## Administration

### Do Not do the Following!

1. **Do not** create a connection to the Db2 database.
2. **Do not** run discovery through Data Curation > Data Discovery on the entire database. We only need to discover three schemas; MORTGAGE, INSURANCE and CUSTOMER. We wil do this as the very last steps in the run book.

If we discover everything it will flood the Default synch catalog with unwanted and unnecessary tables with OpenScale and all System objects. We will then have to go in and delete everything through the Metadata import facility (known as IMAM) and that could mess up IGC if we don’t do it carefully.

### User Permissions

1. Add the experience user to the **Enterprise** catalog as a **Viewer**
2. Add the experience user to the **mortgage-analysis-xxx** project as an **Editor**. They will not be able to view the Data Flow as a **Viewer**.

## Create the Organize Data

This will all be done in the Db2 Server Edition of the cluster.

### Create the Schemas and Tables

1. Run the **Organize Experience.sql** script using the run sql feature of the Db2 Service Edition to create the tables.

It will create 3 schemas - CUSTOMER, INSURANCE and MORTGAGE and 10 tables.

### Load the Data

Load the 10 tables using the load feature of the Db2 Server Edition. You can do it in any order you choose using the provided CSV files:

1. Select the **Auto Insurance Claim.csv** file and load it into the INSURANCE.AUTO\_INSURANCE table.
2. Select the **Auto Insurance Policy.csv** file and load it into the INSURANCE.AUTO\_POLICY table.
3. Select the **Customer.csv** file and load it into the CUSTOMER.CUSTOMER table.
4. Select the **Customer Activity.csv** file and load it into the CUSTOMER.CUSTOMER\_ACIVITY table.
5. Select the **Customer Attrition.csv** file and load it into the CUSTOMER.CUSTOMER\_ATTRITION table.a
6. Select the **Customer Offers.csv** file and load it into the CUSTOMER.CUSTOMER\_OFFERS table.
7. Select the **Mortgage Applicant.csv** file and load it into the MORTGAGE.MORTGAGE\_APPLICANT table.
8. Select the **Mortgage Customer.csv** file and load it into the MORTGAGE.MORTGAGE\_CUSTOMER table.
9. Select the **Mortgage Default.csv** file and load it into the MORGAGE.MORTGAGE\_DEFAULT table.
10. Select the **Mortgage Property.csv** file and load it into the MORTGAGE.MORTGAGE\_PROPERTY table.

**Note** - You can change the script to create and load the data all at once. I just have not had time to create one.

## Prepare the Knowledge Catalogs

### Rename the Default Catalog

This is default synched catalog and necessary so we can create another catalog, but the end user will never see it or use it. But, if we use it in the future I want to rename it so it is more meaningful. After testing, I may choose to delete it all together but won’t know until all testing is done. Therefore, I do have a deletion step at the end if need be.

Go to Organize > All catalogs

1. Click on the **Default Catalog**
2. Go to **Settings**
3. Click on the pencil icon next to the Name and change it to **Governance**.
4. Click the **Apply** button.
5. Click on the pencil icon next to the Description and wipe out what is there and replace it with this **The default synched catalog for sharing information governance assets.** (including the period at the end).
6. Click the **Apply** button.

### Create the Enterprise Catalog

Go to Organize > Catalogs > All catalogs

1. Click on the **New Catalog** button.
2. Enter a Name of **Enterprise**
3. Enter a description of **The enterprise catalog that contains all trusted and governed data and information assets.** (including the period at the end).
4. Check the box to **Enforce data protection rules**.
5. Select the **Create** button.

## Build the Business Glossary

### Create the Governance Artifacts

Import the following governance artifacts in the order listed below:

#### Import Categories

Go to Organize > Data and AI governance > Categories

1. Click on the **Import** button.
2. Click on the **Add file** button.
3. Select the **glossary-organize-categores.csv** file and click **Open**.
4. Click **Next**.
5. Select the **Replace all values** option.
6. Click on **Import**. It should complete successfully with 4 new categories.
7. Click **Close**
8. Select the **Refresh** button on your browser to see them.
9. You should see 2 main categories - **Data Privacy** and **Mortgage Default Analysis**.
10. Click on the **Data Privacy** category and you should see a **PCI** sub-category.
11. Click on the **Morgage Default Analysis** category and you should see a **Sensitive Information** sub-category.

#### Import Business Terms

Go to Organize > Data and AI governance > Business terms

1. Click on the **Import** button.
2. Click on the **Add file** button.
3. Select the **glossary-mortgage-default-analysis-terms.csv** file and click **Open**.
4. Click **Next**.
5. Select the **Replace all values** option.
6. Click on **Import**. It should complete successfully with 28 new drafts.
7. Click **Go to Task** button.
8. Click the **Publish** button. Wait until you see the success message…
9. Go to Organize > Data and AI governance > Business terms
10. From the **Published** tab, scroll to the bottom of the business term list and you should see at the bottom that there are 28 published business terms.

#### Import Policy

Go to Organize > Data and AI governance > Policies

1. Click on the **Import** button.
2. Click on the **Add file** button.
3. Select the **glossary-organize-policies.csv** file and click **Open**.
4. Click **Next**.
5. Select the **Replace all values** option.
6. Click on **Import**. It should complete successfully with 1 new draft.
7. Click **Go to Task** button.
8. Click the **Publish** button. Wait until you see the success message…
9. Go to Organize > Data and AI governance > Policies
10. From the **Published** tab, you should see 1 policy named **Protection of Sensitive Information**.

#### Import Data Governance Rules

Go to Organize > Data and AI governance > Rules

1. Click on the **Import** button.
2. Click on the **Add file** button.
3. Select the **glossary-organize-governance-rules.csv** file and click **Open**.
4. Click **Next**.
5. Select the **Replace all values** option.
6. Click on **Import**. It should complete successfully with 4 new drafts.
7. Click **Go to Task** button.
8. Click the **Publish** button. Wait until you see the success message…
9. Go to Organize > Data and AI governance > Rules
10. From the **Published** tab, you should see 4 governance rules all starting with names like **All…**.

### Assign Classification and Data Classes to Business Terms

Assign the **Sensitive Personal Information** classification to all business terms listed below as you are assigning their data class to the business terms:

This is all done by selecting each business term, scrolling to the bottom and then in the **Related artifacts** section clicking on the **plus sign +** next to **Classifications**.

1. Credit Card Expiration Date
2. Credit Card Number
3. Credit Card Validation Number
4. Email Address
5. Phone Number
6. Social Security Number

Assign the following data classes to their equivelant business terms:

This is all done by selecting each business term, scrolling to the bottom and then in the Related Artifacts section clicking on the + sign next to **Data classes**.

1. Assign the **Identifier** data class to the **Application Id** business term.
2. Assign the **City** data class to the **City** business term.
3. Assign the **Credit Card Expiration Date** data class to the **Credit Card Expiration Date** business term.
4. Assign the **Credit Card Number** data clas to the **Credit Card Number** business term.
5. Assign the **Credit Card Validation Number** to the **Credit Card Validation Number** business term.
6. Assign the **Email Address** data class to the **Email Address** business term.
7. Assign the **Employment Status** data class to the **Employment Status** business term.
8. Assign the **Gender** data class to the **Gender** business term.
9. Assign the **Legal Marital/Civil Status** to the **Marital Status** business term.
10. Assign the **Person Name** data class to the **Name** business term.
11. Assign the **US Phone Number** data class to the **Phone Number** business term.
12. Assign the **US Social Security Number** data class to the **Social Security Number** business term.
13. Assign the **US State Name** data class to the **State** business term.
14. Assign the **US State Code** data class to the **State Code** business term.
15. Assign the **US Street Name** data class to the **Street Address** business term.
16. Assign the **US Zip Code** data class to the **Zip Code** business term.

### Add Artifacts to the Mortgage Default Analysis Category

Go to Organize > Data and AI governance > Category and select the Mortgage Default Analysis category.

In the related Click the **Add to category** button to add relationships to the following data governance artifacts:

1. Classification - Select the **Sensitive Personal Information** classification
2. Data Governance rules - Select the **All Email Addresses Must be Protected, All Phone Numbers Must be Protected and All US Social Security Numbers Must be Protected** data governance rules
3. Policy - Select the **Protection of Sensitive Information** policy

### Create the Data Protections Rules

This is done by going to Organize > Data and AI governance > Rules

1. Click on the **Create Rule** button.
2. Select **Data Protection** rule.
3. Enter a Name of - **Protect Credit Card Information**
4. Enter a Business definition of - **Protect all credit card numbers, expiration dates and validation numbers using the data redaction method.** (including the period at the end)
5. Build the rule as follows:

Criteria = If Data class contains any Credit Card Number, Credit Card Expiration Date, Credit Card Validation Number

Action = Then Redact data in columns containing Credit Card Number, Credit Card Expiration Date, Credit Card Validation Number

1. Select the **Rule** bread crumb, or the **arrow** next to the rule you just created, to go back to Rules main page.
2. Click on the **Create Rule** button.
3. Select **Data Protection** rule.
4. Enter a Name of - **Protect Sensitive Personal Information**
5. Enter a Business definition of - **Protect all Sensitive Personal Information using the data obfuscation method replacing data with similarly formatted but fictional values.** (including the period at the end)
6. Build the rule as follows:

Criteria = If Data class contains any Email Address, US Phone Number, US Social Security Number

Action = Then Obfuscate data in columns containing Email Address, US Phone Number, US Social Security Number

1. Click on your browser’s refresh button to see the new rules. You should see them at the bottom of the list of rules.

### Assign Rules and Business Terms to the Policy

This is done by going to the Organize > Data and AI governance > Policies section and clicking on the **Protection of Sensitive Information** policy:

Go to Organize > Data and AI governance > Policies

1. Select the **Protection of Sensitive Information** policy.
2. Scroll down to the **Rules** section.
3. Click on the **Add rules** button.
4. Click on the 4 data governance rules from the list. They all start with the word **All…**.
5. Select the **Publish** button twice.
6. Scroll down to the **Data protection rules** section.
7. Click on the **Add data protection rules** button.
8. Select the 2 data protection rules you just created to the policy.
9. Scroll down to the **Related artifacts** section.
10. Click on the **plus sign +** next to Business terms.

Enter credit to see all 3 push publish twice

1. Add these business terms all at once by selecting them from the list that pops up. Just scroll through the list, and select each one, its in alphabetical order:

**Credit Card Expiration Date** **Credit Card Number** **Credit Card Validation Number** **Email Address** **Phone Number** **Social Security Number**

1. Click on the **Publish** button at the top and when the dialog appears. Do not enter a comment.
2. From the same **Related artifacts** section, click on the **plus sign +** next to **Classifications**.
3. Click on the **Publish** button at the top.
4. Click on the **Publish** button in the dialog. Do not enter a comment.
5. Select the **Sensitive Personal Information** classification to assign it to the policy.
6. Select the **Publish** button twice.

## Build the Enterprise Catalog

### Create the Enterprise Catalog Data Project

I have created an exported project that contains the connection and tables that will be published to the Enterprise catalog and the connection and 2 tables you will also add to the Governance catalog.

Go to Projects

1. Click on New project
2. Click on **Create a project from a sample or file**
3. Enter a name of - **Enterprise Catalog Data**
4. Enter a Description of - **The connection and 10 tables needed for the CP4D Organize Experience flow.** (including the period at the end)
5. Select the **Enterprise-Data-Catalog.zip** file as the upload file.
6. Select **Create**.
7. Go into the project and select the **Analytics Data Warehouse** connection.
8. Change the **host** and **port** to the host and port of Db2 in the cluster.

I have been using the full host name like this: zen-cpd-zen.apps.icp4d-experiences-lite-08.demo.ibmcloud.com - You can get it from Db2 from the connection section in the consolde by selecting **Open**.

1. Select the **Test** button. When it passes, select **Update**.

### Publish Data Assets to the Enterprise Catalog

Perform the following steps in order to ensure that the Mortgage data is published last and that the **MORTGAGE\_APPLICANT** table is published last so that it appears as the first table in the Recently Added section of the catalog.

**Note** - The Db2 connection will only be published to the catalog during the first set of table publications.

From the Data Assets tab in the Enterprise Catalog Data project, select to view all 11 assets and then sort by Name in ascending order.

#### Publish the Insurance Data

1. Select the **Analytics Data Warehouse** connection.
2. Select the **AUTO\_INSURANCE\_CLAIM** table.
3. Select the **AUTO\_INSURANCE\_POLICY** table.
4. Select the **Publish** button.
5. **Make sure** the **Enterprise** catalog is the selected target catalog.
6. Click the **Publish** button.

#### Publish the Customer Data

1. Select the **CUSTOMER** table.
2. Select the **CUSTOMER\_ACTIVITY** table.
3. Select the **CUSTOMER\_ATTRITION** table.
4. Select the **CUSTOMER\_OFFER\_ACCEPTED** table.
5. Select the **Publish** button.
6. **Make sure** the **Enterprise** catalog is the selected target catalog.
7. Click the **Publish** button.

#### Publish the Mortgage Data

1. Select the **MORTGAGE\_CUSTOMER** table.
2. Select the **MORTGAGE\_DEFAULT** table.
3. Select the **MORTGAGE\_PROPERTY** table.
4. Select the **Publish** button.
5. **Make sure** the **Enterprise** catalog is the selected target catalog.
6. Click the **Publish** button.

Making sure MORTGAGE\_APPLICANT is the last table published

1. Select the **MORTGAGE\_APPLICANT** table.
2. Select the **Publish** button.
3. **Make sure** the **Enterprise** catalog is the selected target catalog.
4. Click the **Publish** button.

### Publish Connection and Tables to the Governance Catalog

The **Analytics Data Warehouse** connection needs to be published to the **Governance** catalog along with the **CUSTOMER** and **MORTGAGE\_APPLICANT** table. These are the only assets we need to publish to this catalog. It has to be done so that **Data Discovery** will find the connection and these two tables in the synched Information Governance Catalog.

From the Data Assets tab in the Enterprise Catalog Data project, select to view all 11 assets and then sort by Name in ascending order.

#### Publish Only the Connection

1. Select the **Analytics Data Warehouse** connection.
2. Select the **CUSTOMER** table.
3. Select the **MORTGAGE\_APPLICANT** table.
4. Select the **Publish** button.
5. **Make sure** the **Governance** catalog is the selected this time as the target catalog.
6. Click the **Publish** button.

### Modify the Governance Catalog Data Assets

We need to add business terms to the 2 tables we published to the **Governance** catalog.

Go to Organize > All catalogs and select the **Governance** catalog.

#### CUSTOMER Table

1. Click on the **CUSTOMER** table.
2. At the table level, click the pencil icon next to the **Business terms** area in the **Overview** section.
3. Place your cursor in the area and start typing the name of following business terms. As you find them, **Select** them:

**Credit Card Number** **Email Address** **Phone Number** **Social Security Number**

1. When you have added all 4, click on the **Apply** button.

#### MORTGAGE\_APPLICANT Table

1. Click on the **MORTGAGE\_APPLICANT** table.
2. At the table level, click the pencil icon next to the **Business terms** area in the **Overview** section.
3. Place your cursor in the area and start typing the name of following business terms. As you find them, **Select** them:

**Email Address** **Phone Number** **Social Security Number**

1. When you have added all 3, click on the **Apply** button.

#### Validate the Information Assets are in IGC

### Modify the Enterprise Catalog Data Assets

Now that all the data is published to the Enterprise catalog we need to modify some of them and add additional metadata and force some behaviors (for a better client expereince).

Go to Organize > All catalogs and select the **Enterprise** catalog

#### Analytics Data Warehouse Connection

1. Click on the **Analytics Data Warehouse** connection.
2. Click the pencil icon next to the **Tags** area and enter the following tags:
3. Enter a tag name of **Analysis** and click on the **plus sign +** to add it.
4. Click in the tag are and select the **Customer** tag and click on the **plus sign +** to add it.
5. Click in the tag are and select the **Insurance** tag and click on the **plus sign +** to add it.
6. Click in the tag are and select the **Mortgage** tag and click on the **plus sign +** to add it.
7. Click on the **Apply** button.
8. Go to the **Review** tab and do the following:

* Give it a rating of **5 Stars** by clicking on the 5th star to the right.
* Enter this text for the review text (including the period at the end) - **Contains all governed, trusted and quality data approved and published by the data governance team to use for the mortgage default analytics, customer attrition and auto insurance claim fraud analytic projects. Some of the data is sensitive but data protection rules are in place to obfuscate it.** (including the period at the end)
* Click on the **Submit** button.

1. Click on the **Enterprise** catalog bread crumb at the top of the page to go back to the catalog.

#### CUSTOMER Table

1. Click on the **CUSTOMER** table.
2. At the table level, click the pencil icon next to the **Business terms** area in the **Overview** section.
3. Place your cursor in the area and start typing the name of following business terms. As you find them, **Select** them:

**Credit Card Number** **Email Address** **Phone Number** **Social Security Number**

1. When you have added all 4, click on the **Apply** button.
2. At the column level, in the **Overview** section:

Click on the icon that looks like an **eye** for each columns below and:

* Assign the **Email Address** business Term to the **EMAIL\_ADDRESS** column
* Assign the **Phone Number** business Term to the **PHONE\_NUMBER** column
* Assign the **Social Security Number** business Term to the **NATIONAL\_ID** column
* Assign the **Credit Card Number** business Term to the **CREDITCARD\_NUMBER** column
* Assign the **Credit Card Expiration Date** business Term to the **CREDITCARD\_EXP** column
* Assign the **Credit Card Valiation Number** business Term to the **CREDITCARD\_CVV** column
* Assign the **US State Code** business Term to the **STATE\_CODE** column

1. Go to the **Review** tab and do the following:

* Give it a rating of **4 Stars** by clicking on the 4th star to the right.
* Enter this text for the review text (including the period at the end) - **Contains the governed and trusted customer data to use for customer attrition and auto insurance claim fraud analysis projects. It can’t be used for mortgage default analysis because applicants have not become customers. It also is a subset of the US master customer table.** (including the period at the end)
* Click on the **Submit** button.

1. Go to the **Profile** tab and do the following:

* **Make sure** the profile has been run. If it has not, click on the buttom to force a run of the profile and make sure it completes and that all columns have been assigned a data class.

If all columns have been assigned a data class, proceed to do the following:

1. Scroll to the Right.
2. Assign the **Credit Card Expiration Date** data class\* to the **CREDITCARD\_EXP** column.
3. Assign the **Credit Card Validation Number** data class to the **CREDITCARD\_CVV** column.

#### MORTGAGE\_CUSTOMER Table

1. Click on the **MORTGAGE\_CUSTOMER** table.
2. Go to the **Review** tab and do the following:

* Give it a rating of **4 Stars** by clicking on the 4th star to the right.
* Enter this text for the review text - **Contains everything needed to perform accurate mortgage default analysis. However, for deeper analysis, it could use more information related to the applicant, so it must be used in conjunction with the MORTGAGE\_APPLICANT data.** (including the period at the end)
* Click on the **Submit** button.

#### MORTGAGE\_APPLICANT Table

1. Click on the **MORTGAGE\_APPLICANT** table.
2. At the table level, click the pencil icon next to the **Business terms** area in the **Overview** section.
3. Place your cursor in the area and start typing the name of following business terms. As you find them, **Select** them:

**Email Address** **Phone Number** **Social Security Number**

1. When you have added all 3, click on the **Apply** button.
2. At the column level, in the **Overview** section:

Click on the icon that looks like an **eye** for each column below and:

* Assign the **Email Address** business term to the **EMAIL\_ADDRESS** column
* Assign the **Phone Number** business term to the **PHONE\_NUMBER** column
* Assign the **Social Security Number** business term to the **SOCIAL\_SECURITY\_NUMBER** column

1. Go to the **Review** tab and do the following:

* Give it a rating of **5 Stars** by clicking on the 5th star to the right.
* Enter this text for the review text - **Contains governed and trusted data related to mortgage applicants to use for the mortgage default analytics project. It contains sensitive information but it is masked, with real data replaced with fictional by contextually correct data that will not affect analytical results.** (including the period at the end)
* Click on the **Submit** button.

1. Go to the **Profile tab** and do the following:

* **Make sure** the profile has been run. If it has not, click on the button to force a run of the profile and make sure it completes and that all columns have been assigned a data class.

This completes the Organize run book. You are **done!** and can now take the Organize experience flow to validate everything works. Log out as admin and login as the experience user.