

# Ocean Ledger Azure Blob Outline v0.0.1

---

## Bronze

Raw data directory

- Landsat

- .tiff, JSON, .csv
- The JSON and .csv are metadata that must be collocated with the data, for the time being.

- Sentinel\_2

- .tiff, JSON, .csv
- The JSON and .csv are metadata that must be collocated with the data, for the time being.

- ICESat-2

- Netcdf4 or hdf5

## Silver

Data that has been pre-processed to some degree and feeds into further workflows/processing/analytics

- Landsat

- .tiff, .csv
- The .csv are metadata that must be collocated with the data, for the time being.

- Sentinel\_2

- .tiff, .csv
- The .csv are metadata that must be collocated with the data, for the time being.

- Bathymetry\_reference

- Apache Parquet file (not geoparquet at the moment)

## Gold

Data that is “customer ready” and can be a stand alone product, but also can be further processed into “tertiary” (e.g. platinum) products

Data formats:

- .tiff
- .gpkg
- Parquet
- Geoparquet

## Platinum

Not often used, but typically abstracted summaries or hotspot type analyses of Gold level products.

Data formats:

- .tiff
- .gpkg
- Parquet
- Geoparquet

## Reference

Data that are external, third party, established reference datasets, standard type datasets like DEMs, framework files like config .yaml, and AOIs polygons.

Example can be the global mangrove watch .gpkg or the GLO30 Copernicus DEM

Data formats:

- .tiff
- .gpkg
- Parquet
- Geoparquet

## Diagnostics\_Meta

Directory for storing diagnostics of model runs, model objects, and other similar. Will likely have subdirectories for further organising, but no more than 1 or 2 sub levels.

- |           |         |
|-----------|---------|
| - Parquet | - .RDS  |
| - .csv    | - .pkl  |
| - .gpkg   | - .yaml |
| - .tiff   |         |