# Cruise Project Database

# To

# MAR-DID

# Requirements

# Background

The cruise project database has been in use since the early 2000s. It is a ‘product’ of what were originally two separate MS Access databases, each with a simple entry form. These databases were not well designed and the merger and migration of the data into Oracle left a poorly designed relational database. Efforts in the past were done in an attempt to clean the data to assist in being able to put proper constraints in place, however conflicting priorities left this work unfinished.

During the summer of 2025 a volunteer was brought in to clean the data.

Without a GUI to facilitate easy data entry, deletion, updates and searches the archivist is forced to perform these operations in Oracle using SQL Developer. Leaving very few guard rails in places to maintain data integrity.

# Database

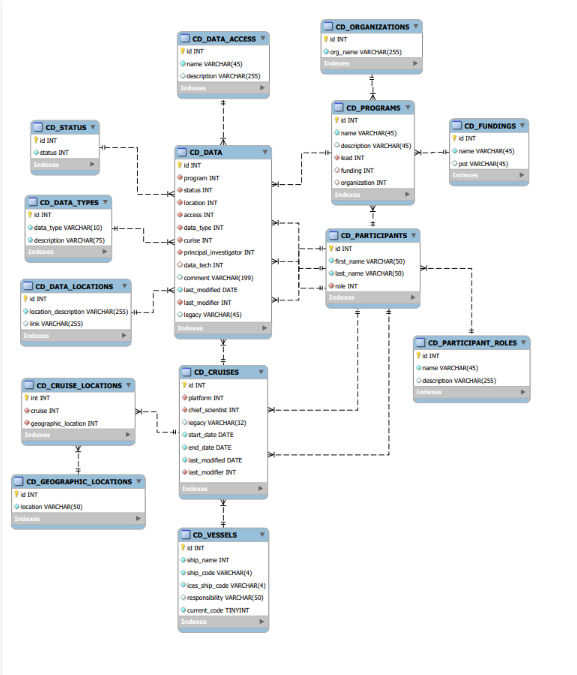
The current database is not properly constrained, that is to say that there are many instances where the value of a column (ex platform) should have a foreign key to a lookup table. This has resulted in issues for searching, for example the platform CCG Alfred Needler could be spelled in a myriad of ways.

Along with improving the database to allow for better data integrity some additional columns would be beneficial. For example Project and Project Funding.

Given past attempts at ‘cleaning’ the data the account has become polluted with unused objects, and possibly legacy objects which fell by the wayside. Given these issues it is advisable to migrate to a new account, MARDID will be created on PTRAN at BIO.

The migration of the current database to a properly constrained database meeting the new requirements will take approximately 2-7 days. This includes the possibility of the discovery of currently unknown data issues requiring remediation on the part of the Datashop.

The current redesigned database:



Note that it is anticipated that there may be changes required as the project moves forward.

# Front End

There should be a front end developed for the database, the current plan is to use Django which will enable for quick development and roll out. The github for this project is <http://ci-bio-mar-did-1:8000/>

There are 2 requirements for the GUI.

1. An entry/update/delete form which will enable managing the metadata without the need for knowledge of the data base schema.
2. A data discovery/search form for end users to access and find what they might be looking for in the data base.

The estimate for the development of the GUI tool is approximately ??????.

Testing of the tool will be performed by Datashop personnel, primarily Shelley and Jeff. 5-10 days.

# Possible Future Enhancement

The system should be designed with the idea of possibly allowing for a new module, one which would allow the end user to down load the files they are seeking. Either from the flat file archive [\\dcnsbiona01a\ARC\Archive](file:///\\dcnsbiona01a\ARC\Archive) or from ODF\_Archive on PTRAN.