Intelligent Contracts Workshop Guide

Contents

AI Ladder & ModelOps

(top)

IBM Cloud Pak For Data provides a stack of capabilities broken down into these 4 areas known as the AI Ladder.
ModelOps is an end to end process for developing and deploying data science assets to production that are monitored for bias.
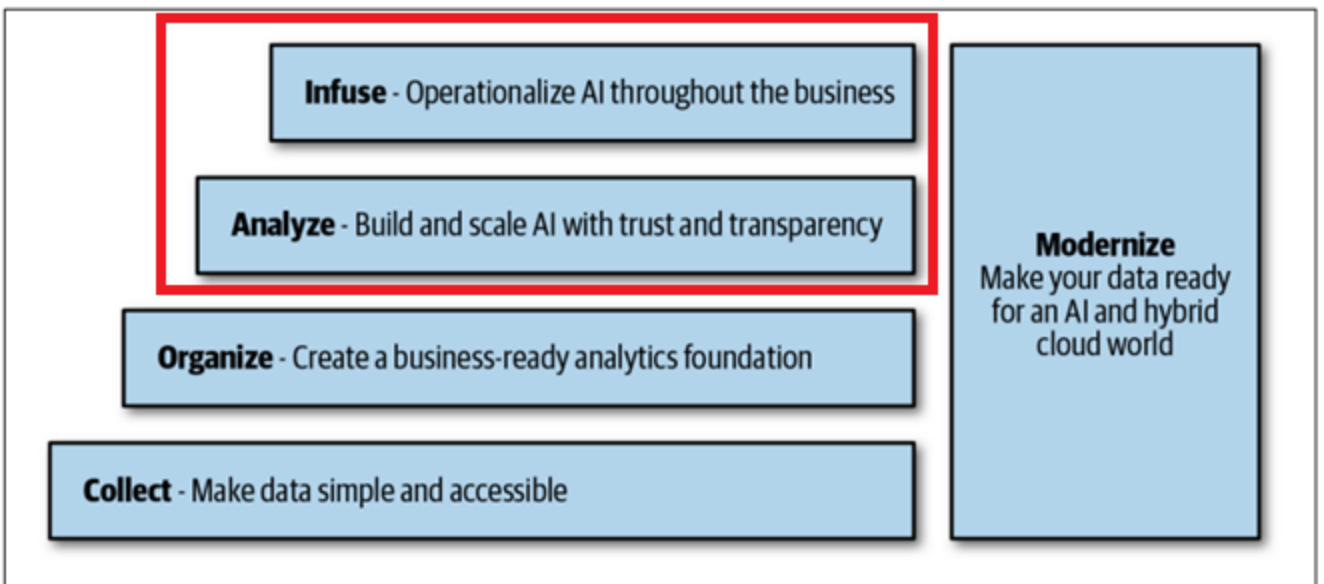

Figure 1 AI Ladder - Collect, Organize, Analyze, and Infuse AI into downstream applications.

Intelligent Contracts Workshop Scenario

(top)

In this workshop, we'll learn how to **infuse** an AI deployed model in an automated business workflow to gain prescriptive suggestions improving the likelihood of contract success.
We'll be using IBM Watson Studio, to **analyze**, build, test and deploy the machine learning models built on structured and unstructured data used to score contracts and make prescriptive suggestions.
Using Watson Discovery, we'll see how we extracted and merged unstructured textual concepts coming from many contracts with structural contract data to improve the accuracy of our AI model.

[ADD STUARTS CONTENT HERE]

# Watson Discovery Contracts Intelligence Workshop

(top)

Watson Discovery Data Flow (top)
Watson Discovery provides multiple methods of acquiring and enriching documents of various type. In the diagram below you can see three phases Acquire, Enrich, Analyze and Search.

In the Acquire stage, data is ingested from one or more data sources, stripped of unnecessary content (such as graphics and formatting) and then passed to the Enrich stage.

In the Enrich stage, the data (typically a mix of structured and unstructured text) is processed using techniques such as Natural Language Processing and Machine learning to provide meaning and context to the raw text. This enriched data is stored as a Collection.

The Analyze and Search stage uses the enriched data from one or more Collections to conduct discovery and exploration or to enable expert assistance through search-based applications.
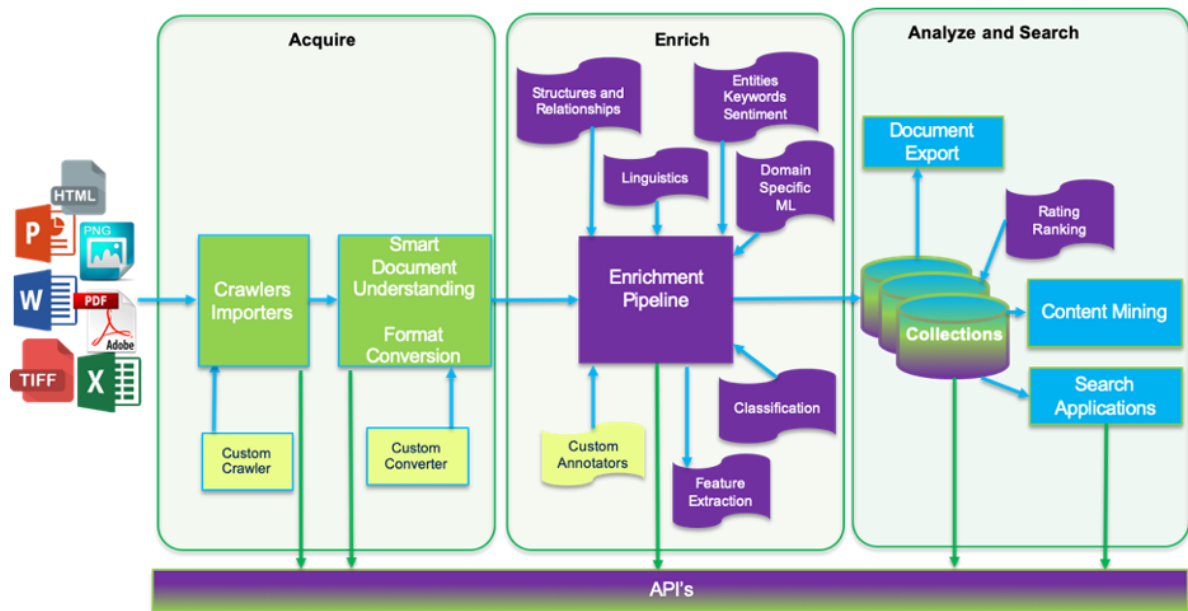
Watson Discovery uses Natural language queries as well as structured queries to find relevant results using passages and highlights across data that has been enriched and ingested. In this particular demo the data ingested is a small number of procurements contracts. We have chosen to apply a pre-built Contracts Intelligence enrichment model during ingestion in order to leverage Watson's understanding of contractual data.

Demo Setup

(top)

NOTE: This use case demonstration calls for Watson Discovery Enterprise Edition to be licensed and used for the purpose of the "Contract Analysis" feature. However, the workshop includes setup and demonstration instructions for the Plus Edition as well if you cannot access the Enterprise Edition.

1. Install Watson Discovery through the Cloud Services Catalog. https://cloud.ibm.com/
2. Once you have launched the Discovery Tooling, look to the right of the initial Project screen for the New project + icon and click on it.

3. Enter a Project name (Contracts Intelligence), Select the "Document Retrieval" project type. Then check the box for "Working with Contracts". Click Next

   NOTE: If you do not have an Enterprise license, then the "Working with Contracts" option will not be available, and you should just continue with the steps outlined below. The optional demo for this will outline steps for a demonstration without the Contracts Model being used

4. Now select "Upload data", then click Next.

5. Name you collection something unique, open the "More processing settings" options drop down and ensure the OCR option is enabled. (some of the content is scanned as an image, so this will enable us to pull out text). Click Next

Let's create a collection for your data

General

Data source
🖥 Upload data

Collection name

Procurement Data

Collection language

English ⌄

Syncing FAQ Content? Beta

Extract pairs of questions and answers
automatically from FAQs (frequently asked
questions). Your application can use the
extracted pairs to give more precise results.

☐ Apply FAQ extraction

⌃ More processing settings

Apply optical character recognition (OCR)   ● On

Back   Next →

6. Now, download the content found in the github repository for this workshop to your computer and then unzip into a local folder…This should give you 6 documents.

You can either click and drag the docs from a folder or click within the box to open a local file explorer popup. Then click Finish in the bottom right corner of the screen.
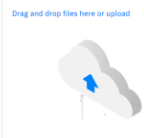
Let's create a collection for your data

Upload your data

Upload files

Max file size is 10MB. Supported file types are PDF,
HTML, JSON, Word, Excel, PowerPoint, PNG, TIFF, JPG,
GIF, TXT, CSV, ZIP, GZIP, TAR.

Drag and drop files here or upload

Favorites
⊘ Recents
🅰 Applicati...
⊕ Downloads
🏠 waltperez
🖼 Pictures
🎵 Music
🎬 Movies

iCloud
☁ iCloud Dri...
🗎 Documents
🖥 Desktop
🖿 Shared

Locations
⊟ Box

< > ☰ ∨ ☷ ∨ 📁 Contracts ⟐ Q Search

Name   Size

🅿 seligence-llc-base-agreement-49130167761.pdf
🅿 p-ntsa-english-v1_1-id-4909035125_0.pdf
🅿 Art Effects Koya Creative Base TSA 2008.pdf
🅿 stanfield-4912049674-tsa.pdf
🅿 0-master-agreement-4905cm0029.pdf
🅿 Epicor-ERP-Manufacturing-Overview-BR-ENS.pdf
🅿 Trademark License Agreement.pdf
🅿 Legal - Apple Media Services - Apple.pdf
🅿 Terms of Service - FIFA.com.pdf
🅿 Art_Effects_Koya_Creative_Base_TSA_2008.rev.pdf

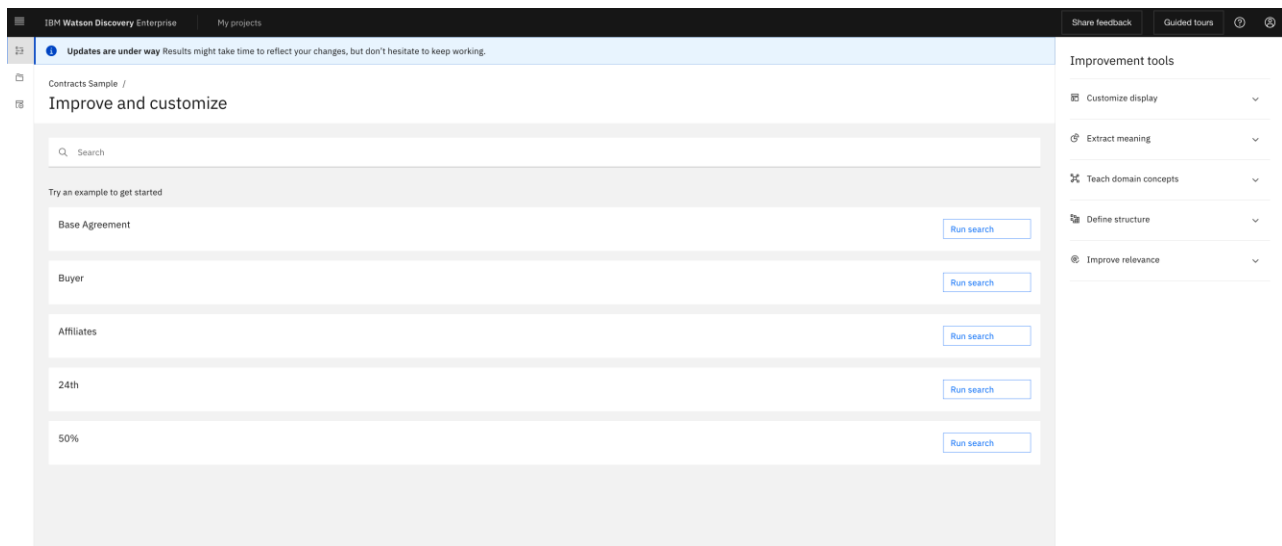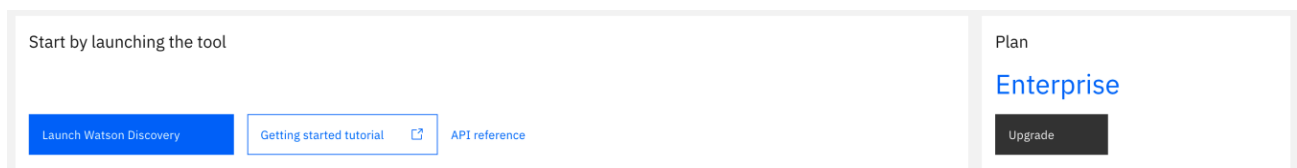Options   Cancel   Open

Back   Finish →

Once the data has loaded, you will see the following screen, which indicates the data has successfully loaded and is ready for the demonstration.

Demonstration (Discovery Enterprise Edition, with Contract Analysis)

(top)

1. Launch Watson Discovery Tooling from Services and Software Resource List in ibm.cloud.com.

2. Select the **Contract Intelligence** Project from the **My Projects** list. You will now be in the **Improve and customize** panel of Discovery Tooling and ready to demo.

3. In the search box enter **Stanfield Intellectual Property** as the query.
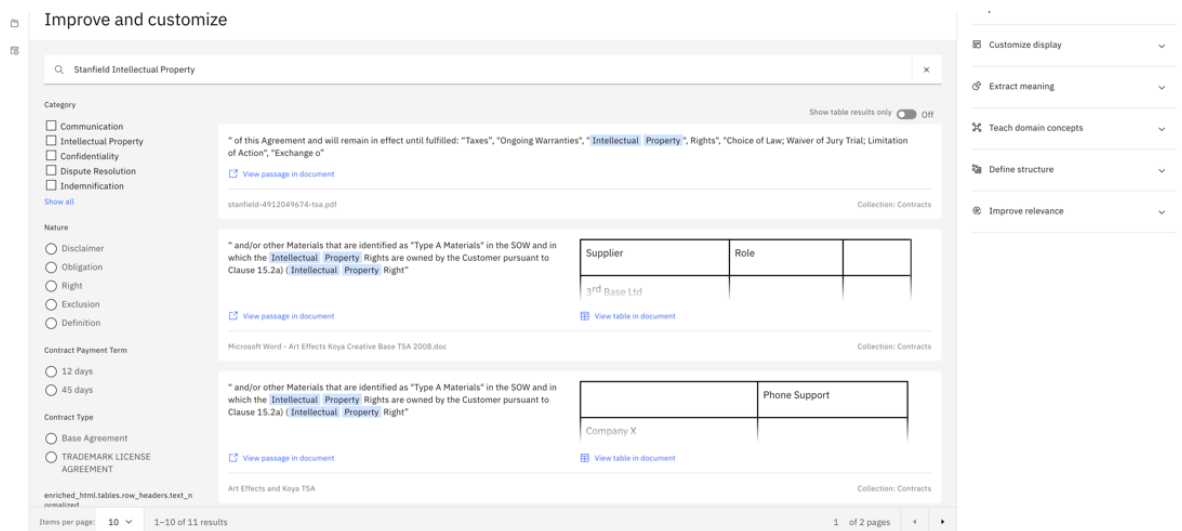
In the return results you will see that not only results have been displayed, but also enriched structure from the document now exists in order to use for digging deep into the exact areas of the contract we are interested in.

4. Click on **view passage in document** for the first result in the list. This will show the natural understanding of the concept we queried highlighted within the original document.

5. Click on **Open advanced view** in the bottom right of the window. This is a Contracts specific User Interface designed to expose the full gambit of the Contracts Intelligence. You can see in this screen, we now have more structured enrichments that have been exposed from the specific document.

6. Click on the **Category**, **Intellectual Property**.

Now every concept from the contract around Intellectual Property has been identified and highlighted in the document. You can use the top navigation to jump from highlighted passage to highlighted passage.

This shows that we have successfully identified the concept of Intellectual Property as an enrichment in the document, which will be crucial as a predictive field with our Machine learning model in Watson Studio.

7. With Intellectual Property still selected, now click on the **Buyer** radio button under the **Party** category. We have now intersected the one area in the contract where Intellectual Property is discussed in regards to the Buyer.
8. Other metadata is extracted as well and can be used to enrich models built with Machine Learning. Click on the **Metadata** tab above the categories.
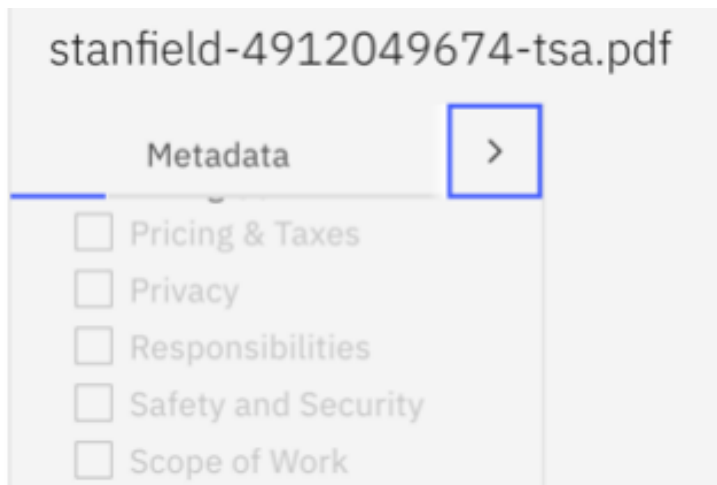
This will expose other metadata, such as the effective date of the contract.

9. You can also add custom enrichments assisted by Watson AI through the Discovery tooling as well. Click on the ╱ *Improve and customize* ╱ hyperlink located at the top left of the window to go back into the main Improve and customize screen. On the right side, click on the **Teach Domain concepts**. This opens more AI assisted tools to further enrich you contracts data by adding custom Dictionary terms, extracting custom entities or even applying custom Machine Learning.

Now that we have shown how we can enrich unstructured content, Watson Studio will be used to leverage existing data fields as well as the newly extracted Watson Discovery enrichment fields to create predictive models.

Demonstration (Discovery Plus Edition)

(top)

This alternative Demonstration will rely on a custom field being created as the pre-trained Contracts enrichments are not available. This path can also be shown in addition to the above path to illustrate the fact that customized fields can also be added to the pre-trained models in Discovery.

1. Launch Watson Discovery Tooling from Services and Software Resource List in ibm.cloud.com.
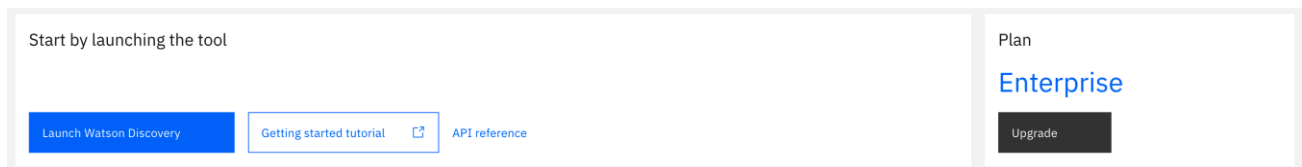
2.
   Select the **Contract Intelligence** Project from the **My Projects** list. You will now be in the **Improve and customize** panel of Discovery Tooling and ready to demo.
3. In the search box enter **Stanfield Intellectual Property** as the query. Notice the Entities being displayed on the left. Since we do not have a pre-trained contracts model listing of entities, we will need to show how we can easily start building a custom entity to easily find the concept of intellectual property.
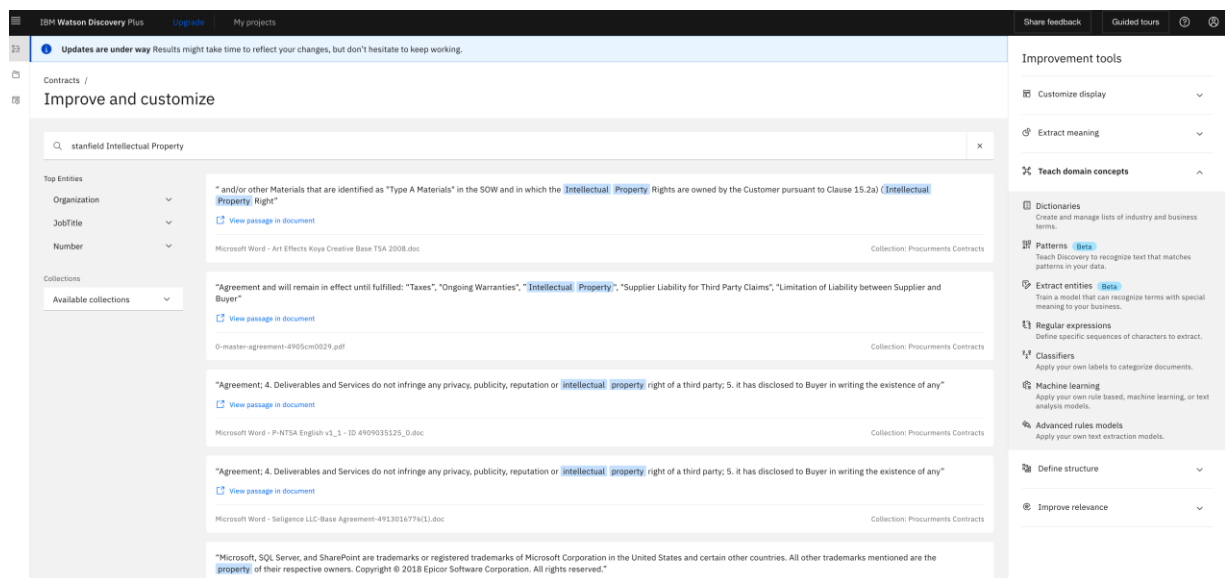
4. On the right side of the screen open the "Teach Domain Concepts" drop down. Now click on the "Dictionaries" option.

Dictionaries

New   +   Upload   +

Name	Used in

○ □ △   **No dictionaries**
△ ○ □   If you add dictionaries, you'll be able to manage them
□ △ ○   here.
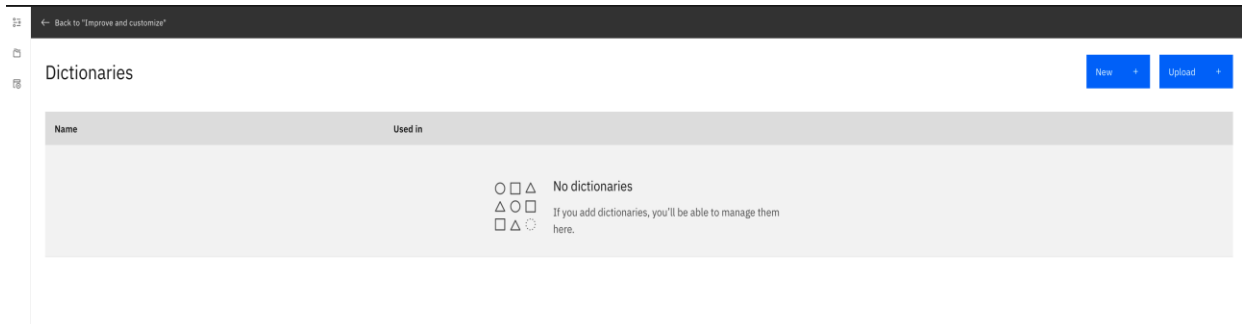
Now, click on "New", then enter the Dictionary name as, "Intellectual Property". For the included terms, enter…intellectual property, inventions, third party code, and developed works. Lastly, click on the "Save Dictionary" located in the upper right corner of the screen.

IBM **Watson Discovery** Plus   Upgrade   My projects

Share feedback   Guided tours   ⑦   ⑧

← Back to "Dictionaries"

Create dictionary

Dictionary name   Language ⓘ

Intellectual Property   English ∨

Save dictionary 🗊

∨ Advanced options

Included terms   **Suggested terms**

Term   Refresh ↻

Enter text   +

Base term	Other terms	Actions
intellectual property		✎ 🗑
inventions		✎ 🗑
third party code		✎ 🗑
developed works		✎ 🗑

Terms per page: 10 ∨   1–4 of 4 terms   1 ∨ of 1 pages   ‹ ›

We're setting up your suggestions.
They should be ready soon.

5. Now click on the field to which this enrichment will be applied to. In our case, click "text".

Where would you like this enrichment to be applied? ×

⦿ Procurments Contracts
  Fields to enrich

  | 1 × Selected fields | ⌃ |

  ☐ table

  ☐ table_of_contents

  ☑ text

  ☐ [text]tle

| Cancel | Apply |

6. The goal is to create an enriched field to be extracted from the unstructured text. Once this has been saved, you will notice a popup that mentions re-processing of the collection has started. It will take a few minutes for the collection to be refreshed with the new field we just created. So, in the meantime, go back to the "Bach to Improve and Customize" screen by clicking on the hyperlink in the upper lefthand corner of the screen.

← Back to "Improve and customize"

Dictionaries   New +   Upload +

Name	Used in			
Intellectual Property	Procurments Contracts	✎	🗑	⤓

Once the processing has finished and enter a query, you should see the following screen, which contains our newly created field of intellectual property. This shows that we have successfully identified the concept of Intellectual Property as an enrichment in the document, which will be crucial as a predictive field with our Machine learning model in Watson Studio.

Improve and customize

You can even go further and show the exact type of Intellectual Property...

Clicking on one of the sub-entity types, will further slim down the results to show only the passages that contain that term. Try clicking "inventions", which should yield one result.

Now that we have shown how we can enrich unstructured content, Watson Studio will be used to leverage existing data fields as well as the newly extracted Watson Discovery enrichment fields to create predictive models.

Using Watson Studio With Watson Discovery To Build And Deploy a Contract Success AI Model
(top)

In our previous demo with Watson Discovery, we learned how it was able to provide Domain Concepts like "Intellectual Property" from many contractual documents.
When building our machine learning model to predict contact success, we're going to merge the unstructured concepts like "Intellectual Property" with the historical contract history to build and deploy a more robust machine learning model to predict contract success.
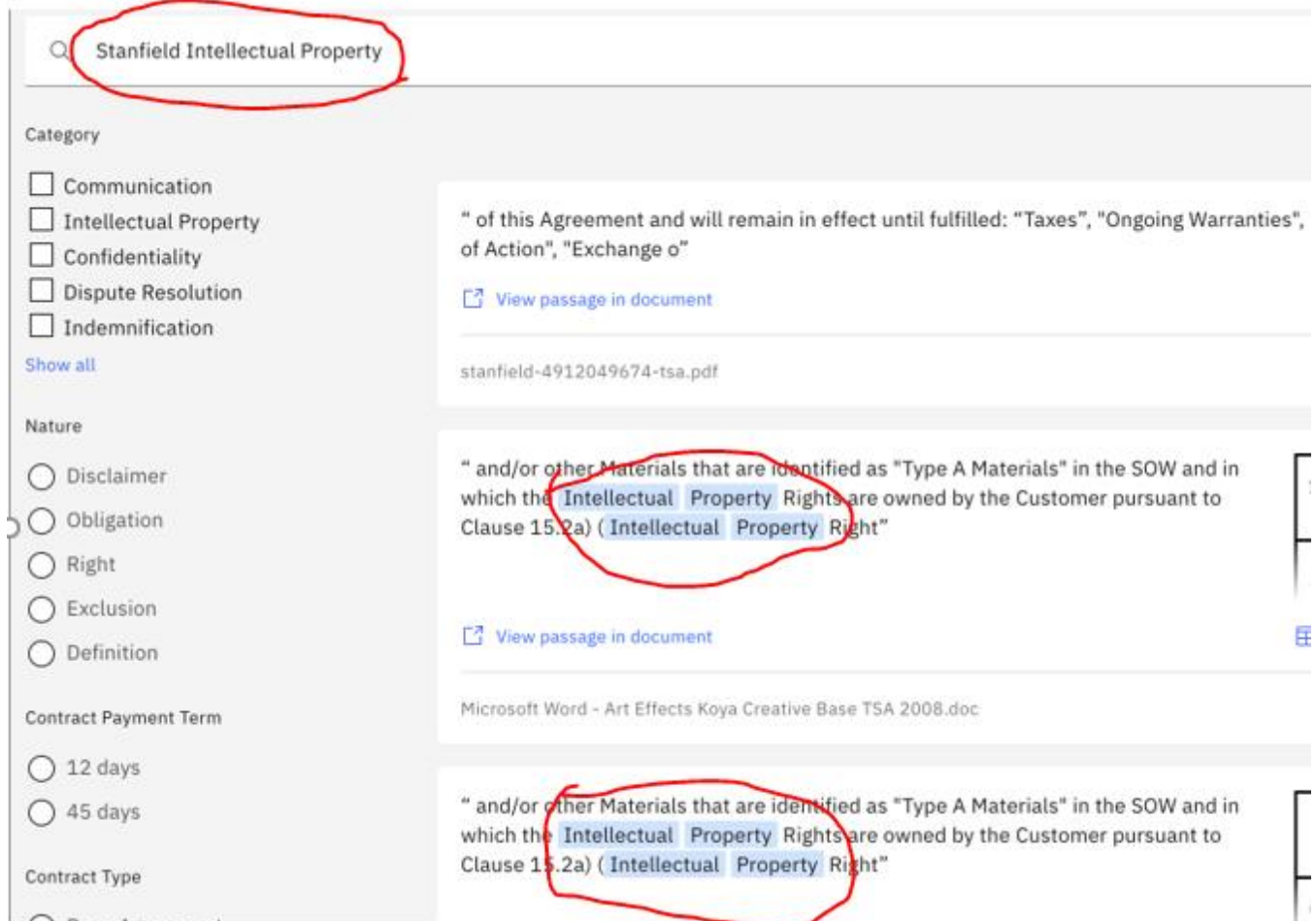
Figure 2 Example of using Watson Discovery To Uncover Common Categories

Required software, access, and files
(top)

1. To complete this lab, you will need a **Cloud Pak for Data as a Service** (**CPDaaS**) account: https://dataplatform.cloud.ibm.com
   If you don't have a CPDaaS account, use the same URL to sign up for a free trial. The account will be activated in approximately 5 minutes.
2. If you already have an **IBM Cloud** account, make sure that you provisioned the required services
   a. **Watson Studio**
   b. **Watson Machine Learning (WML)**
   c. **Watson OpenScale**
   d. **IBM Cognos Dashboard Embedded.**
3. Navigate to your *Resource list* in your **IBM Cloud** dashboard: https://cloud.ibm.com/resources
4. Check if the mentioned services are displayed under **Services**. If not, search for the services in the **Catalog** and add them.
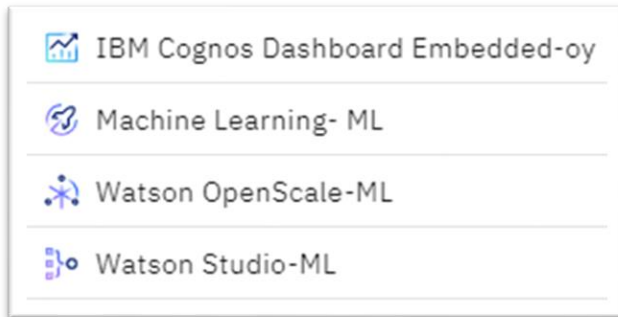
NOTE: You will likely not have the same endings. (IE. '-oy' and '-ML')

5. Download the **project zip file** to your machine. It will be imported into a Cloud Pak For Data Project.

6. Please visit this [git repository](#) to download all the files for this workshop.

7. 🛑 Watch [this recording](#) before you proceed! It will step you through the demo in approx. 9 minutes.

Download These Files From Git
(top)

1) **Intelligent-Contracts.zip**
Import this zip file when creating a new IBM Watson Studio Project
It will include the following project assets:

- 3 CSV files - Scored_Contracts.csv, Contract_Analysis.csv, Unscored_Contracts.csv
- Python Notebook - ContractSuccessRandomForest-Py38 Using Sklearn - Using Open Source SkLearn Python with a Random Forest Model deployed to IBM Platform
- Modeler Flow - Build and deploy the same Python model without the need fof coding
- Dashboard - Historical and Predictive Visualizations using structured and unstructured contract data
- Model - Contract Success Using AutoAI - P3 Snap Random Forest Classifier - This model was generated by AutoAI. We will deploy in the workshop.

2) **Intelligent-Contracts-Workshop-Guide.pdf (CURRENT DOCUMENT)**
Use this step-by-step workbook to:

- Automate Contract Workflow (Ingestion, Analysis, Approvals)
  Uses Cloud Pak For Business Automation
- b) Enrich Intelligence Accuracy (Extracting Information from Contracts)
  Uses Cloud Pak For Data Watson Discovery
- c) Integrate Continuous Improvement (Infusing AI into the Operation)
  Uses Cloud Pak For Data Watson Studio

3) **data_science_modelops_team_sport.png**
ModelOps as a Team Sport image depicting all those involved in delivering AI models and how IBM has them all collaborate together on one platform.
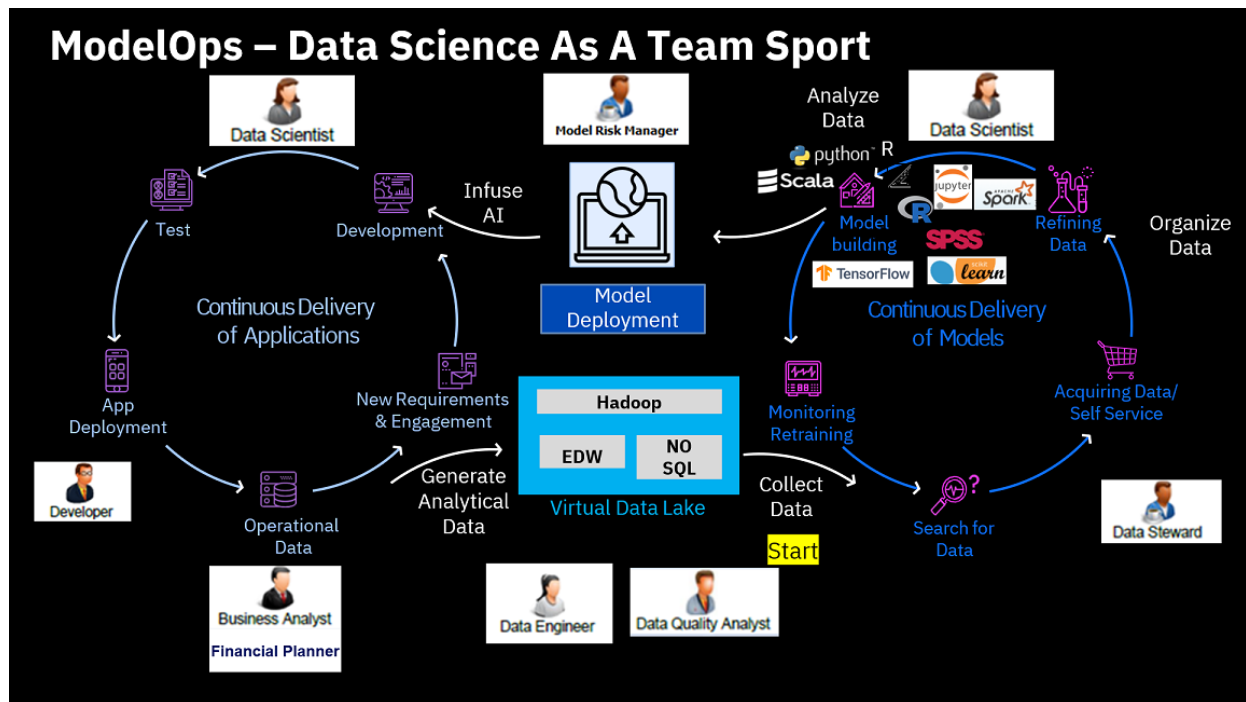
4) **contracts.zip [Already used in previous workshop]**
Sample contracts used in demo with Watson Discovery and used in the modelOps workflow.

5) **modelOpsForContractSuccess.mp4**
A recording of the end to end modelOps demo.

6) **csv_files.zip**
(Optional) Start a new Project in Cloud Pak For Data, import these CSV files and create all the project assets yourself from scratch.
Watch the video as a guide.

7)**Trustworthy AI.mp4**
(Optional) In this 9 minute recording, learn how IBM provides AI guardrails by monitoring key model metrics on and off our platform using OpenScale to avoid data bias or drift with model risk workflow to adhere to regulator guidelines like SR 11–7 that requires model risk management for all models in financial services using OpenPages.

8)**ContractSuccessRandomForest-Py38 Using Sklearn.ipynb**
(Optional) Import this Python notebook using Open Source Sklearn and Random Forest Model.

Required skills
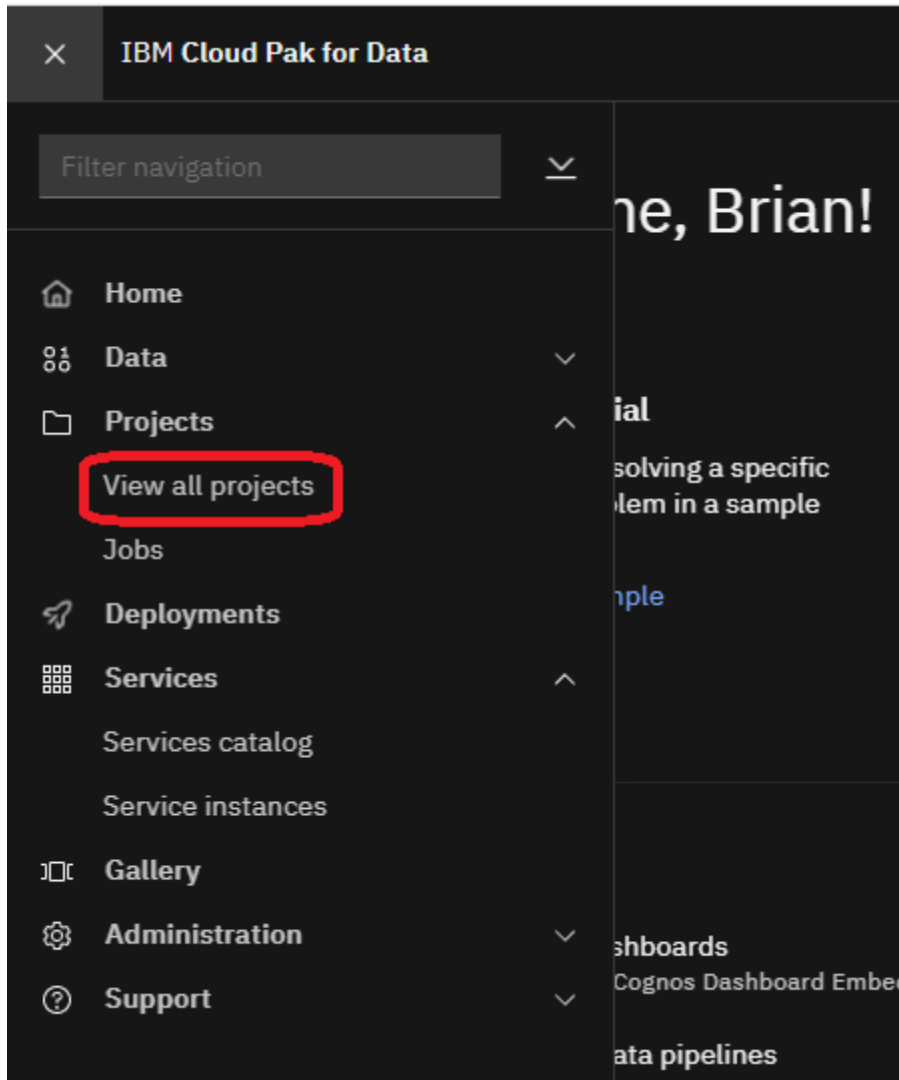(top)

We recommend that users who work through this lab:

- Understand the data science model lifecycle
- Have at least beginner knowledge of different methods for creating models

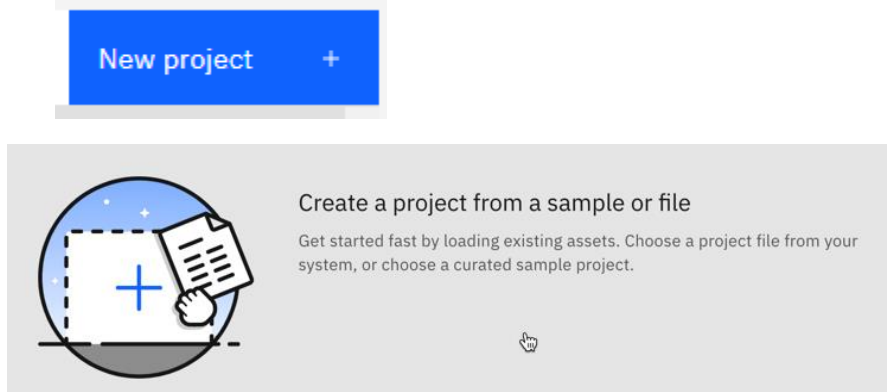## Step 1: Import Intelligent-Contracts.zip As New IBM Watson Studio Project
(top)

1. Sign into [Watson Studio](). From the quick navigation on the left, select **Projects → View all projects**.

2. On the far right, click the **New project**, then select **Create a project from a sample or file**.

Create a project from a sample or file

Get started fast by loading existing assets. Choose a project file from your system, or choose a curated sample project.

3. Drag the *project zip* file (downloaded from Git) from your machine to the **Upload file** section.
    Name your project **Intelligent Contracts**
    Select a **Cloud Object Storage** service
    Select **Dashboard Service** from the dropdown list
    Click **Create**.

4.