

Introduction to Watson Knowledge Catalog

Introduction

This lab will introduce Watson Knowledge Catalog. Watson Knowledge Catalog is a secure enterprise catalog to discover, catalog and govern your data/models with greater efficiency. The catalog is underpinned by a central repository of metadata describing all the information managed by the platform. Users will be able to securely share data with their colleagues more easily, regardless of what the data is, where it is stored, or how they intend to use it. In this way, the intelligent asset catalog will unlock the value held within that data across user groups helping organizations use this key asset to its full potential.

End-to-End Data Science

The general flow of the End to End Data Science PoT will be guided by the activities shown in Figure 1- End to End Flow. This lab will focus on the **Find Governed Data** activity and demonstrate **Connect to Data**.

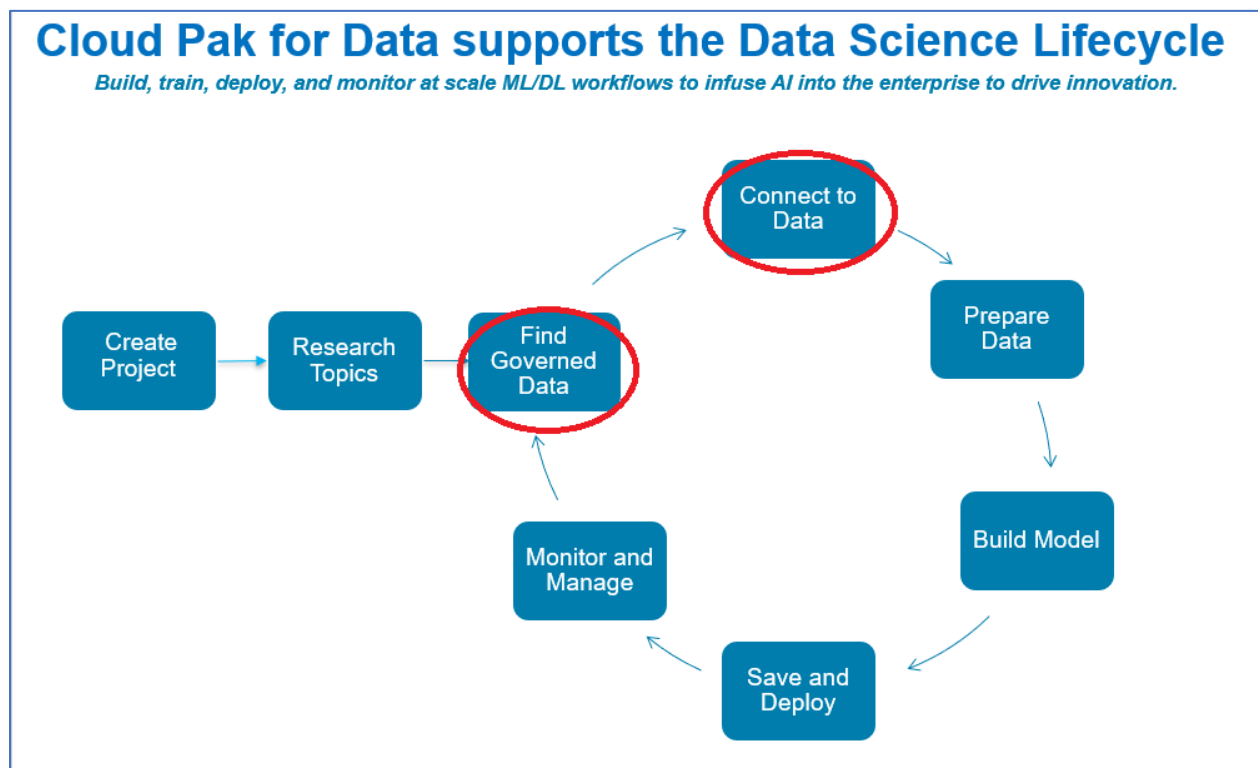


Figure 1- End to End Flow

Objectives

The goal of the lab is for the users to gain familiarity with the features of the Watson Knowledge Catalog. We will perform the following catalog tasks:

1. Create a governed catalog

2. Add a member to the catalog
3. Add Data Assets to the catalog
4. Search the catalog
5. Edit/Review/Profile a structured Data Asset
6. Demonstrate access control features
7. Demonstrate policy creation and enforcement (optional)
8. Push the Data Assets to the project set up in Lab-1

Download Files

Step 1: Download the trafficking_data.zip from the github repository

1. Click [here](#), to download the zipped file.
2. Extract the file contents. You should have two files extracted, (1) Categories.csv, and (2) Occupation.csv.

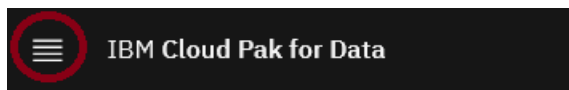
Perform Catalog Tasks

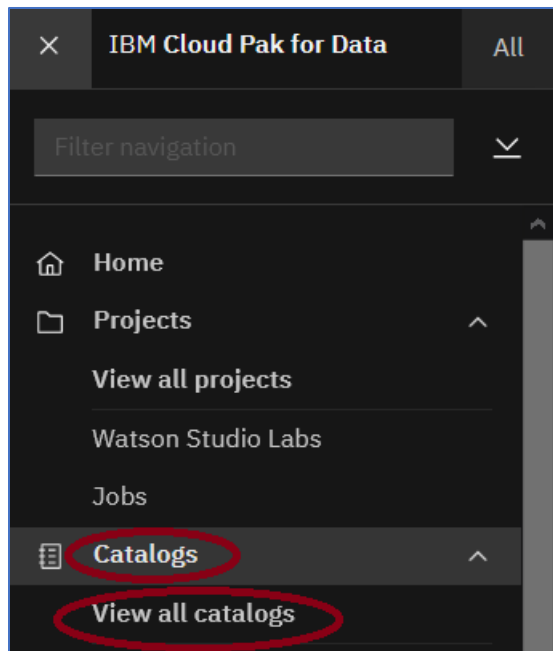
Step 1 – Create a Governed Catalog

1. Click on the **IBM Cloud Pak for Data** browser tab to return to **Cloud Pak for Data**.
2. Click on **Cloud Pak for Data** to go to the home screen.

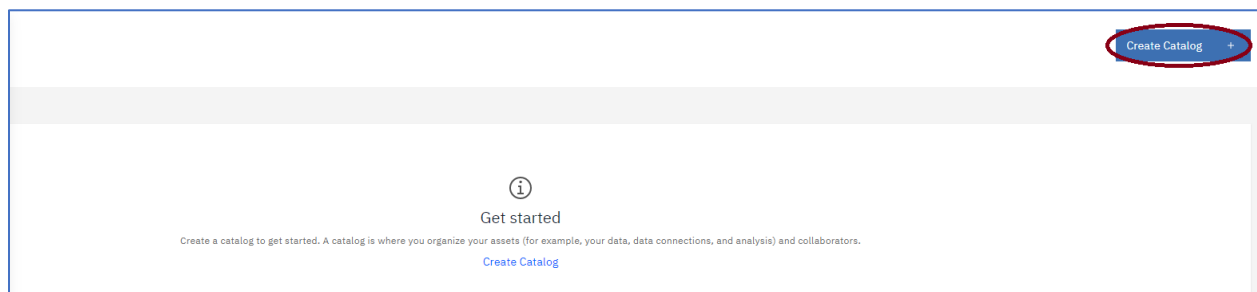


3. Click on  icon, click on **Catalog**, and then **View All Catalogs**.

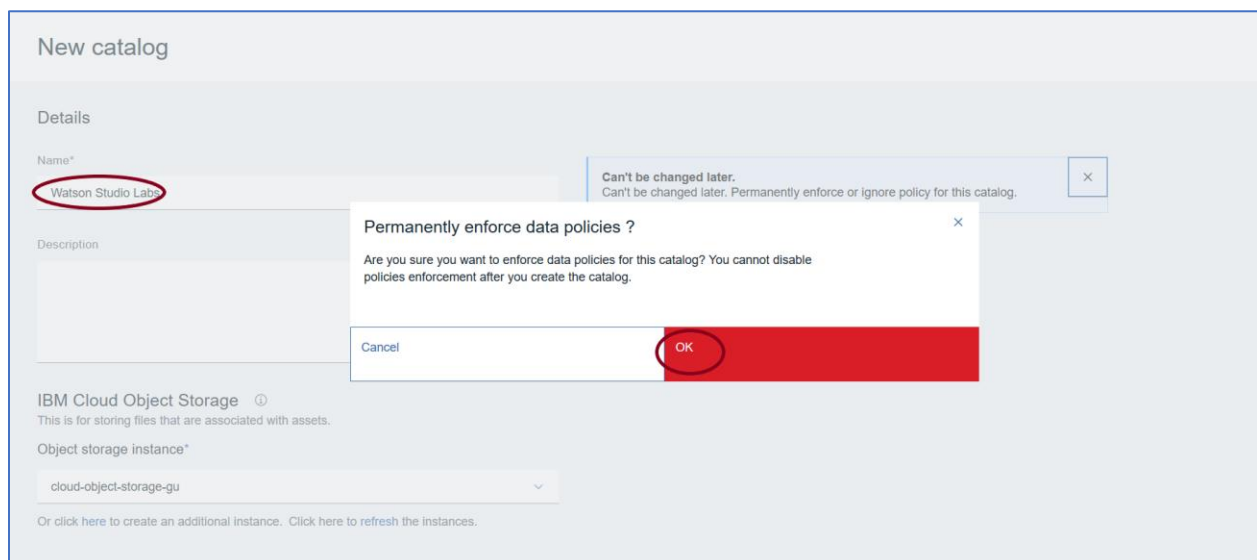




4. Click on **Create Catalog**.



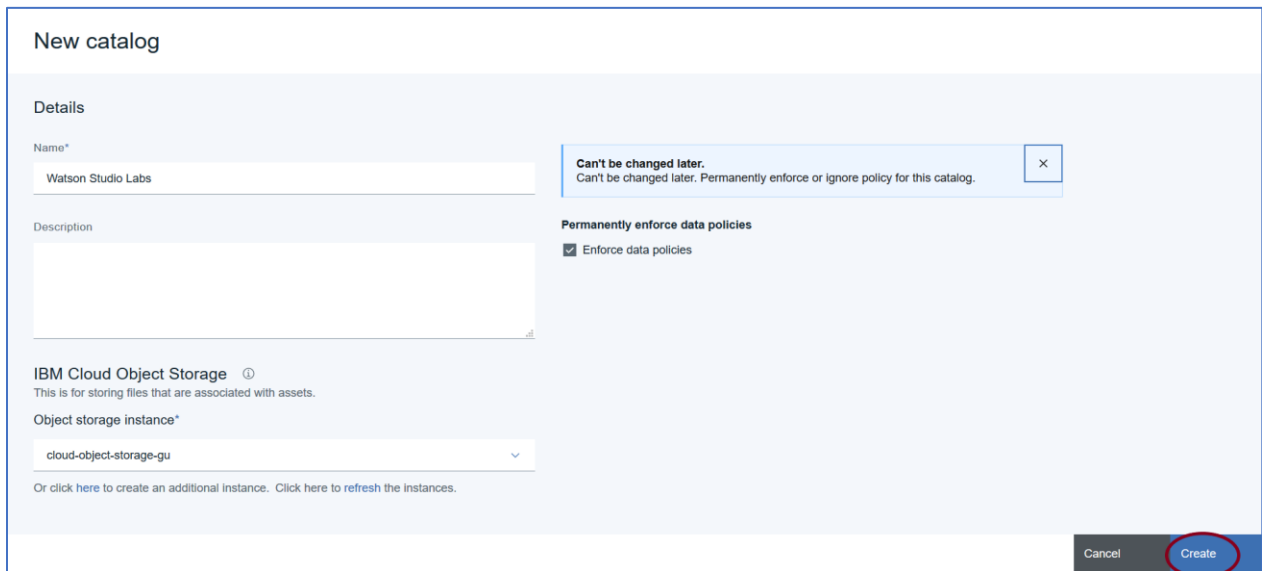
5. Enter **Watson Studio Labs** for the **Catalog** name. Click on **Permanently enforce data policies**. You will get a pop-up as shown below. Click on **OK**.



Note: By default, access to data assets in a catalog are only restricted by the privacy settings of the data assets. Privacy settings and policy rules can limit which members of the catalog can view and use the assets. You can implement data policies to restrict access to data based on the contents of the data. Data policies help you control data access and ensure that the right people can access the right data. Selecting the option to **enforce data policies** enables the enforcement of data policy rules to allow or deny access to a data asset or mask, anonymize and redact data at the data asset field level.

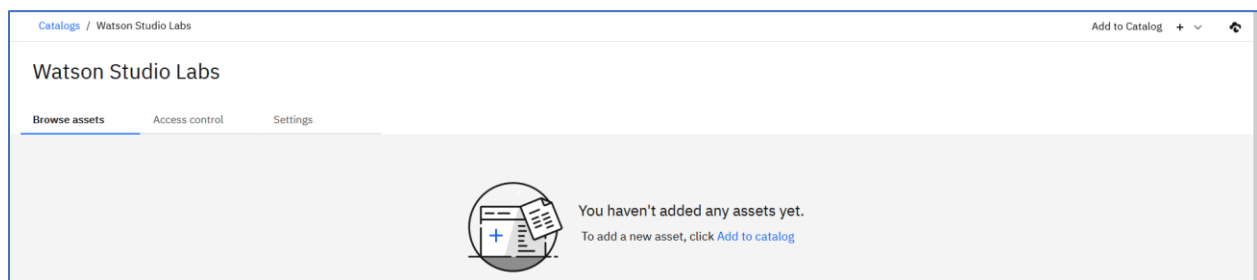
Setting this option for a catalog is a good best practice. Once it is enabled it cannot be undone, but it does not restrict or impede any functionality, it provides additional security measures to protect data assets.

6. Click **Create** to create the catalog.



The screenshot shows the 'New catalog' form in Watson Studio Labs. The form has a 'Details' section with a 'Name*' field containing 'Watson Studio Labs' and a 'Description' field. A warning box states: 'Can't be changed later. Can't be changed later. Permanently enforce or ignore policy for this catalog.' Below the description, there is a section for 'IBM Cloud Object Storage' with a note: 'This is for storing files that are associated with assets.' The 'Object storage instance*' dropdown menu is set to 'cloud-object-storage-gu'. At the bottom right, there are 'Cancel' and 'Create' buttons, with the 'Create' button circled in red. A checkbox labeled 'Permanently enforce data policies' is checked, with the text 'Enforce data policies' below it.

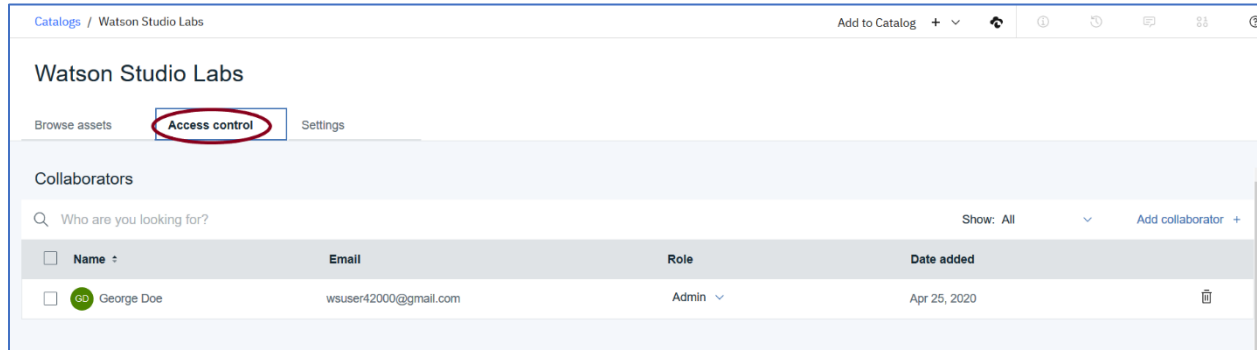
7. Now you are ready to add assets!



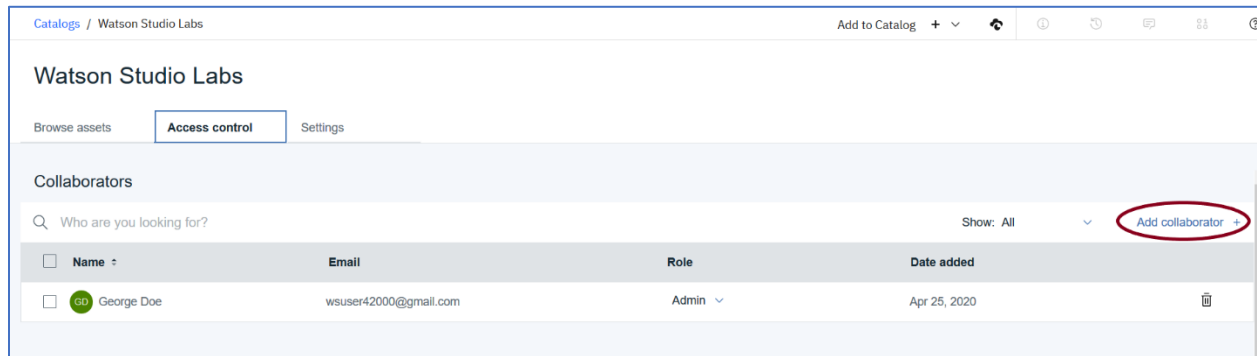
Step 2: Add a Member to the Catalog

To demonstrate the Catalog access control capability, we will add user `ws.catalog.user@gmail.com` as a member of the catalog. Note that the name of this user is “Howard Doe”.

1. Click on **Access Control**.



2. Click on **Add Collaborator**.



3. Leave Viewer as the **Role**. Enter “Howard”. The list should populate with “Howard Doe” and the ws.catalog.user@gmail.com e-mail address.

Add collaborators as Viewers ×

Having trouble adding collaborators? ⓘ ×

Role:

☐ Admin ☐ Editor ☒ Viewer

Collaborators (0)

Howard

Howard Doe	ws.catalog.user@gmail.com
------------	---------------------------

- Click on the **Howard Doe** entry in the list box.

Add collaborators as Viewers ×

Having trouble adding collaborators? ⓘ ×

Role:

☐ Admin ☐ Editor ☒ Viewer

Collaborators (0)

Howard

Howard Doe	ws.catalog.user@gmail.com
------------	---------------------------

- Click **Add**.

Add collaborators as Viewers

Having trouble adding collaborators? ⓘ

X

Role:

☐ Admin
☐ Editor
☒ Viewer

Collaborators (1)

Howard Doe X

Cancel

Add

6. Howard Doe is added as a collaborator on the Watson Studio Labs catalog.

Catalogs / Watson Studio Labs
Add to Catalog +

Watson Studio Labs

Browse assets
Access control
Settings

Collaborators

Who are you looking for?
Show: All
Add collaborator +

<input type="checkbox"/>	Name	Email	Role	Date added	
<input type="checkbox"/>	George Doe	wsuser42000@gmail.com	Admin	Apr 25, 2020	
<input type="checkbox"/>	Howard Doe	ws.catalog.user@gmail.com	Viewer	Apr 25, 2020	

Step 3: Add Data Assets to the Catalog

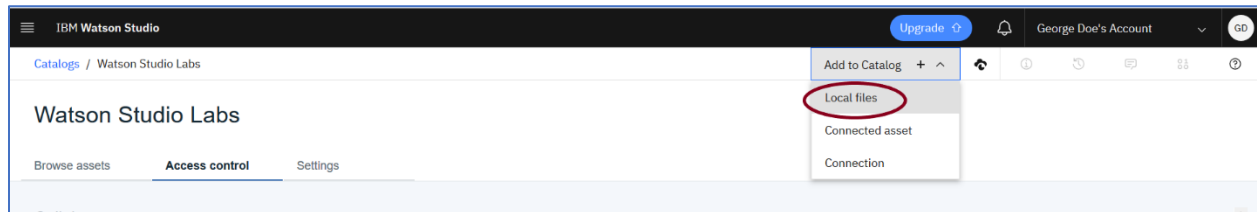
In this step, we will add 4 assets; two local files, a connection, and a connected data asset. The local files we will add are the “Occupation.csv” file, and the “Categories.csv” file. We will set up a connection to a DB2 Warehouse service. We will add a connected data asset that will point to the FEMALE_HUMAN_TRAFFICKING table in the DB2 Warehouse service instance.

This step will introduce you to the three methods available to discover and catalog data assets; **Local files**, **Connected asset** and **Connection**. You will use these methods to catalog

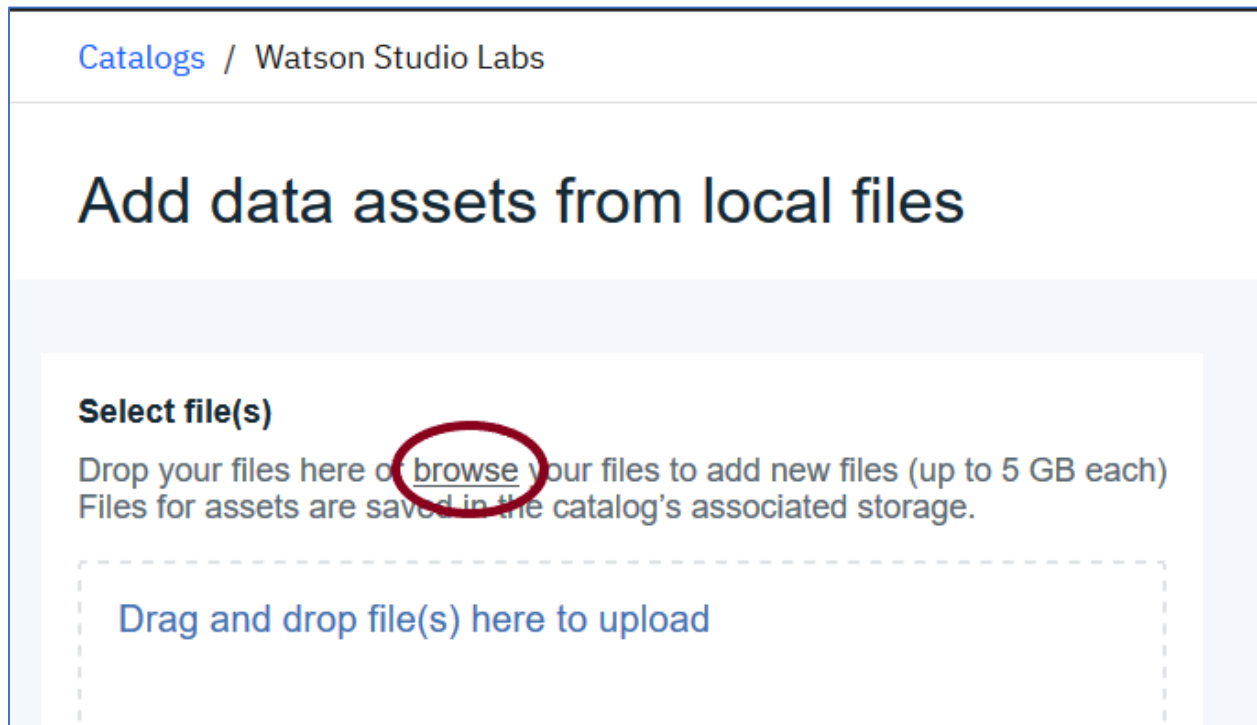
data assets into the newly created Knowledge Catalog and then tag them for users to easily find them, understand their content and make them available throughout IBM Cloud Pak for Data, for use during data preparation and within models, dashboards and notebooks as we will see in our lab exercises.

Add Local Files

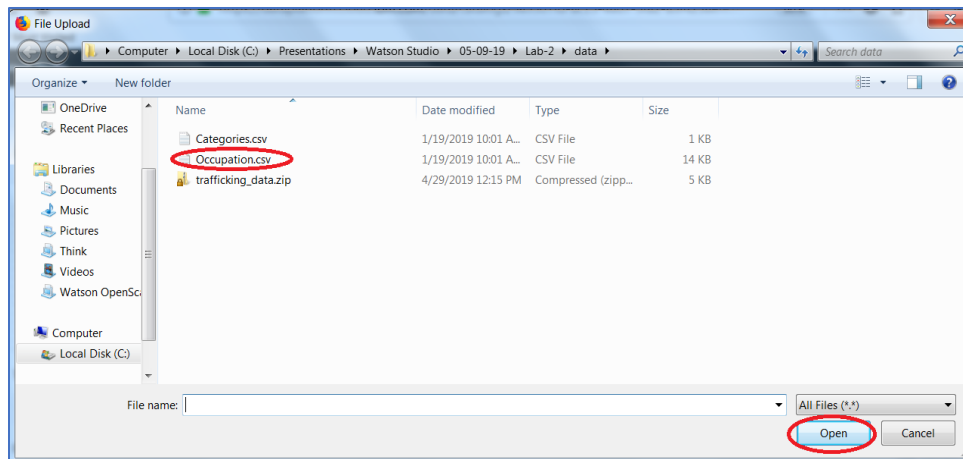
1. Click on **Add to Catalog** and click on **Local Files**.



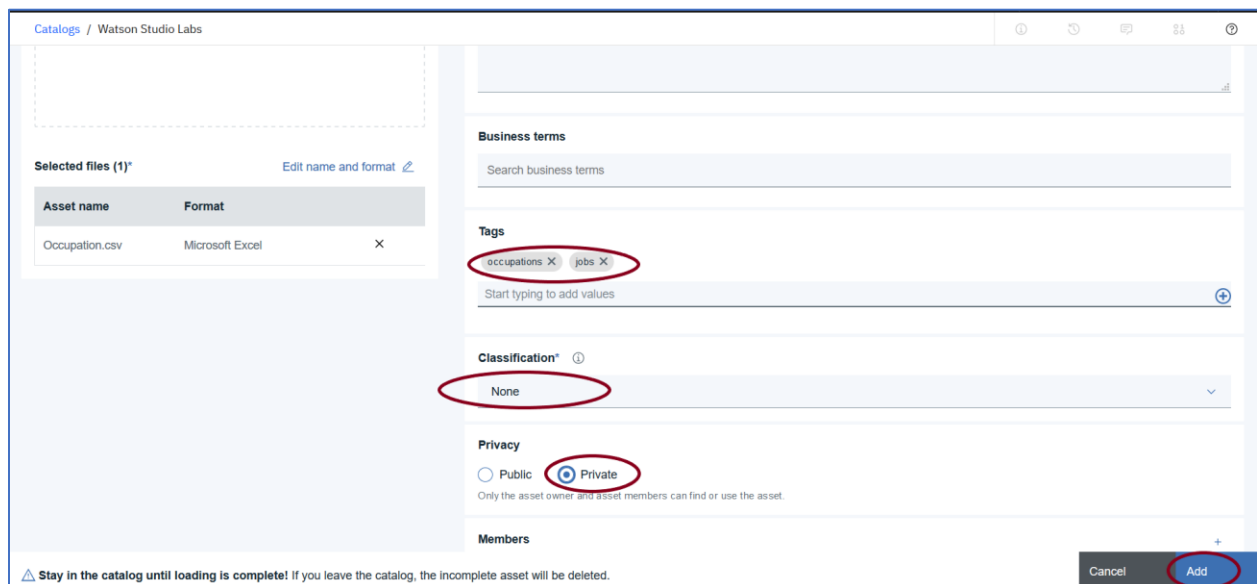
2. Click on **browse** to navigate to the folder where the “Occupation.csv” file is stored.



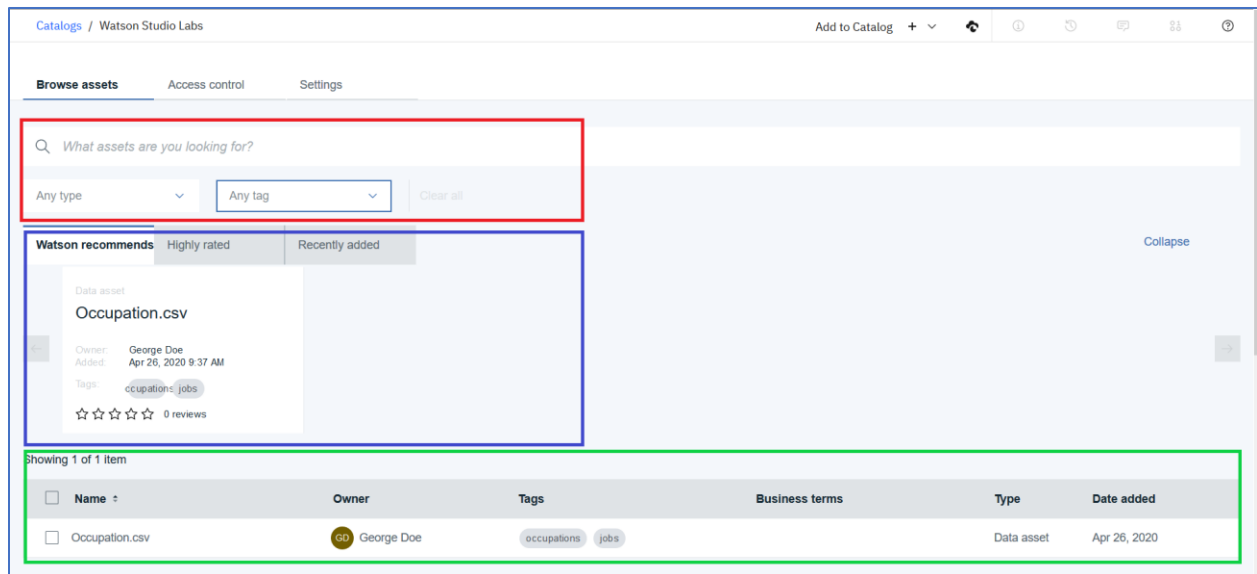
3. Click on **Occupation.csv** and then click **Open**.



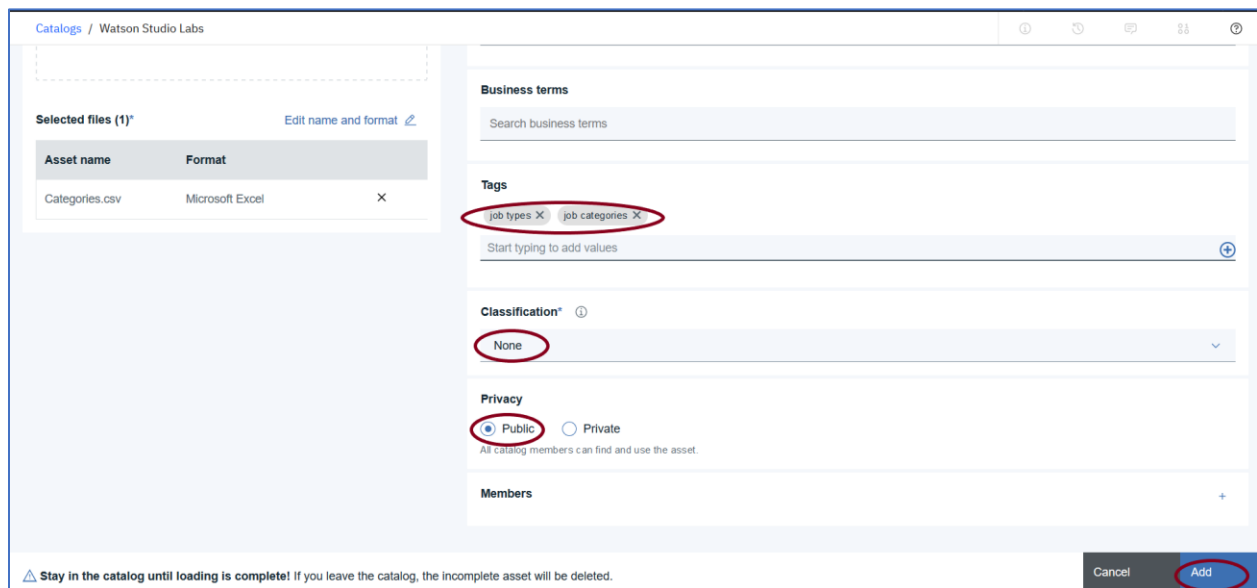
4. We can add tags to the newly added data asset to more easily search for it. Add two tags, enter “occupations”, click on the + icon, and enter “jobs” and click on the + icon. Leave the **Classification** as None. Click on Private for the **Privacy**. Click on **Add**.



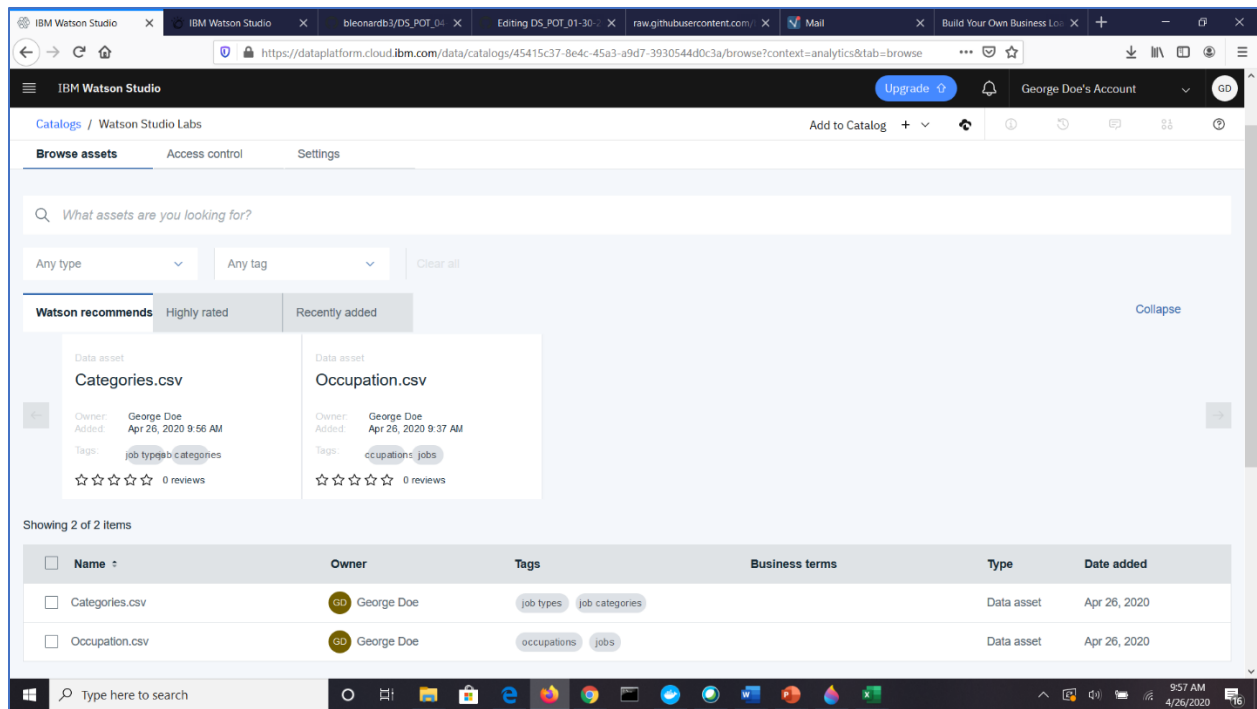
5. The Occupation.csv asset metadata is displayed. Note several features on the panel. The top part of the page contains the search area. This consists of an entry field to type search text and two filters, one by data type, and one by tags (filters shown in red below). Underneath the search area, are 3 tabs: **Watson Recommended**, **Highly Rated**, and **Recently added**. The square below contains assets that are displayed when one of these tabs are selected (shown in blue below). Note this area can be collapsed. At the bottom is the list of assets (shown in green).



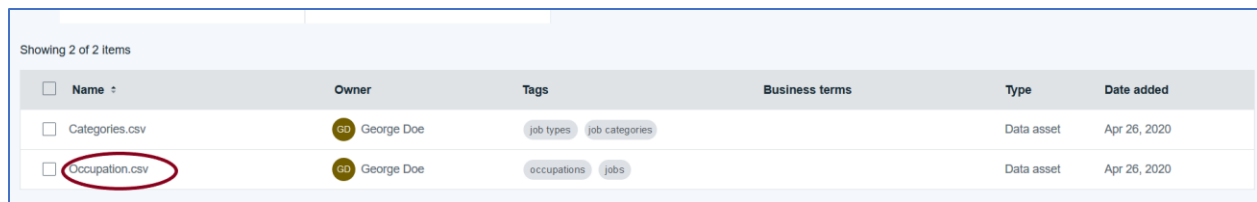
- Follow the same procedure (1-4) above to add the local file “Categories.csv” file to the catalog. For step 4, enter the tag “job types” and then click on the “+” icon, enter the tag “job categories” and then click on the “+” icon, leave the **Classification** as None, leave the **Privacy** as Public, and then click on **Add**.



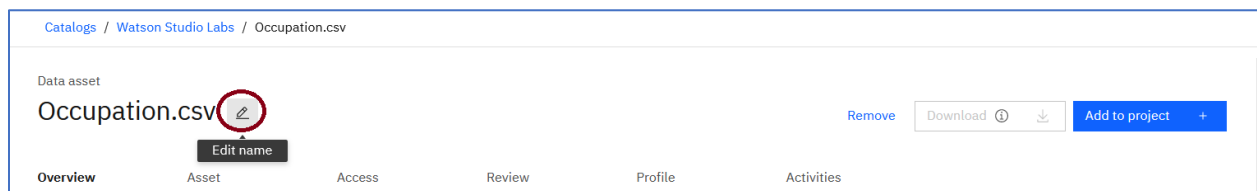
- The Categories.csv asset is displayed



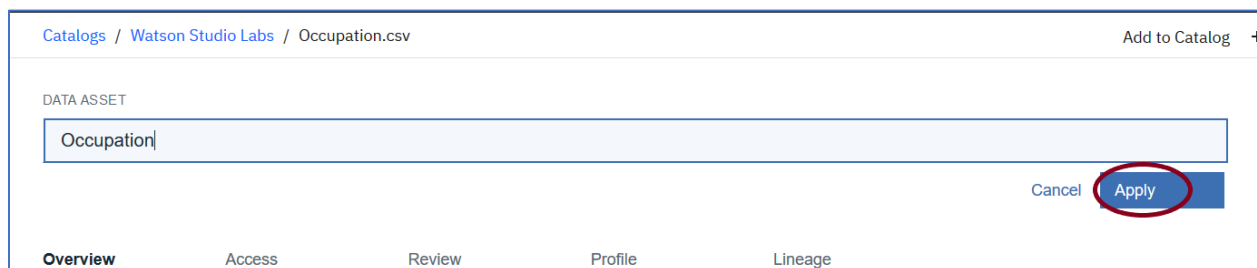
- Change the name of the Occupation.csv asset to be just Occupation by clicking on the **Occupation.csv** asset.



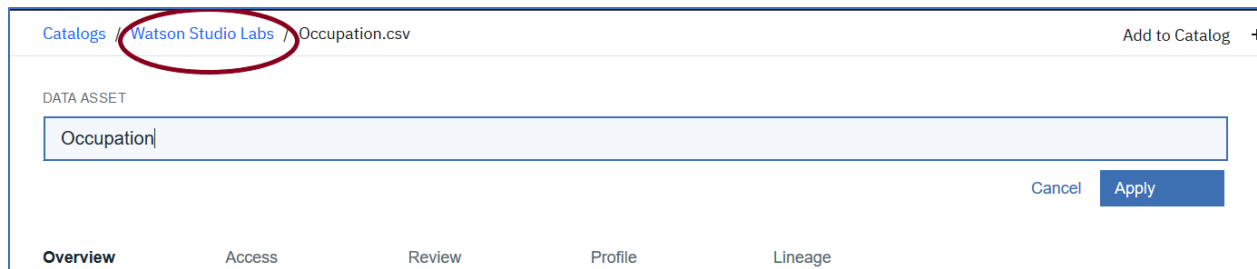
- Click the “pencil” icon adjacent to **Occupation.csv**.



- Change the name to Occupation and click **Apply**.



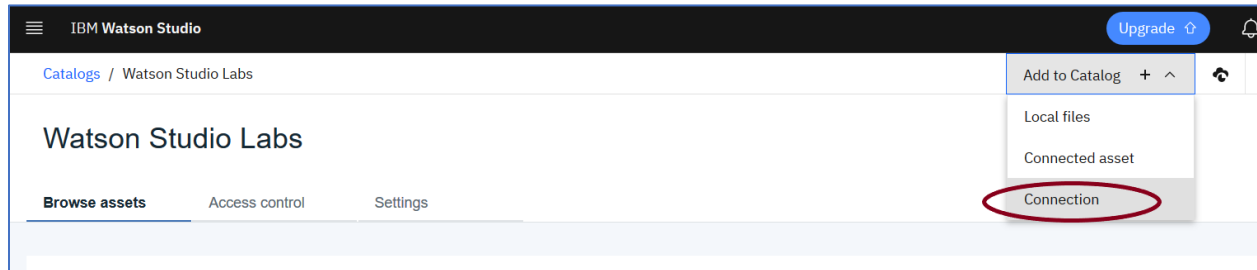
11. Click on **Watson Studio Labs** to return to the catalog display.



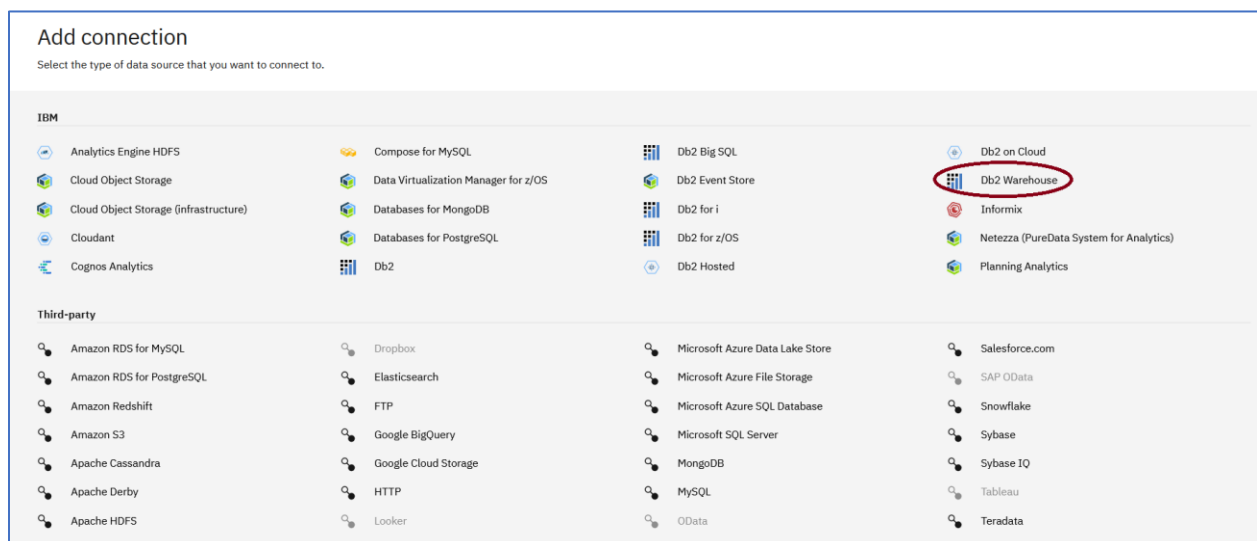
12. Follow steps 8-11 to change the Categories.csv asset name to be Categories.

Add Connection

1. Click on **Add to Catalog**, and then click on **Connection**.



2. The Connection page is displayed with the various connectivity options. Click on **DB2 Warehouse**.



3. Enter “trafficking” for the **Name** of the connection. Leave the **API key** blank. Enter BLUDB for the **Database**. Enter bludmin for the **User name**. Click [here](#) to obtain the DB2 Warehouse Credentials that will be needed for the **Host name** and the **Password**.

Pick a random database connection from the list. Cut and paste the values for hostname and password into the corresponding entry areas on the screen. Click on **Test**.

Enter the information about the data source that you want to connect to.


Connection overview
Name
trafficking
Description
Connection description

Connection details
API Key
Database
BLUDB
Hostname or IP Address
db2w-ojvmspr.us-east.db2w.cloud.ibm.com
Port (optional)
50001
Username
bluadmin
Password

Certificates
Port is SSL-enabled (optional)
☒ The port is configured to accept SSL connections
Secure Gateway
☐ Use a secure gateway
SSL Certificate (optional)
Copy text from a certificate file and paste it in the preceding box.

Cancel Test Create

4. You should receive the following message if the connection is successful. Otherwise, check the credentials to make sure they are entered corrected.

 The test was successful.
Click Create to save the connection information.

5. Click **Create** to create the connection.

Connection overview
Name
trafficking
Description
Connection description

Connection details
API Key
Database
BLUDB
Hostname or IP Address
db2w-ojvmspr.us-east.db2w.cloud.ibm.com
Port (optional)
50001
Username
bluadmin
Password

Certificates
Port is SSL-enabled (optional)
☒ The port is configured to accept SSL connections
Secure Gateway
☐ Use a secure gateway
SSL Certificate (optional)
Copy text from a certificate file and paste it in the preceding box.

Cancel Test Create

6. The connection asset is displayed.

Catalogs / Watson Studio Labs

Add to Catalog +

Any type Any tag Clear all

Watson recommends Highly rated Recently added Collapse

Data asset

Categories

Owner: George Doe
Added: Apr 26, 2020 9:56 AM
Tags: job types categories
0 reviews

Showing 3 of 3 items

Name	Owner	Tags	Business terms	Type	Date added
Categories	George Doe	job types job categories		Data asset	Apr 26, 2020
Occupation	George Doe	occupations jobs		Data asset	Apr 26, 2020
trafficking	George Doe			Connection	Apr 26, 2020

Add Connected Data

1. We have added a connection to the DB2 Warehouse service. We will now add a Connected Data Asset to point to the FEMALE_HUMAN_TRAFFICKING table in the warehouse. Click on **Add to Catalog** and click on **Connected asset**.

Catalogs / Watson Studio Labs

Add to Catalog +

Any type Any tag Clear all

Watson recommends Highly rated Recently added

Local files
Connected asset
Connection

2. Click on **Select Source**.

Add asset from connection

Source*

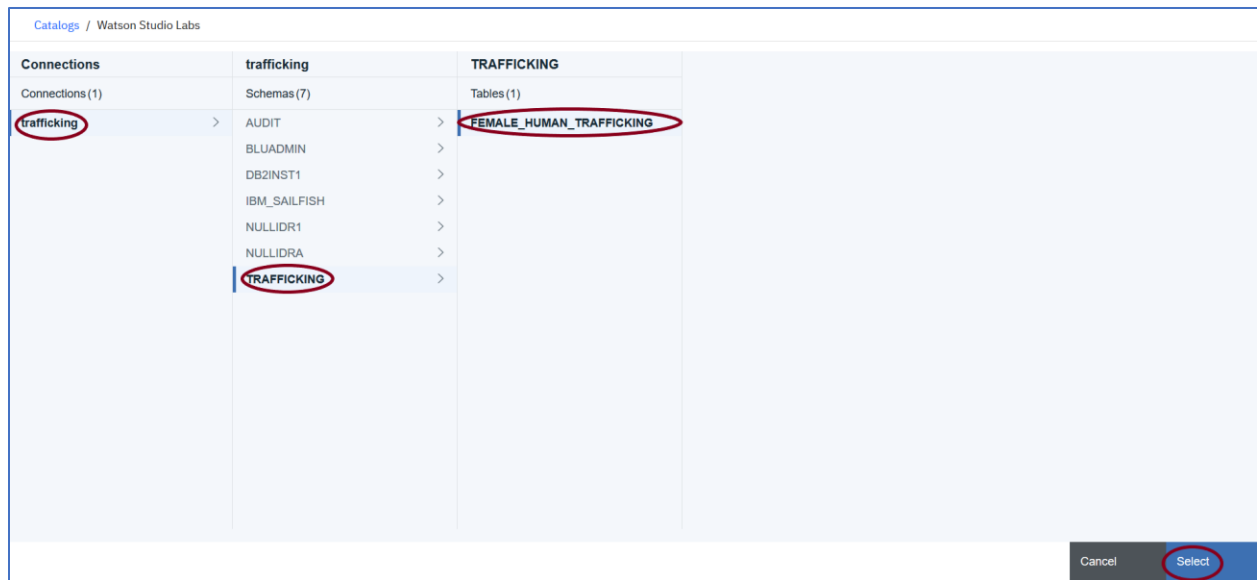
Select source

Name*

Asset name

Description

3. Click on the **trafficking** connection, click on the **TRAFFICKING Schema name** click on **FEMALE_HUMAN_TRAFFICKING**, and then click **Select**.



4. Enter female_human_trafficking for **Name**, enter “trafficking” for **Tags** , click the “+” icon, enter “female human trafficking” for **Tags**, click the “+” icon. Leave **Classification** as None. Scroll down as needed to click on Private, for **Privacy**, and click on + icon to the right of **Members**. The privacy setting of “Private” specifies that only catalog members on the access control list for this data asset will have access to the data asset.

The screenshot shows the 'Catalogs / Watson Studio Labs' interface with a form for configuring a data asset. The 'Source' is 'Db2 Warehouse'. The 'Name' field is 'female_human_trafficking'. The 'Description' field is empty. The 'Business terms' field is empty. The 'Tags' field contains 'trafficking' and 'female human traffi'. The 'Classification' is set to 'None'. The 'Privacy' is set to 'Private'. The 'Members' field has a '+' icon circled in red. At the bottom right, there are 'Cancel' and 'Add' buttons.

5. Enter Howard in the **Members** field. The list of matching users is displayed in the member list. Click on **Howard Doe**.

Add members ×

Members 0

Howard

Howard Doe	ws.catalog.user@gmail.com
------------	---------------------------

Cancel Add

6. Click on **Add** to add Howard Doe to the access control list for the female_human_trafficking asset.

Add members ×

Members 1

Howard Doe ×

Cancel **Add**

7. Click on **Add**.

Name*

female_human_trafficking

Description

Business terms

Search business terms

Tags

trafficking X female human traffi... X

Start typing to add values

Classification* ⓘ

None

Cancel Add

8. All of the assets have now been added to the catalog. The catalog display should appear as below.

Catalogs / Watson Studio Labs

Any type Any tag Clear all

Add to Catalog +

Watson recommends Highly rated Recently added Collapse

Data asset

Categories

Owner: George Doe
Added: Apr 26, 2020 9:56 AM

Tags: job types job categories

☆☆☆☆☆ 0 reviews

Connection

trafficking

Owner: George Doe
Added: Apr 26, 2020 11:48 AM

☆☆☆☆☆ 0 reviews

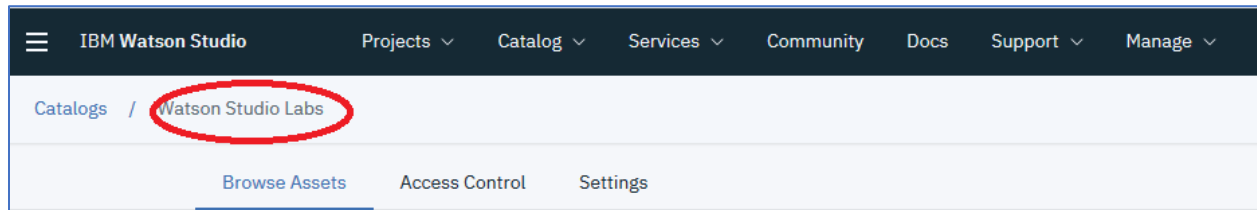
Showing 4 of 4 items

<input type="checkbox"/>	Name	Owner	Tags	Business terms	Type	Date added
<input type="checkbox"/>	Categories	George Doe	job types job categories		Data asset	Apr 26, 2020
<input type="checkbox"/>	Occupation	George Doe	occupations jobs		Data asset	Apr 26, 2020
<input type="checkbox"/>	female_human_trafficking	George Doe	trafficking female human trafficking		Data asset	Apr 26, 2020
<input type="checkbox"/>	trafficking	George Doe			Connection	Apr 26, 2020

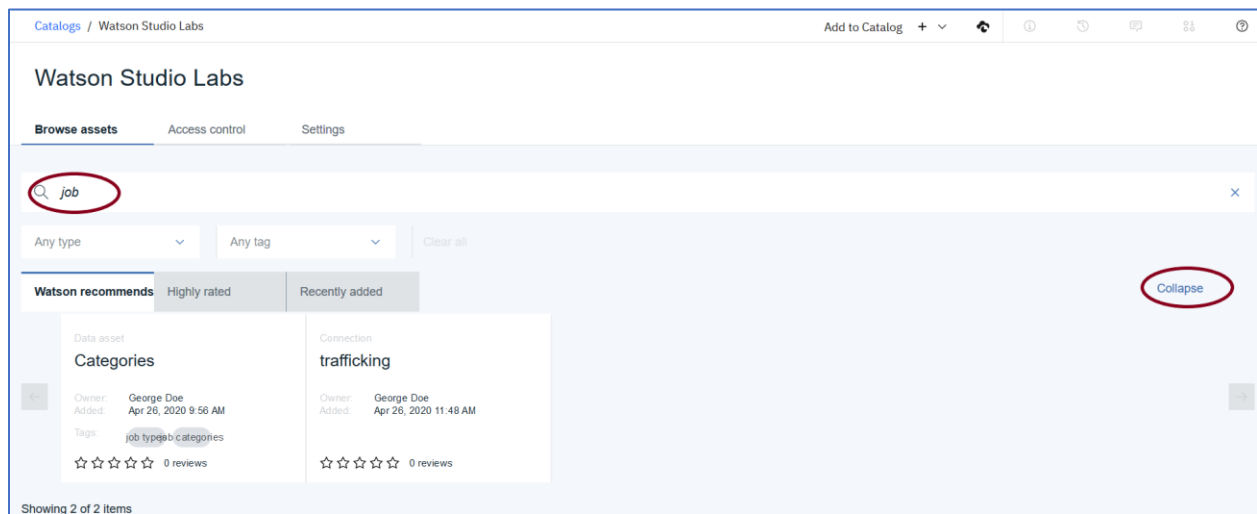
Step 4: Search the Catalog

Several options exist for searching the assets in the catalog. We can enter text in the search input text field and the system will match against asset names and tags. We can then further filter the search results by selecting the Asset Type, and/or the Tags. Or we can directly select the Asset Type or Tags without entering anything in the search text field.

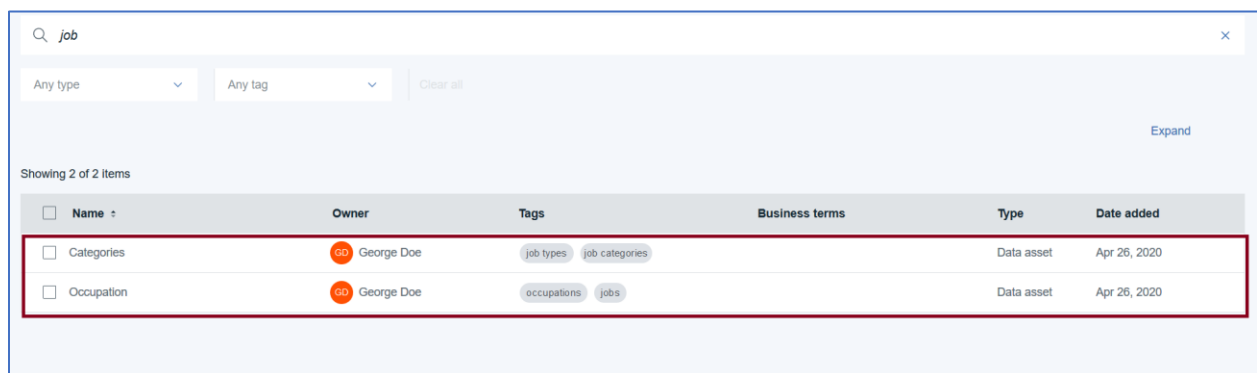
1. Click on **Watson Studio Labs** to return to the Catalog home screen.



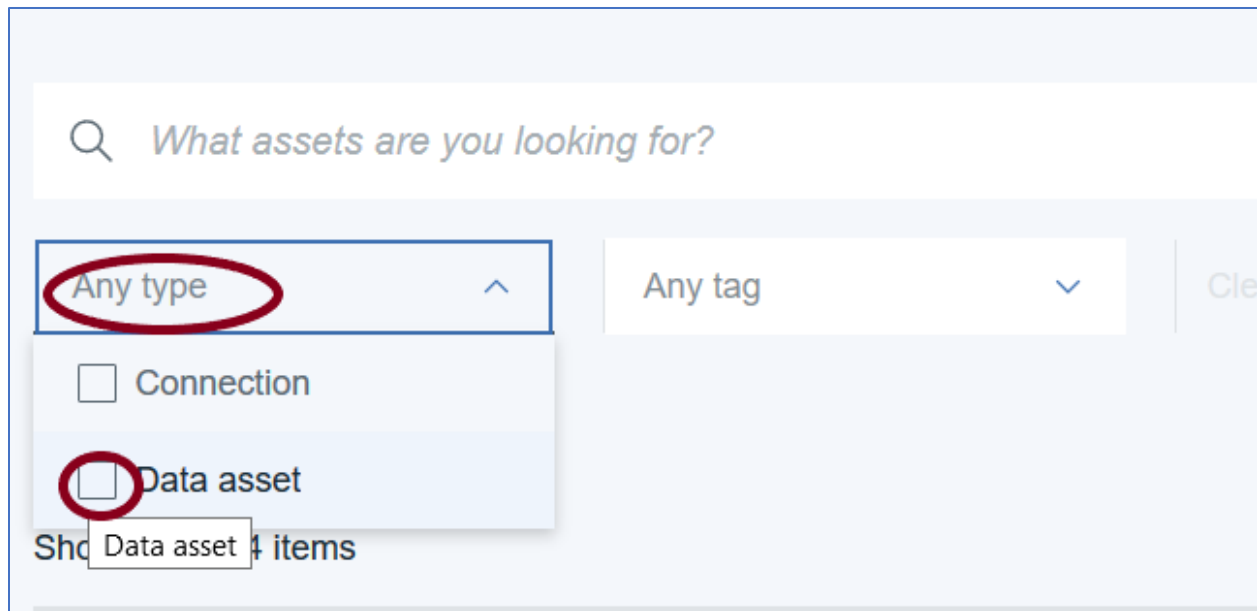
2. Enter **job** in the **search text field** and click on **Collapse**.



3. Data assets are displayed that contain the word or letters **job** in their Name, Description, or tags.



4. Clear the **search text field**. Click on **Any Type**. Click on the check box next to **Data Asset**.

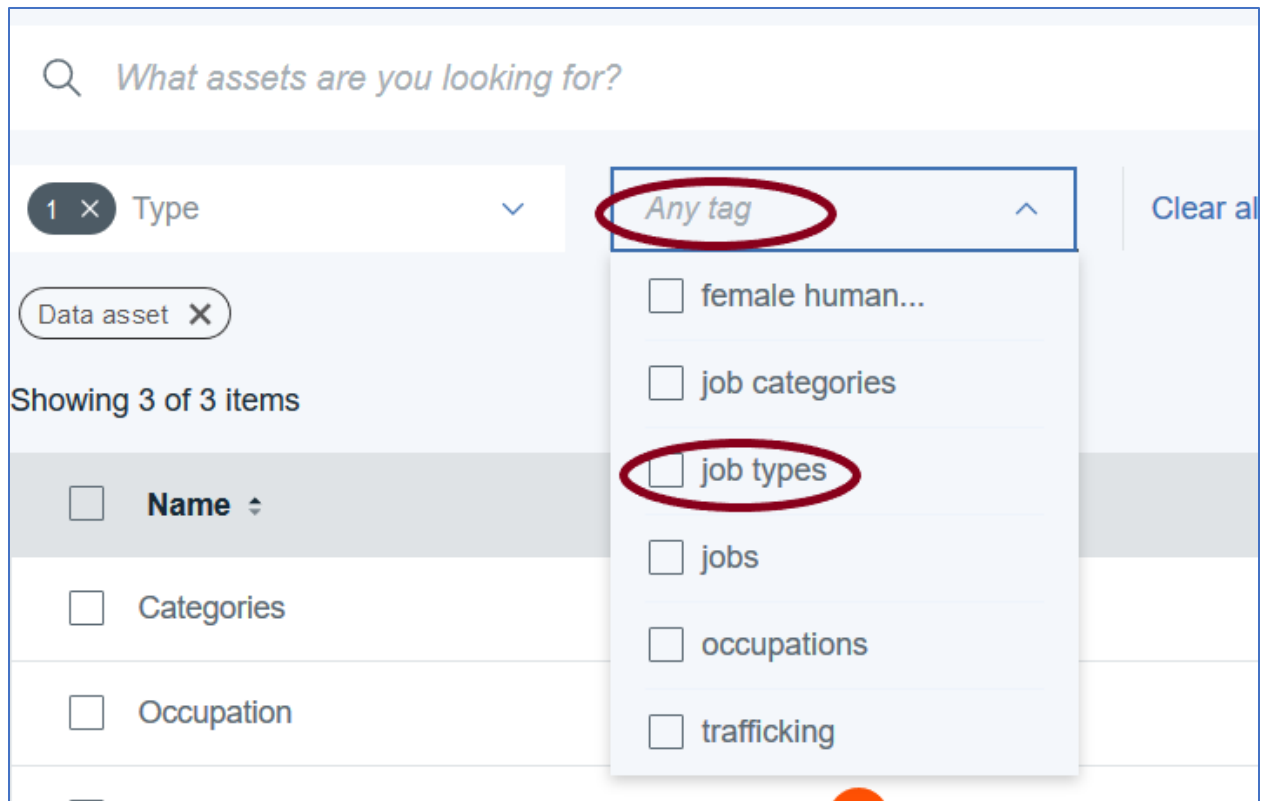


- The search results list the 3 Data Assets. The 1 Connected asset, and the 2 uploaded files.
Note: Other asset types that can be cataloged include, Connections, Models, Notebooks, and Dashboards.

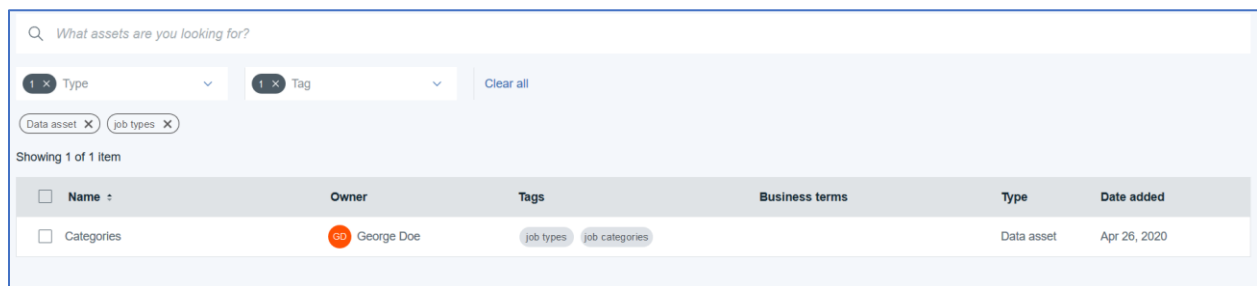
The screenshot shows the search results table. The table has 7 columns: Name, Owner, Tags, Business terms, Type, and Date added. There are 3 rows of data. The first row is for "Categories", the second for "Occupation", and the third for "female_human_trafficking". Each row has a checkbox in the "Name" column. The "Owner" column shows "George Doe" for all three rows. The "Tags" column shows "job types" and "job categories" for "Categories", "occupations" and "jobs" for "Occupation", and "trafficking" and "female human trafficking" for "female_human_trafficking". The "Type" column shows "Data asset" for all three rows. The "Date added" column shows "Apr 26, 2020" for all three rows.

<input type="checkbox"/> Name	Owner	Tags	Business terms	Type	Date added
<input type="checkbox"/> Categories	George Doe	job types job categories		Data asset	Apr 26, 2020
<input type="checkbox"/> Occupation	George Doe	occupations jobs		Data asset	Apr 26, 2020
<input type="checkbox"/> female_human_trafficking	George Doe	trafficking female human trafficking		Data asset	Apr 26, 2020

- Click on **Any Tag**. Click on the check box next to **job types**.



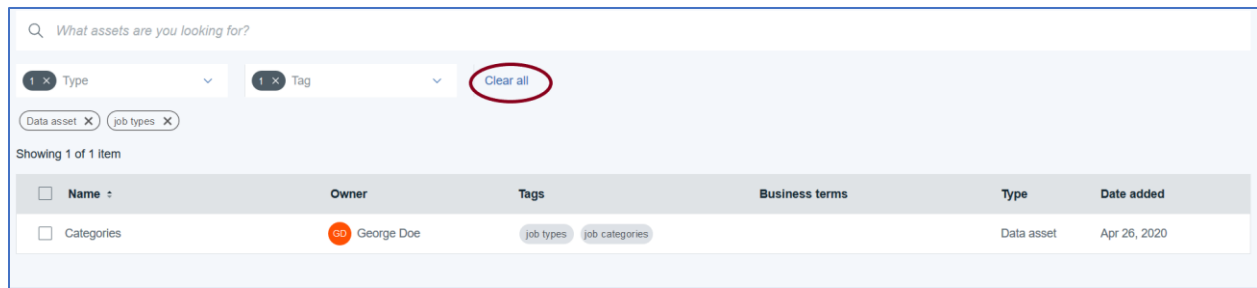
7. The Categories Data Asset is listed.



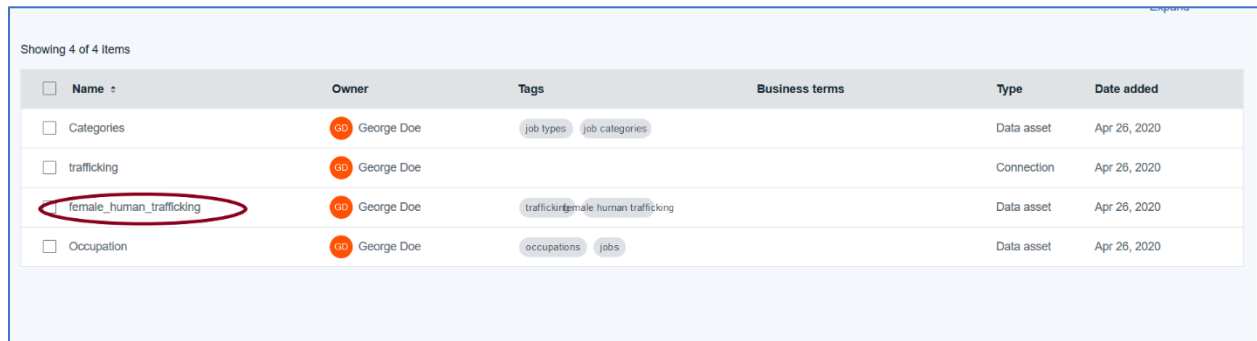
Step 5: Edit/Review/Profile Data Asset

In this step, we will drill down on the female_human_trafficking data asset to show how to edit the metadata describing the asset, to provide a rating and review of the asset, and to profile the asset.

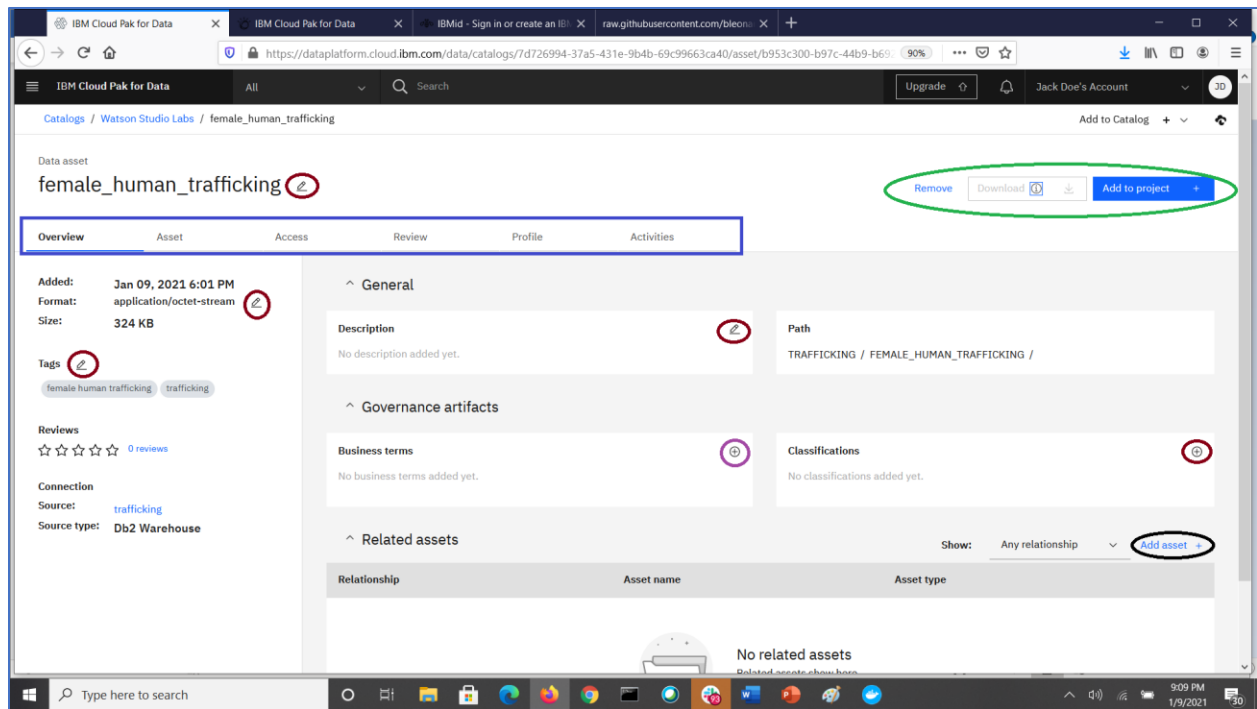
1. Click on **Clear all**.





2. Click on the female_human_trafficking asset.

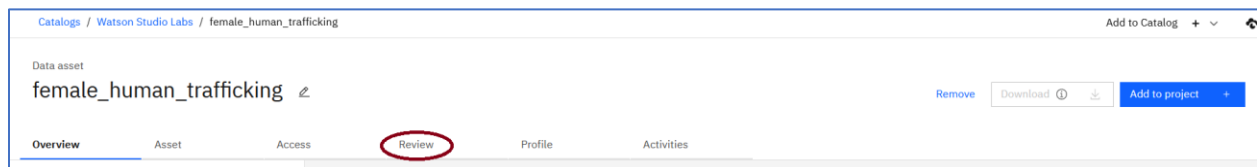


3. The metadata and data contents of the asset are displayed as shown below. Toward the top of the screen, are 6 tabs, highlighted in blue below. The panel shown below the 6 tabs corresponds to the **Overview** tab. The **Asset** tab displays the contents. The **Access** tab allows the user to edit the access control list for this asset. The **Review** tab enables the user to provide a rating and descriptive review of the asset. These social features aid in searching for relevant and useful assets. The **Profile** tab invokes an auto-profiling capability where the columns of the data asset are categorized into one of many classifiers that are built into the platform. This capability will be shown below. Finally, the **Activities** tab tracks the changes made to the data asset over its lifecycle. This capability is not available in the lite catalog offering so will not be exercised in this lab.

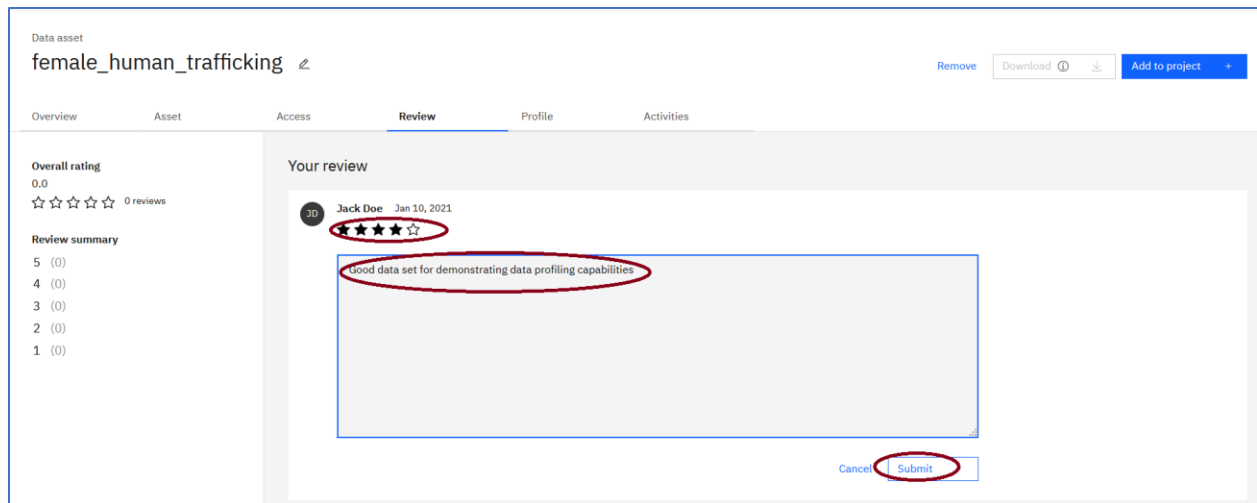


The actions circled in green will **Remove** the asset from the catalog, **Download** the asset to the local file system (grayed out), or add the asset to a catalog enabled project (**Add to Project**). The **Name** of the data asset, the **Format**, the **Description**, or the **Tags** can be edited by clicking on the  icon. Clicking on the  icon next to the **Business Terms** or the **Classification** will bring up a list from which you can select entries.

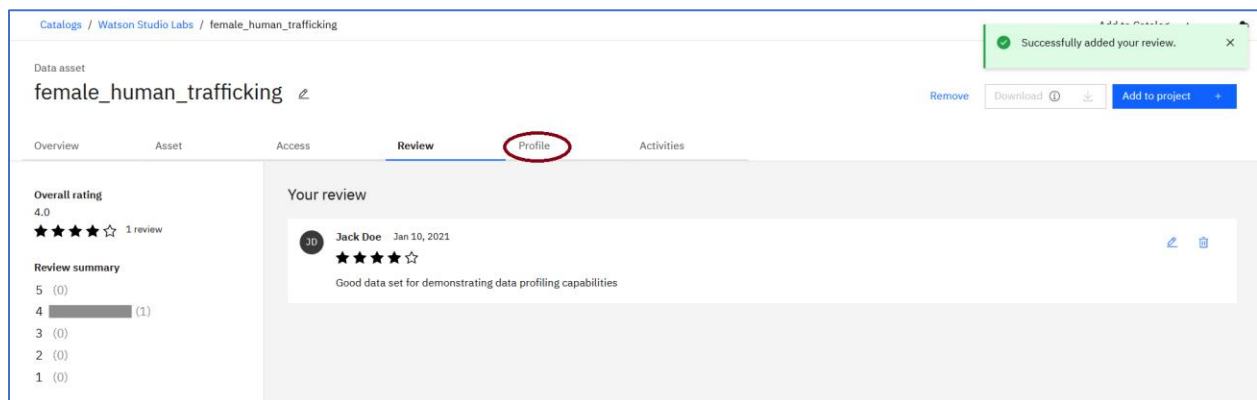
4. Click on the **Review** tab.



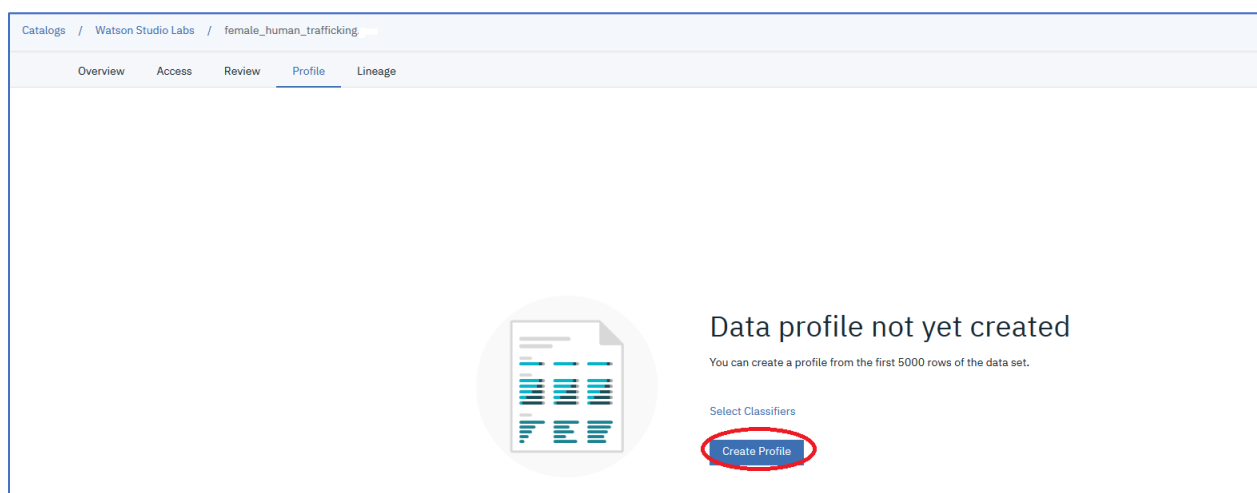
5. Click on the fourth star to rate this asset. Enter “Good data set for demonstrating data profiling capabilities” in the **Your Review** text field. Click on the **Submit** button.



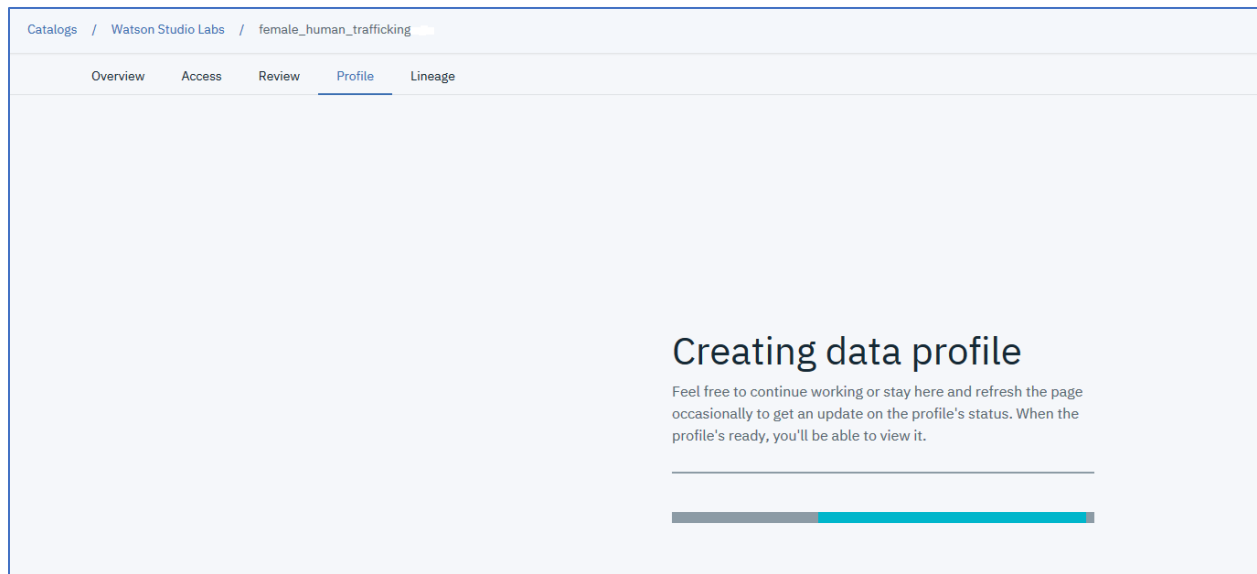
6. The review of the asset is displayed. Click on the **Profile** tab.



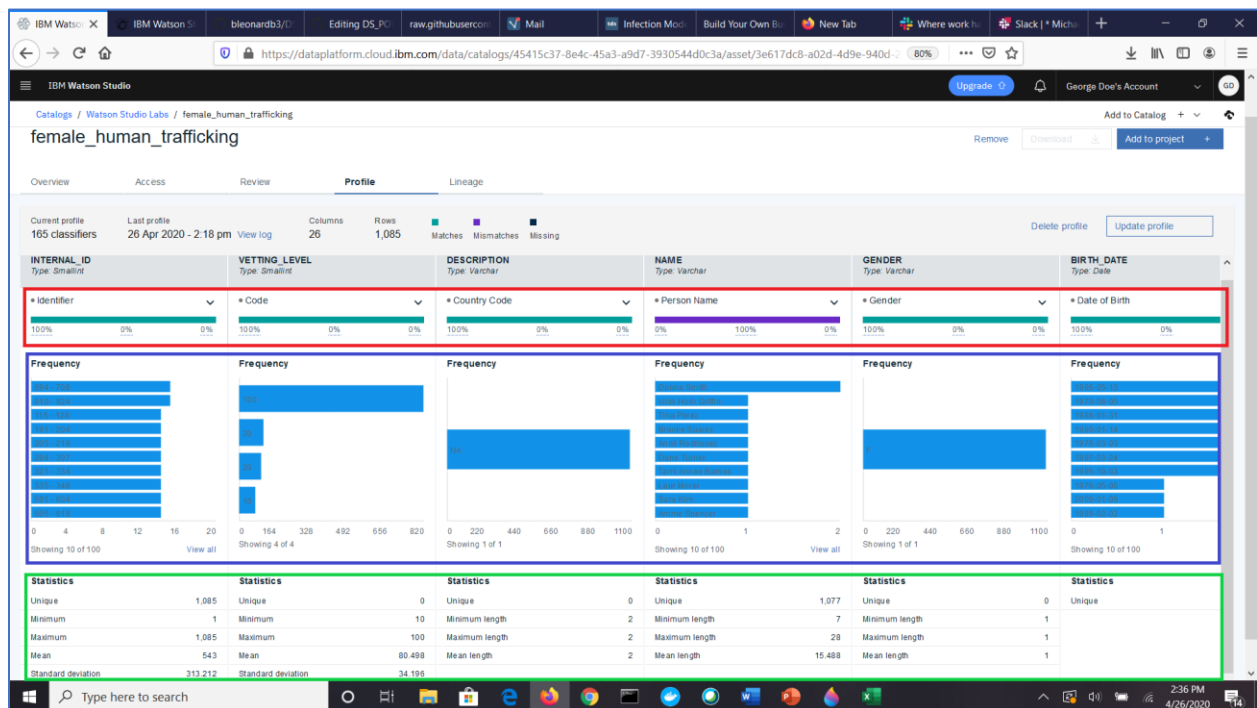
7. Click on **Create Profile**. If a profile was already created, it would be displayed instead of this prompt screen. In that case, skip to step 9.



8. This step could take some time. Please wait a minute or so, and then click on browser refresh.

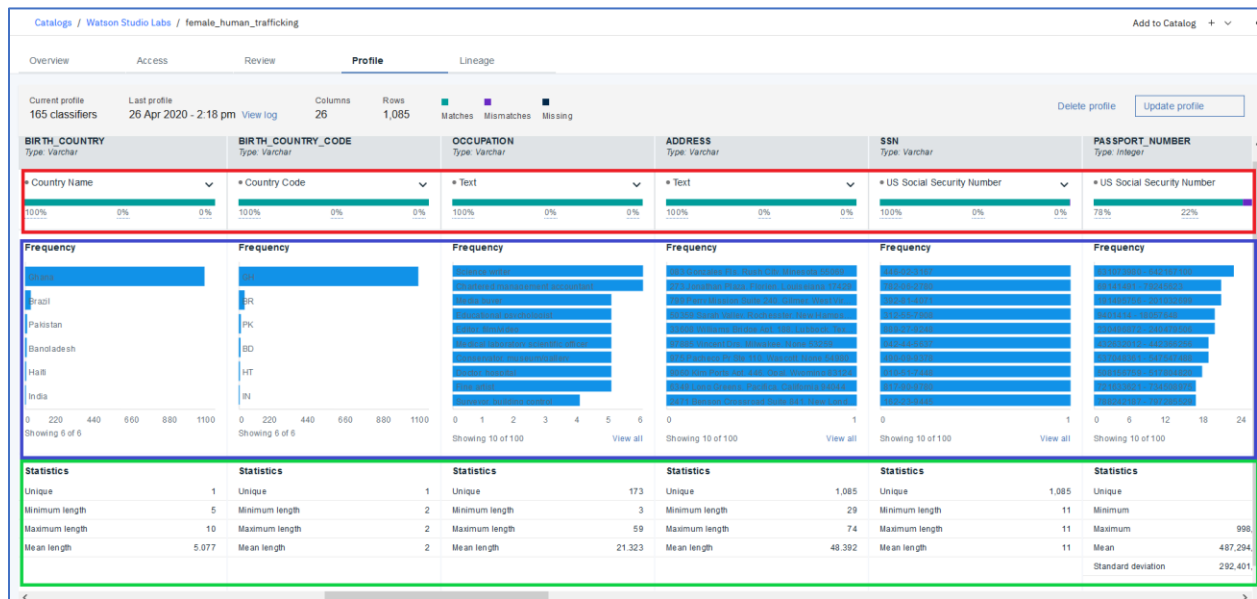


9. The profile information is displayed below using several screenshots, based on horizontal scrolling, to display many of the columns of the data asset. The highlighted red box contains the inferred column classifications. The highlighted blue box contains counts of the number of occurrences of each value (for discrete fields), and a histogram for continuous fields. The green box contains statistics calculated from the contents of the columns.

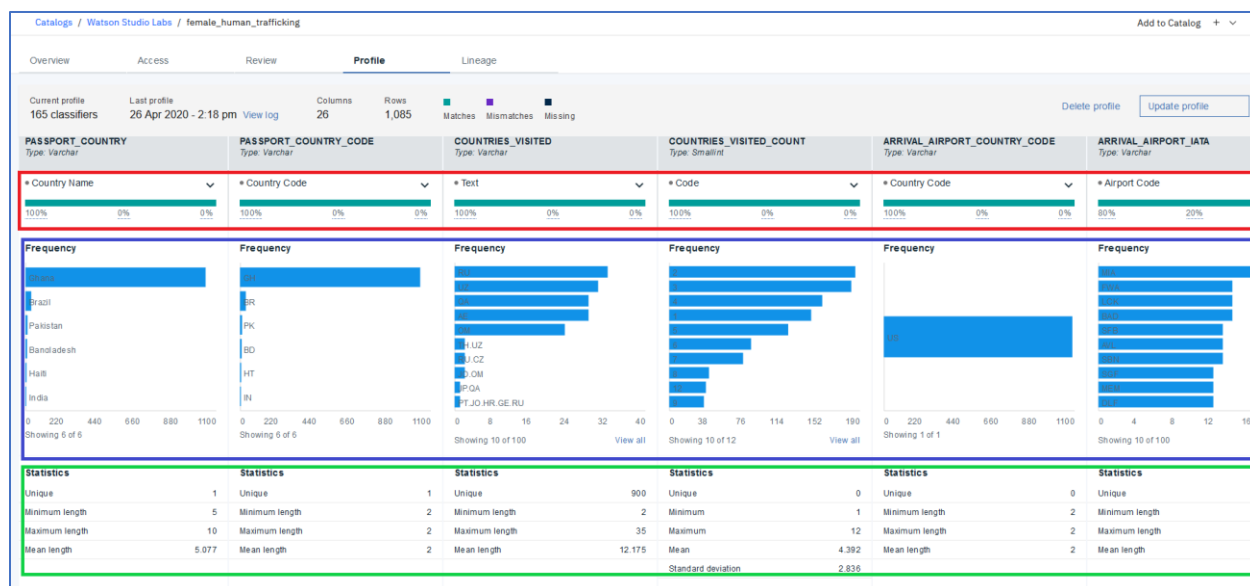


The classifications of the columns are listed below:

- INTERNAL_ID – Identifier
- VETTING_LEVEL – Code
- DESCRIPTION – Country Code – this is an error due to the description input as NA. We can correct this misclassification.
- NAME: Person Name – the classification accurately determines that this field is the name of a person.
- GENDER: Gender – the classification accurately determines that this field is the gender of a person
- BIRTH_DATE: Date of Birth – the classification accurately determines that this field represents a date of birth.




- BIRTH_COUNTRY: Country Name- correctly classified as a country
- BIRTH_COUNTRY_CODE – Country Code - correctly classified
- OCCUPATION – Text – there is not a classifier for occupation, so it determines that this is a text string.
- ADDRESS: Text
- SSN – US Social Security Number – correctly classifies the column as a social security number.
- PASSPORT_NUMBER – incorrectly classifies as a social security number. This can be corrected as we will see below.



- PASSPORT_COUNTRY - Country Name- correctly classified as a country
- PASSPORT_COUNTRY_CODE: Country Code - correctly classified
- COUNTRIES_VISITED – Text
- COUNTRY_VISITED_COUNT – Code – this is an error. It’s not a code. This can be re-classified if necessary.
- ARRIVAL_AIRPORT_COUNTRY_CODE – Country Code – correctly classified.
- ARRIVAL_AIRPORT_IATA: Airport Code – the classification accurately determines that this field represents an Airport Code.

There are more columns in the data set but this gives a representative sample of the profiling capability. The classification of the columns becomes important when we create and enforce security policies.

10. Let’s fix the classification for passport. Click  icon next to the **US Social Security Number** underneath the **PASSPORT_NUMER** column. Click on **View All**.

PASSPORT_NUMBER
Type: Integer

- US Social Security Number

^

✓ US Social Security Num... 78% matches

Not classified 21% matches

Canadian Social Insura... 7% matches

Routing Transit Number 3% matches

View all

9401414 - 18057648

11. Type “Pass” in the **Search** field. Click on **Passport Number** radio button. Click on **Add**.

Edit data class

Q Pass

x

All data classes

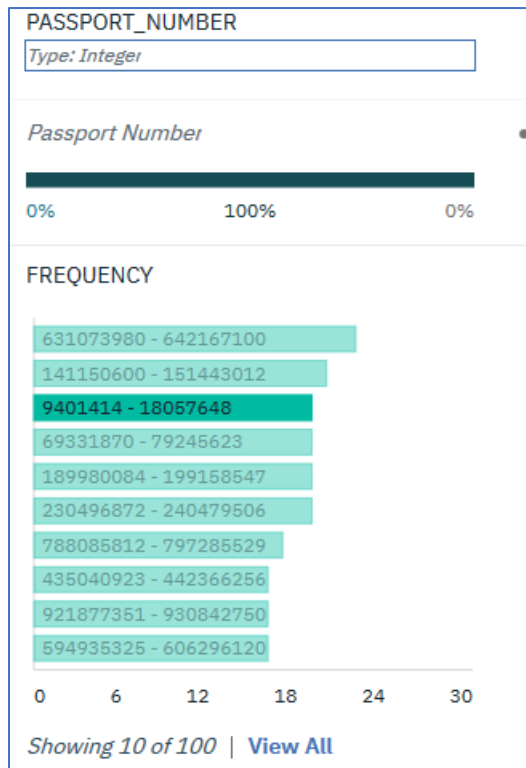
☐ • US Social Security Number 78% matches

☒ Passport Number -

Cancel

Add

12. The column classification has been changed.



Step 6: Access Control

All Catalog members can access assets with a privacy setting of Public. Assets that have a privacy setting of Private are secured based on the asset access control list. A summary of the access control settings for our Watson Studio Labs catalog and the contained Data Assets is shown below.

Watson Studio Lab Catalog members:

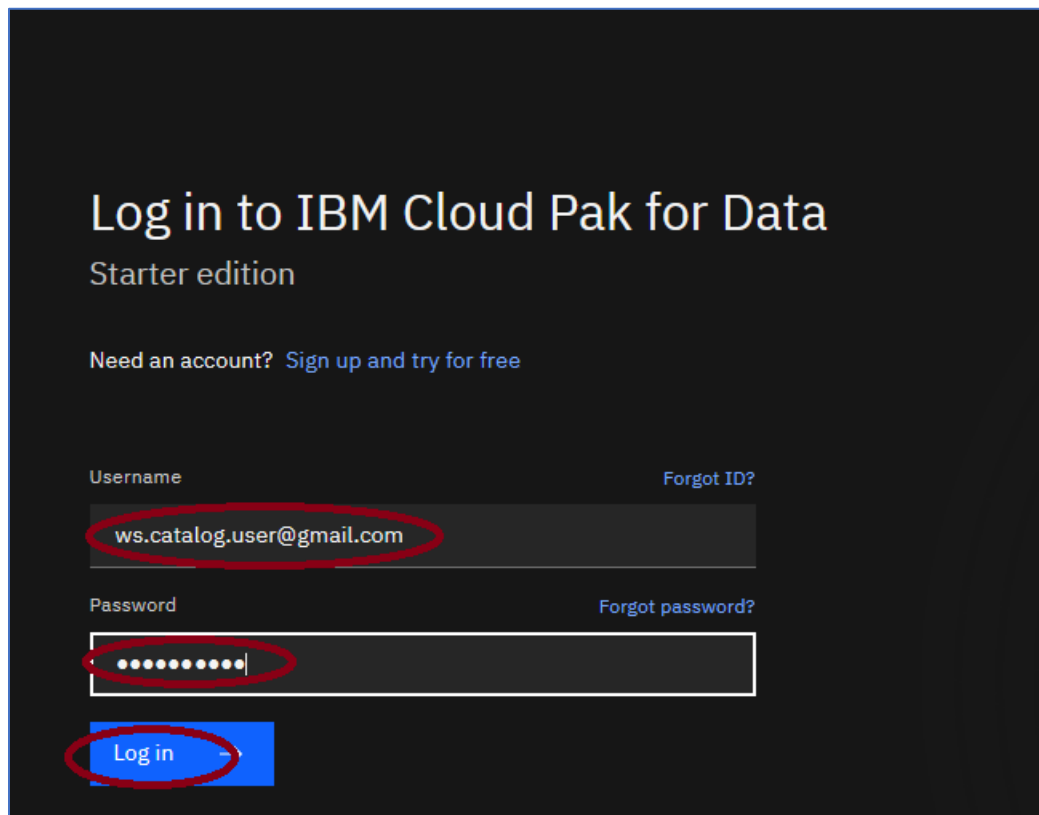
1. Catalog owner (logged on user)
2. ws.catalog.user@gmail.com

Data Asset Privacy:

1. female_human_trafficking (Private) ACL Members:
 - a. asset owner (logged on user)
 - b. ws.catalog.user@gmail.com
2. Occupations (Private) ACL Members:
 - a. asset owner (logged on user)
3. Categories (Public)
4. trafficking (Connection) (Public)

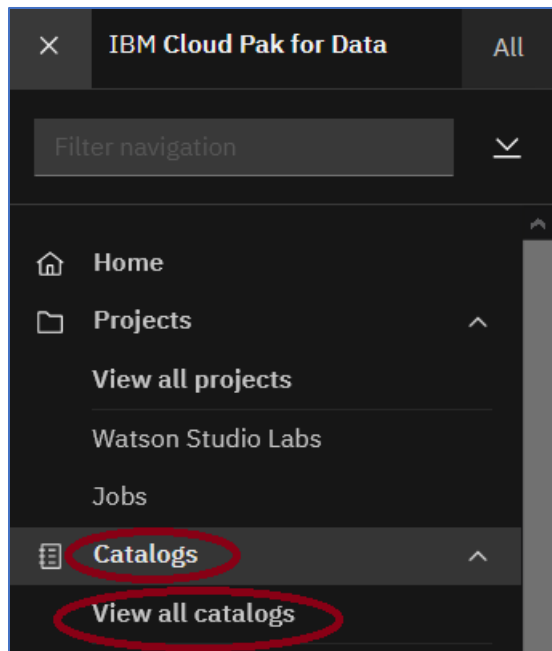
We would expect that if we log in as `ws.catalog.user@gmail.com`, that only 3 Data Assets would appear. The two public Data Assets, and the Private Data Asset where `ws.catalog.user@gmail` is on the access control list. This user would not have access to the Occupations asset, since it is a Private asset, and the user is NOT on the access control list for this asset. To verify complete the following steps.

1. Open a different browser (if you are using FireFox, then open Chrome or vice versa).
2. Type in **`datapatform.cloud.ibm.com`** for the URL to navigate to the Cloud Pak for Data sign in.
3. Login using ws.catalog.user@gmail.com for the **Username** and **wsPassw0rd** for the **Password**, and click **Log in**.

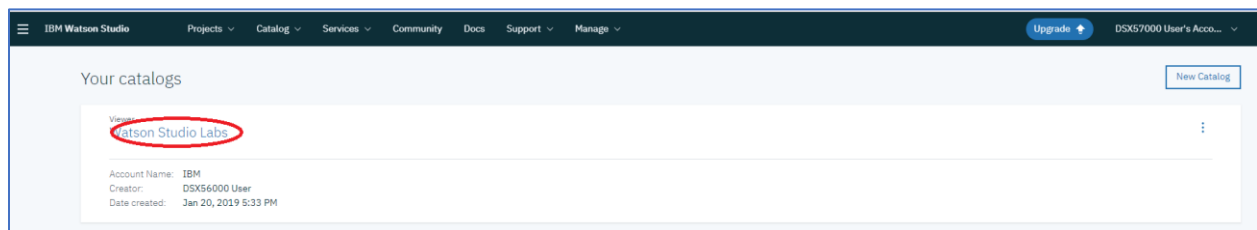


4. Click on  icon. Click on **Catalog** and then **View All Catalogs**

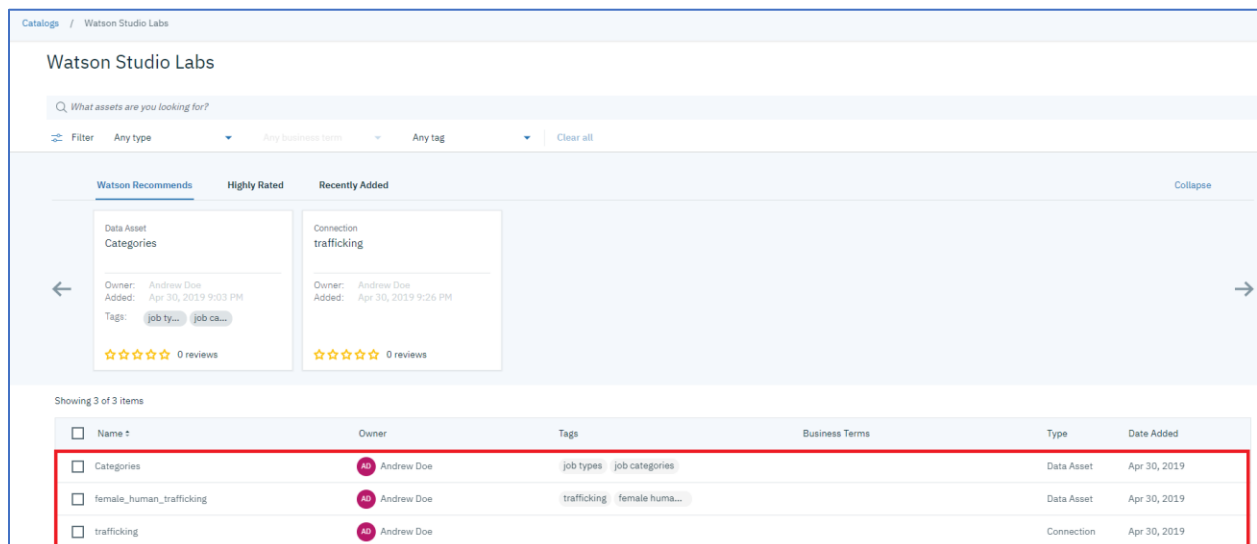




5. Scroll down to find the **Watson Studio Labs** catalog that corresponds to your name. You can also type <Ctrl><F> and enter your name to search. Note that there will be 1 catalog per lab participant. Click on the **Watson Studio Labs** catalog corresponding to your user name.




6. The Watson Studio Labs catalog is displayed. Three Data Assets are listed as expected.

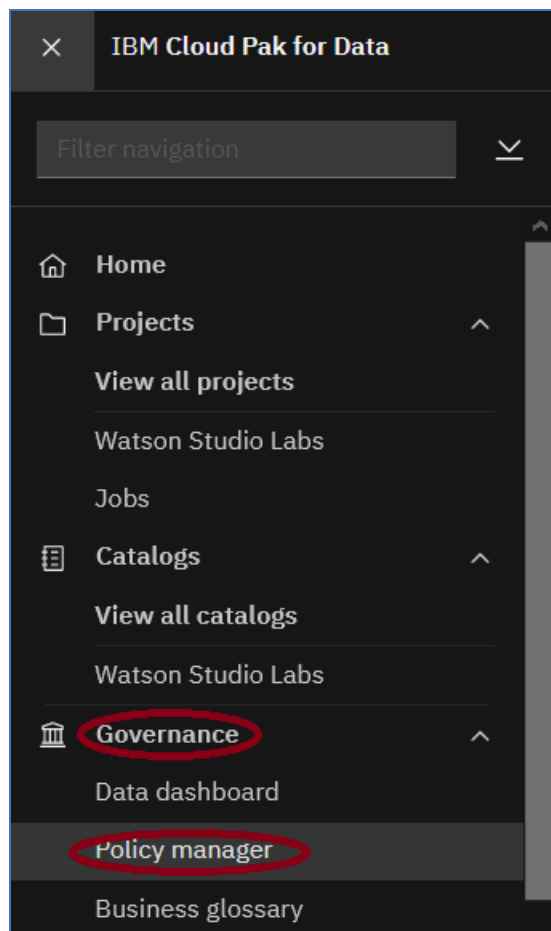
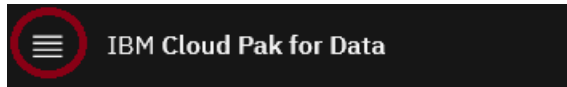


7. Leave this browser at this screen for Step 7 below.

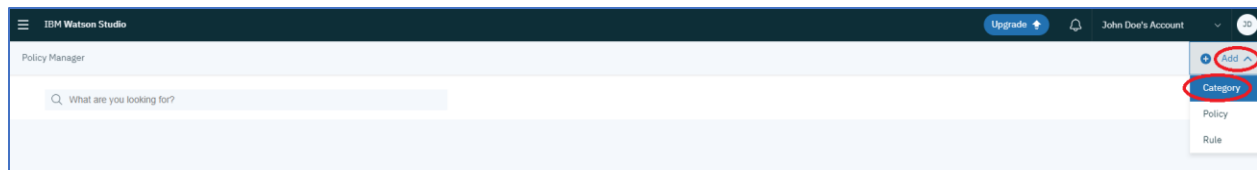
Step 7: Create and Enforce Policy (optional)

In this step, we will create a policy that will anonymize the SSN and Passport fields when user ws.catalog.user@gmail.com tries to access the female_human_trafficking data asset.

1. **RETURN TO THE BROWSER WHERE YOUR** Cloud Pak for Data id is logged in.
2. Click on the  icon. Click on **Governance** and then **Policy Manager**.

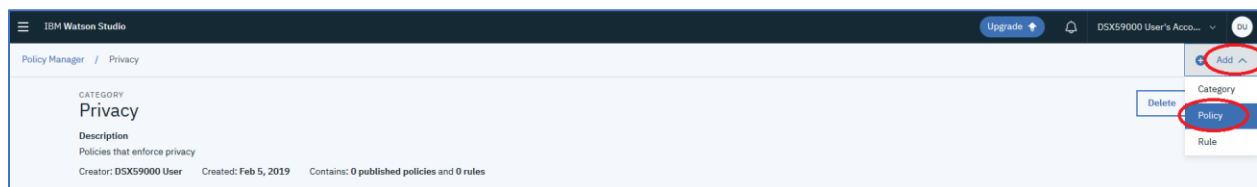


3. Click on **Add** and then **Category**.



4. Type in **Privacy** for the **Name** of the Category, and optionally a **Description**, and click on **Create**.

5. Click on **Add** and then **Policy** to add a policy to the Privacy Category.



6. Type in PII for the Policy **Name**, type in “Limit access to PII data” for the Policy **Description**, click on **create a new rule**.

New policy

Details

Name*

PII

Type*

Access

Category*

Privacy

Description

Limit access to PII data

Rules 0

No rules in policy!

To add rules to this policy,

Create a new rule

Select from existing rules

Cancel

Save as draft

Publish

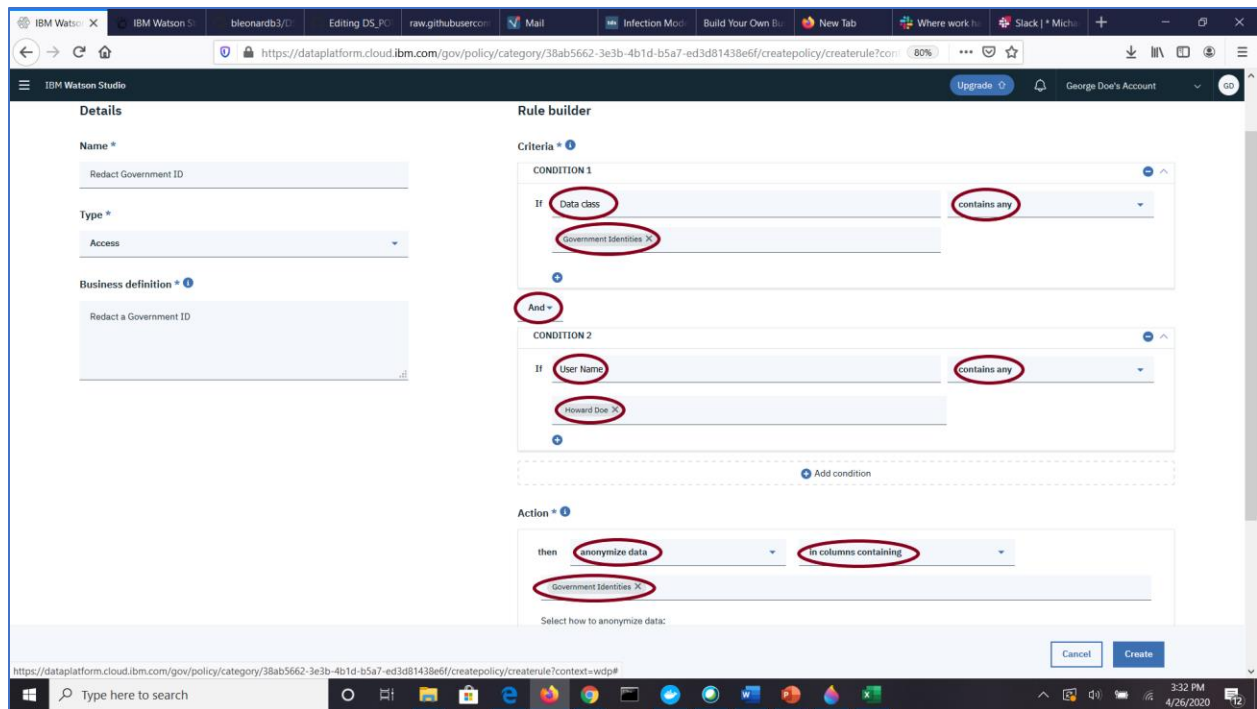
- Enter “Redact Government ID” for the rule **Name**, enter “Redact a Government ID” for the **Business Definition**, for the rule click in the entry fields to select the following:

If **Data Class** contains any
Government Identities

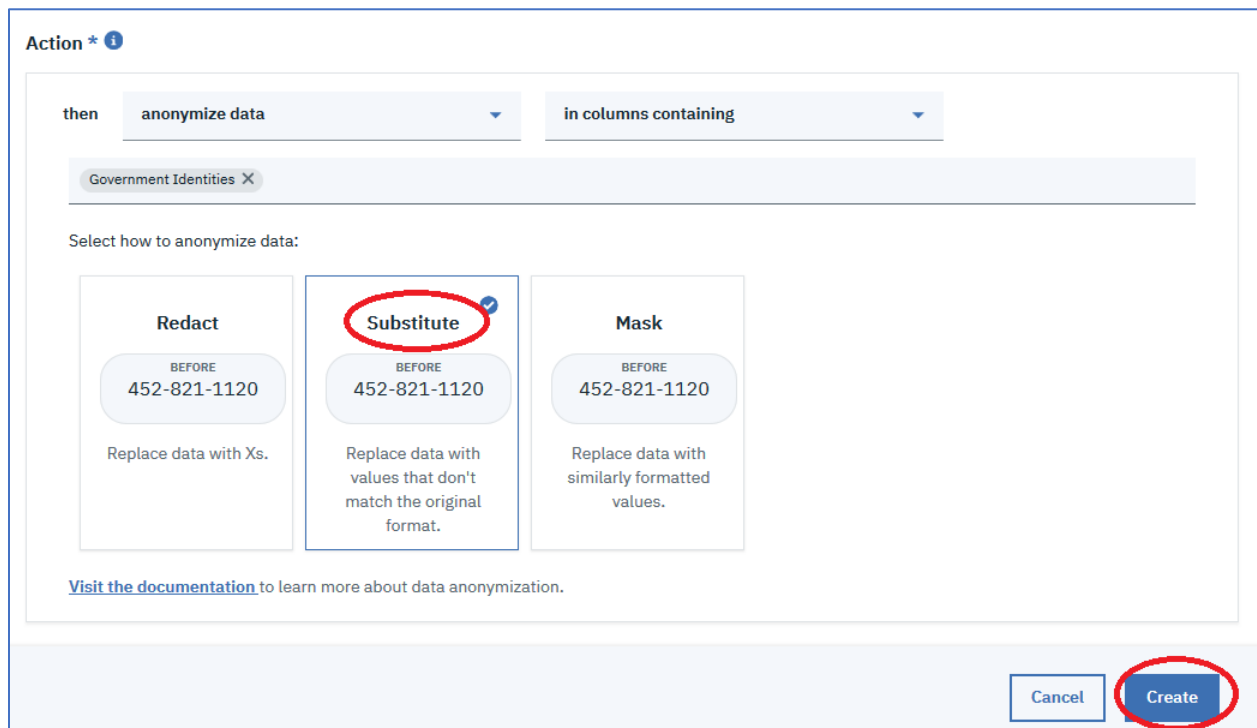
And

If **User Name** contains any
Howard Doe (make sure you select Howard Doe from the list!)

then **anonymize data** where **in columns containing**
Government Identities



8. Scroll down and select how to anonymize data: **Substitute**, and then click **Create**.



9. Click on **Publish**.

New policy

Details

Name*

PII

Type*

Access

Category*

Privacy

Description

Limit access to PII Data

Rules 1

Redact Government ID

Redact a Government ID

Cancel Save as draft **Publish**

10. Click on **Publish**.

Attention!

Are you sure you want to publish this policy?

Cancel **Publish**

11. The PII policy has been added.

Policy Manager / Privacy

CATEGORY

Privacy

Description

Policies that enforce privacy

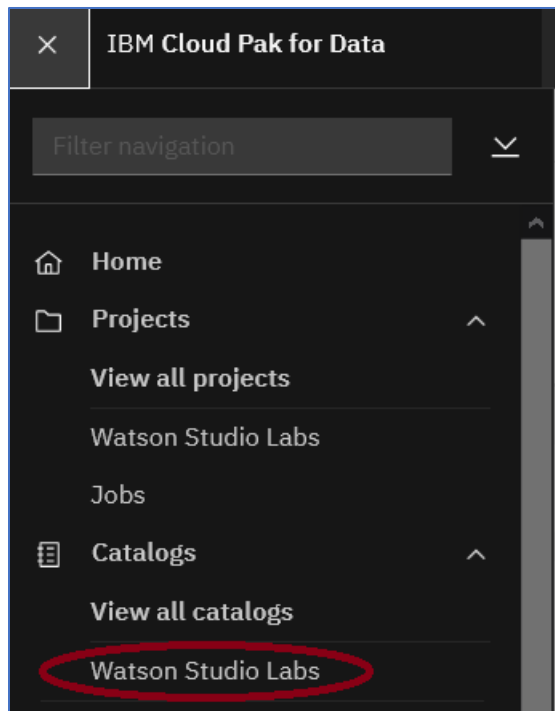
Creator: DSX59000 User Created: Feb 5, 2019 Contains: 1 published policies and 1 rules

Q What are you looking for?

NAME	TYPE	STATUS +	CONTAINS	LAST MODIFIED
> PII	Policy	Published	1 Item	Feb 5, 2019

12. Click on the  icon, then click on **Catalog** and then **Watson Studio Labs**.





13. Switch back to the browser with the ws.catalog.user@gmail.com user logged in. Click on the **female_human_trafficking** asset.

Available Assets

Showing 3 of 3 assets

<input type="checkbox"/> NAME	OWNER	TAGS	TYPE	DATE ADDED
Categories.csv	DSX59000 User		Data Asset	Jan 31, 2019
female_human_trafficking	DSX59000 User	trafficking female human trafficking	Data Asset	Jan 31, 2019
trafficking	DSX59000 User		Connection	Jan 31, 2019

14. Click on **Asset**.

Data asset

female_human_trafficking

Remove Download Add to project

Overview **Asset** Access Review Profile Activities

Added: Jan 09, 2021 6:01 PM
Format: application/octet-stream
Size: 324 KB

Tags: female human trafficking trafficking

Reviews: ★★★★★ 1 review

Connection: Source: trafficking Source type: Db2 Warehouse

General

Description: No description added yet.

Path: No path

Governance artifacts

Business terms: No business terms added yet.

Classifications: No classifications added yet.

Related assets

Show: Any relationship Add asset

Relationship	Asset name	Asset type
--------------	------------	------------

15. The female_human_trafficking asset is displayed. Wait for the Data Masking to complete.

Catalogs / Watson Studio Labs / female_human_trafficking

DATA ASSET

female_human_trafficking

Remove Download Add to project

Overview Access Review Profile Lineage

Description
There is no description available for this asset.

Added: Apr 26, 2020 2:18 PM
Format: application/octet-stream
Size: 321 KB

Business terms
There are no terms available for this asset.

Tags
female human trafficking trafficking

Reviews
★★★★☆ 1 review

Classification

Schema: 26 Columns | 1085 rows | 2 Columns masked

Last refresh: just now

INTERNA... Smallint	VETTING_L... Smallint	DESCRIP... String	NAME String	GENDER String	BIRTH_D... Date	BIRTH_COU... String	BIRTH_COUNTRY_C... String	OCCU String
Identif...	Code	Countr...	Perso...	Gender	Date o...	Country N...	Country Code	Text

Data masking in progress
This asset is being masked by the data enforcement rule: PII. You can wait here to see a preview of the asset or we can notify you when the preview is ready.

Notify Me

16. Note that 2 columns are anonymized.

Mail Red Hat Connect for Business Pa... Red Hat OPEN - BEEKMANB (Ber... IBM Watson Studio

datapatform.cloud.ibm.com/data/catalogs/45415c37-8e4c-45a3-a9d7-3930544d0c3a/asset/3e617dc8-a02d-4d9e-940d-284735b6753f?context=wdp

IBM Watson Studio Upgrade Howard Doe's Account

Catalogs / Watson Studio Labs / female_human_trafficking

Added: Apr 26, 2020 2:18 PM
Format: application/octet-stream
Size: 321 KB

Business terms
There are no terms available for this asset.

Tags
female human trafficking trafficking

Reviews
★★★★☆ 1 review

Connection
Source: Watson Studio Labs_DataCatalog
Source type: Cloud Object Storage

Classification
None

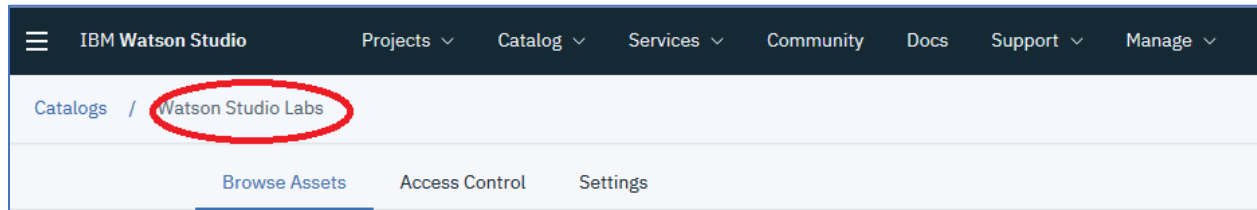
OCCUPAT... String	ADDRESS String	SSN String	PASSPORT_NUMB... String	PASSPORT_COU... String	PASSPORT_COUNTRY_... String	COUNTRIES_VIS... String
Text	Text	US So...	Passport Number	Country Name	Country Code	Text
Art therapist	7089 Jack Trail, I	fb7d4f9691eba8	306d01cd075d507b4e45	Ghana	GH	RU
Engineering geol	02820 Hart Ville,	2fbcd08f718a87	86e62336a88cf0162d0f	Ghana	GH	ME,PE,UZ,UA
Engineer, civil (c	7161 James Via,	ddda0cbe571d3	cdba7a80fca4ed910e968	Ghana	GH	TN,QA,BE,VN,PA
Teacher, adult ec	724 Morrow Lodj	af5fe28b6e8b2e	e9eac57f67f58e3d3f81f6	Ghana	GH	OM
Ecologist	36678 Henderso	5c884e47ceb68	3cc44ebad035a1dc65b4f	Ghana	GH	LB,AE,SN,PH
Geoscientist	376 Bowers Ran	73dd27ae9d2d5	9f5c9998d97f12919c09	Ghana	GH	HK,KH,OM,PG,LT
Bookseller	2448 Autumn Co	d0dbd68afe433	a17651e3085bc2179149	Ghana	GH	HR,RU,UY,CA,EG,GB
Medical laborato	1140 Arnold Trai	b41260f74d0c2	2248c250bf21d96a2e55	Ghana	GH	UZ,NZ,IJ
Biomedical scien	38712 Griffith Or	fe27fd2e3d87c1	11be85abf2b588bcb89c	Ghana	GH	MX,CO,CL,OM,UA,UZ
Dealer	143 Solis Roads,	0ba4ea52625a6	fc2f6c5a7eb557903b2dcl	Ghana	GH	BZ,UZ,IE,LT
Doctor, hospital	30788 Chan Vlg,	87c98549d6f7al	9195ef4244e18d002adfl	Ghana	GH	RU
Architectural tec	3930 Murray Mal	612723dd767d4	e0b0d542c081ea093129	Ghana	GH	MX,IL,RU,IT,BE,CA,EG,UN
Glass blower/det	3206 Singleton T	2eb1e3b53abda	ec27e923ed2e4da79b94	Ghana	GH	BE,IL,PK,OM
Field seismologi	17407 Fitzgerald	86bd9bc45c84	33e80462082cc9031190	Ghana	GH	AE
Plant breeder/ge	42740 Adam Bro	d5ed650eea722	94ece1d9368d01027c00	Ghana	GH	GH,BG,RU

17. Close the browser tab for user ws.catalog.user@gmail.com.

Step 8: Push the data assets to the project

This step will add the cataloged data assets to our Watson Studio project to be used in subsequent labs. Return to your Watson Studio account.

1. You should be at the Catalog home screen. If not, click on **Watson Studio Labs** to go to the Catalog home screen.

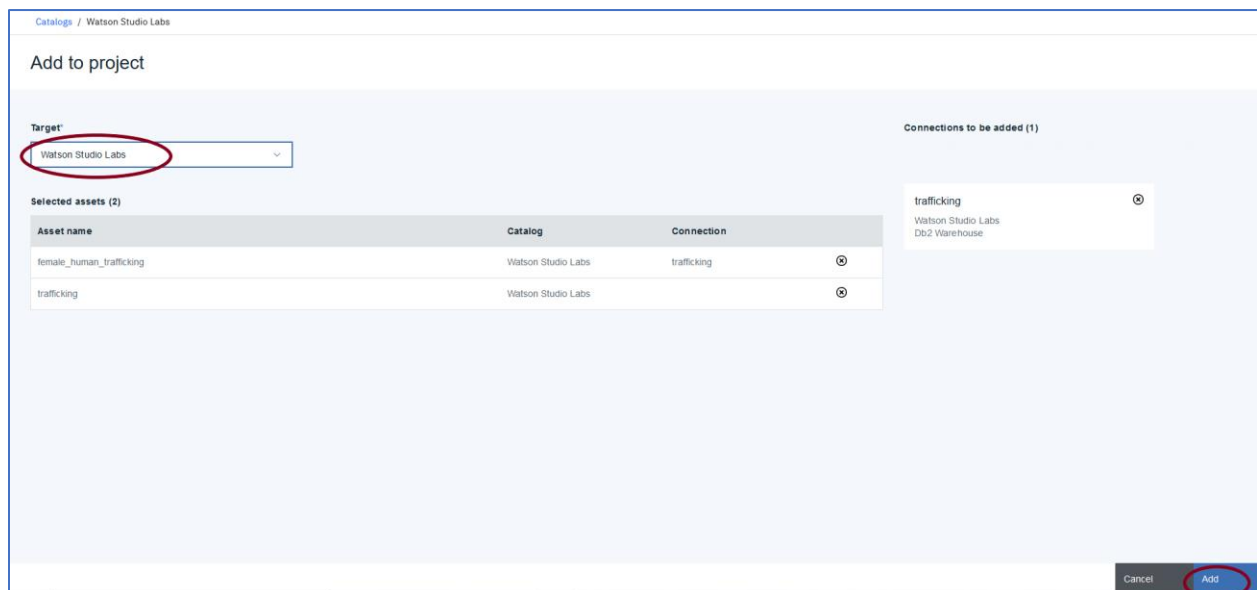


2. Click on the checkbox to the left of the **female_human_trafficking** asset and the **trafficking** asset. Click on **Add to Project**.

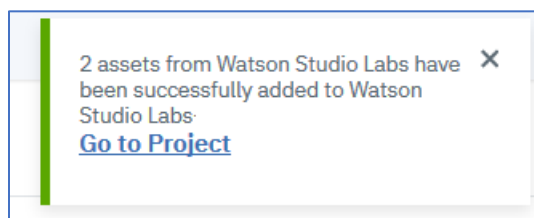
Add to project Remove 2 of 4 items selected Cancel

Name	Owner	Tags	Business terms	Type	Date added
<input type="checkbox"/> Categories	George Doe	job types job categories		Data asset	Apr 26, 2020
<input type="checkbox"/> Occupation	George Doe	occupations jobs		Data asset	Apr 26, 2020
<input checked="" type="checkbox"/> female_human_trafficking	George Doe	trafficking female human trafficking		Data asset	Apr 26, 2020
<input checked="" type="checkbox"/> trafficking	George Doe			Connection	Apr 26, 2020

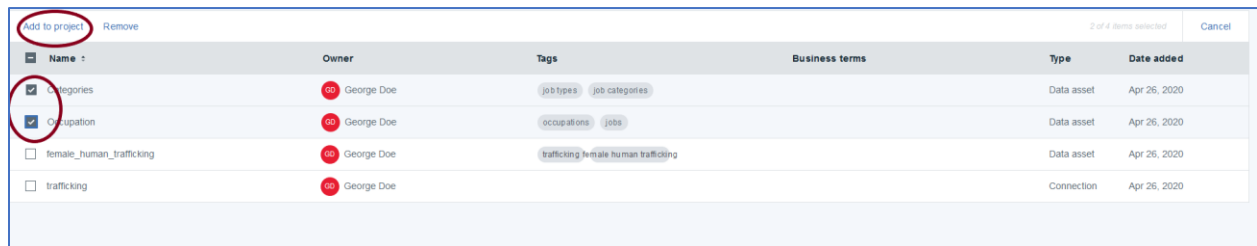
3. Select the Watson Studio Labs as the **Target**. Click on **Add**.



4. Wait for the message that 2 assets have been added to the project in the top right.

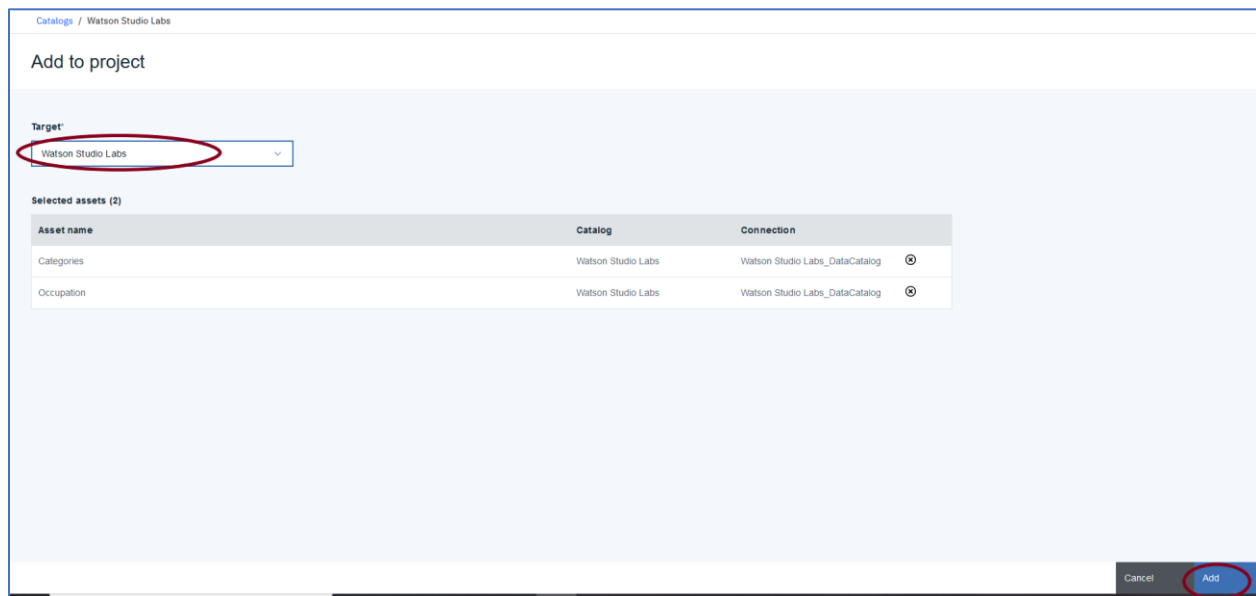


5. Click on **Categories**, and **Occupation**, and then click again on **Add to Project**.



Name	Owner	Tags	Business terms	Type	Date added
<input checked="" type="checkbox"/> Categories	George Doe	job types job categories		Data asset	Apr 26, 2020
<input checked="" type="checkbox"/> Occupation	George Doe	occupations jobs		Data asset	Apr 26, 2020
<input type="checkbox"/> female_human_trafficking	George Doe	trafficking female human trafficking		Data asset	Apr 26, 2020
<input type="checkbox"/> trafficking	George Doe			Connection	Apr 26, 2020

6. Select **Watson Studio Labs** as the **Target** and click **Add**.

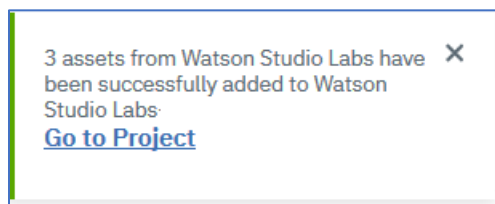


Target: Watson Studio Labs

Asset name	Catalog	Connection
Categories	Watson Studio Labs	Watson Studio Labs_DataCatalog
Occupation	Watson Studio Labs	Watson Studio Labs_DataCatalog

Cancel Add

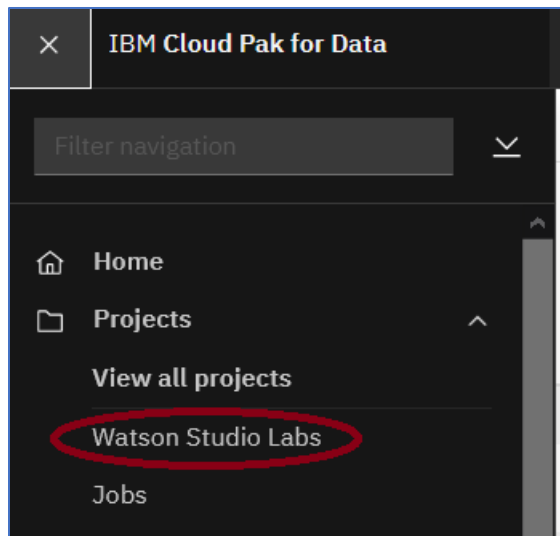
7. Wait for the message that **3** assets have been added to the project in the top right. Two data assets and a connection to the catalog object storage.



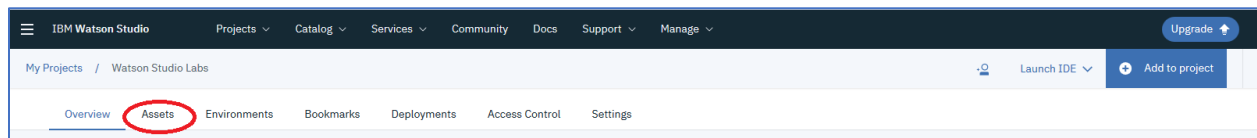
8. Click on the hamburger icon .



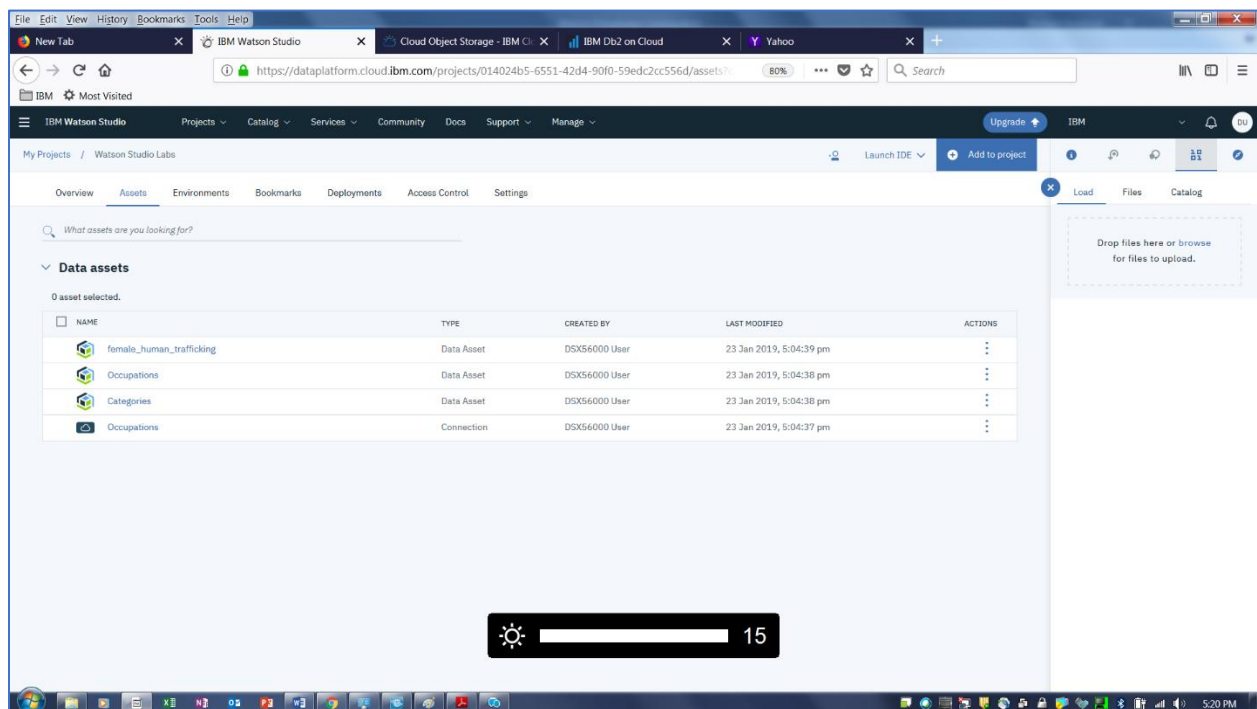
9. Click on **Watson Studio Labs** under **Projects**.



10. Click on the **Assets** tab.



11. The Data assets are now incorporated in the project.



You have completed Lab-2!

- ✓ Created a governed catalog
- ✓ Added a member to the catalog
- ✓ Added Data Assets to the catalog
- ✓ Searched the catalog
- ✓ Edited/Reviewed/Profiled a structured Data Asset
- ✓ Demonstrated access control features
- ✓ Demonstrated policy creation and enforcement (optional)
- ✓ Pushed the Data Assets to the project set up in Lab-1