Vera C. Rubin Observatory
Data Management

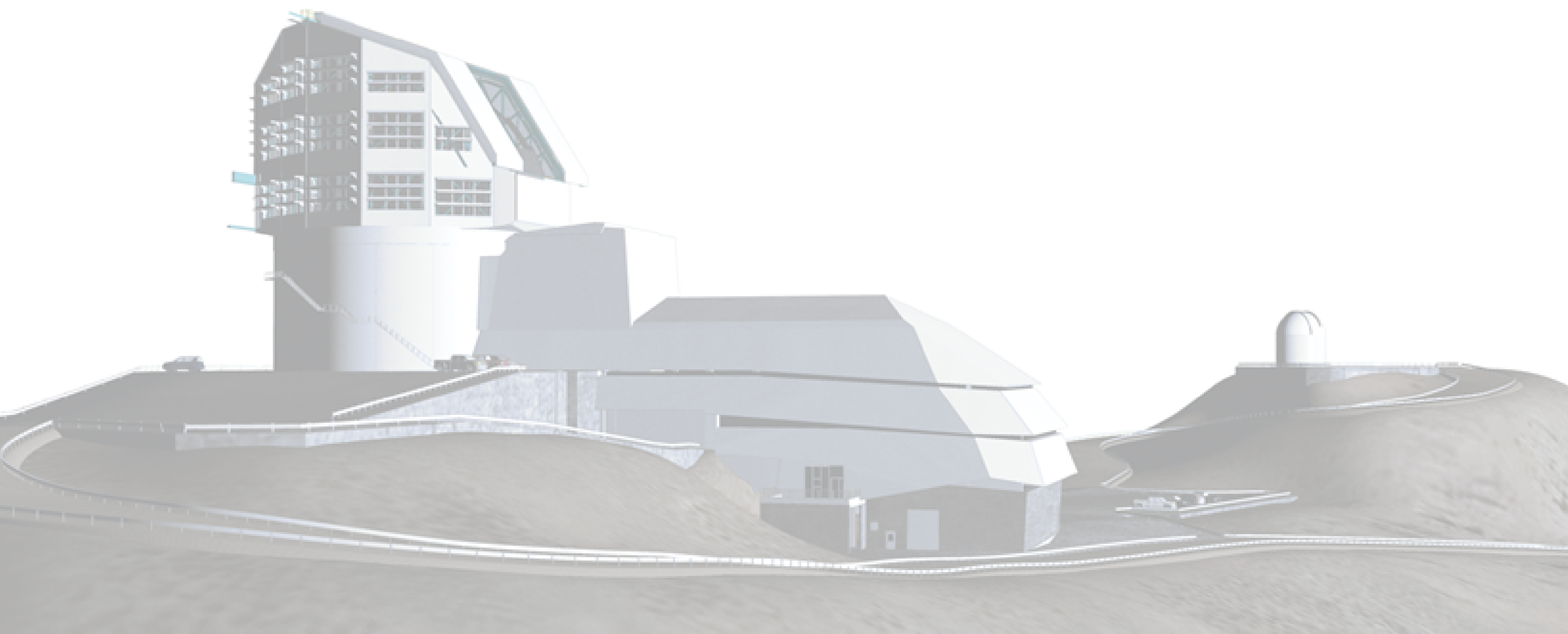
Community Science Use Cases

Melissa Graham, Christina Adair, Jim Annis, Jeff Carlin, Alex Drlica-Wagner, Leanne Guy, Greg Madejski, Robert DePeyster

RTN-002

Latest Revision: 2023-04-11

DRAFT



Abstract

The Community Science Team (CST) in the Rubin System Performance department is generating a suite of use-cases to capture the full range of anticipated interactions regarding community support for Rubin software and data products.

Change Record

Version	Date	Description	Owner name
1	YYYY-MM-DD	Unreleased.	Melissa Graham

Document source location: `https://github.com/rubin-observatory/rtn-002`

Contents

Community Science Use Cases

1 Introduction

In this document we gather use-cases for community engagement between the Rubin Observatory and its science users, and model them in a format defined by model-based systems engineering (MBSE) practices.

1.1 Template

Use Case: *unique identifier number name (short active verb phrase)*

Characteristic Information
Trigger: *the action upon the system that starts the use case, may be time event*
Goal in Context: *a longer statement of the goal, if needed*
Primary Actor: *a role name for the primary actor, or description*
Scope: *what system is being considered black-box under design*
Level: *one of: Summary, Primary task, Subfunction*
Preconditions: *what we expect is already the state of the world*
Success End Condition: *the state of the world upon successful completion*
Failed End Condition: *the state of the world if goal abandoned*

Main Success Scenario
put here the steps of the scenario from trigger to goal delivery, and any cleanup after
Step 1: *action description*
Step 2: *action description*
Etc.

Extensions
put here there extensions, one at a time, each referring to the step of the main scenario
Alteration 1: IF *condition* THEN *action or sub-use case*
Alteration 2: IF *condition* THEN *action or sub-use case*
Etc.

Sub-Variations
put here the sub-variations that will cause eventual bifurcation in the scenario
Variation 1: IF *condition* THEN *list of sub-variations*
Variation 2: IF *condition* THEN *list of sub-variations*

Related Information (Optional)
Priority: *how critical to your system / organization*

Performance Target: *the amount of time this use case should take*
Frequency: *how often it is expected to happen*
Superordinate Use Case: *optional, name of use case that includes this one*
Subordinate Use Cases: *optional, depending on tools, links to sub.use cases*
Channel to primary actor: *e.g. interactive, static files, database*
Secondary Actors: *list of other systems needed to accomplish use case*
Channel to Secondary Actors: *e.g. interactive, static, file, database, timeout*

Open Issues (Optional)
list of issues about this use cases awaiting decisions

Schedule
Due Date:

2 Use Cases

2.1 Generic Community Self-Help Using CST Resources

Use Case: 001 Generic Community Self-Help Using CST Resources

Characteristic Information
Trigger: Community member posts a request for help in the forum hosted by the CST.
Goal in Context: *a longer statement of the goal, if needed*
Primary Actor: Community member.
Scope: CST Online Resources
Level: *one of: Summary, Primary task, Subfunction*
Preconditions: An online web forum (or open help desk) with active participation from experienced community members exists.
Success End Condition: The community helps itself and solves the issue.
Failed End Condition: The community cannot solve the issue and the SP-CST must ingest it.

Main Success Scenario
Step 1. A user makes a post on the LSST Community web forum detailing their issue.
Step 2. Community members reply with suggestions.
Step 3. The user solves their issue and marks their post solved.

Extensions
none

Sub-Variations
Variation 1. IF the community does not have the relevant expertise THEN the issue is submitted to the SP-CST.

Related Information (Optional)
Priority: High (the system must be able to accommodate this activity)
Performance Target: Days
Frequency: High

Superordinate Use Case:
Subordinate Use Cases:
Channel to primary actor:
Secondary Actors:
Channel to Secondary Actors:

Open Issues (Optional)
list of issues about this use cases awaiting decisions

Schedule
Due Date:

2.2 Generic Sequence of Events from a CST Perspective

Use Case: 002 The CST Solves an Issue

Characteristic Information
Trigger: The issue is identified by the Primary Actor.
Goal in Context: User submits Help Desk request for a potential Camera issue.
Primary Actor: A scientist within the project and/or the scientific user community of the Rubin Observatory.
Scope: Community Engagement Team
Level: Primary Task
Preconditions: Operations phase has started.
Success End Condition: The issue is resolved and the scientific performance of the Rubin Observatory is maintained or increased.
Failed End Condition: The issue is not resolved and the scientific performance of the Rubin Observatory is negatively impacted.

Main Success Scenario
Step 1: Identify individuals with the relevant expertise.
Step 2: Understand the issue and how it threatens scientific performance.
Step 3: Define the importance and urgency of the issue.
Step 4: Develop a strategy to resolve the issue.
Step 5: Work to implement the resolution.
Step 6: Review and close the issue.

Extensions
Alteration 1.1. IF the issue evolves THEN new individuals with expertise are identified.
Alteration 1.2. IF all identified individuals needed to resolve the issue are from the science user community THEN they may proceed without formally completing steps 2-5, but should still review and close the issue.
Alteration 2.1: IF the issue evades understanding THEN new individuals with expertise are identified.
Alteration 2.2. IF the issue does not threaten scientific performance THEN it may not require

a resolution, and could be reviewed and closed after skipping steps 3, 4, and 5.
Alteration 3.1. IF the issue is deemed unimportant or non-urgent THEN it may not require a resolution, and could be reviewed and closed after skipping steps 3, 4, and 5.
Alteration 4.1: IF a strategy cannot be developed THEN return to steps 1-3.
Alteration 5.1: IF work to implement a resolution is not completed THEN return to step 4.
Alteration 6.1: IF the resolution does not pass review THEN return to step 4.

Sub-Variations
Variation 5.1: As work proceeds, new (related) issues may be spawned as children.

Related Information (Optional)
Priority: *TBD as part of the main success scenario*
Performance Target: *the amount of time to resolve issues will be very issue-dependent*
Frequency: *the frequency of different issue types will depend on the issue itself*
Superordinate Use Case:
Subordinate Use Cases:
Channel to primary actor: interactive communication
Secondary Actors: Rubin Observatory scientists and Rubin science community users
Channel to Secondary Actors: interactive communication

Open Issues (Optional)

Schedule
Due Date:

2.3 A Camera Fault Requires Modification of the Alert Production System

Use Case: 003 Observatory Personnel Detect a Fault, Leads to Change in Alert Production Pipeline

Characteristic Information
Trigger: A fault in one of the Raft Electronics Boards (a sensor).
Goal in Context: This fault causes and abnormally high number of DIASource detections for that sensor, most of which are appropriately flagged as artifacts by the real/bogus algorithm and do not become Alerts, but some are not flagged and are released.
Primary Actor: The System Performance Verification and Validation (SP-VV) Lead Scientist
Scope: alert production, broker community
Level:
Preconditions: Operations
Success End Condition: The sensor anomaly is solved by modifying the Alert Production (AP) pipeline.
Failed End Condition: Bogus DIASources from this sensor fault continue to be released as Alerts.

Main Success Scenario
Step 1: Rubin Observatory Operations team monitors verification and validation outputs during the night.
Step 2: Night staff notice an anomalous spike in the number of DIASources detected.
Step 3: Night staff run diagnostics to isolate the excess DIASources to a single sensor, identify the underlying fault, and determine that the hardware cannot be quickly/easily fixed (i.e., sensor will have to be replaced).
Step 4: Night staff summarize their findings for the SP-VV Lead Scientist.
Step 5: SP-VV team runs further diagnostics and coordinates with Data Production (DP) to plan a fix.
Step 6: DP implements the fix (e.g., retraining the real/bogus characterization).
Step 7: SP-VV and DP coordinate with the CST to summarize the issue, it's fix, and potential science impacts for the community.

Extensions
Alteration 5.1: IF a significant amount of the anomalous DIASources were released as Alerts THEN SP-VV team coordinates with the CST to prepare an initial statement describing the fault and it's impact on the Alert stream, to be posted to the Community Forum and sent to a community brokers email list.

Sub-Variations
Variation 3.1: IF the sensor anomaly is fixed with hardware THEN a software fix might not be necessary

Related Information (Optional)
Priority: High
Performance Target: identify scope of issue within one day; apply fix to AP within a week
Frequency: Low
Superordinate Use Case:
Subordinate Use Cases:
Channel to primary actor: Internal discussions via Slack channels
Secondary Actors: CST, DP, science users
Channel to Secondary Actors: posts to the science users via the Community Forum

Open Issues (Optional)

Schedule
Due Date:

2.4 Large Queries: Community Self-Help Using CST Resources

Use Case: 004 Large Queries: Community Self-Help Using CST Resources

Characteristic Information
Trigger: A user's request for help posted on the LSST Community web forum.
Goal in Context: A user receives assistance in executing large queries on a Rubin data product.
Primary Actor: The System Performance Community Engagement Team (SP-CST)
Scope: SP-CST
Level:
Preconditions: In-Operations, with users accessing the Rubin Science Platform Notebook Aspect, for which training tutorials exist and are served by the SP-CST
Success End Condition: The user has a path towards solving their issue.
Failed End Condition: The user does not know how to solve their issue.

Main Success Scenario
Step 1: A user makes a post on the LSST Community web forum detailing issues they are having in querying the Object catalog from the latest data release (their Jupyter Notebook will not run.)
Step 2: Other users and an SP-CST member respond to the post with advice about their query.
Step 3: Discussion thread on the post realizes the user needs more computational resources.
Step 4: The user applies to the Resource Allocation Committee for additional processing.

Extensions
Step 4 Alteration: Optional addition, the SP-CST add this user's particular case to documenation about large queries to avoid such issues in the future.

Sub-Variations

Related Information (Optional)
Priority: Low
Performance Target: a few days at most
Frequency: probably pretty often
Superordinate Use Case:
Subordinate Use Cases:
Channel to primary actor: Interactive (web forum)
Secondary Actors: other community users
Channel to Secondary Actors: Interactive (web forum)

Open Issues (Optional)

Schedule
Due Date:

2.5 Science Platform Bug: A Help Desk Submission

Use Case: 005 Science Platform Bug: A Help Desk Submission

Characteristic Information
Trigger: Help desk submission
Goal in Context: address bug in Rubin Science Platform (RSP)
Primary Actor: CST member
Scope: the science platform
Level:
Preconditions: The RSP is operating and has users. Some ticketing system is operating. The ticket has been submitted.
Success End Condition: The ticket is resolved by either: (1) solving user issue that is not actually a bug, or (2) identifying the bug and providing information to the RSP. I guess the real success end condition is correcion of the bug, but more relevant for the CST team is probably providing a short-term workaround.
Failed End Condition: Bug is not able to be reproduced and no work around is available.

Main Success Scenario
Step 1: CST confirms bug is reproducible.
Step 2: CST communicates bug to RSP developers.
Step 3: CST provides work-around to user.
Step 4: RSP developers fix bug.
Step 5: New RSP version rolled out.

Extensions

Sub-Variations

Related Information (Optional)
Priority: most will probably be low priority
Performance Target: days to weeks
Frequency: infrequent
Superordinate Use Case:
Subordinate Use Cases:
Channel to primary actor: interactive via Help Desk
Secondary Actors: Rubin Science Platform developers in Data Production Dept.
Channel to Secondary Actors: interactive, e.g., JIRA

Open Issues (Optional)

Schedule
Due Date:

2.6 Survey Strategy Alteration for Photometric Redshifts

Use Case: 006 Survey Strategy Alteration for Photometric Redshifts

Characteristic Information

Trigger: DESC photo-z and weak lensing working groups determine that their projected photo-z statistics fall short of the LSST SRD year 10 target values for photo-z quality, and the DESC spokesperson communicates this formally to the SP-CST lead via email.
Goal in Context: As the Rubin Observatory intends to meet the science requirements of its LSST servey, staff will need to study and implementent changes to the survey strategy to aquire the right data to meet the photometric redshift statistics requirements.
Primary Actor: SP-CST lead, upon receipt of email, opens an issue in an issue tracking system.
Scope: We are designing the interaction of the survey performance team with the observatory operations team as a formal request for a change comes from the science community.
Level: Summary
Preconditions: This scenario envisions the Observator is in full operations and that data has been released to the science collaborations. Likely this is year 4 of the survey and year 3 data is under study by the DESC. It takes time for science collaborations to understand the limitations of the data and the likely lag from data release to formal feedback is a year. The scenario envisions that most scientists believe that the limitations of the photo-zs can be improved by changing the kind of data being acquired as opposed to changing the algorithms.
Success End Condition: A year's worth of new data improves the photo-z statistics as expected, and the DESC photo-z and weak lensing working groups can project the 10 year data meeting requirements.
Failed End Condition: The photo-z statistics remain stubbornly high. The SRD photo-z requirements are not met, and the leading systematic on the weak lensing cosmology results is the photo-z statistics. This limits the precision of the resulting cosmological parameters.

Main Success Scenario

0. The DESC determines there is an issue in the photo-zs and communicates this to the SP-CST.
1. SP-CST creates an issue ticket assigned to the Community Scientist with expertise in cosmology adding relevant watchers from the SP Survey Scheduling Team (SP-SST), Data Production Algorithms and Pipeline Team (DP-AP), and Observatory Operations Observatory Science

Team (OO-OS).
2. A meeting is organized with DESC scientists, SP-CST, SP-SST, DP-AP, OO-OS representatives, and members of the Survey Cadence Optimization Commitee (SCOC) with the agenda of deciding what the next steps are. It is decided to perform simulations of a cadence that increases u and y band cumulative exposure times.
3. The SP-SST publishes a Community post describing the issue and the actions being taken and responds to the DESC spokesperson.
4. The SP-SST generates the OpSim results with standard metrics.
5. The DESC (and other science collaborations) evaluate the simulations to determine impact on the issues with the weak lensing and photo-z as well as on all other science cases of concern and sumbits reports back to the SCOC.
6. The SCOC meets (≈6 months) and decides to recommend a u-band focus for year 9 to the Directorate. The written summary of their discussion is provided to the SP-CST.
7. The SCOC written summary is made publicly available by the SP-CST.
8. The issue ticket is closed.

Extensions

Sub-Variations

Related Information (Optional)
Priority: This issue is of high priority but is not particularly urgent.
Performance Target: The time from opening to closing the ticket should be <1 year.
Frequency: Issues affecting the wide field cadence are likely every couple of years.
Superordinate Use Case:
Subordinate Use Cases:
Channel to primary actor: interactive (email)
Secondary Actors: SP-SST, DP-AP, OO-OS, SCOC
Channel to Secondary Actors: interactive (JIRA)

Open Issues (Optional)

Schedule

Due Date:

2.7 A Science-Driven Modification to a Data Product

Use Case: 007 A Science-Driven Modification to a Data Product

Characteristic Information

Trigger: A manuscript is published in the literature that highlights and issue with Rubin Observatory Data Release products. This paper is seen by a CST member, who passes it along to the Community Scientist and Science Collaboration with the relevant expertise.
Goal in Context: While the Data Releases will be subject to verification and validation before release, it is inevitable that sub-optimal outcomes for certain science goals will be identified as a result of data processing decisions that were taken (or weren't even considered). The CST and the DRP team need to implement processes to (1) assess the possible impacts of alterations to the code base that solve the issue, and (2) if possible, incorporate code changes that mitigate the emergent scientific issue while not compromising overal DRP results.
Primary Actor: The CST member with primary expertise in transients.
Scope: Data Release Production
Level:
Preconditions: A Data Release has occurred, with sufficient time passed for scientists to carry research projects through to publication.
Success End Condition: The DRP team implements and tests and algorithmic change that solves the problem.
Failed End Condition: A solution is either not identified, or is deemed deterimental to other aspects of data quality, and the systematic offset in nuclear transient occurrence rates with galaxy host types remains unchanged.

Main Success Scenario

1. A manuscript is published in the literature that highlights a systematic offset in the detection efficiency of nuclear transients in different types of host galazies with Rubin Observatory Data Release products.
2. This paper is brought to the attention of a CST member, who creates an Issue Ticket assigned to the Community Scientist(s) and Science Collaboration(s) with the relevant expertise (in this case, predominantly time-domain science and the TVS SC).
3. The Community Scientist and other relevant parties investigate the issue and identify a possible cause in the DM pipelines.

4. The ticket is discussed with DM developers, who identify possible improvements to mitigate the issue, and consider the possible side effects of implementing those changes.
5. DM developers implement the code changes on a ticket branch, and the relevant CST and/or TVS-SC verify that the new algorithm decreases (or removes) the systematic offsets with host galaxy type from the delivered detection efficiencies, while not adversely affecting other pipeline outputs.
6. The changes are merged to the science pipelines, and one of the relevant parties defines a metric that can be monitored each time a new version of the Science Pipelines and DRP is released.
7. The CST member details the mitigation and demonstrates its outcome in a public-facing document (e.g., a DM Tech Note or similar), and the issue is closed.

Extensions

Altered Step 5.1: Relevant CST and/or TVS-SC members test the new algorithm to assess whether it improves the formerly low detection efficiencies in certain types of host galaxies. During this verification, it is discovered that the changes that were made (e.g., lowering detection thresholds near resolved galaxies) adversely affected other scientific data products.
Altered Step 5.2: It is decided that it is not acceptable to sacrifice data quality of other pipeline products to solve this issue with galactic transients. As an alternative, a method is devised that does not require new or altered measurements, but instead uses existing measurements such as offset from the nearest galaxy, color, morphology, etc. to derive a classification scheme that more accurately captures the subtlety of identifying transients in galaxy nuclei.

Sub-Variations

Related Information (Optional)

Priority: Not urgent, but important to have a process to deal with emergent science effects such as this. The amount of effort to be spent will depend on how frequency issues like this come to our attention.
Performance Target: 1-3 months
Frequency: 4-8 times per year (most likely all clustered in a period shortly after data releases)
Superordinate Use Case:

Subordinate Use Cases:
Channel to primary actor: Personal communication (email, Slack, or during regular meetings) to CST or individual Community Scientist
Secondary Actors: Data Management developer(s), interested Science Collaboration members
Channel to Secondary Actors: Email or Slack (perhaps first passing through the DM-SST?)

Open Issues (Optional)

Schedule
Due Date:

A References

B Acronyms

Acronym	Description
AP	Alert Production
CST	Community Science Team
DESC	Dark Energy Science Collaboration
DM	Data Management
DM-SST	DM System Science Team
DP	Data Production
DRP	Data Release Production
LSST	Legacy Survey of Space and Time (formerly Large Synoptic Survey Telescope)
MBSE	Model Based Systems Engineering
OS	Operating System
OpSim	Operations Simulation

RSP	Rubin Science Platform
RTN	Rubin Technical Note
SC	Science Collaboration
SCOC	Survey Cadence Optimization Committee
SP	System Performance
SRD	LSST Science Requirements; LPM-17
SST	Subsystem Science Team
TBD	To Be Defined (Determined)
TVS	Transients and Variable Stars Science Collaboration
photo-z	photometric redshift