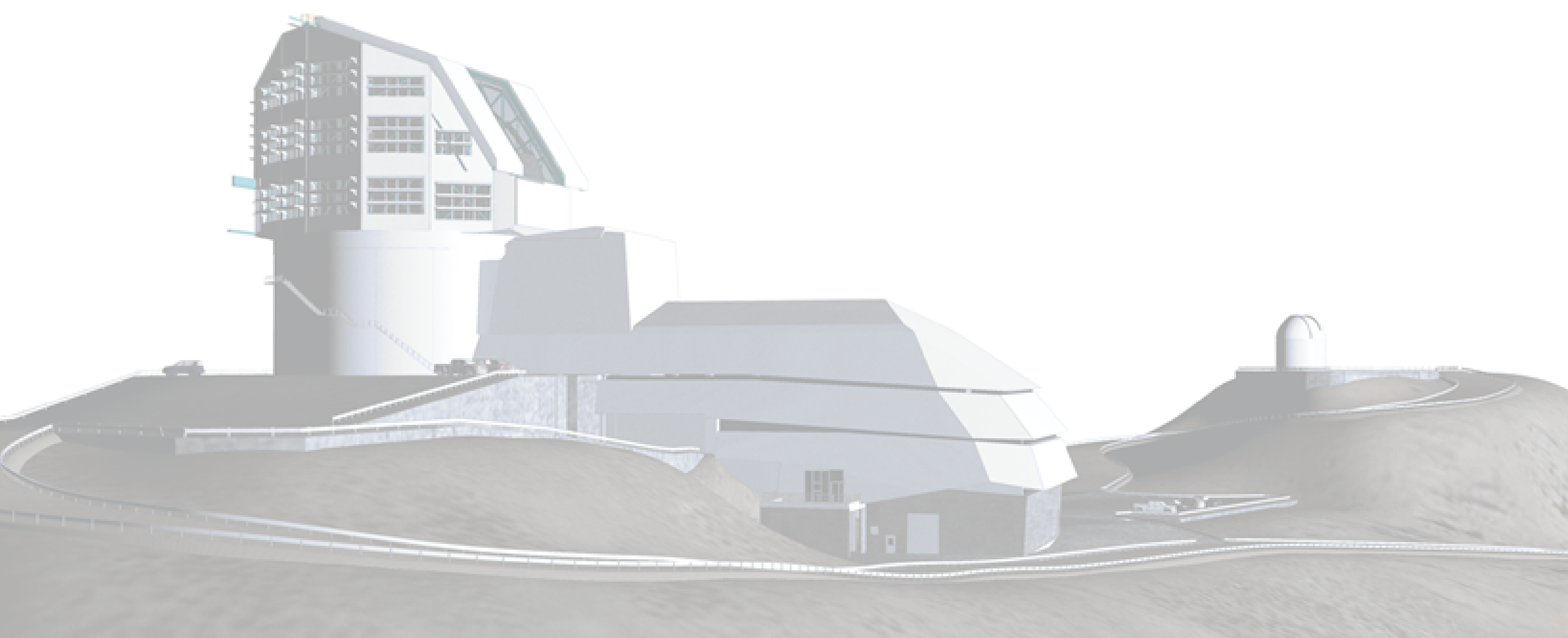
Vera C. Rubin Observatory
Data Management

Rubin Observatory Plans for an Early Science Program

Leanne P. Guy, Eric Bellm, Bob Blum, Melissa L. Graham, Željko Ivezić, Michael Strauss.

RTN-011

Latest Revision: 2021-11-11



Abstract

Rubin Observatory is committed to enabling high-impact science during the first year of operation of the Legacy Survey of Space and Time (LSST). This document provides the plans for a dedicated Early Science Program designed to realize that goal. It outlines the strategy for obtaining observations during commissioning to enhance opportunities for Early Science and presents plans to implement incremental template generation to augment alert production in the early phases of the survey. The Rubin Operations team will work closely with the science community to design the Early Science Program so as to maximize the time-domain and solar system science achievable in the first year of operations. This is a living document; both it and the Early Science Program will continue to evolve over the course of commissioning and pre-operations in response to the state of the as-built system and to community guidance.

Change Record

Version	Date	Description	Owner name
1	2020-10-30	First draft	Leanne Guy
2	2020-12-16	Draft 1.1	Bob Blum
3	2021-10-08	Rework structure	Leanne Guy
4	2021-10-21	Add timeline	Leanne Guy
5	2021-11-05	Edits throughout	Eric Bellm
6	2021-11-09	Global edits and consolidation	Leanne Guy

Document curator: Leanne Guy
Document source location: `https://github.com/rubin-observatory/rtn-011`

Contents

Rubin Observatory Plans for an Early Science Program

1 Summary

Rubin Observatory is putting in place a dedicated *Early Science Program* to ensure high-impact science during the first year of operation of the Legacy Survey of Space and Time (LSST). This Early Science Program is motivated by the fact that in the current baseline there will be no science-ready data products released before Data Release 1 (DR1) .

Time-domain astronomy is a key component of LSST's four science pillars and is enabled by alerts on LSST detections of transient, variable, and/or moving objects. Alerts are the only data product that will be immediately available (within 60 seconds of image readout) and publicly shareable, i.e not subject to a proprietary period (Jurić et al., LSE-163), (Blum & the Rubin Operations Team, RDO-013). The worldwide community is actively preparing to process the LSST alert stream and use it to generate groundbreaking scientific results. Additionally, for many science goals, time-sensitive follow-up observations after discovery are crucial to take full advantage of the Rubin data.

A key component of the Early Science Program is the capability to build *incremental* templates from on-sky imaging as it becomes available during commissioning and the early phases of the survey. Such templates will be built periodically as images accumulate to allow for partial alert generation over an incomplete sky footprint. Where possible, templates will be built from all available commissioning data before the start of year one and used to generate alerts during year one. How extensive these templates are at the start of full survey operations will be influenced by the overall success of commissioning. During year 1, templates will be built progressively from data obtained during year one (e.g., on a monthly timescale), and used to generate alerts during year one, either instead of, or in addition to using commissioning data to build templates.

The data acquired during commissioning will be released to the Rubin data rights community via Data Previews and will include data products for both static-sky science and time-domain science.

2 Introduction

This note describes the plan for ensuring the Rubin community will have the data products and services necessary to produce early science during the time between commissioning and the first data release in year 1. The current planned start date for full survey operations is 1 April 2024, with considerable uncertainty. As of November 2021, the construction project schedule shows completion between December 2023 and July 2024 (i.e. planned finish and finish with contingency). It is unlikely the project will finish in December 2023 and highly likely it will finish by July 2024.

2.1 Definition of Early Science

Early Science (ES) is defined as any science enabled by Rubin for its community through and including the first data release, Data Release 1 (DR1). DR1 will be based on the first 6 months of data and is scheduled for release 12 months after the start of full survey operations, (Marshall, RDO-011). Community expectations for early science are high due to the transformative nature of the Rubin data and the extensive amount of on-sky time planned in commissioning and science validation.

2.2 Motivations for an Early Science Program

The Early Science Program is motivated by the fact that in the current baseline there will be no science-ready data products released to the community before Data Release 1 (DR1). This is due to the fact that the templates are needed for Difference Imaging are produced in Data Release Production (DRP) and initially released as part of Data Release 1 (DR1). Alerts cannot be issued until a template image for a given field and filter is available, hence in the current baseline, no alerts can be issued until after Data Release 1 (DR1).

The success of Early Science then depends on various scenarios coming out of commissioning as we transition into operations, as described in § 2.3.

2.3 Early Science scenarios

Recent planning on the construction project has led to a reduced amount of on-sky time in commissioning, including a reduction in the time dedicated to final science validation of the as-built system compared to earlier draft plans. The total amount science validation time is currently planned for 8 weeks. As Rubin construction moves through the challenging phase of System Integration, Test and Commissioning (SIT-COM), on–sky time could be further reduced. The Operations team is thus planning for various outcomes that might require special attention to producing Early Science opportunities in the first part of regular operations to ensure the community has prompt access to science-ready data products while the survey begins its relentless coverage of the sky leading to DR1.

In all cases, it is assumed that the Rubin Construction project delivers an integrated system that can capture, transfer and process science-quality data at the time Operations begins. Planning will then consider the following three high-level options to ensure Early Science:

- **Option A**: Science validation is completely successful. Begin the LSST working towards DR1 while providing Early Science data products.
- **Option B:** On-sky time in commissioning was reduced such that fewer science-ready data products were produced before the LSST begins but the Rubin system can capture and produce science-quality data. Instigate a dedicated Early Science observing campaign for a limited period (3-6 months) at the start of full operations that is different to the regular survey cadence.
- **Option C:** The Rubin System passes the construction completion requirements and can capture and produce science-quality data but the operations team is not yet ready to begin the LSST. Further shakedown of operations procedures and data taking is required, which will result in a delay to the start of full survey operations. This option will encompass elements of options B and C as appropriate.

All options will include alert generation of some type, with the major distinction being the relative availability of templates in time, sky position, and filter. The principle aspect of this strategy both in SIT-COM and year 1 operations is incremental template generation. In full survey operations, template images for difference image analysis and alert generation are constructed as part of the annual DRP. In order to support alert generation in year 1, Rubin will incrementally generate templates during SIT-COM and year 1 using the best images available and covering as much sky in as many filters as possible. Details of the current strategy for alert generation with incremental templates are given in § 5.

Rubin pre-operations is working with the construction project to provide a series of Data Previews based on the data acquired during the SITCOM Science Validation Surveys, as well as any additional science-quality data taken throughout the full commissioning period (Marshall, RDO-011).

These options are current at the time of writing, but as this document is ”living”, we expect the plans to mature as we approach full survey operations and the extant SIT-COM program emerges and is executed. At some future point, a single option will be adopted and executed, and at that time, the details will be more fully specified.

3 Roadmap and Timeline

This timeline provides a list of key dates related to the Early Science program. It will be updated with more dates related to the development of the Early Science Program as they are defined.

2021-11-12: Issue first version of the Rubin Observatory Plans for an Early Science Program.

2023-01-20: Commissioning Camera (ComCam) on sky, engineering first light.

2023-07-01: LSST Science Camera (LSSTCam) on sky.

2023-10-01: Start of the Science Validation Surveys.

2023-12-18: Current earliest completion date for construction.

2024-07-01: Current latest completion date for construction.

2024-04-01: Planned start of Operations Phase.

2024-04/09: Current forecast date for Data Preview 2, LSSTCam On-Sky Data.

2025-02/07: Current forecast date for Data Release 1.

4 Science Drivers

It will not be possible to survey the whole sky in all filters and generate templates by the end of the commissioning period. A strategy for template generation in the early phases of the survey, which will require balancing a tradeoff between various factors such as smaller area with multiple filters vs a single filter over a large area, must be devised. Different science drivers naturally lead to different prioritization strategies, e.g. milky way science would prefer templates that cover the galactic plane, time domain science would prefer templates in multiple bands rather than a single band for a larger area. Supernova, transient and variable science strongly advocate for templates for all bands in the Deep Drilling Fields to be prioritized. Rubin Operations will work closely with the science community to develop a science-driven approach to template generation in the early phases of the survey that will benefit the maximum number of science cases.

4.1 Time Domain

The Transients and Variable Stars Science Collaboration (TVSSC) reviewed the opportunities for Early Science for non time-critical and time-critical science cases in (Hambleton et al., 2020) and (Street et al., 2020) respectively. In both cases, they recommend the prioritization of template acquisition in multiple bands as the preferred strategy rather than single-band coverage over a larger area of sky.

4.2 Solar System

The Solar System Science Collaboration (SSSC) reviewed opportunities for Early Science in (Schwamb et al., 2021) for several high impact solar system science opportunities that would be enabled by accelerated template generation and alert production in year 1. They find that template generation options that maximize the sky coverage in year 1 where LSST Solar System Processing can run daily are strongly preferred, even if the templates result in noisier image subtraction compared to later years.

4.3 Static Science

Datasets for static science will flow from the Science Validation Surveys carried out during commissioning and released as Data Preview 2 (DP2). The commissioning team are planning to acquire on-sky observations that would enable science validation studies for the four LSST science drivers. Guidance is being sought from the community to enhance opportunities for science validation and early science based on commissioning data.

4.4 Target of Opportunity

Rubin Observatory will be prepared to take advantage of Targets of Opportunities (TOO) in the first year of operations (and hopefully SIT-COM). [RTN-008] describes potential data processing scenarios for TOO observations in the early operations era.

5 Alert Production in Commissioning and Early Operations

5.1 Processing Overview

The DPDD summarizes the pipelines which will be used during Prompt Processing to produce alerts as well as other prompt data products (§7), including Solar System Processing. In brief, raw images have instrument signatures removed and are photometrically and astrometrically calibrated. When template images for the corresponding region of the sky are available, the template is subtracted from the new processed visit image and sources are detected on the image difference. Alerts are then generated for all DIASources detected at five sigma in the difference. At the end of the night, DIASources without a history of previous detection are input into Solar System Processing, which attempts to link them with other past detections and identify new Solar System objects.

Both Alert Production and Solar System Processing thus depend on the presence of template images. During steady-state operations, these templates will be constructed during the annual Data Releases and will built from the best available subset of images taken. The input images for DRP-produced templates will accordingly have very good seeing and comprehensive spatial coverage. Coadding multiple images enables artifact rejection [DMTN-080] and reduces noise. All of these template characteristics all help to ensure that image differencing is highly complete and highly pure.

To enable alert production to proceed during commissioning and early operations, it is necessary to accept templates of lower quality. Because we have a smaller set of input images to choose from and uncertain knowledge about future observations, on-the-fly (or incremental) template generation necessarily must balance the trade off of earlier template availibility against template quality and spatial completeness. A template constructed today will enable alerts tomorrow, but that template might produce fewer or lower-quality alerts than one constructed from more data in a week or a month's time. Substantial validation will be required to determine when to build incremental templates to maximize the net throughput of Early Science. Nevertheless our goal is to enable Alert Generation to begin as soon as the data are scientifically useful.

Coadding multiple images is formally required due to the noise-level requirements placed on the DM system. Additionally, the LSST survey is heavily dithered, so without coadding many images onto a common sky plane it is both difficult and inefficient to obtain image differences for a new pointing from past single images. Finally, single-image templates do not permit removal of artifacts, transients, and moving objects from the template, creating additional false positive sources in the resulting differences.

Scientifically it is important that once a template is constructed for a given region of sky, it is used exclusively until it can be updated in the next Data Release. Repeated changes to the template make it extremely difficult to construct usable lightcurves for objects from individual difference image sources: transient objects such as supernovae will be contaminated by changing flux levels from the evolving template, and variable objects such as variable stars and AGN will require repeated corrections for different template flux levels as well.

5.2 Supporting Incremental Template Generation

The Rubin Construction Data Management (DM) Science team (DM-SST), carried out a study of several options for Alert Production in Year 1, reported in DMTN-107 : Options for Alert Production in LSST Operations Year 1. Representatives of the Rubin Project Science Team (PST), DM-SST and Operations reviewed the proposed DM-SST options and converged on a the following strategy for Alerts in year 1:

- Commissioning Data Templates: Build templates, where possible, from all commissioning data before the start of year one, and use them to generate alerts during year one.
- Year One Data Templates: Build templates progressively from data obtained during year one (e.g., on a monthly timescale), and use them to generate alerts during year one,

either instead of, or in addition to using commissioning data to build templates.

To handle alert generation outside the template building process attached to the annual DRP, the Construction project initiated a change request to include incremental templates in the DM system workflow. This change has been accepted and is now part of the baselined DM project in constructiom. A summary of the changes is the following:

- LCR-2273: Construct Image Differencing Templates Outside DRP, new requirement 1.4.6 Template Coadds ID: DMS-REQ-0280, The DMS shall periodically create Template Images in each of the u,g,r,i,z,y passbands. Templates may be constructed as part of executing the Data Release Production payload, or by a separate execution of the Template Generation payload. Prior to their availability from Data Releases these coadds shall be created incrementally when sufficient data passing relevant quality criteria is available.
- To enable artifact rejection, templates will be built with at least three images in year one, and five in subsequent years. (Rubin OSS-REQ-0158)
- Once a template is produced for a sky position and filter it will not be replaced until the next Data Release to avoid repeated baseline changes.
- Templates are not necessarily built from the first N images that are collected.

6 Rubin Observatory Commissioning

The baseline schedule for on-sky observations during SIT-COM includes six months of technical integration and testing, and concludes with an 8-week period of sustained observing in the form of two Science Validation Surveys (Claver et al., SITCOMTN-005).

6.1 Template generation during commissioning

By the end of the commissioning period, coadd templates for use in difference imaging will only be available for $\approx 10\%$ of the sky. Generating templates over a wide area is not an explicit goal of commissioning; however, where possible, if commissioning observations are agnostic to pointing and filter, we would endeavour to choose a pointing and filter that maximizes building templates to enable early science. During LSSTCam commissioning we intend to incrementally generate templates over the maximal sky area supported by the available observations.

The LSST SRD places well-defined criteria on the quality of the difference image and the amount of noise that a template can contribute to a difference image. These criteria result in a minimum of three images being needed to construct a template for use in year one. The commissioning period provides an excellent opportunity to investigate how many visits in a given band are sufficient to construct a usable template. Given the desire to maximize the science harvest prior to the Data Release 1 (DR1), relaxing these criteria is an option to be explored.

6.2 Alert generation during commissioning

Due to the need to verify the instrument characteristics, template quality, and image differencing and Real/Bogus performance, real-time alerts will not be immediately available during the commissioning period. Where the accumulated ComCam data is sufficient for alert generation, we expect to provide alerts at high latency (weeks–months). The goal for these *canned* alerts is to enable alert brokers and science users to understand their characteristics and to help to validate their quality rather than to enable rapid followup and Early Science per se. Templates generated during commissioning will be used for Alert Production, with the goal of delivering real-time alert distribution to community brokers by the time of the Science Validation Surveys at the end of LSSTCam commissioning.

6.3 Data Previews based on commissioning data

Data acquired during the Science Validation Surveys is expected to be of science-quality and will be released to the Rubin data rights community via two Data Previews, Data Preview 1 (DP1) for data from the commissioning camera (ComCam) and Data Preview 2 (DP2) for data from the LSST science camera (LSSTCam) and all previous commissioning data. Data Previews will be produced using the DRP pipeline and will include data products for both static sky science and time domain science.

7 Early Science Data Products

Here we provide a summary of the data products that are expected to be made available as part of the Early Science Program. The definitive source for all LSST data products is the Data Products Definition Document, (Jurić et al., LSE-163). The Rubin data rights policy is described in Blum & the Rubin Operations Team (RDO-013).

7.1 Prompt data products

Prompt data products are described in detail in the Data Products Definition Document (DPDD). Alert packets are triggered by difference image source detections and transmitted to community alert brokers[1] and are publicly available. Similarly, daily Solar System Processing identifies new Solar System Objects from difference image sources and reports those publicly to the Minor Planet Center.

Catalog and image products, as well as services for running user-generated processing on the data, are available to Rubin Data Rights holders after 24 hours through the Rubin Science Platform (§ 7.3). `DIASource`, `DIAObject`, and `SSObject` catalogs are queryable using VO interfaces to the Prompt Products Database.

7.2 Data Release data products

Static science datasets for Early Science will flow from the Science Validation Surveys in commissioning. Images and catalogs from the DRP of the commissioning data will be made available to data rights holders in the form of Data Previews via the access mechanisms described in 7.3. Due to the relatively short time periods available for commissioning observations (§2.3), these Data Previews will necessarily be limited in their area and temporal coverage relative to full Data Releases, however all Data Preview data products will be in the same science data model format as for future Data Releases.

[1] See `https://www.lsst.org/scientists/alert-brokers` for the list of selected brokers.

7.3 Access to Early Science data products

Alerts will be accessible by the community via one or more of the nine Rubin-endorsed Community Brokers. The Rubin Data Rights community will access the Rubin data products via the Rubin Science Platform, (Jurić et al., LSE-319).

8 Survey Cadence

Early Science observations should align as closely as possible with main survey and ultimate long-term science goals Details on the implications for the Survey Cadence of the Early Science Program will be added in future as the outcome of commissioning becomes clear.

9 Community Engagement

Rubin Observatory will work closely with the community on the detailed design of the Early Science Program.

9.1 Survey Cadence Optimization Committee

The Survey Cadence Optimization Committee (SCOC) is an advisory committee to the Rubin Observatory Operations Director consisting of 10 members drawn almost entirely from the science community. The SCOC was convened in 2020 and will be a standing committee throughout the life of Rubin Observatory operations.

Early Science observations should align as closely as possible with the main survey and ultimate long-term science goals; the SCOC will be involved in all aspects of development of the Early Science Program. Specifically, the SCOC will make specific recommendations for Early Science observations, based on the plans for commissioning and the realized performance of the telescope and software.

9.2 Community Forum

The Rubin Observatory Community Platform has a dedicated category for Early Science[2], where community members are encouraged to open discussions on the topic of early science.

9.3 Community Input

A process will be put in place to formally solicit input from the community. Several science collaborations have already been pro-active in providing input on considerations for template generation in year one on both the community forum and as research notes.

A References

[DMTN-080], AlSayyad, Y., 2018, *Coaddition Artifact Rejection and* `CompareWarp`, DMTN-080, URL `https://dmtn-080.lsst.io`,
LSST Data Management Technical Note, draft version

[RTN-008], Bellm, E.C., 2020, *LSST Processing of Gravitational Wave TOO Data in the Early Operations Era*, RTN-008, URL `http://RTN-008.lsst.io`

[RDO-013], Blum, R., the Rubin Operations Team, 2020, *Vera C. Rubin Observatory Data Policy*, RDO-013, URL `https://ls.st/RDO-013`

[SITCOMTN-005], Claver, C., Bauer, A., Bechtol, K., et al., 2021, *Construction Completeness and Operations Readiness Criteria*, SITCOMTN-005, URL `https://sitcomtn-005.lsst.io`,
LSST SITCOM Technical Note

[DMTN-107], Graham, M.L., Bellm, E.C., Slater, C.T., the DM System Science Team, 2019, *Options for Alert Production in LSST Operations Year 1*, DMTN-107, URL `https://dmtn-107.lsst.io`,
LSST Data Management Technical Note

[2]See `https://community.lsst.org/t/about-the-early-science-category/5775`

Hambleton, K., Bianco, F., Clementini, G., et al., 2020, Research Notes of the AAS, 4, 40, URL `https://doi.org/10.3847/2515-5172/ab8129`, doi:10.3847/2515-5172/ab8129

[LSE-319], Jurić, M., Ciardi, D., Dubois-Felsmann, G., 2017, *LSST Science Platform Vision Document*, LSE-319, URL `https://ls.st/LSE-319`

[LSE-163], Jurić, M., et al., 2017, *LSST Data Products Definition Document*, LSE-163, URL `https://ls.st/LSE-163`

[RDO-011], Marshall, P., 2020, *Release Scenarios for LSST Data*, RDO-011, URL `https://docushare.lsstcorp.org/docushare/dsweb/Get/RDO-11`

Schwamb, M.E., Jurić, M., Bolin, B.T., et al., 2021, 5, 143, URL `https://doi.org/10.3847/2515-5172/ac090f`, doi:10.3847/2515-5172/ac090f

Street, R.A., Bianco, F.B., Bonito, R., et al., 2020, Research Notes of the AAS, 4, 41, URL `https://doi.org/10.3847/2515-5172/ab812a`, doi:10.3847/2515-5172/ab812a

B Acronyms

Acronym	Description
AGN	active galactic nuclei
ComCam	The commissioning camera is a single-raft, 9-CCD camera that will be installed in LSST during commissioning, before the final camera is ready.
DM	Data Management
DM-SST	DM System Science Team
DMS	Data Management Subsystem
DMS-REQ	Data Management System Requirements prefix
DMTN	DM Technical Note
DP1	Data Preview 1
DP2	Data Preview 2
DPDD	Data Product Definition Document
DR1	Data Release 1
DRP	Data Release Production
ES	Early Science
LCR	LSST Change Request
LSE	LSST Systems Engineering (Document Handle)

LSST	Legacy Survey of Space and Time (formerly Large Synoptic Survey Telescope)
MAF	Metrics Analysis Framework
OSS	Observatory System Specifications; LSE-30
PST	Project Science Team
RDO	Rubin Directors Office
RTN	Rubin Technical Note
SCOC	Survey Cadence Optimization Committee
SIT	System Integration, Test
SITCOM	System Integration, Test and Commissioning
SRD	LSST Science Requirements; LPM-17
SST	Subsystem Science Team
TOO	Target of Opportunity
VO	Virtual Observatory