



Data Virtualization Use Cases at SNL

Marquita Apodaca

Sandia National Laboratories

NLIT 2022

Albuquerque, NM

Introduction

Marquita Apodaca

I have a passion for all things relating to Enterprise Architecture and Data Architecture! I have been with Sandia for 12 years and overall experience within the Department of Energy including Los Alamos National Labs and National Nuclear Security Administration for the past 20 years in various roles in IT Project Management and Enterprise Architecture. I hold certifications in The Open Group Architecture Framework (TOGAF), Scaled Agile Framework (SAFe) Product Owner/Product Manager and am a Certified Scrum Master.

What is Data Virtualization

Data Virtualization allows business users to access enterprise data across individual systems regardless of the type of data source, combine any type of data, and consume the data in any mode.

Data Sources: data bases, data warehouses, cloud applications, excel files

Data Formats: data bases, web services, APIs, files

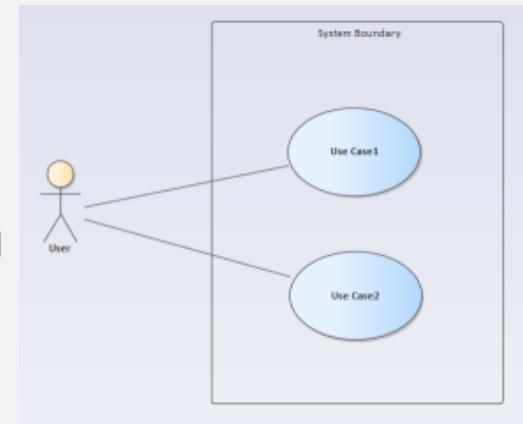
Tools Used

- Denodo – Data Virtualization Tool
- Microsoft SQL Server Management Studio (SSMS)
- Microsoft PowerBI
- Salesforce Tableau
- Atlassian Confluence
- Microsoft SharePoint

The Denodo logo consists of the word "denodo" in a bold, black, sans-serif font, followed by a red icon of a cluster of dots.The Tableau logo features a stylized icon of a cluster of dots in orange, yellow, and blue, followed by the word "tableau" in a blue, sans-serif font.The Microsoft logo is the four-colored square (red, green, blue, yellow) followed by the word "Microsoft" in a grey, sans-serif font.The Confluence logo features a blue icon of a stylized 'X' or 'C' shape, followed by the word "Confluence" in a blue, sans-serif font.

Use Cases

1. Creating API's from Data Sources in Denodo
 - Automating capturing evidence in our Solution Delivery Lifecycle (SDL)
Evidence sheets for Project Management information
2. Using Confluence Confiforms data as a Data Source in Denodo
 - Making Technical Business Management (TBM) Data accessible at an enterprise level, financial reports, TBM reporting
3. Identifying duplicate data across many sources
 - Project Management data and identifying a source of truth



Use Case 1 – Denodo API

Our Solutions Delivery Lifecycle (SDL) Process identified several areas to automate in Evidence Sheet Submissions

- Project Evaluation Application determines if a project is a Low, Medium or High Category
- The SDL Evidence sheet requires evidence of the Category

PRE Application

Project Evaluations

☐ All PrEs

☒ My PrEs

☐ All Revisions

☒ Current Revisions

CLEAR FILTERS

EDIT

DELETE

CLONE

REVISE

DOWNLOAD PDF

EXPORT SUMMARY

INITIATE REVIEW

REVIEW

UNLOCK

ADMIN CHANGE

PrE ID

Rev

Project Name

Current Status

Locked for Editing By

Approved Project Category

Portfolio

Program

Any

Any

Any

SDL Evidence Sheet

<div><div>Determine Project Evaluation Category (1.j)</div><div>Provide a link to PrE entry, project artifact, and/or location in another artifact where PrE Category result has been documented.</div><div>For Medium and High PrE Category solutions: if consideration of the PPDR creates artifacts either in the PPDR tool or a PPDR template, include link(s) to either here. (As described in Step 1.k)</div></div>
--

Confluence and Denodo API Use Case

- Creating Data Sources in Denodo from other systems Databases for easy access
- Join or Union data from data sources in Denodo to virtualize the data
- Easy to create a REST API to be consumed by other systems

PRE Data with Project Manager

Created by Marquita Apodaca, last modified just a moment ago


Locate your PRE Record and take a screen shot to be saved in this field:

!!!PASTE SCREEN SHOT HERE!!!

Search under *projectname* for your project. *sl_name* for the PrE record Project Manager or *sl_operid* for the Project Managers SNL User Account. You can clear @currentUser to show more users.

 approvedprojectcategory = pmname = projectname = status = userid = 

pmname	userid	projectname	status	approvedprojectcategory
Apodaca,Marquita	mapodac	Sandia Speed Test Epic Driven	APPROVED	LOW

 Like Be the first to like this

Reasons for Creating an API in Denodo

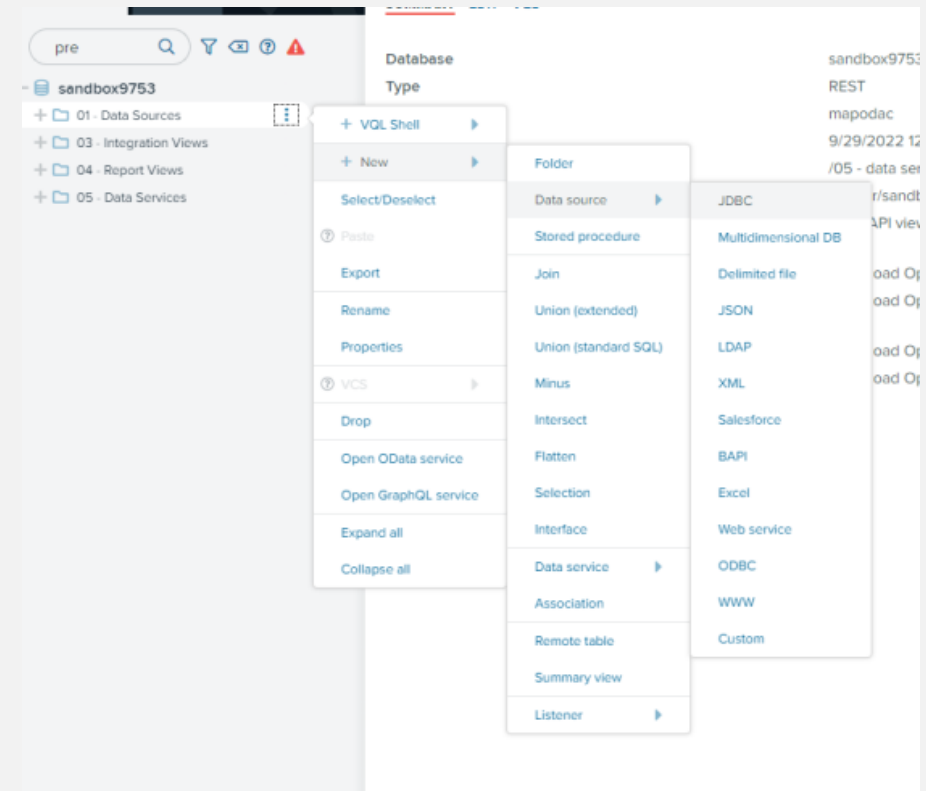
- Access to Data
- Ease of Use to create Denodo API
- No extra resources were needed to code an API from scratch
- API is reusable and can be easily modified
- Automation of data entry into SDL Evidence Sheet



How to Create a REST API in Denodo

Establish access to Data Sources in Denodo

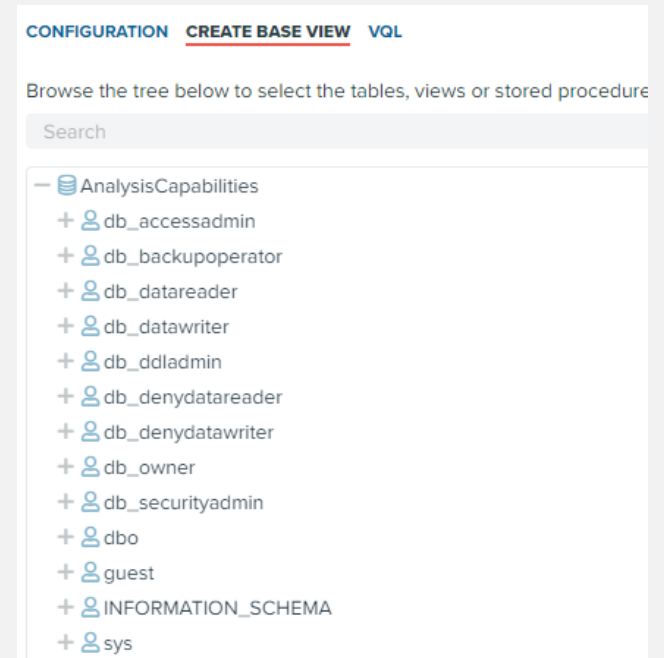
- For this API, both Data Sources needed were SQL Server Databases
- Create new JDBC Data Source



Create Base View from Data Source

Once you have established connection to your data source, you can create base views

- Standard practice is to name your base view: bv_(name of base view)
- Selecting tables in Denodo is simple using the GUI



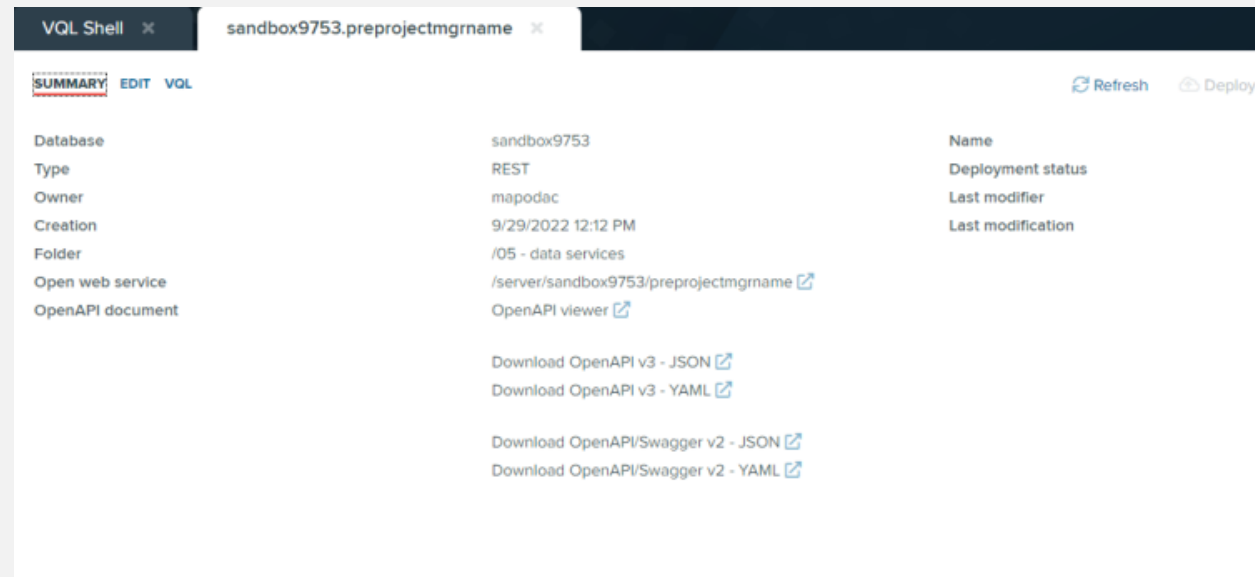
Create an Integration View

- Use Integration Views when you need to join data from multiple data sources
- Integration views allow you to create custom queries to get the data for your business needs
- In this case we needed to join our person data with the data from our Project Rigor Evaluation tool

Create a Report View and Rest API

Create a Report View from your Integration View

Once you have your report view, you can create a new Data Service (REST API)



Denodo API in Confluence

In this example I used the following Macros:

- Table from JSON Macro
- Table Filter Macro

PRE Data with Project Manager

Locate your PRE Record and take a screen shot to be saved in this field:

!!!PASTE SCREEN SHOT HERE!!!

Search under projectname for your project, sl_name for the PrE record Project Manager or sl_operid for the Project Managers SNL User Account. You can clear @currentUser to show more users.

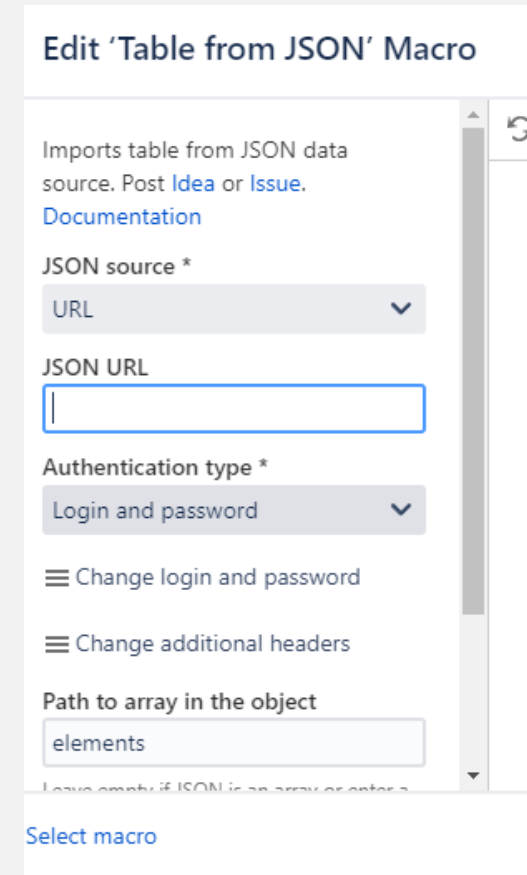
 **Table Filter** | column = approvedprojectcategory,pmname,projectname,status,userid | isOR = AND | separa...

 **Table from JSON** | url = https://rhtdvdevs01.s...

Table from JSON Macro

To use this macro you will need to have

- A URL with JSON Formatting
 - Denodo creates a Swagger link
- The array you are using, located in JSON URL
- Login Credentials
 - For testing I used my Sandia Credentials
 - Entity Account using metagroups



The screenshot shows the 'Edit 'Table from JSON' Macro' configuration window. It includes a description: 'Imports table from JSON data source. Post [Idea](#) or [Issue](#). [Documentation](#)'. The 'JSON source *' dropdown is set to 'URL'. The 'JSON URL' field is empty. The 'Authentication type *' dropdown is set to 'Login and password'. Below this are links for 'Change login and password' and 'Change additional headers'. The 'Path to array in the object' field contains 'elements'. A note at the bottom says 'Leave empty if JSON is an array or object'. At the bottom left is a 'Select macro' link.

Table Filter Macro

In order to filter the results, wrap the Table from JSON Macro with the Table Filter Macro

- Set filter Default values to return a specific subset of data
- Configure settings to hide columns and change column names



The Final Result

- The filter will automatically display results for the current user
- The results are searchable by Project Name, Username, or Project Manager

Determine Project Evaluation Category (1.j)

Provide a link to PrE entry, project artifact, and/or location in another artifact where PrE Category result has been documented.

For Medium and High PrE Category solutions: if consideration of the PPDR creates artifacts either in the PPDR tool or a PPDR template, include link(s) to either here. (As described in [Step 1.k](#))

Search under *projectname* for your project, *sL_name* for the PrE record Project Manager or *sL_operid* for the Project Managers SNL User Account. You can clear @currentUser to show more users.

approvedprojectcategory =

pmname =

projectname =

status =

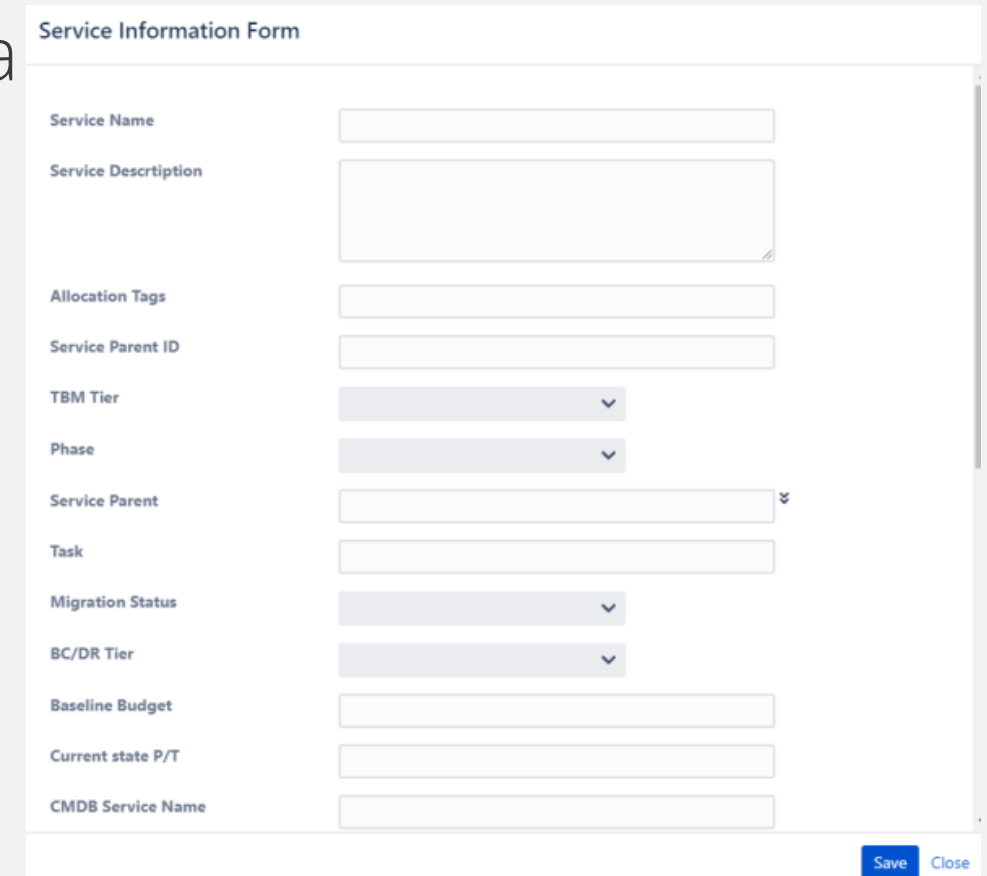
userid =

pmname	userid	projectname	status	approvedprojectcategory
Apodaca,Marquita			APPROVED	LOW

Use Case 2 – Confluence as a Data Source

In this example we used Confluence data through the Conforms Macro to create a data source in Denodo

- Conforms is a plug in for Confluence to create data entry forms



The screenshot displays a 'Service Information Form' with the following fields and controls:

- Service Name:** A single-line text input field.
- Service Description:** A multi-line text area.
- Allocation Tags:** A single-line text input field.
- Service Parent ID:** A single-line text input field.
- TBM Tier:** A dropdown menu with a downward arrow.
- Phase:** A dropdown menu with a downward arrow.
- Service Parent:** A single-line text input field with a double-check icon on the right.
- Task:** A single-line text input field.
- Migration Status:** A dropdown menu with a downward arrow.
- BC/DR Tier:** A dropdown menu with a downward arrow.
- Baseline Budget:** A single-line text input field.
- Current state P/T:** A single-line text input field.
- CMDB Service Name:** A single-line text input field.

At the bottom right of the form, there are two buttons: 'Save' (in blue) and 'Close' (in light blue).

Confiforms Features

- Data can be displayed in various formats including in a table like the one below
- A user can submit their responses by clicking on the “Create New Service Entry” button

Service information self service

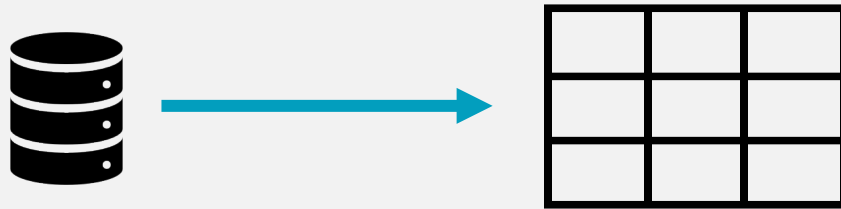
809668419

Create New Service Entry

Service Name	Service Description	Allocation Tags	Service Parent ID	TBM Tier	Phase	Service Parent	Task	Migration Status	BC/DR Tier	Baseline Budget	Current stage P/T
Stakeholder Relations											

Exporting Data out of Confluence Confiforms

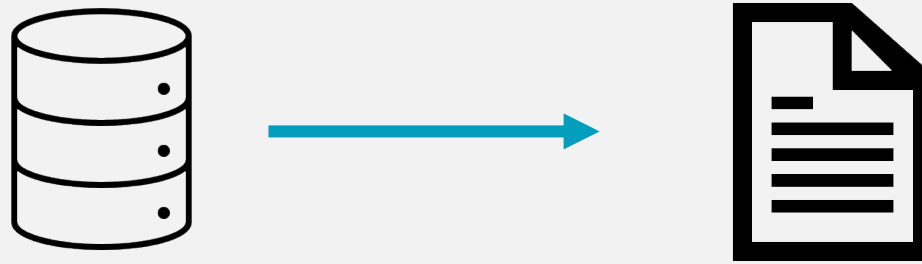
- Once your form is established and has collected data, you can easily export the data in several formats including JSON which is needed to create a REST API



Confluence as a Data Source Continued

Once you have your data exported in JSON format, you can create a Data Source in Denodo

- The Data Source can be used to create reports in reporting tools like Tableau or PowerBI
- Tableau and PowerBI do require installing an ODBC Connector in order to access the Data Source from Denodo



Use Case 3 – Identify Duplicate Data

Denodo is a good tool to do a comparison of data across an organization

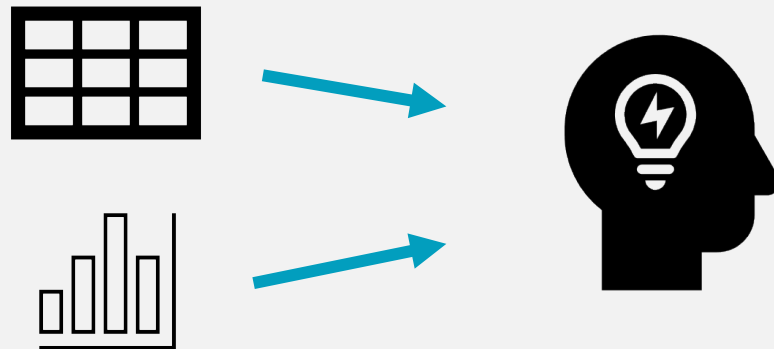
- You can create data sources from several platforms such as SQL, Oracle, Confluence, SharePoint, and CSV to name a few
- You can virtualize the data and run queries against other data sources within the Denodo tool



Analysis of Data

Although it is easy to run a query in Denodo, I find it useful to have Microsoft SQL Server Management Studio available to look at data in a particular database.

If the data source I am using is from Conforms, I will export the data into a CSV file or excel file.

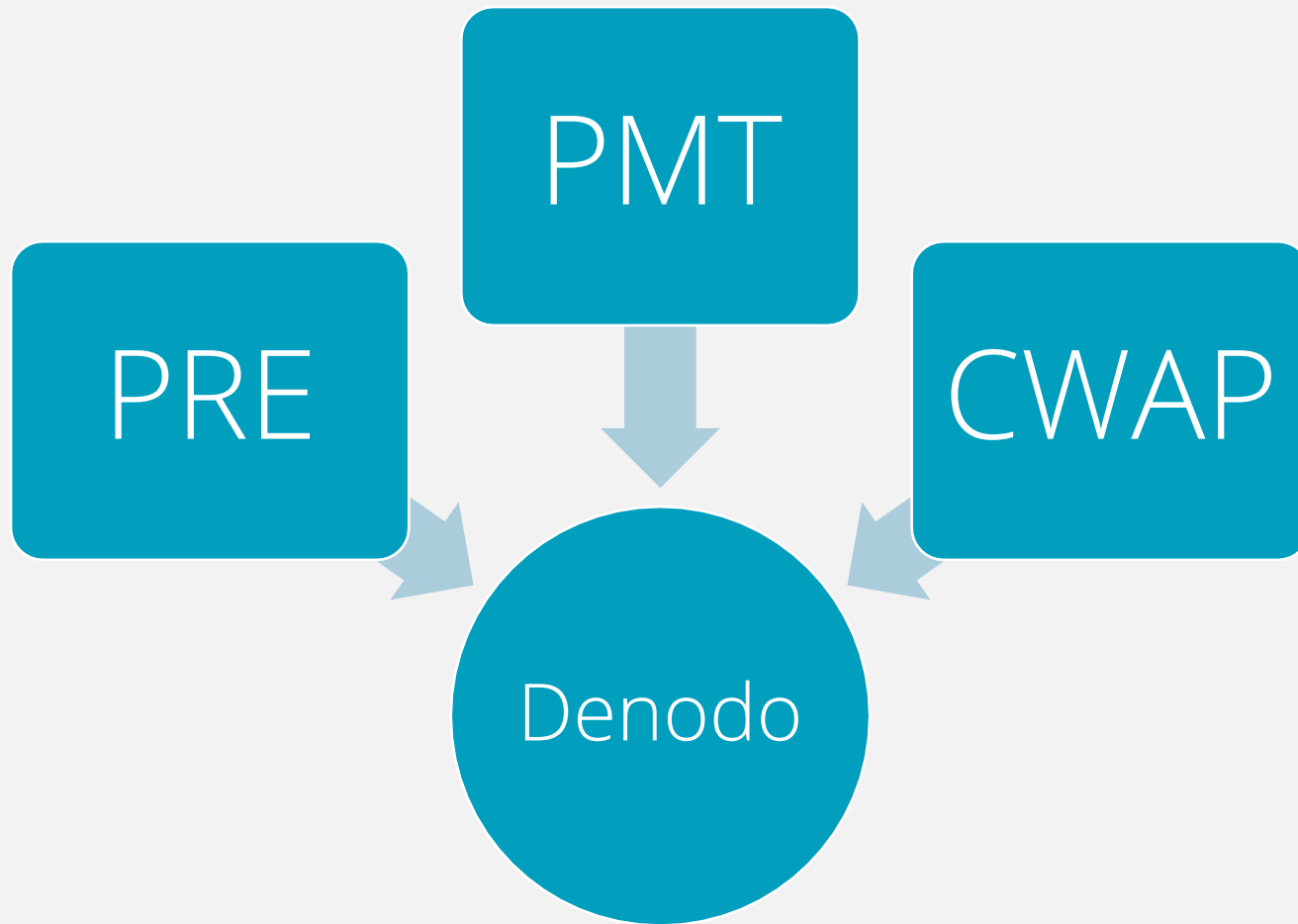


Project Management Data and Systems

Project Management data is used at an Enterprise level to determine cost, scheduling, and status

- There are several enterprise systems at Sandia that track similar data for project management
- We choose 3 systems
 - Center Work & Authorization Process (CWAP) – Project Authorization
 - Project Evaluation (PRE) – Categorize Projects into High, Medium, or Low
 - Portfolio Management Tool (PMT) – Track Financials

Common Project Management Data Fields



- Project Manager Name
- Project Name
- PRE
- Project UID

Reuse of PM Data

Once we establish a source of truth for Project Management Data, we can reuse it across other PM systems. This would help:

- Automate data entry
- Eliminate duplicate data
- Reduce human error in data entry
- Validate data

Questions?