# Centralized Cruise Database - Technical Documentation

## Overview:

The Centralized Cruise Database is used to track information about each PIFSC research cruise including activities, chief scientists, regions, etc. to remove the need for each division/program to manage this information. This centralized database is available for all PIFSC database users to reference with their various division data sets. The Cruise Database also references non-confidential personnel information from the PEEPS database.

## Resources:

* Centralized Cruise Database Version Control Information:
  + URL: [git@pichub.pifsc.gov:centralized-data-tools/centralized-cruise-database.git](mailto:git@pichub.pifsc.gov:centralized-data-tools/centralized-cruise-database.git)
  + Database: 0.11 (Git tag: cen\_cruise\_db\_v0.11)
* [Comments for all centralized cruise database Views](centralized_cruise_DB_view_comments.xlsx)
* [Cruise Data Management Application Technical Documentation](../CRDMA/docs/Cruise%20Data%20Management%20Application%20-%20Technical%20Documentation.docx)
* [Cruise Database Diagram Documentation](Centralized%20Cruise%20Database%20Diagram%20Documentation.docx)
* [Database Naming Conventions](Centralized%20Cruise%20Database%20-%20DB%20Naming%20Conventions.docx)
* [PL/SQL Coding Conventions](Centralized%20Cruise%20Database%20-%20PLSQL%20Coding%20Conventions.docx)
* [Data Integration SOP](Centralized%20Cruise%20Database%20-%20Data%20Integration%20SOP.docx)
* CCD\_CRUISE\_PKG Package (documentation is defined in the package specification and package body)
  + [Testing Documentation](test%20cases/CCD_CRUISE_PKG/CCD_CRUISE_PKG%20Package%20Testing%20Documentation.docx)
* Data Validation Module (DVM)
  + [Testing Documentation](test%20cases/DVM_PKG/Centralized%20Cruise%20Database%20DVM%20Testing%20Documentation.docx)

## Database Setup:

* Grant the Cruise database schema permissions
  + Execute the sections of the [grant\_info.sql](../SQL/queries/grant_info.sql) file using the schemas based on the comments to grant the CEN\_CRUISE schema the necessary permissions
  + \*Note: the permissions granted to the CEN\_CRUISE schema are listed in [CEN\_CRUISE\_permissions](CEN_CRUISE_permissions.xlsx)
* [Installing or Upgrading the Database](Installing%20or%20Upgrading%20the%20Database.docx)
* Cruise/reference data can be purged and reloaded for development purposes using [refresh\_ref\_data.sql](../SQL/queries/refresh_ref_data.sql)
* Grant external schemas permissions to the Centralized Cruise Database
  + Modify the Centralized Cruise Database's [grant\_external\_schema\_privs.sql](../SQL/queries/grant_external_schema_privs.sql) to replace the [EXTERNAL SCHEMA] placeholders with the given schema name and execute using the CEN\_CRUISE schema

## Features:

* The database requires the Centralized Utilities to be deployed on the CEN\_UTILS schema in order for the database views to work properly when querying the cruise and cruise leg information
  + Version Control Information:
    - URL (Git): [git@pichub.pifsc.gov:application-development/centralized-tools.git](mailto:git@pichub.pifsc.gov:application-development/centralized-tools.git) in the centralized\_utilities folder
    - Application: 0.5 (Git tag: cen\_utils\_v0.5)
    - Database: 0.8 (Git tag: cen\_utils\_db\_v0.8)
* The Data Validation Module (DVM) is used to perform QC validation on the Centralized Cruise Database data managed in this database. Custom data validation criteria were developed for this operational data set.
  + Version Control Information:
    - URL (Git): [git@pichub.pifsc.gov:centralized-data-tools/data-validation-module.git](mailto:git@pichub.pifsc.gov:centralized-data-tools/data-validation-module.git)
    - Application: 1.2 (Git tag: DVM\_v1.2)
    - Database: 0.3 (Git tag: DVM\_db\_v0.3)
* The Database Version Control Module is used to track the database version installed on a given database schema.
  + Version Control Information:
    - URL (Git): [git@pichub.pifsc.gov:application-development/centralized-tools.git](mailto:git@pichub.pifsc.gov:application-development/centralized-tools.git) in the DB\_version\_control folder
    - Application: 0.11 (Git tag: db\_vers\_ctrl\_v0.11)
    - Database: 0.2 (Git tag: db\_vers\_ctrl\_db\_v0.2)
* The Database Logging Module is used to log CDIM execution information in the database.
  + Version Control Information:
    - URL (Git): [git@pichub.pifsc.gov:application-development/centralized-tools.git](mailto:git@pichub.pifsc.gov:application-development/centralized-tools.git) in the DB\_log folder
    - Application: 0.1 (Git tag: db\_log\_v0.1)
    - Database: 0.1 (Git tag: db\_log\_db\_v0.1)
* The Authorization Application Module was originally designed to manage application access and permissions within the application. This is a flexible method that allows users and permission groups to be defined that will determine if a user has enabled access to the application and what permission(s) they have in the application.
  + Version Control Information:
    - URL (Git): git@pichub.pifsc.gov:application-development/centralized-tools.git in the auth\_app folder
    - Application: 0.1 (Git tag: auth\_app\_v0.1)
    - Database: 0.7 (Git tag: auth\_app\_db\_v0.7)
* The PIFSC APEX custom error handling function has been implemented on the application to suppress sensitive error information within the database application to satisfy Security Control: SI-11.
  + Version Control Information:
    - URL: [git@pichub.pifsc.gov:application-development/apex\_tools.git](mailto:git@pichub.pifsc.gov:application-development/apex_tools.git) in the "Error Handling" folder
    - Application: 1.1 (Git tag: APX\_Cust\_Err\_Handler\_v1.1)
    - Database: 0.2 (Git tag: APX\_Cust\_Err\_Handler\_db\_v0.2)

## Data Flow:

* [Data Flow Diagram (DFD)](DFD/Centralized%20Cruise%20DFD.pdf)
* [DFD Documentation](DFD/Centralized%20Cruise%20Data%20Flow%20Diagram%20Documentation.docx)

## Business Rules:

* The business rules for the Centralized CTD Database are defined in the [Business Rule Documentation](Centralized%20Cruise%20Database%20-%20Business%20Rule%20Documentation.docx) and each specific business rule listed in the [Business Rule List](Centralized%20Cruise%20Database%20-%20Business%20Rule%20List.xlsx) with a Scope of "Cruise DB" apply to the underlying database and rules with a Scope of "Data QC" apply to the QC criteria used to evaluate Cruise data in the underlying database.

## DVM QC Validation Criteria:

* Technical documentation for the DVM can be found in the DVM git repository in the docs\Data Validation Module.docx document and instructions for how to define the QC criteria can be found in the How to Define Criteria in Data Validation Module.docx document.
* The specific QC criteria developed for the Centralized CTD database can be found in the Centralized CTD repository in the docs\QA\_QC Validation Criteria.xlsx which is exported into SQL\queries\load\_DVM\_rules.sql
* To view a specific cruise's associated QC validation issues connect to the Oracle Enterprise database and query the CTD\_FILE\_CAST\_ERR\_RPT\_V and specify the value of :cruise\_name which is a valid cruise alias for a given research cruise (e.g. SE-15-01):
  + SELECT \* FROM CENTRAL\_CTD.CTD\_FILE\_CAST\_ERR\_RPT\_V WHERE regexp\_like (CRUISE\_ALIASES\_DELIM, '^'||:cruise\_name||',', 'i') OR regexp\_like (CRUISE\_ALIASES\_DELIM, ', '||:cruise\_name||'$', 'i') OR regexp\_like (CRUISE\_ALIASES\_DELIM, ', '||:cruise\_name||', ', 'i');
* Annotating QC Validation Issues:
  + TBD (An APEX application will be developed for viewing/annotating errors with appropriate security controls and auditing features)

## Cruise Database Reference Data:

* Cruise Legs and Cruise Leg Aliases:
  + [Cruise Leg Name Alias Documentation](Cruise%20Leg%20Name%20Alias%20Documentation.docx)
    - There is no limit on the number of cruise leg aliases that can be defined for a given cruise leg
    - The information for the defined cruise leg aliases can be viewed by querying the CEN\_CRUISE.CCD\_CRUISE\_LEG\_DELIM\_V view
  + <Cruise_Leg_DDL_DML_generator.xlsx> contains sheets labeled "Cruises", "Cruise Legs", "Cruise Leg Aliases" that defines the cruises (CCD\_CRUISES), cruise legs (CCD\_CRUISE\_LEGS), and cruise leg aliases (CCD\_LEG\_ALIASES) respectively for each research cruise defined in the Centralized Cruise database. The DML to load this reference data is generated in labeled columns.
    - These DML statements can be exported to a DML file so these values can be easily loaded into a given database schema.
  + An [APEX application](../CRDMA/docs/Cruise%20Data%20Management%20Application%20-%20Technical%20Documentation.docx) is available to manage the cruises, cruise legs, and leg aliases