Lasair Software Management Plan for LSST:UK Phase C



|  |  |
| --- | --- |
| **Project Acronym** | LUSC-C |
| **Project Title** | UK Involvement in the Large Synoptic Survey Telescope |
| **Document Number** | LUSC-C-?? |

|  |  |
| --- | --- |
| **Submission date** | day/month/year |
| **Version** |  |
| **Status** | Draft |
| **Author(s) inc. institutional affiliation** |  |
| **Reviewer(s)** | Person 1 (Org. short name),  Person 2 (Org. short name) |

|  |  |
| --- | --- |
| **Dissemination level** | |
| <Public/ Confidential/ Restricted> | *<Choose from: Public; Confidential – only for members of the LSST:UK Consortium (including sponsors)>; Restricted – other constraints to be specified* |

# Version History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Comments, Changes, Status** | **Authors, contributors, reviewers** |
| 0.1 | DD/MMM/YY |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Table of Contents

[Version History 2](#_Toc496866495)

[1 Executive Summary 2](#_Toc496866496)

[2 Introduction 2](#_Toc496866497)

[2.1 Purpose 2](#_Toc496866498)

[2.2 Glossary of Acronyms 2](#_Toc496866499)

[3 Section Heading 2](#_Toc496866500)

[3.1 Subsection Heading 2](#_Toc496866501)

[3.1.1 Sub-subsection heading 2](#_Toc496866502)

[‘Small Heading’ to break up longer sections of pros 2](#_Toc496866503)

[3.2 Header level 2 2](#_Toc496866504)

[3.2.1 Header level 3 2](#_Toc496866505)

[4 References 2](#_Toc496866506)

[Annex A. Name of Annex 2](#_Toc496866507)

Index of Figures

[Figure 3 Example figure inserted using [Insert][Caption][Figure] 2](#_Toc432517668)

Index of Tables

[Table 1 Example table inserted using [Insert][Caption][Table] 2](#_Toc432517669)

# Executive Summary

All deliverables, with more than four pages of content, should include an Executive Summary. The Executive Summary shall include principal findings and conclusions.

Recommended length of this section: not more than 2 pages.

# Introduction

## Purpose

This is the Lasair Software Management Plan for Phase C of the STFC-funded LSST:UK project.

This differs from the Phase B version in that it is formatted as an LSST:UK Technical Report and more importantly, since Lasair has been accepted as a Rubin Observatory Community Broker and as an In-kind Contribution (id. UKD-UKD-S4), the contents of the plan have been updated to reflect compliance with the Coding Guidelines as expressed in the Rubin Observatory “Manual for In-kind Contributors and Recipients” (ls.st/RDO-41).

Under LSST:UK’s Software Development Requirements (see <https://lsst-uk.atlassian.net/wiki/spaces/LUSC/pages/624918539/Software+Development+Requirements>), software produced by the UK team is required to have a software management plan produced at its outset that is subsequently maintained and reviewed throughout its development.

The previous version of the Software Management Plan was created at the outset of the LSST:UK Phase B. This was based on the suggested checklist template from the Software Sustainability institute (see <https://lsst-uk.atlassian.net/wiki/download/attachments/624918539/SoftwareManagementPlanChecklist.pdf?version=1&modificationDate=1575389139881&cacheVersion=1&api=v2>). This checklist comprises a list of questions. Section 3 contains the answers to the questions from that checklist. These answers have been updated to reflect the Lasair activities in Phase C of LSST:UK.

The contents of Section 4 indicate how the Lasair software development in Phase C complies with each of the 12 items in the Rubin Observatory RDO-41 Coding Guidelines.

## Glossary of Acronyms

# Software Management Plan Checklist

## What software will you develop?

* We will develop a broker, called Lasair, which ingests high-bandwidth streams of transient astronomical observations and filters out the interesting substreams that can advance science, as well as adding useful additional information.
* There is an advanced prototype (lasair.lsst.ac.uk) that serves the ZTF survey, and a new version will serve the LSST survey when it starts in 2025.
* Lasair is composed of a set of ingestion modules, coordinated via a Python script, with web pages, web services and Kafka services. Lasair is built with the following high-level components:
  + web/kafka server,
  + ingestion pipeline,
  + database,
  + context classifier.
* These services are built from third-party tools including: Apache, Django, Kafka, MariaDB, Galera, Cassandra , plotly, jQuery A fuller list is maintained by the project team.

## Who are the intended users of your software?

* The Lasair system is open to the public, but we focus on professional astronomers in the UK.
* Users wish to extract what their science wants from a wide and diverse stream of astronomical transients.
* There are two levels of authentication: public and free, self-signup. The second requires only a valid email address.
* We have built a "citizen science" project to identify super-luminous supernovae.

## How will you make your software available to your users?

* The Lasair software is available open source from github repositories [https://github.com/lsstuk/lasair-lsst and lasair4], if anyone would like to replicate the system.
* However, we do not expect many users to run the software themselves; rather they will access an instance of the software that we will run on their behalf.
* There is a client library that enable users to interact with the Lasair API.
* The system presents as a website at which filters and queries can be built by users; these can be executed in "pull" mode where a user (or API) makes a request, or where interesting observations are "pushed" to the user as soon as possible.

## How will you support those who use your software?

* We have built documentation in many forms: quickstart, how-to as text and video, FAQ, etc.
* Channels available to users are: email to Lasair helpdesk, and the Rubin Community forum (see (<https://community.lsst.org/c/support/support-lasair/55>). The preferred route is the latter.
* Those who have accounts with Lasair will get occasional emails about upgrades / new releases.

## How will your software contribute to research?

* The Lasair system will enable astronomers to get what they want from a billion dollar science survey (LSST), by filtering and storing the firehose of data.
* Specifically, Lasair will improve understanding of supernovae, active galaxies, stellar outbursts, planet formation, and solar system objects.

## How will your software relate to other research objects?

* Lasair is useful only because it consumes an astronomical survey (ZTF, LSST), which is itself a large software infrastructure combined with a large telescope.
* Lasair was originally described in Smith, Williams, et. al, https://iopscience.iop.org/article/10.3847/2515-5172/ab020f
* The software is described in <https://lasair.readthedocs.io/>.
* How will you measure your software's contribution to research?
* The Readthedocs at its front page asks users “If you make use of this, please cite our paper:…”

## Where will you deposit your software to guarantee its long-term availability?

* The scripts that can configure and deploy an instance of Lasair, are held on the third party GitHub service, which we expect will persist for the long term, due to the volume of important software that it hosts

# Compliance with Rubin Observatory RDO-41 Coding Guidelines

In this section, the Rubin RDO-41 guidelines are shown below in *blue italics*. Below each of these guidelines, the text explains if and how these are being addressed by the Lasair team in Phase C of LSST:UK.

1. Be designed in collaboration with the recipients, following the recipients’ standard practices.
   1. Use cases, algorithm choices, interfaces (to other packages and target datasets), and (as needed) code structure should all be discussed using the recipients’ communication channels (e.g. LSSTC Slack, recipient Jira or GitHub, working group meetings, etc).

The Lasair team holds occasional review meetings with its user community, the most recent being in Oxford, UK on September 13th and 14th, 2023 (see https://lsst-uk.atlassian.net/wiki/spaces/LUSC/pages/3210018983).

1. Be developed collaboratively, in a shared version-controlled repository accessible to the recipient group, such that the recipients can follow progress, comment and query, and make contributions of their own.
   1. Commits should represent atomic (i.e. small and indivisible) changes in functionality.
   2. The code should compile, and its tests should pass, before the change is committed.
   3. The commit message should be an informative summary of the change.
   4. Pull requests should be limited to a single feature, and code review practices should be agreed in advance with the recipients (e.g. to prevent reviewer overload)
   5. New contributors should be encouraged and assisted, as a way of ensuring the maintenance and reusability of the code.
   6. Ideally the repository should be open (i.e. publicly visible or available to the Rubin Community) or, as a minimum accessible to the recipient group(s) during package development. One of these must be agreed on with the recipient(s) and noted in the contribution work plan.

Section 3.3 above, explains how the Lasair is developed using an open Github repository. The Lasair team have agreed a pull request process for software commits with developers strongly encouraged to add tests that are executed as a part of the Lasair Jenkins CI set-up.

1. *Be packaged using common, easily used tools (such as setuptools and pip/pypi in Python, Autotools and cmake in C++, etc).*

Lasair itself is not a software package as such rather it is a whole system hence it is not expected that others will build a Lasair instance. However its API is packaged using setuptools/PyPi.

1. *All functions, packages, dependencies and datasets of or related to the software and necessary for it to function, should also be easily available to the community without licensing or other restriction.*

The Principles of Operation lists the Lasair software package dependencies (see <https://lsst-uk.atlassian.net/wiki/spaces/LUSC/pages/2763161628>).

1. *Adhere to reasonable standards, such as those already adopted by the recipients’ related packages, or PEP 8 (Python) and LSSTDM (C++), with style optimized for readability.*

No specific standard is currently followed since it is a small team but there are comments, doc strings, tests and the previously mentioned Principles of operation. As yet no problems have been encountered as result of the lack of a specific style guide.

1. *Include a test suite that uses common testing tools, and which is either used by a continuous integration service (like GitHub or travis-CI) or could be used by one.*

Test suite is activated using Lasair’s the Jenkins CI set-up at https://lasair-jenkins.lsst.ac.uk. There is an ambition to further extend test coverage.

1. *Be demonstrated and validated following the recipients’ recommendations, preferably via reusable notebooks or scripts that are checked into the package repository.*

Collab notebooks are documented in the Lasair readthedocs at ,

https://lasair.readthedocs.io/en/develop/core\_functions/python-notebooks.html

In addition the lasair-examples Github reporsitory s a Github repo (https://github.com/lsst-uk/lasair-examples) contains example code for Lasair users.

1. *Be fully documented at all times, such that at any point in time the package can be picked up and contributed to by any skilled developer in the recipient group. In addition, a simpler User Guide must be provided for those not accessing the software for development but as end users.* 
   1. *The set of function, class and module docstrings should contain everything the user needs to understand in order to use the code. The set of in-line comments should provide a complete explanation of (and citations to) the algorithms implemented.*
   2. *The README should enable the recipients to understand what is being developed and how to get involved.*
   3. *Tutorials should show what the package does, by leading the user through the set of use cases that define the goals of the package. Jupyter notebooks are particularly powerful for teaching Python packages in this way, and can even be used as system tests.*
   4. *Technical notes (and potentially a journal paper) describe the problem that the package is intended to solve, and what algorithms it implements to solve it.*

The user functionality is fully documented in readthedocs (see https://lasair.readthedocs.io/en/main/). Skilled users have the previously mentioned API, colab notebooks and Lasair examples. There are also some videos.

Whilst there is developer documentation, this is limited in its scope.

1. *Be versioned semantically (using e.g. GitHub releases) to improve communication about the code and reproducibility of its outputs.*

Versioning is used in the API and the API client code.

The development team do use versioning but are reluctant to use the conventions of semantic versions elsewhere since previous attempts did not offer value and did not improve communication and reproducibility.

1. *If appropriate, be published in a suitable journal (such as the Astronomical Journal, which encourages method papers with associated software packages), with the recipients who made significant contributions to the development of the code as co-authors, in accordance with the recipients’ publication policy. The main developers are responsible for publishing the code, but can optionally delegate the paper writing to a recipient collaborator(s). In this case the developers should be credited as early authors on the paper based on the code.*

Lasair was originally described in Smith, Williams, et. al, (see https://iopscience.iop.org/article/10.3847/2515-5172/ab020f)but it would benefit from a follow-on. Expect to produce a new paper when LSST is ready.

1. *Be released publicly no later than the time of its first application in a journal paper, unless the recipients have scientific reasons for postponing and successfully petition the Rubin Operations Director (via their IPC) for an exception. (The Rubin Operations Director will take the CEC’s advice on the requested exception, and may delegate their ruling to the Rubin data policy committee or publication board.)Be versioned semantically (using e.g. GitHub releases) to improve communication about the code and reproducibility of its outputs.*

Lasair is available publicly as explained in section 3.3

1. *Carry a permissive open source license such as BSD-3-Clause. This is to maximize the re-useability of in-kind generated code by the LSST Science Community (and has the added benefit of high visibility for the code’s authors)Be versioned semantically (using e.g. GitHub releases) to improve communication about the code and reproducibility of its outputs.*

Lasair has an Apache 2.0 License, see

https://github.com/lsst-uk/lasair-project-management/blob/master/LICENSE

# References

References should be formatted as a numbered paragraph, which can then be referenced dynamically from text using [Insert][Cross Reference][Numbered Item]

1. Title of Deliverable, Project Deliverable Dx.x.x.
2. Title of Deliverable, Deliverable Dx.x.x of another project.
3. Author 1, Author 2, "Title of article" Publisher, ISBN xxxxxx.
4. Name of Annex
   1. Annex section
      1. Annex subsection