.g. Date_regex:
(0[1-9]|[12]\[0-9]|3[01])

```

2.3 Breve descripción de los componentes de la aplicación

2.3.1 Componente de extractor de datos personales:

Personal Data Extractor component is the controller which controls the flow of data between all the components. It also integrates with NLU.

2.3.2 Componente regex

El componente regex analiza el texto de entrada utilizando las expresiones regulares proporcionadas en los archivos de configuración para extraer datos personales. Las expresiones regulares se utilizan para extraer datos personales para aumentar la salida de NLU

2.3.3 Componente score:

El componente score calcula la puntuación de un documento, que está entre 0 y 1, en función de los datos personales identificados y los datos de configuración. Utiliza el siguiente algoritmo:

```

Let score be 0
For each category{
    cat_weight = weightage for the category
    cat_entity_types = list of entity types for the category
    for each cat_entity_types{
        score = score + ( ( cat_weight/100 ) * ( 100 - score ) )
    }
}
score = score / 100; // to make it between 0 and 1

```

2.3.4 Componente Viewer:

El componente Viewer es el componente de la interfaz de usuario de la aplicación. El usuario puede examinar un archivo, que contiene la transcripción del chat y enviarlo al componente de extracción de

datos personales. Después de que los datos personales procesados se muestren en una vista de árbol, junto con la puntuación de confianza general.

3. Despliegue de aplicación

3.1 Desplegar una aplicación de Java Liberty en IBM Cloud

Puede desplegar la aplicación Java Liberty utilizando el botón deploy to IBM Cloud o utilizando pasos manuales.

3.1.1 Desplegar utilizando "Deploy to IBM Cloud"

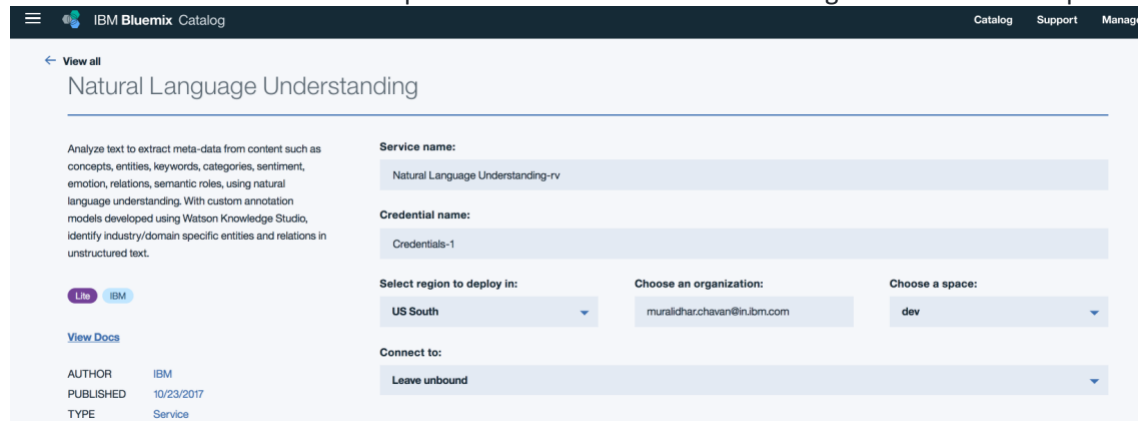
Haga clic en desplegar en el botón de IBM Cloud para desplegar la aplicación en IBM Cloud. Se le presentará una vista de cadena de herramientas y se le pedirá que "despliegue" la aplicación. Continúe y haga clic en el botón deploy. La aplicación debe implementarse. Asegúrese de que se ha iniciado la aplicación y de que se crea un servicio NLU y se enlaza a la aplicación que acaba de implementarse.



3.1.2 Implementar mediante pasos manuales

Si ha utilizado el botón deploy to IBM Cloud para desplegar la aplicación, omita esta sección y salte a la sección "4. Desarrolle el modelo Watson Knowledge Studio ". Si no ha utilizado deploy en IBM Cloud Button para desplegar la aplicación, complete las secciones "3.1.2.1 Create NLU Service Instance" y "3.1.2.2 implemente la aplicación Java en IBM Cloud" a continuación.

1. Paso 1: haga clic [aquí](#) para crear el servicio NLU
2. Paso 2: aparece la siguiente pantalla

A screenshot of the IBM Bluemix Catalog page for the "Natural Language Understanding" service. The page has a dark header with the IBM Bluemix Catalog logo and navigation links. The main content area is light blue and contains a description of the service, a "View Docs" link, and a table with metadata. To the right, there are form fields for "Service name", "Credential name", "Select region to deploy in", "Choose an organization", "Choose a space", and "Connect to".

AUTHOR	IBM
PUBLISHED	10/23/2017
TYPE	Service

3. Paso a paso: edite el campo "nombre del servicio:" para decir NLUGDPR y deje los otros valores predeterminados. Haga clic en crear

IBM Bluemix Catalog

View all

Natural Language Understanding

Analyze text to extract meta-data from content such as concepts, entities, keywords, categories, sentiment, emotion, relations, semantic roles, using natural language understanding. With custom annotation models developed using Watson Knowledge Studio, identify industry/domain specific entities and relations in unstructured text.

Life IBM

View Docs

AUTHOR IBM
PUBLISHED 10/23/2017
TYPE Service
LOCATION US South, Germany, Sydney, United Kingdom

Service name: NLUGDPR

Credential name: Credentials-1

Select region to deploy in: US South

Choose an organization: muralidhar.chavan@in.ibm.com

Choose a space: dev

Connect to: Leave unbound

Features

- Concepts
- Entities

Need Help? Contact Bluemix Sales

Estimate Monthly Cost Cost Calculator

Create

4. Paso a paso: la instancia del servicio NLU debe crearse.

3.1.2.2 desplegar la aplicación de Java en IBM Cloud

Paso a paso: Clone el [repositorio](#)

Paso 6: Abra el símbolo del sistema. Inicie sesión en el espacio de IBM Cloud utilizando el comando siguiente. Asegúrese de que usted inicia sesión al mismo espacio donde la instancia del servicio NLU fue creada en la sección "3.1.2.1 crea la instancia del servicio NLU"

```
cf login
```

Paso 7: Cambie el directorio al directorio raíz del repositorio clonado

Paso 8: encontrará el archivo manifest. yml en la carpeta raíz del proyecto. Verifique si el nombre del servicio NLU es lo mismo que el que está creado en el Paso 1. Si no es así, actualice el nombre del servicio NLU al que se creó anteriormente

```
manifest.yml
1 declared-services:
2   NLUGDPR:
3     label: natural-language-understanding
4     plan: free
5 applications:
6 - instances: 1
7   domain: mybluemix.net
8   buildpack: liberty-for-java
9   path: target/PersonalDataScorer.war
10 services:
11 - NLUGDPR
12 env:
13   Categories: Very_High,High,Medium,Low
14   Very_High_Weight: 50
15   High_Weight: 40
16   Medium_Weight: 20
17   Low_Weight: 10
18   Very_High_PII: MobileNumber
19   High_PII: Person,DOB,EmailID
20   Medium_PII: Name,DOJ
21   Low_PII: Company
22   regex_params: DOB,DOJ
23   DOB_regex: ([0-9]{12})[0-9]{13}[0-9]{1}[- /](Jan|Feb|Mar|Apr|May|Jun|Jul|Aug|Sep|Oct|Nov|Dec)[- /](19|20)\d\d
24   DOJ_regex: ([0-9]{12})[0-9]{13}[0-9]{1}[- /](Jan|Feb|Mar|Apr|May|Jun|Jul|Aug|Sep|Oct|Nov|Dec)[- /]\d\d\d
25   wks_model: 10:b7a9e265-5325-45bd-b1c7-38a063097179
26
```

Paso 9: Construir archivo de guerra usando el comando

```
mvn clean package
```

Paso 10: Desplegar la aplicación de Java Liberty mediante el comando siguiente. Proporcione un nombre de aplicación único para que la ruta no se haya tomado ya en IBM Cloud.

```
cf push <unique-application-name> -p target/PersonalDataScorer.war
```

Paso 11: en IBM Cloud Dashboard, asegúrese de que la aplicación se ha implementado correctamente y se está ejecutando.

IBM Bluemix Dashboard

Search Items

Dashboard

Cloud Foundry Apps (1) 2 GB/16 GB Used

NAME	ROUTE	MEMORY (MB)	INSTANCES	RUNNING	STATE	ACTIONS
gdpr-personaldata-scorer	gdpr-personaldata-scorer.mybluemix.net	1024	1	1	Running	

Paso 12: En el panel de IBM Cloud, haga clic en la aplicación que se implementó en el Paso 9. En los enlaces de navegación del lado izquierdo haga clic en conexiones. Verifique que el servicio NLU creado en el paso a paso esté enumerado.

IBM Bluemix Cloud Foundry Apps

Cloud Foundry apps / gdpr-personaldata-scorer

gdpr-personaldata-scorer Running Visit App URL

Routes

Connect existing Connect new

NLUGDPR
Natural Language Understanding
free

View credentials Docs

4. Desplegar el modelo de Watson Knowledge Studio

4.1 Importar artefactos

4.1.1 Type Systems

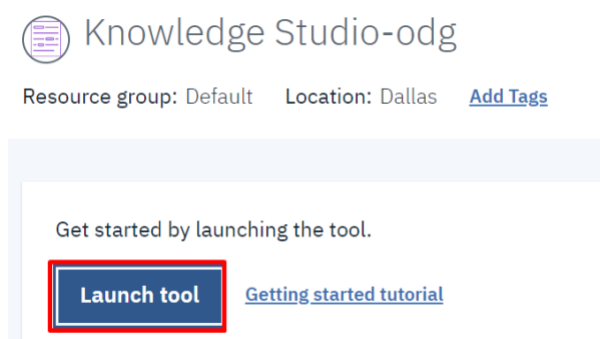
Puede obtener más información sobre los Type systems [aquí](#) los sistemas de tipos pueden crearse o importarse desde un archivo JSON de sistemas de tipos ya creado. Se deja al usuario crear sus propios sistemas de tipo o utilizar un archivo JSON de sistemas de tipo proporcionado en este repositorio. Si desea importar el archivo JSON de sistemas de tipos, descargue el archivo denominado `typesystems.JSON` en la carpeta WKS de este repositorio en el sistema de archivos local. El archivo JSON tiene tipos de entidad como nombre, PHONENO, EMAILID, Address. Puede editar/agregar/eliminar tipos de entidad para satisfacer sus necesidades.

4.1.2 Documentos

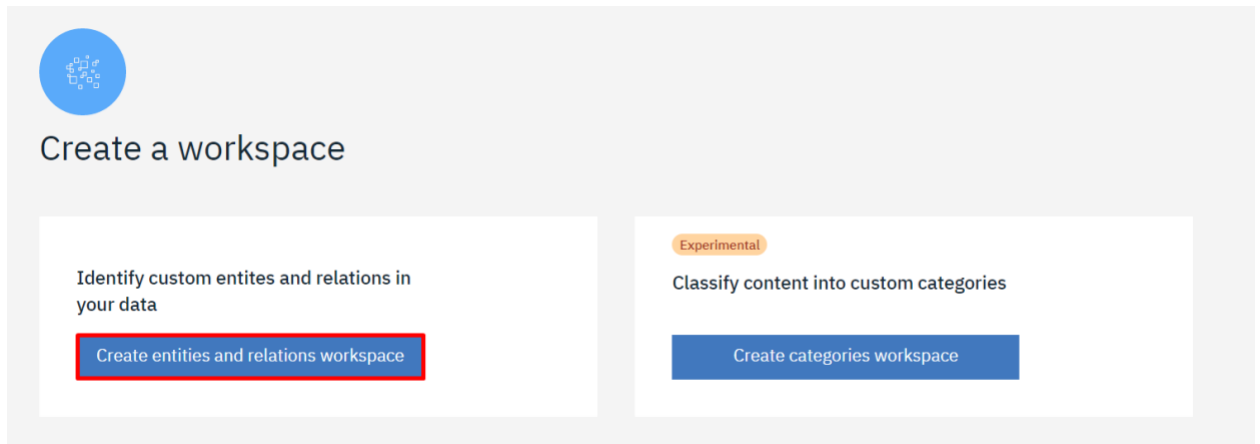
Puede obtener más información sobre los documentos [aquí](#) necesitaremos un conjunto de documentos para entrenar y evaluar el modelo WKS. Estos documentos contendrán el texto no estructurado del que identificaremos los datos personales. Consulte algunos de los archivos de documento de ejemplo en la carpeta `samplechattranscripts`. Para entrenar el modelo WKS, se necesita un amplio y variado conjunto de documentos. Para completar este ejercicio, vamos a considerar un conjunto más pequeño de documentos.

4.2 Create Workspace

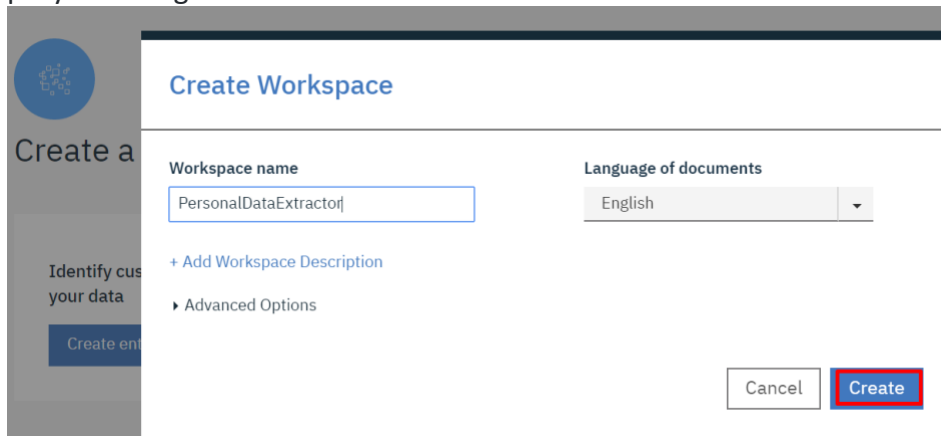
1. Haga clic en **Launch tool** desde la página de administración.



2. Hacer Click en **Create entities and relations Workspace**



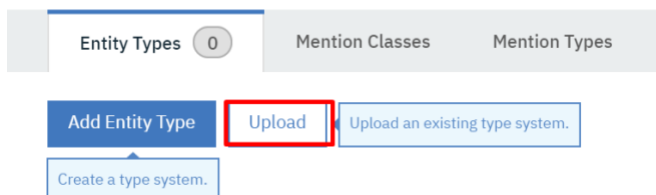
3. En la ventana emergente crear nuevo espacio de trabajo, escriba el nombre del nuevo proyecto. Haga clic en create.



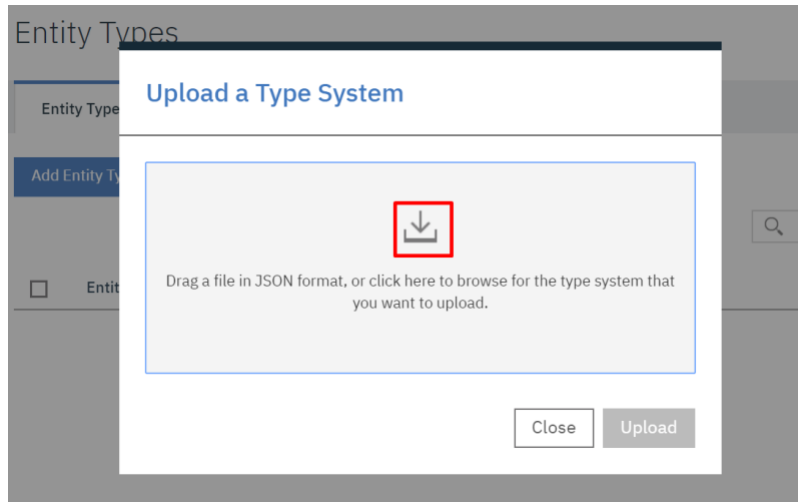
4.3 Import type system

1. Una vez creado el proyecto, haga clic en Type Systems en la barra de navegación.
2. Seleccione la pestaña Entity Types y haga clic en Upload

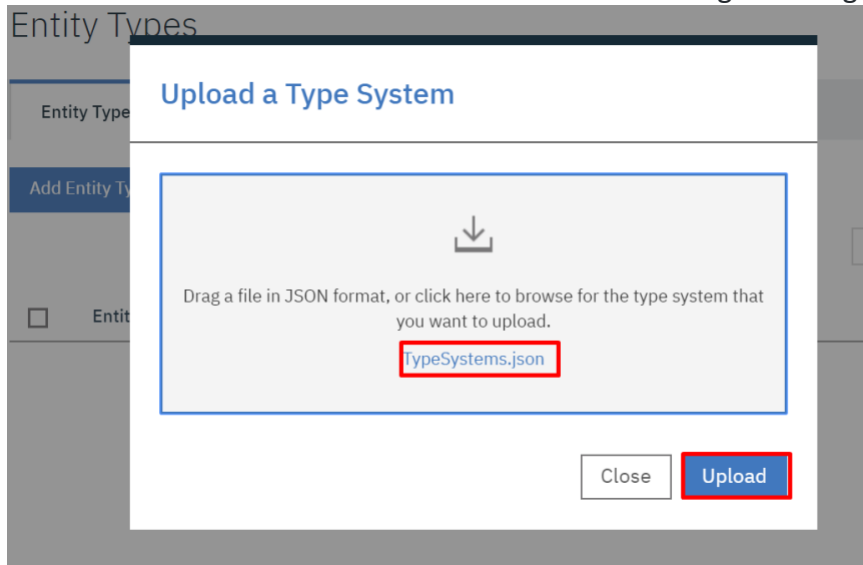
Entity Types



3. Haga clic en el icono cargar y busque el archivo typesystems.json que se descargó del repositorio git



4. El archivo seleccionado se muestra en la ventana emergente. Haga clic en Upload



5. Los documentos se enumeran a continuación. Anote los tipos de entidad o palabras clave que nos interesan. Puede Agregar o editar entidades.

Entity Types

Entity Types

8

Mention Classes

Mention Types

Add Entity Type

Upload

Download Types

Entity Type Name

Roles

Subtypes

Action

DOJ

DOJ

Edit

Delete

MobileNo

MobileNo

Edit

Delete

EmployeeId

EmployeeId

Edit

Delete

Company

Company

Edit

Delete

EmailId

EmailId

Edit

Delete

Address

Address

Edit

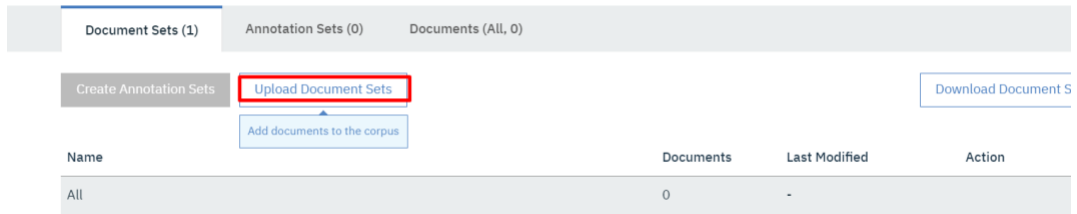
Delete

4.4 Importar documentos

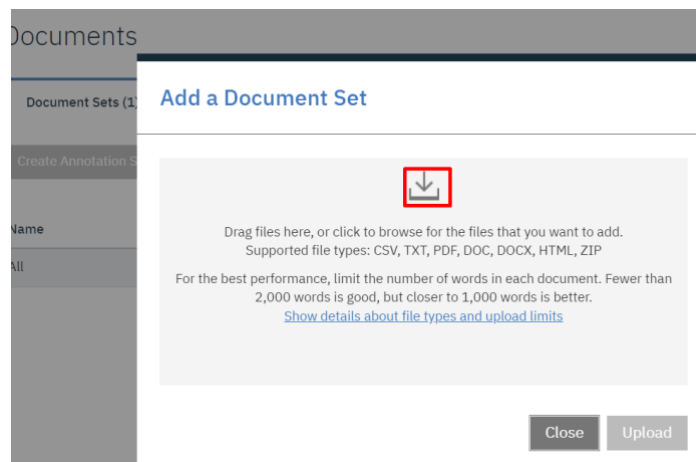
1. Haga clic en documents en la barra de navegación izquierda bar

IBM Watson Knowledge Studio				
Back to Workspaces				
Assets				
Documents				
Entity Types				
Relation Types				
Dictionaries				
Rule-based Model				
Machine Learning Model				
Settings				
Help				
Documents				
Document Sets (1) Annotation Sets (0) Documents (All, 0)				
Create Annotation Sets		Upload Document Sets		Download Document Sets
Add documents to the corpus				
Name	Documents	Last Modified	Action	
All	0	-		

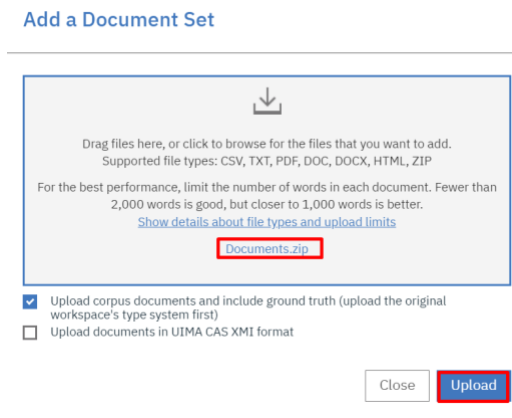
- Haga clic en Upload Document Sets Documents



- Haga clic en el botón Upload en la ventana emergente



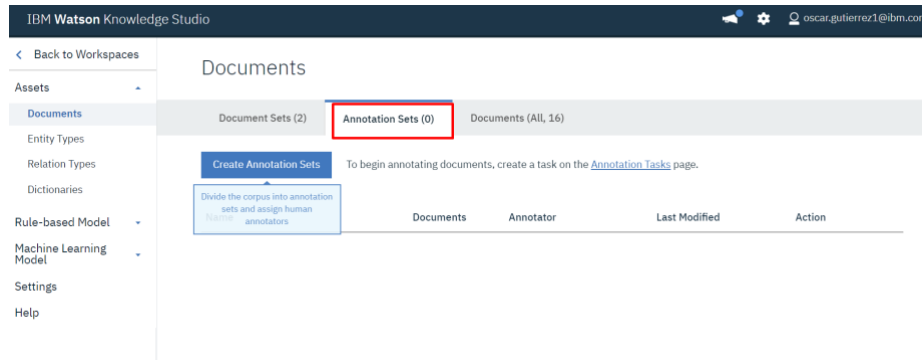
- Busque y seleccione Documents . zip archivo que se descargó desde el repositorio GitHub anterior
- Haga clic en upload



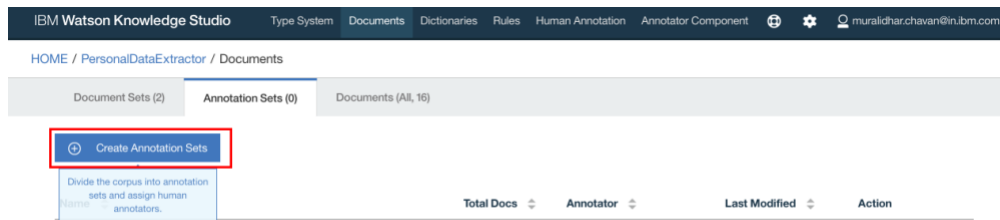
- Documents are now imported.

4.5 Crear y asignar conjuntos de anotaciones

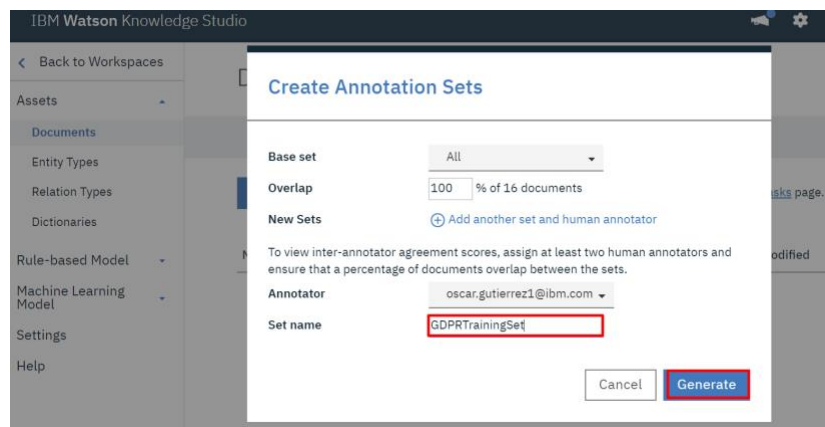
1. Haga clic en `annotation sets` para crear conjuntos de anotaciones.



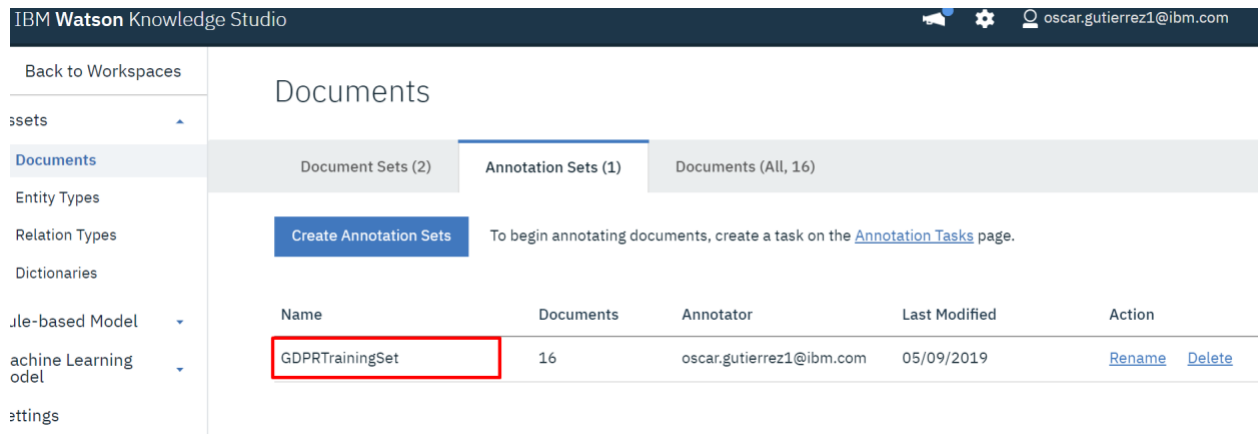
2. Click Create Annotation Sets



3. Escriba el nombre del conjunto de anotaciones y haga clic en `generate`



4. El conjunto de anotaciones esta creado.

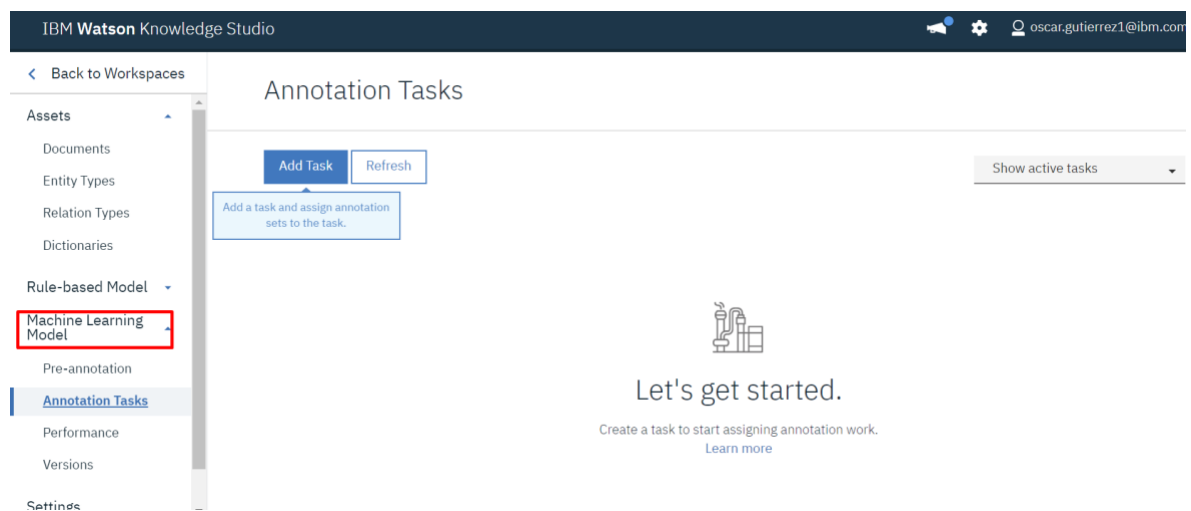


The screenshot shows the IBM Watson Knowledge Studio interface. The top navigation bar includes the logo, a notification bell, a settings gear, and the user profile 'oscar.gutierrez1@ibm.com'. The left sidebar contains a menu with 'Back to Workspaces', 'Assets' (expanded), 'Documents' (selected), 'Entity Types', 'Relation Types', 'Dictionaries', 'Rule-based Model', 'Machine Learning Model', and 'Settings'. The main content area is titled 'Documents' and has three tabs: 'Document Sets (2)', 'Annotation Sets (1)', and 'Documents (All, 16)'. Below the tabs is a blue button 'Create Annotation Sets' and a text instruction: 'To begin annotating documents, create a task on the [Annotation Tasks](#) page.' A table lists document sets with columns: Name, Documents, Annotator, Last Modified, and Action. The first row is 'GDPRTrainingSet' with 16 documents, annotated by 'oscar.gutierrez1@ibm.com' on '05/09/2019'. The 'Name' cell is highlighted with a red box. Action links 'Rename' and 'Delete' are provided for each row.

Name	Documents	Annotator	Last Modified	Action
GDPRTrainingSet	16	oscar.gutierrez1@ibm.com	05/09/2019	Rename Delete

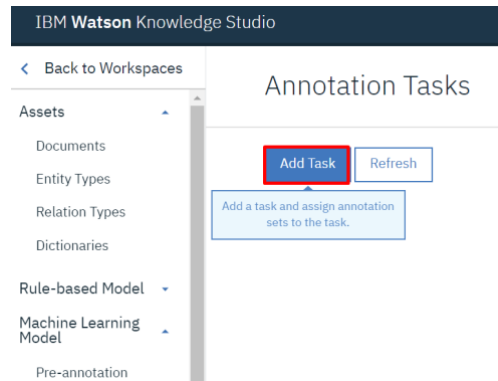
4.6 Modelo de Machine Learning

1. Haga clic en Machine Learning Model en la barra de navegación izquierda.

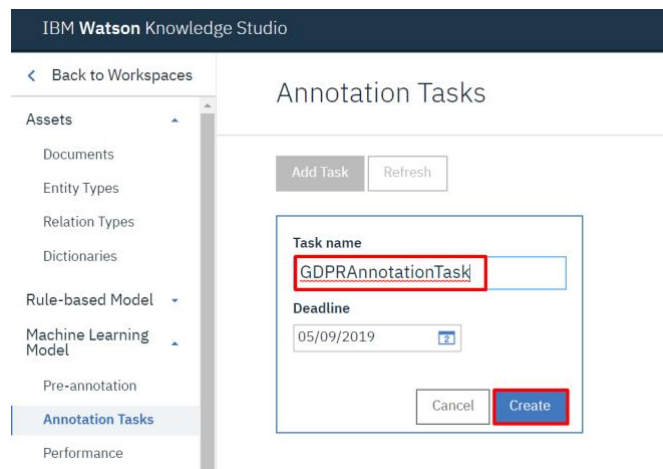


The screenshot shows the IBM Watson Knowledge Studio interface for the 'Annotation Tasks' page. The top navigation bar is the same as the previous screenshot. The left sidebar menu is identical, but 'Machine Learning Model' is highlighted with a red box. The main content area is titled 'Annotation Tasks' and features an 'Add Task' button and a 'Refresh' button. A tooltip points to the 'Add Task' button with the text: 'Add a task and assign annotation sets to the task.' On the right, there is a 'Show active tasks' dropdown menu. The main area contains a large illustration of a factory and the text 'Let's get started.' followed by 'Create a task to start assigning annotation work.' and a 'Learn more' link.

2. Haga clic en Add Task



3. Escriba un nombre para la tarea y haga clic en Create



4. En la ventana emergente, seleccione el conjunto de anotaciones que se creó anteriormente

Add Annotation Sets to the Task

Cancel Create Task

Annotation Set Name	Annotator Name	Task Name
<input checked="" type="checkbox"/> GDPRTrainingSet	oscar.gutierrez1@ibm.com	

5. Haga clic en `Create Task`

Add Annotation Sets to the Task

Cancel Create Task

Annotation Set Name	Annotator Name	Task Name
<input checked="" type="checkbox"/> GDPRTrainingSet	oscar.gutierrez1@ibm.com	

6. La tarea debe ser creada, haga clic en la tarea.

IBM Watson Knowledge Studio

Back to Workspaces

Assets

Documents

Entity Types

Relation Types

Dictionaries

Rule-based Model

Machine Learning Model

Pre-annotation

Annotation Tasks

Performance

Annotation Tasks

Add Task

Refresh

GDPRAnnotationTask

Deadline: 05/09/2019

Progress

0%

7. A continuación, tenemos que anotar asignación de entradas de documento con tipos de entidad definidos en `Type systems`

8. Haga clic en Annotate

IBM Watson Knowledge Studio

Back to Workspaces | GDPRAnnotationTask

Assets

- Documents
- Entity Types
- Relation Types
- Dictionaries

Rule-based Model

Machine Learning Model

- Pre-annotation
- Annotation Tasks**
- Performance
- Versions

Deadline: 05/09/2019

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.

Edit Annotation Sets Refresh Calculate Inter-Annotator Agreement Check Overlapping Documents for Conflicts

Apply Type System Updates

Annotation Set Name	Annotator Name	Status	Action
GDPRTrainingSet	oscar.gutierrez1@ibm.com	IN PROGRESS	Annotate

9. Haga clic en OK para cualquier mensaje de alerta que aparezca

10. Se abre el editor de la verdad terrestre. Aquí puede seleccionar cada documento uno por uno para anotar todos los documentos. Haga clic en cualquiera de los documentos

IBM Watson Knowledge Studio

Back to Workspaces

Assets

Documents

Entity Types

Relation Types

Dictionaries

Rule-based Model

Machine Learning Model

Pre-annotation

Annotation Tasks

Performance

Versions

Settings

Select Document

Annotation Set: GDPRTrainingSet

Annotator: oscar.gutierrez1@ibm.com

Task: GDPRAnnotationTask

Submit All Documents

Showing 1-16 of 16

Document Name	Status	Last Modified	Actions
Chat001.txt	Ready	Oct 30, 2017 2:38:28 AM	Open
Chat002.txt	Ready	Oct 30, 2017 2:38:28 AM	Open
Chat011.txt	Ready	Oct 30, 2017 2:38:28 AM	Open
Chat012.txt	Ready	Oct 30, 2017 2:38:28 AM	Open
Chat013.txt	Ready	Oct 30, 2017 2:38:28 AM	Open
Chat014.txt	Ready	Oct 30, 2017 2:38:28 AM	Open
Chat015.txt	Ready	Oct 30, 2017 2:38:28 AM	Open

Close

- En los documentos, seleccione una entrada que desee extraer del documento como entidades.
A continuación, haga clic en el tipo de entidad en la parte derecha de la pantalla
- De forma similar, haga esto para todas las palabras clave del documento

Knowledge Studio

Back to Task

Open document list

Alphab...

14pt

1

View Details

Replace

Concordance

Attribute View

Ready

Save

Mention

Relation

Coreference

Chat001.txt

1 Rep: This is Thomas.

2 How can I help you?

3 Caller: This is Alex.

4 I want to change my plan to corporate plan Rep: Sure, I can help you.

5 Do you want to change the plan for the number from which you are calling now?

6 Caller: yes Rep: For verification purpose may I know your date of birth and email id Caller: My data of birth is 10-Aug-1979 and my email id is alex@gmail.com Rep: Which plan do you want to migrate to Caller: Plan 450 unlimited Rep: Can I have your company name and date of joining Caller: I

Entity

Mention

Type

Subtype

Role

Address

Company

DOB

DOJ

EmailId

EmployeeId

MobileNo

Name

13. Una vez que todas las palabras clave se asignan a tipos de entidad, seleccione **completado** en el menú desplegable de estado

The screenshot shows the Knowledge Studio interface with the document 'Chat001.txt' open. The document content is as follows:

1 Rep: This is Thomas.
2 How can I help you?
3 Caller: This is Alex.
4 I want to change my plan to corporate plan Rep: Sure, I can help you.
5 Do you want to change the plan for the number from which you are calling now?
6 Caller: yes Rep: For verification purpose may I know your date of birth and email id Caller: My data of birth is 10-Aug-1979 and my email id is alex@gmail.com Rep: Which plan do you want to migrate to Caller: Plan 450 unlimited Rep: Can I have your company name and date of joining Caller: I work for IBM and doj 01-Feb-99 Rep: Ok.. I have taken your request to

The status dropdown menu is open, showing the following options: Ready, Ready, In Progress, and **Completed** (highlighted with a red box). The 'Save' button is also visible.

On the right side, the 'Entity' table is visible, showing the following columns: Type, Subtype, and Role. The table contains the following rows:

Type	Subtype	Role
Address		
Company		
DOB		
DOJ		
EmailId		
EmployeeId		
MobileNo		
Name		

14. Haga clic en **Save** para guardar los cambios

The screenshot shows the Knowledge Studio interface with the document 'Chat001.txt' open. The document content is the same as in the previous screenshot. The status dropdown menu is now closed, and the 'Save' button is highlighted with a red box.

On the right side, the 'Entity' table is visible, showing the same columns and rows as in the previous screenshot.

15. Repita los pasos anteriores para todo el documento. Todos los documentos deben ser anotados y completados

IBM Watson Knowledge Studio

Back to Workspaces

Assets

Documents

Entity Types

Relation Types

Dictionaries

Rule-based Model

Machine Learning Model

Pre-annotation

Annotation Tasks

Performance

Versions

Settings

Help

Select Document

Annotation Set: GDPRTrainingSet

Annotator: oscar.gutierrez1@ibm.com

Task: GDPRAnnotationTask

Submit All Documents

Showing 1-16 of 16

Document Name	Status	Last Modified	Actions
Chat001.txt	Completed	May 9, 2019 11:02:57 AM	Open
Chat002.txt	Completed	May 9, 2019 11:17:37 AM	Open
Chat010.txt	Completed	May 9, 2019 11:19:43 AM	Open
Chat011.txt	Completed	May 9, 2019 11:21:24 AM	Open
Chat012.txt	Completed	May 9, 2019 11:25:04 AM	Open
Chat013.txt	Completed	May 9, 2019 11:26:00 AM	Open
Chat014.txt	Completed	May 9, 2019 11:27:29 AM	Open
Chat015.txt	Completed	May 9, 2019 11:28:54 AM	Open
Chat016.txt	Completed	May 9, 2019 11:30:47 AM	Open

16. Si el estado se muestra in progress, haga clic en el botón refresh

← GDPRAnnotationTask

Deadline: 05/09/2019

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.

Edit Annotation Sets

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"/> GDPRTrainingSet	oscar.gutierrez1@ibm.com	SUBMITTED	View

17. Ahora el estado debe cambiar a submitted

18. Seleccione el nombre del conjunto de anotaciones y haga clic en el botón Accept

IBM Watson Knowledge Studio

Back to Workspaces

GDPRAnnotationTask

Assets

- Documents
- Entity Types
- Relation Types
- Dictionaries

Rule-based Model

Machine Learning Model

Pre-annotation

Annotation Tasks

Deadline: 05/09/2019

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.

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

Accept Reject

Annotation Set Name	Annotator Name	Status	Action
<input checked="" type="checkbox"/> GDPRTrainingSet	oscar.gutierrez1@ibm.com	SUBMITTED	View

19. Haga clic en ok en la Ventana emergente de confirmación

Confirmation

Are you sure that you want to accept the selected annotation sets?

Cancel OK

20. El estado de la tarea ahora cambia a completed

IBM Watson Knowledge Studio

Back to Workspaces

GDPRAnnotationTask

Assets

- Documents
- Entity Types
- Relation Types
- Dictionaries

Rule-based Model

Machine Learning Model

Pre-annotation

Annotation Tasks

Deadline: 05/09/2019

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.

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

Annotation Set Name	Annotator Name	Status	Action
GDPRTrainingSet	oscar.gutierrez1@ibm.com	COMPLETED	View

21. Haga clic en Annotator Component en la barra de navegación superior

IBM Watson Knowledge Studio Type System Documents Dictionaries Rules Human Annotation **Annotator Component** muralidhar.chavan@in.ibm.com

HOME / PersonalDataExtractor / Human Annotation

Tasks Ground Truth Editor Settings IAA Settings Annotation Guidelines

Go Back to Task List

GDPRAnnotationTask

Deadline: 10/30/2017

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.

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

Annotation Set Name	Annotator Name	Status	Action
GDPRTrainingSet	Muralidhar Chavan	COMPLETED	View

22. Crearemos el anotador de Machine Learning Model y hacemos clic en bajo Machine learning y luego en Train and Evaluate

Back to Workspaces

Assets Documents Entity Types Relation Types Dictionaries Rule-based Model Rules Versions Machine Learning Model Pre-annotation Annotation Tasks **Performance** Versions Settings

Performance

PersonalDataExtractor
Language of documents: English

Number of documents per set

- 0 Training Set
- 0 Test Set
- 0 Blind Set

Training Set [View Ground Truth](#)

Test Set [View Ground Truth](#) [View Decoding Results](#)

Last trained on:

Last evaluated on:

[Train and evaluate](#) You have not trained the machine learning model. Click here to train and evaluate it.

Document set evaluation

Model over time [View Log](#)

1

0.80

Mention
Precision: --
Recall: --

Relation
Precision: --

23. Seleccione el conjunto cuya anotación se completó en los pasos anteriores. Haga clic en Train and Evaluate

IBM Watson Knowledge Studio oscar.gutierrez

Back to Workspaces Training / Test / Blind Sets

Select Training/ Test/ Blind Sets

[Train](#) [Train & Evaluate](#)

Document Set	Task Status
<input type="checkbox"/> All	
<input type="checkbox"/> GDPR_set	
<input checked="" type="checkbox"/> GDPRTrainingSet	COMPLETED

Create new sets by splitting the selected document sets

Ratio
Enter the percentage of documents to include in each set.

70 Training Set (70% Recommended)

23 Test Set (23% Recommended)

7 Blind Set (7% Recommended)

Performance

Train processing... Time elapsed: 40 seconds

PersonalDataExtractor

Language of documents

English

Number of documents per set

11

3

2

Training Set

Test Set

Blind Set

Training Set

View Ground Truth

Test Set

View Ground Truth

View Decoding Results

Last trained on:

Last evaluated on:

Train and evaluate

You have not trained the machine learning model. Click here to train and evaluate it.

24. El proceso de entrenamiento y evaluación tiene lugar. Este paso llevará unos minutos completar

IBM Watson Knowledge Studio

Train processing... Time elapsed: 2 minutes

Back to Workspaces

Training / Test / Blind Sets

Assets

Documents

Entity Types

Relation Types

Dictionaries

Rule-based Model

Machine Learning Model

Pre-annotation

Annotation Tasks

Performance

Versions

Last trained on:

Last evaluated on:

Train

Evaluate

Train & Evaluate

Edit Settings

Set Name	Number of Documents	Action
Training Set	11	View Ground Truth
Test Set	3	View Ground Truth View Decoding Results
Blind Set	2	

25. El proceso se mostrará como Success cuando este Completado.

IBM Watson Knowledge Studio

Success: May 9, 2019 12:18:14 PM

Machine learning model evaluation complete.

Back to Workspaces

Training / Test / Blind Sets

Assets

Documents

Entity Types

Relation Types

Dictionaries

Rule-based Model

Machine Learning Model

Pre-annotation

Annotation Tasks

Performance

Versions

Settings

Last trained on:

Last evaluated on:

Train

Evaluate

Train & Evaluate

Edit Settings

Set Name	Number of Documents	Action
Training Set	11	View Ground Truth
Test Set	3	View Ground Truth View Decoding Results
Blind Set	2	

5. Desplegar un modelo de WKS en Watson Natural Language Understanding

1. Entrar a WKS, usando el URL señalado abajo del paso de los requisitos previos, navegue al **Annotator Component** y haga clic en el **NLU**

IBM Watson Knowledge Studio

Back to Workspaces

Assets

- Documents
- Entity Types
- Relation Types
- Dictionaries

Rule-based Model

Machine Learning Model

- Pre-annotation
- Annotation Tasks
- Performance
- Versions**

Settings

Help

Versions

Machine Learning Model

Machine learning models, once trained, can be used to annotate new documents prior to human annotation.

Run this model

Export the current version of your model to use in other Watson applications such as Watson Explorer.

Export current model

Version History and Deployment

Version	Base	Creation Date	Entity Scores	Relation Scores	Description	Action
1.1	Current Version		0.86 (1.00 / 0.75)	N/A		Create Version
1.0		05/14/2019	0.86 (1.00 / 0.75)	N/A		Promote Delete Deploy

Deployed Models (1)

Model ID: 7c6e9b17-cd0a-4048-b27b-63c78abe692b	Service ID: 7fa692fa-88a3-4f04-86db-4576a879-bc7b	Undeploy Status
--	---	-----------------

Deploy Model v1.0

Deploying to Natural Language Understanding
You must have a subscription to the IBM Watson™ Natural Language Understanding service, and know the names of your IBM Cloud space and service instance. [Learn more](#)

IBM Cloud Information
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.

Region
Dallas

Space or resource group
☐ Spaces ☒ Resource groups
Default

Service name
NLUGDPR

Cancel Deploy

2. Hacer clic en Create Version.
Versions

Machine Learning Model

Machine learning models, once trained, can be used to annotate new documents prior to human annotation.

[Run this model](#)

Export the current version of your model to use in other Watson applications such as Watson Explorer.

[Export current model](#)

Version History and Deployment

Version	Base	Creation Date	Entity Scores	Relation Scores	Description	Action
2.1		<i>Current Version</i>	0.86 (1.00 / 0.75)	N/A		Create Version

3. Click en Ok en la Ventana emergente

Take a snapshot

Create a snapshot of the current annotator component artifacts. This action creates a new version of the model, and keeps a copy of the artifacts that were used to build it.

Description (Optional):

Enter a description

Cancel

OK

4. Una vez creada la nueva versión hacemos clic en Deploy.

Versions

Machine Learning Model

Machine learning models, once trained, can be used to annotate new documents prior to human annotation.

[Run this model](#)

Export the current version of your model to use in other Watson applications such as Watson Explorer.

[Export current model](#)

Version History and Deployment

Version	Base	Creation Date	Entity Scores	Relation Scores	Description	Action
2.2		Current Version	0.86 (1.00 / 0.75)	N/A		Create Version
2.1		05/26/2019	0.86 (1.00 / 0.75)	N/A		Promote Delete Deploy

5. Seleccione Natural language Understanding y luego, haga clic en Next.

Deploy Model v2.1

Select a service to deploy to.
All services require a subscription. [Learn more](#)

☒ Natural Language Understanding

☐ Discovery

[Cancel](#) [Next](#)

6. Seleccione las instancias de servicio de IBM Cloud region, Space y NLU. Haga clic en Desplegar.

Deploy Model v1.0

Deploying to Natural Language Understanding

You must have a subscription to the IBM Watson™ Natural Language Understanding service, and know the names of your Bluemix space and service instance. [Learn more](#)

Bluemix Information

IBM Bluemix® is IBM's cloud platform. Click [here](#) to open Bluemix and create an account or look up details for an existing service.

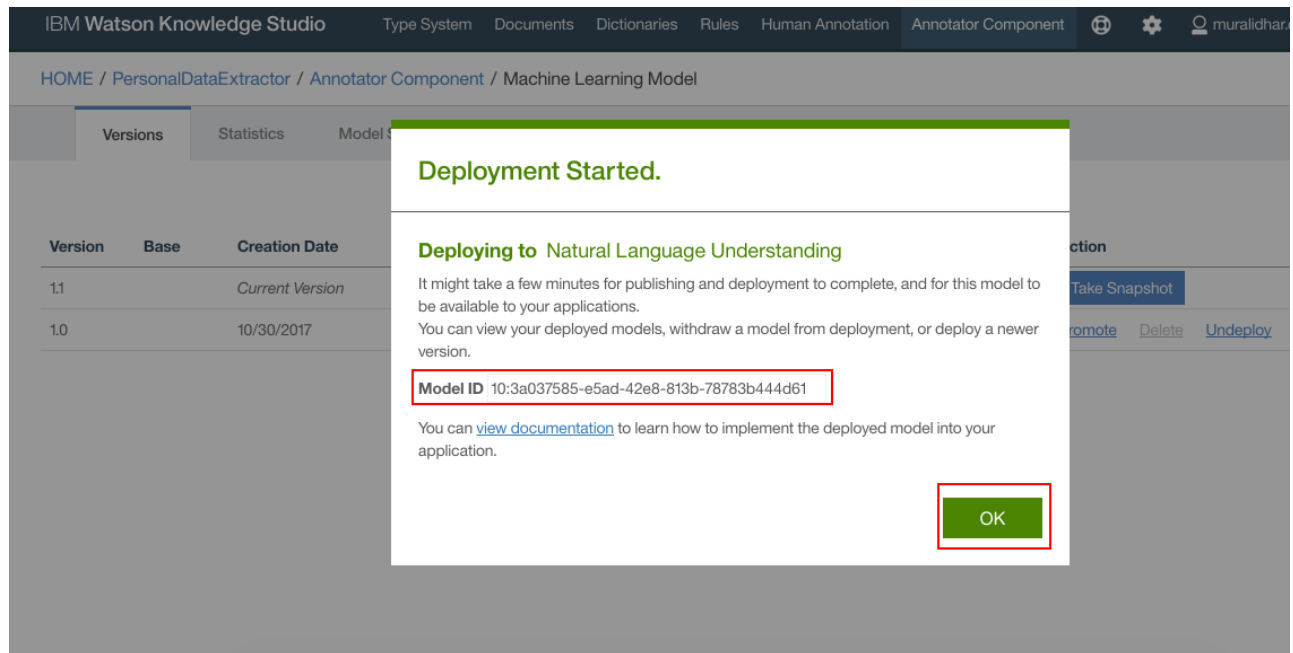
Region

Space

Service instance

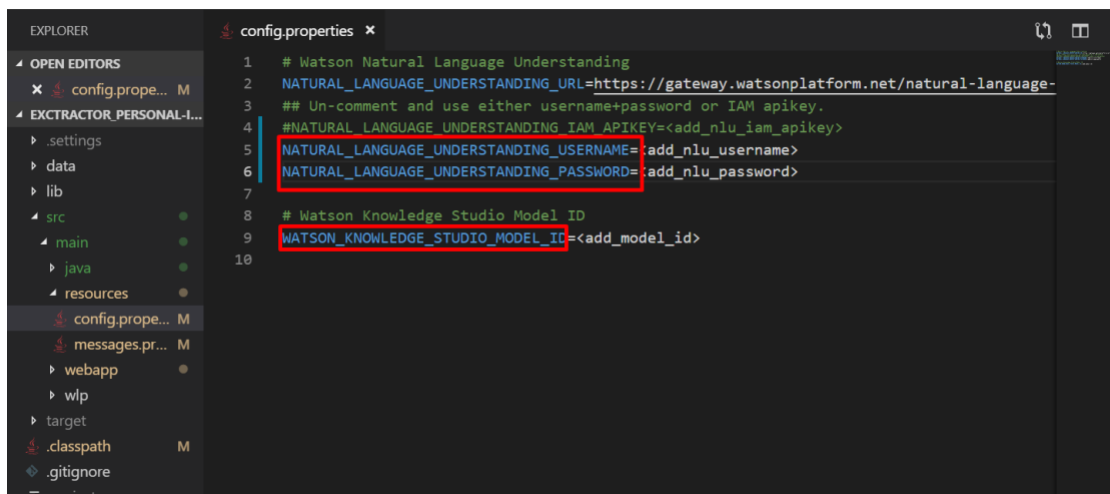
[Cancel](#) [Deploy](#)

7. El modelo WKS debe ser desplegado en el NLU. Anote el ID del modelo. Haga clic en OK

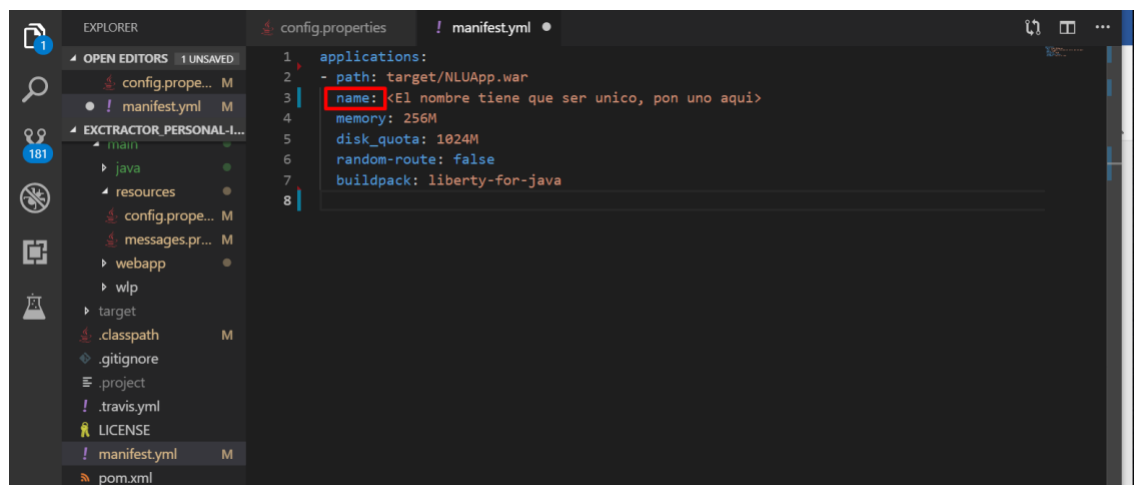


Despliegue de la app

- Ir a la carpeta descargada del repositorio y abrir el archivo “config.properties” (src/main/resources)
- Poner las respectivas credenciales del servicio de Natural Language Understanding las cuales ya habíamos tomado nota.
 - El username es: “apikey” (Sin las comillas)
 - El password es: <el valor de esa “apikey”>
 - Model_ID: <El ID del modelo personalizado que se creó en Knowledge Studio>







- Abrir el archivo “manifest.yml” (root) y poner un nombre



- En la terminal posicionarse en el archivo descargado de GitHub y correr el siguiente comando
 - Mvn clean package
- Correr el siguiente comando para desplegar la app en la nube
 - lbmcloud cf push
- Ahora podemos irnos a la parte de servicios de IBMcloud para poder disfrutar de nuestra app que acabamos de crear

Resource list

[Create resource](#)[Collapse all](#) | [Expand all](#)

Name ▲	Group	Location	Status	Tags
Q Filter by name or IP address...	Filter by group or org...	Filter...	Q Filter...	Filter...
> Devices (0)				
> Kubernetes Clusters (0)				
▼ Cloud Foundry Apps (1)				
 jabv-asset	AlejandroB.Villasenor@ibm.com / dev	Dallas	● Stopped	--
▼ Cloud Foundry Services (0)				
▼ Services (5)				
 Continuous Delivery	Default	Dallas	Provisioned	--
 Internet Services-mj	Default	Global	Provisioned	--
 Knowledge Studio-ed	Default	Dallas	Provisioned	--

FEEDBACK

Escribe el texto que quieres analizar

Empleado: Hola! Soy Juan. Como puedo ayudarle

Cliente: Me llamo Alex. Quiero cambiar mi plan al
plan

Empleado: Claro. El numero que quiere cambiar al
plan es con el que está llamando

Cliente: Si

Empleado: Para propositos de verificacion le haré
algunas preguntas, está de acuerdo

155 words

Despejar

Extraer