**Meeting Notes** - Thursday, October 26, 2023 (Initial meeting)

**Attendees**

·         Ye, Gordon, Collom, Aaron, Connolly, Christian, Lane Jackson, Jayna Thanki

**Meeting Minutes**

* Aaron - Uploaded corrosion data into Geomart with help from Travis but does not have ECDA or other important CIS related data elements
* TIMP Risk table and spatial features [map]  is complete dataset on Oracle server - ESRI system database hosted on Oracle
* PODS schema is separate but similar and has asset data
  + CIS is inspection data [not asset data] for transmission pipelines
  + Transmission pipeline risk marked within certain distance on Route names - needs logic from business side to connect PODS Pipe table to CIS Risk data using Route Name column, which is common between PODS and ESRI GIS system
  + Old CIS data had been renamed so may not connect to new based on name values
* Potentially, may have CIS inspection data on distribution side but now transmission pipelines
* SharePoint site is possibility but involves working with OPS
* One time pull to get all historical data
* Add UXG0 [Garcia, Uriel] from Risk to meeting series
* What file type can foundry consume
  + Oracle preferred since CIS data is originally in Oracle system
* CIS data import to be done as a full import every time - replace past data with newly imported
* 2 tables generated by DA and Corrosion Engineering and has 1m-2m rows
  + Patrick Chu is Corrosion Engineer
* Get DA (Direct Assessment) (technic of assumption gathering, etc.) engineer - led by M3la [Manuel] but Manuel said Corrosion Engineer representation is fine for now

**Action Items**

* **Aaron** to send details of CIS Risk data Oracle connectivity
  + Oracle Server/Service Name, Database name, Schema name, Table name[s]

·        **Jayna** to send link to Ontology Design template for business review and input as applicable - done

o    [Ontology Object Template Gas TIMP CIS Risk Data.xlsx](https://pge.sharepoint.com/:x:/r/sites/ODGTeam/Shared%20Documents/Conceptual%20Model/Gas%20Ontology%20Federated%20Team/Gas%20TIMP%20CIS%20Read%20Public%20Ontology/Ontology%20Object%20Template%20Gas%20TIMP%20CIS%20Risk%20Data.xlsx?d=wced714e05d7c4a57b5f6ff5b7e23ebf3&csf=1&web=1&e=FNBep1)

·         **Jayna** to follow up within OIT team on options for CIS Risk data ingestion onto Foundry using Oracle source system

·        **Business** to complete necessary Foundry trainings to get familiar with Foundry system and use

o    Use below link to get guidance on Foundry access and various trainings that are available

·          <https://wiki.comp.pge.com/display/ARAD/Palantir+Foundry>

* **Business** to review BDS Onboarding PPT if useful. Provides information on CDE [Critical Data Element] identification process as well as many other data steward and SME related tasks as they apply to Ontology

·         [BDS Onboarding.pptx](https://pge.sharepoint.com/:p:/r/sites/dma/Shared%20Documents/DM%26A%20Asset%20Registry/PwC%20-%20Asset%20Registry%20Docs/Asset%20Registry%20CDE%20Inventory/Cross-team%20Collaboration/BDS%20Onboarding%20+%20Kickoff/BDS%20Onboarding.pptx?d=w324055c8973c43c48844a760de4e7a75&csf=1&web=1&e=JpbgZ6)

11/8/2023 Changes

**OIT team Ontology Lead change**

Please know **Marcella Kwan** from my [OIT] team will lead this effort as an Ontology Designer. Marcella and I to touch base directly as and when we need to step in for each other to ensure this project keeps moving forward.

**Business team changes**

Added **Uriel Garcia** and removed **Jason Klekcha** as requested during last meeting and changed series to once weekly to accommodate schedules.

Also, changed **Christian Connolly** as optional – he informed me via Teams chat that he will only join as and when questions or requirements input needed from **DA [Direct Assessment]** engineering team.

**11/9/2023 Meeting Notes**

* Kiana Go to join just to listen in
* CISRead contains same data but coming from two orgs - **ECDA/DA** [External Corrosion Direct Assessment] and **Corrosion engineering**
* Vendors capture the data and send it on to DA and Corrosion engineering who upload it on to their risk server
* Risk Server is an Oracle database, which may retire in near future
* **Aaron [Corrosion side]** copies that and sends it to Travis to upload it on to GMCloud, but it's an exploratory process - not a permanent solution as of now. Need Travis to help make this solution permanent and reliable
* **DA** does not do upload their version of CIS data to GMCloud
* **Aaron would like to consolidate DA and Corrosion engineering CIS data into one table** with a column like Organization to specify if data belongs to DA or Corrosion engineering ORG since structure for both is same and add some clean-up/additional transformation to it
* For **Corrosion Engineering dataset** - **Object id is the unique ID, Route and Line\_Num** (both may be duplicated or have same/similar data so need to clean up and probably use only one of these) is the common/business key between GIS and CIS data, but needs some clean up and manual work to ensure a match
* Route and Line\_Num has numerical values like 400
* There may be a reference table that describe what this numerical value means, but per Aaron, descriptions do not need to be included for Route/Line\_Num data to be meaningful or valuable to business in Foundry

Action Items

* Marcella to meet directly with Aaron on data mapping
* Aaron to meet with Travis to work out a plan to upload both ECDA and Corrosion engineering CIS datasets onto GMCloud - meeting set up at 2pm today - 11/9

**11/15/2023 Meeting Notes**

* Have identified the correct dataset in Geomart
  + Problems with units unification due to different vendors having different standardizations, e.g. Volt vs. mV.
* In addition to unit unification, some missing data cleanup should be done before ingesting into Foundry
* Linked shapefile should be in WGS84 (Aaron)
* Should fill in Long/Lat for all the data points
* How does business wants to work with the CISData set? Is there any augmentations to the final dataset?
  + Ask Steven Liu

Action Items

* Data cleaning:
  + Add the additional columns Aaron suggested:
    - “Organization”: whether data comes from DA or CorrosionEng
    - “GUID”: Unique ID for each data upload
    - Ask Travis to change the dataset values.
  + Unit standardization (V vs. mV)
    - Aaron’s group will re-upload to Geomart after making changes