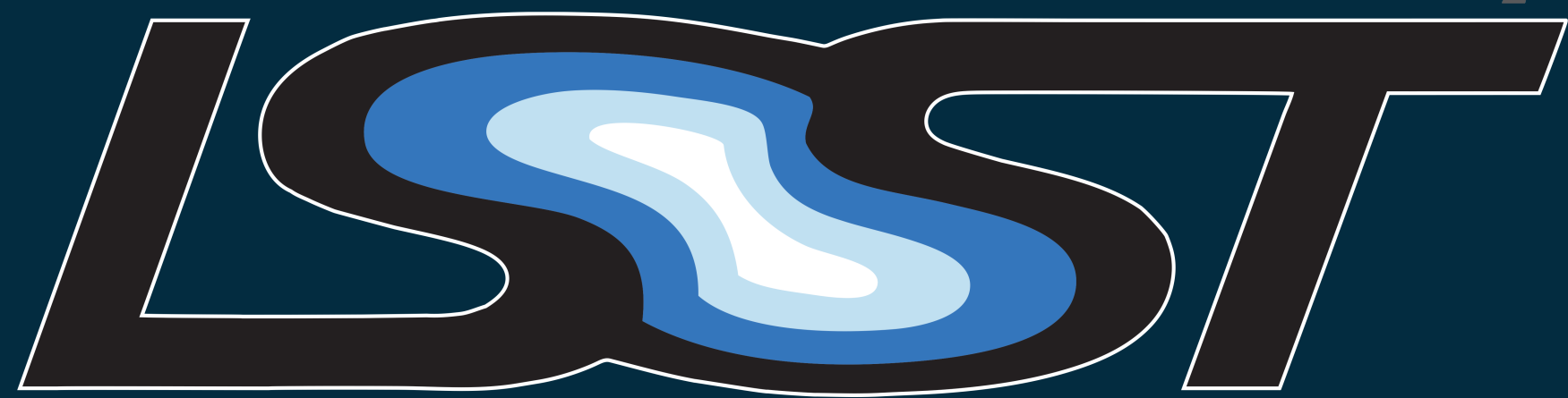
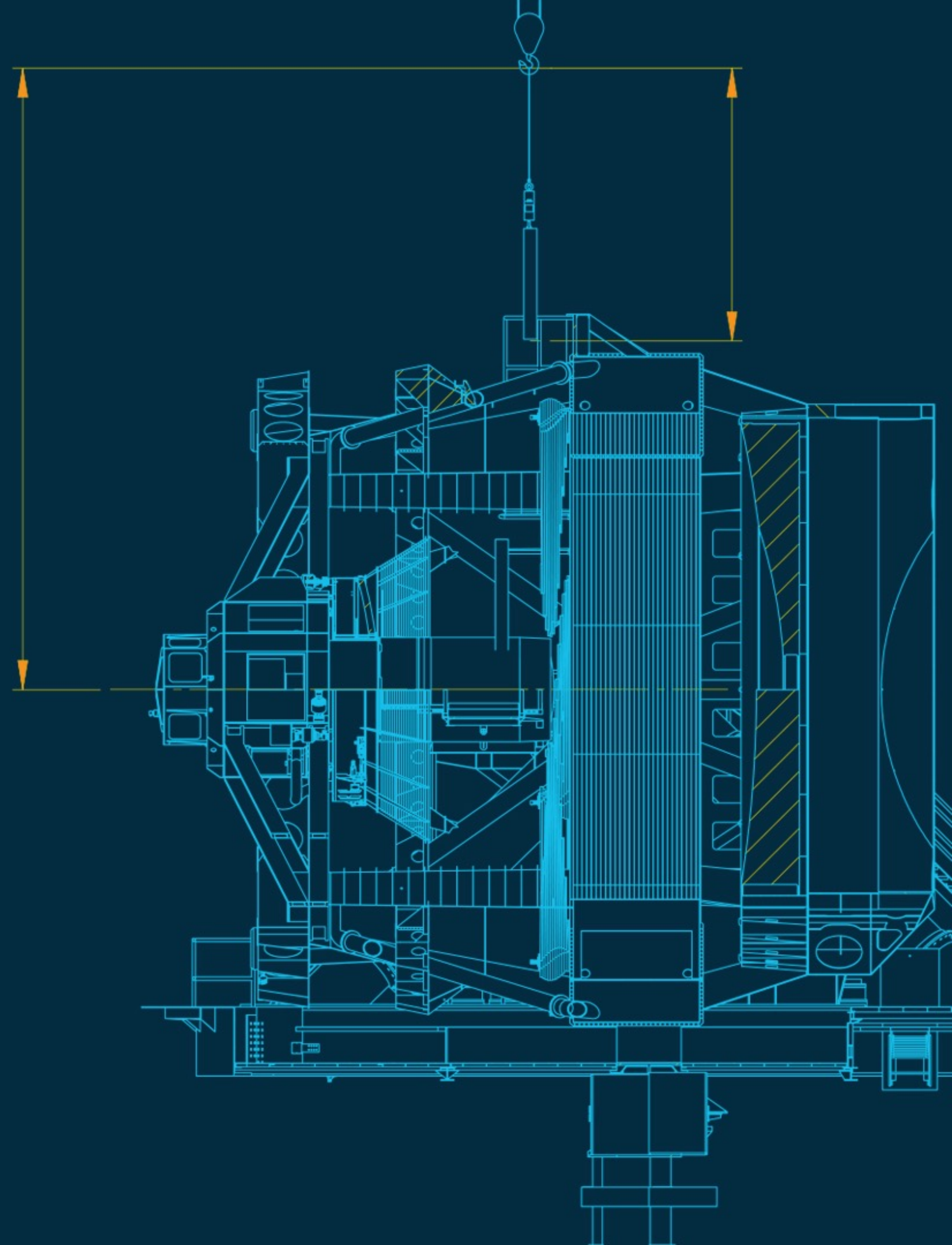


# How do we manage calibration products?

Rubin Observatory



*Legacy Survey of Space and Time*





## What do we mean by calibration products?

- Data products used by the Science Pipelines to perform ISR/calibration on science data.
- Some are image data: biases, darks, flats, etc.
- Some are not: defect lists, crosstalk matrices, etc.
- Some are outputs of "cp\_pipe" (the calibration products pipeline) based on data taken on the mountain; some come from other sources (e.g. lab measurements).



## What's the current status?

- Some calibration products are versioned in Git repositories.
- See, for example, RFC-595.
- Some usually live on filesystems.
- Generally, things which “look like images”.
- Presumably because they are too big for Git(?).
- There has to be some logic involved in choosing what calibration product to use when processing data – has it been sanity-checked? Is it applicable over the time range of the observation?
- I'm aware of a bunch of discussions about “blessing” products (please, let's not use that word) and validity ranges, but not of a coherent policy (am I missing something?)



## What's coming in the future?

- Does everything live on the Data Backbone / in the Butler?
- If so, do the Git repositories vanish? How do we manage a transition period?
- What capabilities do we need for ingesting / managing / versioning products that don't come from cp\_pipe?
- What capabilities do we need in terms of manual checking or editing of calibration products, whatever their origin?
- Before we can make sensible plans, we need to understand the big-picture workflow.



## Why is John worried?

- As far as I can establish, there is no documented plan for how all this fits together.
- There are a whole bunch of implicit requirements of the form “I was told that such-and-such a product needs to be editable with a text editor”, which are neither written down anywhere nor have clear provenance.
- Talking to  $N$  folks who know something about the topic, I'm pretty sure there are  $N+1$  ways in which this “obviously” has to work.
- Not clear who is responsible for figuring out what actually happens during operations. Architecture? Science Pipelines? Data Facility? The Ops Team?
- Working groups take a bunch of time...!
- There's some intermediate state between now and ops, and we're not ready for it...



How do we manage calibration products?

## What next?

- Is this all written down somewhere I've not spotted?
- Do we need a working group?
  - If so: who's on it?
- Does it come out of operations rehearsals?
  - Or planning for ops rehearsals?

