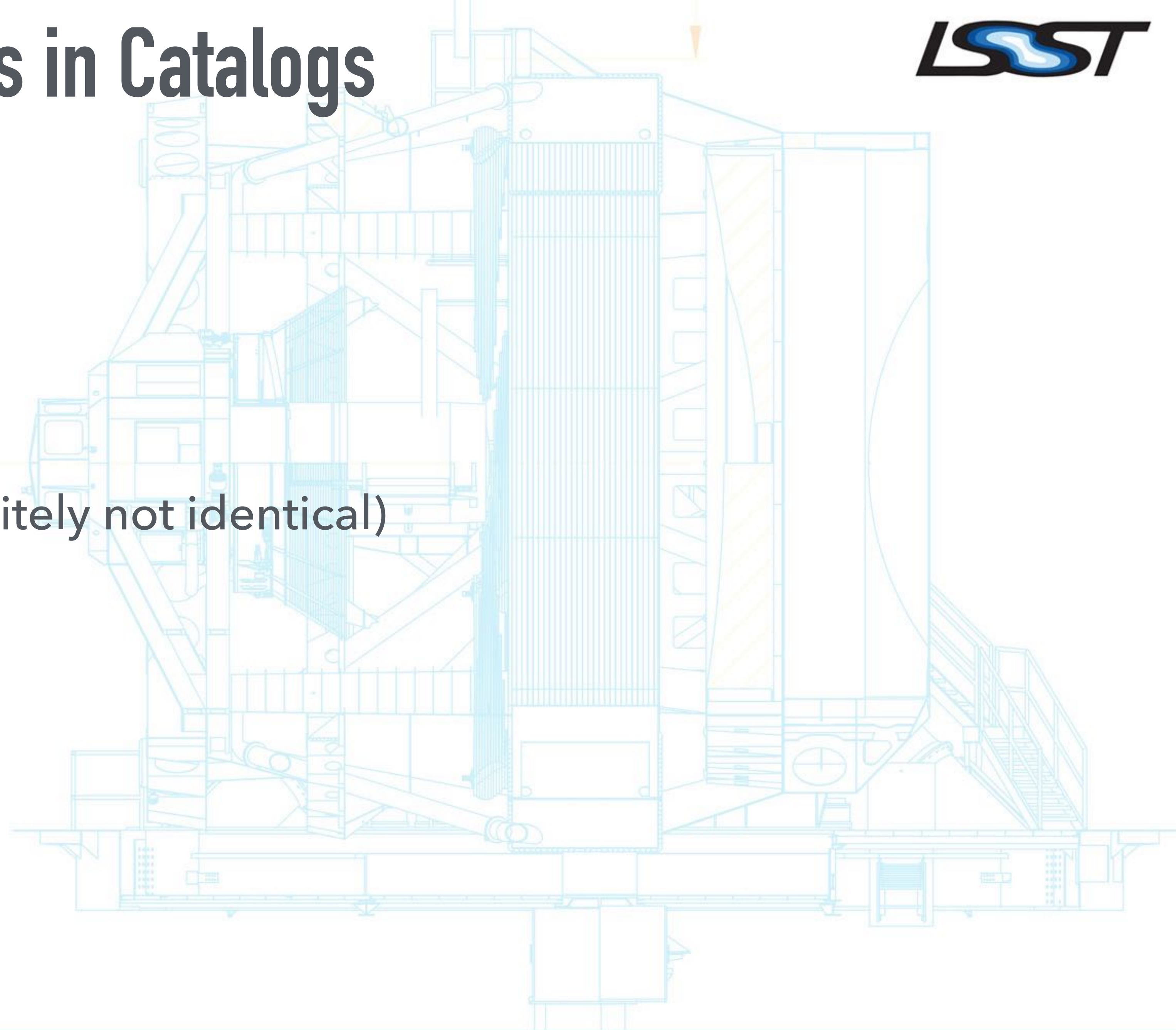# Metadata

Kian-Tat Lim

LSST

Large Synoptic Survey Telescope

Pipelines Produce Values in Catalogs

- Computed by algorithms

- Needed for Science Validation

- Compatible with DPDD

(These three overlap but are definitely not identical)

Values Must Be Documented

- Columns (not varying from row to row) with user-understandable names

- Short descriptions

- References to longer descriptions/algorithms/DPDD/"help text"

- Units

- UCDs

- Relationships (RA to dec, both to errors)

- More?

(Also, need a single reference position for an object, not one per filter)

Metadata Propagation

- Metadata goes from pipeline outputs to database tables during ingest

- Metadata goes from database tables to VO clients by TAP query of meta-tables

- Metadata is also packaged with query results by the TAP server

  - Metadata for derived columns is much more limited
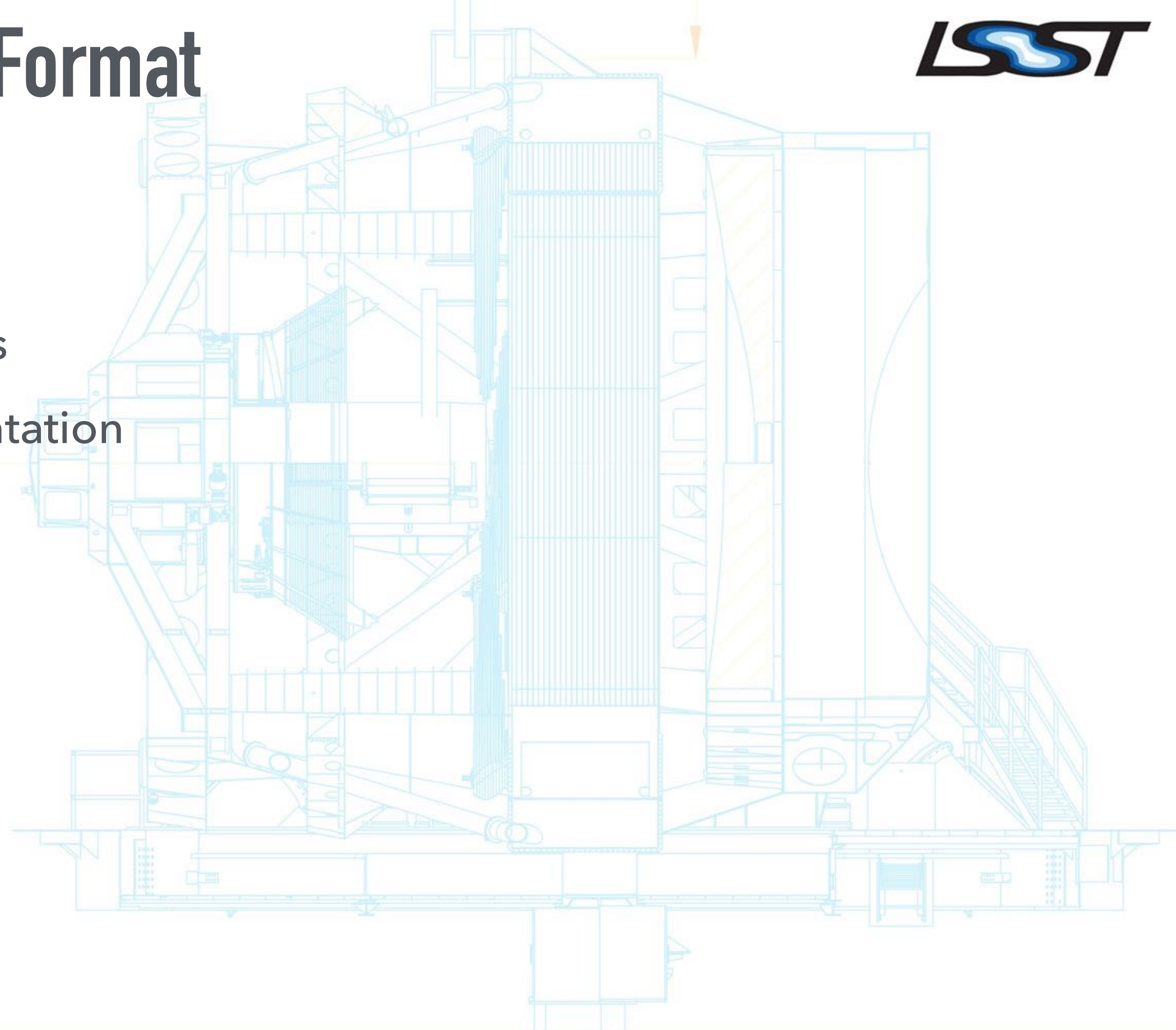
Question 1: Metadata Definition

Options:

- Capture metadata at point in code where column is created

- Allow columns to be created without metadata but require external metadata

Question 2: Persistence Format

Options:

- Embedded metadata

- Link to external description files

- Link to external prose documentation

Question 3: Database Ingest

Options:

- Load metadata once

- Load metadata once, check against inputs

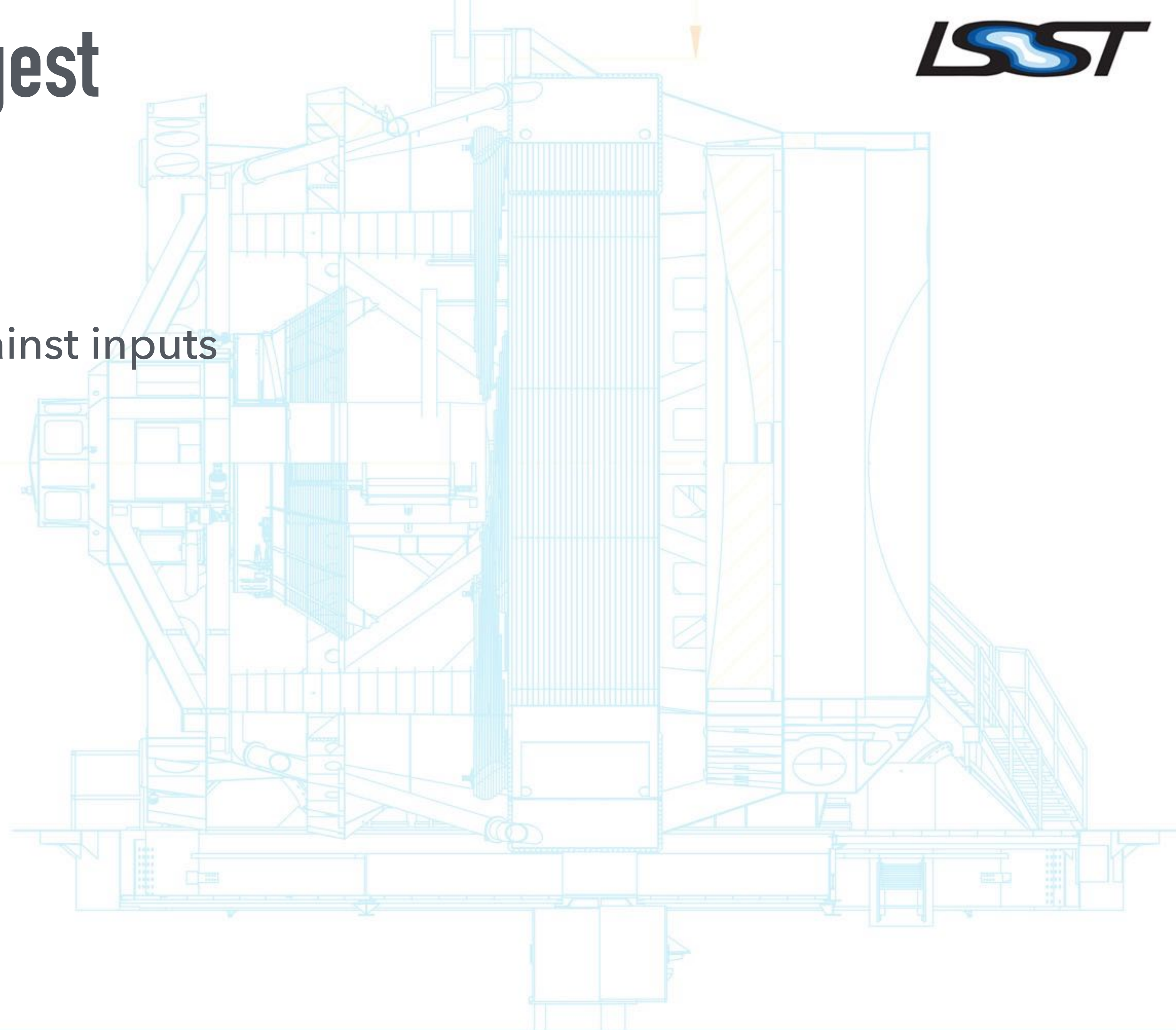- Load metadata with each input

Image Metadata

- DPDD specifies very little; no explicit tables

- Need data model (CAOM agreed, but not formally RFCed; ObsCore also for SIA)

- Instead, `ExposureInfo` and `VisitInfo` have been evolving independently

- Neither are ordinarily ingested into table in nominal schema

- WCS, PSF, and calibration information are needed too

- Interactions with the Gen2/Gen3 Butler Registry tables

DMTN-067: Catalog Data Model

- Written to capture agreements at the end of the LSP Workshop in December 2017

- Virtually no movement has occurred

- Need to turn statements into activities

DMTN–067 Catalog Meta–Tables

Activities:

- Generate meta-tables for pipeline outputs

- Add UCDs and relationships to lsst.afw.table schema interface

- Audit column metadata

- Load CI run catalogs and meta-tables into SQLite database
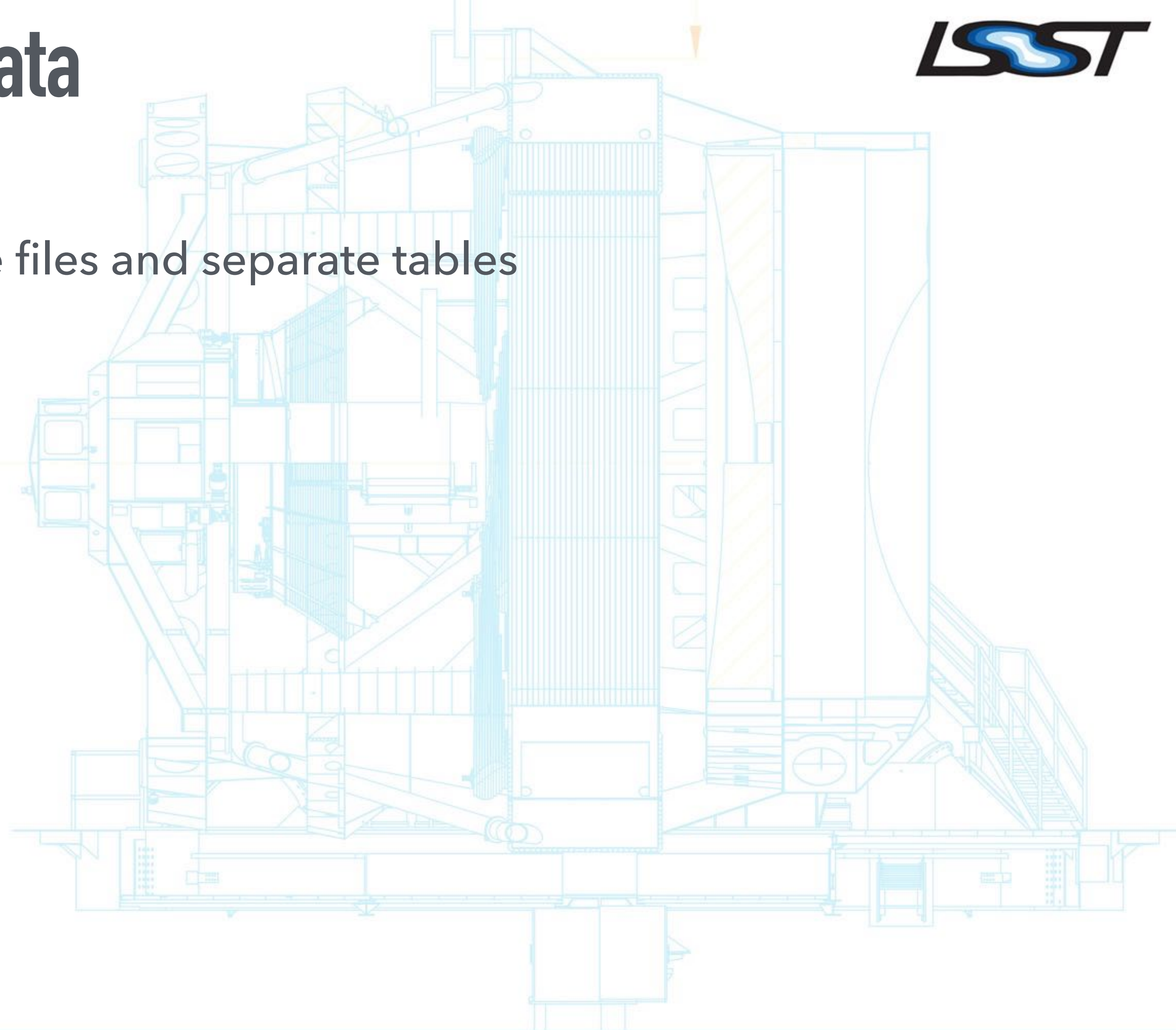
DMTN–067 DPDD Catalogs

Activities:

- Generate YAML files for DPDD-required outputs

- Generate DPDD tables from YAML files

- Begin testing pipeline outputs against YAML files (which will fail)

- Create verification task to ensure persisted format for each DPDD output

DMTN-067 Image Metadata

Activity:

- Load visit metadata from image files and separate tables

DMTN-067 Research

Activity:

- Continue to monitor VODML