2022-04-12 Science Verification - Planning Meeting Agenda and Meeting notes

Faro planning meeting - details to come.

Date

12 Apr 2022

Location

Zoom alias : https://ls.st/rsv

Zoom: https://noirlab-edu.zoom.us/j/96120781994?pwd=bCtpT3Q3b2RsU1ZBRUFaSnUrZXo3Zz09#success

Attendees

- Yusra AlSayyad
- Robert Lupton
- Erik Dennihy
- Keith Bechtol
- Michael Martinez
- Peter Ferguson
- Colin Slater
- Jeffrey Carlin
- Leanne Guy

Regrets

Useful links

Metric tracking dashboards

- DC2 and HSC RC2 monthly reprocessing (Gen3 and Gen2/3, faro): https://chronograf-demo.lsst.codes/sources/2/dashboards/75
- Nightly CI (Gen3, faro): https://chronograf-demo.lsst.codes/sources/2/dashboards/70

Development

- Spring 2022 epic: DM-33385 - 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.

Rubin science pipelines weekly change log and Github repo: https://github.com/lsst-dm/lsst_git_changelog

Status of jenkins jobs and slack channel for notifications: #dmj-s_verify_drp_metrics slack channel

Discussion items

Item	Priority (vote here for your top three with your initials)	Minutes	Meeting Notes

Relationship with visualization	ED, LPG, JLC, CTS, PSF: high	Group these together with others below, and set up a team from pipelines, SST, commissioning-verification and bring in someone from architecture (KT) to develop design for a visualization framework CTS: I'd spin this slightly differently; we should figure out how users want to interact with the faro results and build what's necessary for that. (In practice that will mostly fall under the "visualization" umbrella, but that's not itself the goal I'm interested in). KB: I think it would be helpful to break down "visualization" into a least two categories in order to make the problem more tractable: 1. Visualizing the **results of metric calculation** • metric values vs. science pipelines version (how are Science Pipelines changing in sense of catching regressions or improvements) • metric values vs. exposure number of on-sky day (how is observatory system changing) • metric values vs. sky position 2. Visualizing the **inputs to metric calculation** • Understanding / interpreting / presenting our metric values • Why did the metric value change (e.g., in response to changes in Science Pipelines and /or a change in data quality)? • Catching data anomalies that might not be readily apparent in our defined scalar metrics Ticket: DM-29753 - Jira project doesn't exist or you don't have permission to view it.	• Different technical solutions to the 2 types of vis. that KB presents, e.g what is the spatial structure of the Stellar Locus width and why did the Stellar Locus blow up? • YAS: would out how pipelines are changing in 'inputs' - KB, yes, this belongs in both suggested categories • CTS: 3 diff things here - will need to make choices about the things to do and priorities - what do we need when, what tooling do we need • Starting point would be to collect use cases from many perspectives including Science Pipelines, commissioning, • JLC: Should we organize a design workshop to flesh out the workshop • All in favour of an in-person workshop on this. - ◦ After DP0.2 (30 June) ◦ Within next 3 months
Diagnostic / drill-down capability	LPG - this is related to visualization. PSF	High priority to architect a design and prototype in next 6 months Start with a set of use cases - e.g 2pt correlation function	
Where does common /shared (with other stack packages) code live? e.g selectors, algorithms.	ED, KB, LPG, JLC: high	Set up session with DRP to address in next cycle KB: faro and analysis_drp have similar but different implementations of selectors. This could be a good place to start.	• Must be using same stars for the same studies - is some restructuring needed to make this happen? • Factor so that we have a common library to do this? • Selectors could be a good concrete place to start all of this - analysis_drp had some, inspired faro for another set of selectors - 2 implementations of the same idea- different base classes so not clear how the full packages could be merged (or if they should be merged). Good place to start. • Decide on technical grounds on where the best place to put the code is . • Plan to get together developers in analysis_drp and faro to propose and implement technical solution
How to handle situations when we have multiple summary statistics that we would like to compute with the same underlying computations	LPG: low priority. KB: low priority CTS: medium risk	LPG: If the cost of underlying computing is small, I prefer to stick to fully independent metrics. No computation/resource issues around this at the moment. Matching is an exception to this. A matched catalogue should be computed once for all metrics that use it CTS: My concern is that this makes each new metric seem like more of a burden and raises the threshold for "is this worth writing a new metric". E.g. if I want to count the fraction of pixels with a given mask bit set, is it "worth" writing a new task for each type of mask bit?	

Metric dataset type naming convention and metric package specifications	KB: High – this is technical debt that will catch up with us. Coupled to some extent with lsst.verify.Measurement
	
	LPG: high but also should be fast to resolve.
		LPG: Not doing this now will create technical debt (and nightmare code)
		
		metric package name, metric names, variants (dataset type)
		
		JLC: I'm not sure this will be fast or easy, but it will only get more difficult the longer we wait. <--- +1 PSF
		
		Previous thinking on this topic at Metric Calculation Package Reorgnization
		
		Ticket:
		DM-27101 - Jira project doesn't exist or you don't have permission to view it.
			• Every metric is a new dataset type in repo
			• Dataset type names are global across all repos and are difficult to change
			  ◦ Can't change dimensions for example
			• Current convention is metricvalue_<verification_package>_<metric_name>
			• Verification packages are coupled to this discussion because we have been including them in the dataset type names
			• Is there support for namespacesing Butler? RFC to change DS definitions after the fact withdrawn:
			RFC-804 - Jira project doesn't exist or you don't have permission to view it.
			• CTS - symptom of every metric having a separate class.
			• There are three different technical issues that are related because of the current dataset type naming convention we have adopted
			  ◦ namespacing
			  ◦ verification package organization
			  ◦ naming of individual metrics
			• RHL – should we be taking a step back to think about how the structure for holding metric values
			• lsst.verify.Measurement is not currently being maintained
			• Priority to resolve in the next 6 months
			• Computing metrics on a per-ccd scale in the current structure is very poor performance.
			• Naming related to how we are storing metrics so no point solving the naming if we are chaning the date structures - break down in to constituent parts

Review of documentation	LPG: High - essential to enable wide contribution of metrics from a broad community (à la MAF) JLC: medium – once we get all the basics for various analysis contexts in place, I think providing straightforward guidance about how to implement metrics should be a top priority to enable others to get involved.	JLC: It may make sense to sort out where "shared" code and modules will live before doing this, as it may change the recommendations for implementing metrics. KB: One concrete actionable task would be to update the README file in the faro github repo Relevant Tickets: • DM-30144 - Jira project doesn't exist or you don't have permission to view it.	
Review of unit tests / coverage	LPG: medium. PSF	We should look regularly at unit test coverage.	
List of metrics to be implemented / review of metric specifications	LPG: low - we have a good list for now that we need to complete. These are by no means a complete set. Additional Non SRD metrics will be driven by need (and probably out of commissioning) ED	LPG: Exception is for requests from on-sky observing to implement a metric to support commissioning.	

Metrics using SSI (Fake Injection)	LPG: related to above	JLC Note: the synthetic source package in the Stack is currently being refactored, so it may make sense to wait a while.	

Persisting metrics as lsst.verify. Measurement objects or something different??	KB: high in the sense that will be needed for commissioning.
	
	LPG: high in that should we feed this back now.
		KB: Separate from faro development. More in the FAFF domain or lsst.verify.
		
		LPG: Performance issue - it is prohibitive to load 1000s into memory to do aggregations across metrics. 10K takes 30 mins to load into memory.
		
		LPG: Suggest we create a ticket incl. use case + timescale of need for the lsst.verify developers.
		
		CTS: This is closely related with item #3 (each metric is an independent task).
			• discussed above as part of "Metric dataset type naming convention and metric package specifications"
			• KB: I think we want a way of making metric values more easily searchable like a relational database that can be correlated with EFD, etc.
			  ◦ YA: +1 There are visitSummary and CcdVisit Table type things that would be good for that.
			• JC: chronograph is expecting to take lsst.verify.Measurement as input
			  ◦ LPG: Thi is becase we designed it to do so - we can change this.

Matching routine for use with external catalogs	LPG: High JLC: medium PSF: Medium KB: medium	We have simple algorithms available/implemented for matching with external catalogs, but not yet a more general solution. KB: I think we have a adequate solutions for applications such as absolute astrometry where we would be associating bright isolated stars (e.g., 20th magnitude). More thought is needed for applications such as evaluating detection completeness at 27th magnitude. PSF: I think this needs some communication with other parts of DM, maybe try and focus on API?	
Rendezvous with image metadata	KB: high in the sense that we need this capability, but not immediately clear that faro is the right tool LPG: Is this for faro?	LPG: Commissioning to provide use cases to evaluate if faro is appropriate	
Currently, the spatial scales of metric calculation in faro are directly tied to the spatial scales that are used for data processing (e.g., patch, tract, detector, visit)		KB: The faro design has been that metric values are stored as scalars with a specific associated data id that is related to the units of data processing. There is a potential risk that the spatial scales of data processing (e.g., patch, tract, detector, visit) are not matched to the spatial scales of interest for science verification and validation. This is potential limitation of the current design.	

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