2021-07-27 Science Metric Development Agenda and Meeting notes

Date

27 Jul 2021

Location

Browser	Room System	Phone Dial-in
https://bluejeans.com/690803707 or https://ls.st/rsv	1. Dial: 199.48.152.152 or bjn.vc 2. Enter Meeting ID: 690803707 -or- use the pairing code	Dial-in numbers: • +1 408 740 7256 • +1 888 240 2560 (US Toll Free) • +1 408 317 9253 (Alternate Number) Meeting ID: 690803707

Attendees

- Jeffrey Carlin
- Colin Slater
- Erik Dennihy
- Tony Tyson
- Keith Bechtol
- Robert Lupton

Regrets

- Leanne Guy (NSF review)
- Simon Krughoff (vacation)

Metric Tracking Dashboard

URL:

Discussion items

Item	Who	Pre-meeting notes	Notes and Action Items
News, announcements and task review	Leanne Guy	Welcome to Erik Dennihy	

Bugs, issues, topics, etc of the week	All	• JSON/lsst.verify NaN issue in faro was resolved by turning faro off. We need to
		  ◦ a) get a new validation dataset in place and turn it back on and
		  ◦ b) work out how to deal with NaN in reporting to squash (we will have to deal with NaN)
		• validate_drp is also failing due to NaN. Recommend leaving to pipelines to decide what to do as they are still using it for G2/3 parity testing.
		• RFC-793 : use PyTest in science pipelines. Support! We don't have enough unit tests in faro.
			Failure due to a change in json reader. Python requests model does not want shape invaild json over http. The datasets being used for validate_drp. Don't compute valid metrics due to changes in implementation. Disable in nighly builds. Longer term solution is to develop a better CI dataset.
			
			Should we stop using NaN as a sentinel value when attempted to compute the metric but insufficient data.
			
			Multiple issues
			
			• Transport layer
			• Do we want to have more classes of failure status? For example:
			  ◦ Not enough data
			  ◦ Algorithm failed
			• What happens when NaN goes to influxDB? Currently treats as those metrics do not exist.
			
			There is a distinction between visualization and what is stored in the database. NaN do not show up as entries in table in Chronograph. We want to be aware of results when.
			
			☑ Jeffrey Carlin make a proposal to define failure status for metrics. 31 Aug 2021
			
			Are we concerned that we do not currently have valid metrics values from nightly CI to look on dashboard? Breakages are on CI, not RC2 dataset. Currently thinking about the nightly CI mainly for checking breakages, not looking at performance for nightly CI. We do care about this for larger datasets.
			
			Simon has been compiling larger dataset that could be a middle ground.
			
			Simon met with DRP folks. Once Dan Taranu is back from vacation, create dashboard for DRP metrics monitoring meeting. This should change soon. faro development group may need to help with dashboard creation.

Reprocessing status and metrics review	Jeffrey Carlin Keith Bechtol	Dashboards can be found here: https://chronograf-demo.lsst.codes/sources/2/dashboards Monthly reprocessing: • DRP faro metrics (DRP metrics nightly CI Gen3 (validation_data_*)) ◦ (new validation datasets should be added to this dashboard) • DRP validate_drp metrics monthly Gen 2 (HSC RC2)	w30 processing in progress. Few errors. Mostly done. Metrics should have been calculated, but are not yet being displayed on a dashboard ☑ Jeffrey Carlin Check with Yusra on status of faro dashboard for RC2 03 Aug 2021

Development status	Leanne Guy	• Fall 2021 epic
		
		⚠ DM-30748 - Jira project doesn't exist or you don't have permission to view it.
		
		• Backlog epic:
		
		⚠ DM-29525 - Jira project doesn't exist or you don't have permission to view it.
			• ⚠ DM-26214 - Jira project doesn't exist or you don't have permission to view it.
			
			there has been a lot of discussion on this PR. Summarize the main takeaways
			   ○ Add DM license preamble to all python files
			      ■ ⚠ DM-31135 - Jira project doesn't exist or you don't have permission to view it.
			      ■ Started this for all the files touched so far
			   ○ Create / improved doc strings
			   ○ Homogenize run method signatures across all the measurement tasks. Suggest in general to inherit run method from CatalogMeasurementBaseTask base class, which can accept and pass keyword arguments. We should avoid passing dataIds to measurement tasks were possible. For example, we can use FilterLabel or list of FilterLabels instead of passing the dataIds if we just need to know which bands are being used.
			      ■ ⚠ DM-31061 - Jira project doesn't exist or you don't have permission to view it.
			   ○ Added most of the utility for loading external reference catalogs into the CatalogMeasurementBaseTask base class
			      ■ See https://github.com/lsst/faro/blob/tickets/DM-26214/python/lsst/faro/base/CatalogMeasurementBase.py
			   ○ Add unit tests for task code in faro
			      ■ ⚠ DM-31158 - Jira project doesn't exist or you don't have permission to view it.
			      ■ Started this https://github.com/lsst/faro/blob/tickets/DM-26214/tests/test_tasks.py
			   ○ Use consistent DM class naming convention for base classes
			      ■ ⚠ DM-31136 - Jira project doesn't exist or you don't have permission to view it.
			      ■ Started this
			   ○ Development of reference catalog loading capability in faro with proper configuration options led to discovery and resolution of a bug in LoadReferenceCatalogTask. Thanks to Eli.
			      ■ ⚠ DM-31188 - Jira project doesn't exist or you don't have permission to view it.
				:

AOB	Leanne Guy	Next meeting 03 Aug 2021		

List of tasks (Confluence)

Description	Due date	Assignee	Task appears on

☐ Leanne Guy to talk to Science Pipelines (Yusra) about when do this transfer 19 Oct 2021	19 Oct 2021	Leanne Guy	2021-09-28 Science Metric Development Agenda and Meeting notes
☐ Leanne Guy arrange to disucss at a future meeting if there are metrics from PDR3 & this paper that we might want to include in faro. 26 Oct 2021	26 Oct 2021	Leanne Guy	2021-08-31 Science Metric Development Agenda and Meeting notes
☐ Colin to ask about capturing ideas for improvement to the stellar locus algorithm 30 Nov 2021	30 Nov 2021		2021-11-09 Science Metric Development Agenda and Meeting notes
☐ Colin Slater to make a preliminary draft agenda for a workshop to clarify visualization use cases for science verification and validation		Colin Slater	2022-04-19 Science Metric Development Agenda and Meeting notes
☐ Jeffrey Carlin to review metric specification package organization and the relationship to formal requirements documents		Jeffrey Carlin	2022-04-19 Science Metric Development Agenda and Meeting notes
☐ Keith Bechtol Schedule a time to have focused discussion on verification package, potentially next status meeting		Keith Bechtol	2021-09-14 Science Metric Development Agenda and Meeting notes
☐ Keith Bechtol to make a ticket to better understand mapping of these camera and calibration products characterization efforts to verification documents and the focus of these efforts. Discuss with the SCLT		Keith Bechtol	2021-09-14 Science Metric Development Agenda and Meeting notes