

# 2021-06-01 Science Metric Development Agenda and Meeting notes

## Date

01 Jun 2021





## Attendees

- [Keith Bechtol](#)
- [Simon Krughoff](#)
- [Colin Slater](#)
- [Jeffrey Carlin](#)

## Goals

## Discussion items

Time	Who	Pre-meeting Notes	Notes and Action Items
News, announcements and task review			
CMR		Current draft of <a href="#">DMTR-311</a>	<ul style="list-style-type: none"><li>• Jeff made a ticket to use update columns and has BPS job running now that should finish soon (~30 min). The change was made in the default rather than the yaml configuration.</li><li>• Should we be using "slot" designations in general? Depends. One does have to be careful where the slots are pointing.</li><li>• Where does documentation exist for the columns? Not an immediate answer. We expect several of the names/columns to change.</li><li>• The ellipticity residuals were the big change in reported metric values.</li><li>• Need to specify high memory (40 G) when running the BPS jobs to compute metric on full tract.</li><li>• Even with largest container in RSP, cannot load full tract catalog from object table into memory.<ul style="list-style-type: none"><li>◦ Getting the schema is a good ticket for middleware. Should be able to retrieve the schema.</li><li>◦ Check to see if there is composite dataset type for the schema</li></ul></li><li>• faro should be transitioning to use object table and loading a subset of columns</li><li>• Need to figure out how to specify the subset of columns to load in the connections class.<ul style="list-style-type: none"><li>◦ Might have to make a runQuantum function to do this, though it would be nice if could specify as argument in the input in connections class</li></ul></li><li>• Considering to use w18 instead of w14 for v22 pipelines release<ul style="list-style-type: none"><li>◦ One issue that would need to be backported, but might change conda dependencies. Release management reasons.</li><li>◦ There is a w18 RC2. Should we run faro on w18, Monika included. Jeff would need to re-run TE1 and TE2. Then would have w18 results in hand.</li><li>◦ w22 processing in underway</li></ul></li></ul>
Status of reprocessing		<ul style="list-style-type: none"><li>• w_2021_14 (RC2)</li><li>• w_2021_16 (DC2 gen3) (thanks Dan)</li></ul>	<ul style="list-style-type: none"><li>• Now running faro with BPS, need documentation</li></ul>

Bugs, issues, etc of the week		<ul style="list-style-type: none"> <li>Apparent TEx performance degradation in bands other than i</li> </ul> <div>  <a href="#">DM-30359</a> - Jira project doesn't exist or you don't have permission to view it. </div>	
Development status		<p>Review progress on the current development epic</p> <div>  <a href="#">DM-26993</a> - Jira project doesn't exist or you don't have permission to view it. </div> <p>Backlog epic:</p> <div>  <a href="#">DM-29525</a> - Jira project doesn't exist or you don't have permission to view it. </div>	<ul style="list-style-type: none"> <li>Spring 2021 epic was intended to be finished 31 May 2021</li> <li>Next week could be an opportunity to do a comprehensive review of current epic and make priorities for the next several months <ul style="list-style-type: none"> <li>One high-priority item would be using object table and source table (parquet format)</li> </ul> </li> </ul>
AOB		<ul style="list-style-type: none"> <li>Reference catalogs interface development: <div>  <a href="#">DM-30054</a> - Jira project doesn't exist or you don't have permission to view it. </div> <ul style="list-style-type: none"> <li><a href="https://github.com/lst/pipe_tasks/blob/tickets/DM-30054/python/lst/pipe/tasks/loadReferenceCatalog.py">https://github.com/lst/pipe_tasks/blob/tickets/DM-30054/python/lst/pipe/tasks/loadReferenceCatalog.py</a></li> </ul> </li> <li>What feedback would we like to receive from science pipeline developers from faro presentation on Tuesday 2 June? <ul style="list-style-type: none"> <li>What metrics do we expect pipelines to be producing in situ?</li> <li>Visualization</li> <li>Dedicated data structure to hold metric values in butler repo?</li> <li>What inputs should be used for metric calculation (e. g., parquet tables vs. direct outputs of pipelines)?</li> <li>External reference catalogs <ul style="list-style-type: none"> <li>Adding reference catalogs</li> </ul> </li> <li>External calibrations</li> <li>Spatial scales for computing metrics (e.g., patch/tract versus other spatial partitioning scheme)</li> <li>Communication between faro development and science pipeline development</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>Simon is updating validation_data_hsc CI dataset. Need a Gen3 dataset to use for getting started. Central 6 sensors from something like 10 visits. Calibrations and reference catalogs.</li> <li>For science pipelines meeting, <ul style="list-style-type: none"> <li>Not necessarily trying to necessarily propose solutions, or get into design discussions, but try to find out who to connect with in DM to make progress</li> <li>"What can faro do for you?"</li> </ul> </li> </ul>

## Action items

