Data security for Rubin communication channels

William O'Mullane, Frossie Economou

2024-04-10

1 Introduction

As outlined in DMTN-199 only the Rubin team have access to pixel data for an embargo period of 80 hours in operations and 30 days during commissioning. We assume this also means PNG or other format images captured from screens, etc., which may show pixel data, e.g., from RubinTV on the summit.

Secure or authenticated links to pixel images are mentioned in several instances below. These could be any pointer to an image which one can only access by login - hence verifying one has access to the image. Examples could be a butler dataId, a jupyter notebook to be run at USDF, an unsigned S3 URL[1].

2 General policy

Many scenarios are discussed in SITCOMTN-076 concerning information sharing in commissioning. In general we may state:

No pixel images seen by the team in any communication space should be shared outside the space at anytime. Images should not be left open on laptops at meetings or public spaces. Images should not be shown in in presentations.

This is even more important before System First Light where we need to carefully control the public first light images.

In general, we are trying to maintain a broadly open system for Rubin. This implies that team members and collaborators have access to a wealth of other information about Rubin, aside from pixel data. During commission, much of this information will be about problems. We do not wish to hide the issues and problems, and we expect the community to be respectful of

[1] starting with s3:// but not usable in a browser

this information and consider the state of the system, which in commissioning is not finished.

3 Systems that need policies

In various discussions the list of tools we need to have polices for are:

1. Confluence
2. Jira
3. github (e.g., draft technotes, Construction paper drafts, notebooks with plots)
4. technotes on lsst.io
5. Slack
6. summit tooling (i.e RubinTV)
7. USDF tooling
8. email lists

We discuss each in a section below. Tech notes and Slack are possibly most heavily used so lets start there.

3.1 Slack

We have a wide variety of very open channels on Slack, these include large numbers of science collaborators and member of the astronomy community. So far, AuxTel and Simonyi star tracker images have been shared frequently on certain Slack channels which are not private. There is the possibility to have private channels on Slack — we have few and tend to avoid them for several reasons: they are not discoverable, some bots don't work in them, etc. Even using Private channels, we would face a problem of members potentially inviting non-staff to the channel where images may be shown.

The current (LSSTC) Slack workspace was intentionally conceived as a community (including LSST Science Collaborations) inclusive space. This is at odds with a number of goals:

- All channels would be available to all staff in the space — no problem sharing images as long as ALL agree they never leave the space. This allows us a great project space to enjoy successes as they arise.
- Fail quietly: We need a space where we can openly discuss problems (e.g. "omg what if we crack the AuxTel array" discussion in an open channel)
- Embargo: Freely post data without having to stop and think whether it is embargoed.
- GitOps: Slack is "the UNIX command line" for highly distributed teams. SQuaRE offers (and plans on expanding its offerings) of slackbots that actively manipulate project resources, report on project data, issue alerts on operational services etc. Being assured that only trusted staff have access allows us to expand what these services can do.
- High Priority: This gives a high-importance lower volume workspace that we can less disruptively monitor out of hours, during holidays and vacations and overall busy times, reducing the communication attack surface.
- Privacy: We already pin phone numbers to certain channels, and there are concerns about freely sharing staff phone numbers, vacation schedules etc
- On/Off-boarding: Even if someone is off-boarded from the project, there are legitimate reasons they should still maintain their wider community slack access. A staff-only slack can be tightly controlled together with other high value project access.
- Slack Culture: We have issues where slack cultures class, eg. community folks at-channel just because a seminar is about to start, respecting quiet days etc. A separate workspace can maintain a more ops-oriented slack culture. It is also easier to respond to inappropriate behavior when it's your own staff engaging in it.

There are some reasons we may not with to have a private Slack:

- "We are at risk of abandoning the community": The motivation behind LSSTC was to establish good working relationships with non-project staff, particularly DESC. This was at a time when Slack had very poor support for multiple workspaces and guest channel access. Moreover the LSSTC workspace is now well established for this purpose and ops leaders will remain reachable on it.

- "It's fine, we have private channels": The proliferation of private channels have been a bit of a nightmare. We're always wondering who should or should not have access to them, forget to add new people, forget to off-board departing people, are unclear what channels should or should not be private, etc. It also has led to reduced transparency within the project — a number of channels that were open in the old slack space were made private when we moved to the LSSTC slack.
- "Too hard to determine out who is staff": It's true that the proliferation of in-kind, grad students and other participants have muddied the waters. However this is a problem we still have in the current set-up - it's just less obvious. The most clear heuristics include "are we paying (corollary: can we "fire") someone who has violated project rules; are they on the builder's accrual list; do they have summit access; etc.
- "We already say we don't do community user support on Slack": This is true, we do say it, but we should recognize that it's emotionally hard work when someone is asking a question on Slack to determine whether they are staff or no, and if not to tell them to go elsewhere. Sometimes we just answer without checking, which muddies our stance.
- "We're too busy for this kind of change": True, but this will only get worse. There is also a plan being prepared to provide more consistent naming for certain types of channels (support, status, etc.) and update default status semantics, so this would be a good time to implement that.

Much as we dislike having ANOTHER Slack space resurrecting the old LSST Slack space adding only the team members with image access to it and using it for all the the nighttime summit channels would be a clean solution. Whether we make private channels or switch to a private space — this needs to be done preferably before July 2024. We need to start a concrete plan for this.

3.1.1 Proposed policy

See also section 2. Resurrect lsst.slack.com Slack org for restricted use of the Rubin team, at least during commissioning.

Reuse of lsst.slack.com as some advantages:

- This is already set up grandfathered as free so we don't risk applying and being turned

down for another free workspace

- Can deliberately blur construction / commissioning / ops lines since it is free
- Rubin would be in total control of configuration, access and any paid features
- Workspace is already configured for our use which would speed up any transition

We could revisit after commissioning.

3.2 Technotes on lsst.io

The lsst.io site is intentionally public.

Getting auth protection for some notes for some time is not very consistent and will lead to problems.

SITCOMTN-076 goes in more detail on technotes.

3.2.1 Proposed policy for technotes

Continue current default of public technotes with development occurring on branches. The review process described in SITCOMTN-076 occurs at the stage that a development branch is being merged to the main branch via a Pull Request. The content of a technote is considered to be approved for release, once merged to the main branch. During development, embargoed pixel images can only be referenced in technotes as authenticated links — pixel images (e.g., PNGs) must NOT appear in technotes until specifically approved for release

3.3 Github

There may be many things in repos in github, e.g., draft technotes, Construction paper drafts, notebooks with plots. These could be in private repos. Even so, private repos are visible to all team members on github so care needs to be taken.

An alternative would be to again use authenticated links to images until they are not embargoed and keep all notes public.

3.3.1 Proposed policy for Github

Rendered notebooks containing image data must not appear in public repos. Notebooks should be linked via Timesquare with authenticated access.

3.4 Confluence

Much of Confluence is public (all of DM is public by choice). SITCOM is restricted to login but not more than that. The main area where one may expect to see images would be performance analysis — an admittedly cursory glance at all attachments suggests there are not many. There are many links to notebooks — this is fine as notebooks require execution/access.

SITCOM would like a Confluence space where they could share presentations with potentially embargoed images. This could be allowed by making all such meeting pages private and ensuring the SITCOM group is appropriately restricted.

3.4.1 Proposed policy for Confluence

Do not up load any pixel images or parts of pixel images or screenshots of images to confluence. Use an authenticated link to any image rather than the actual image if it is needed in a page.

3.5 Jira

As for confluence all DM tickets are public in Jira. Other projects are password protected but anyone with access to Jira may view such tickets. Its not obvious we could even secure such a system.

3.5.1 Proposed policy for Jira

Do not up load any pixel images or parts of pixel images to Jira. Use an authenticated link to any image rather than the actual image if it is needed in a ticket. It is acceptable to put a screenshot of an effect on a CCD if it is pertinent to the issue. Some details of this type of

image data are given in Section 3.4 of SITCOMTN-076.

3.6 summit tooling (i.e RubinTV)

This all has to be secured for commissioning use. We should aim to put this in place around ComCam on sky July 2024.

3.6.1 Proposed policy for summit tooling

Summit tooling needs to go behind the Summit 2FA VPN which is in place as of 6 March 2024. RubinTV needs to have authentication added and be restricted to the same Summit IPA groups.

3.7 USDF tooling

USDF tooling should only be accessible to USDF account holders who are in the commissioning or operations teams.

3.7.1 Proposed policy for USDF tooling

Ensure only account holders in allowed groups have access to USDF tooling. Allowed groups are those involved in commissioning plus DM and SP staff involved in QA and essential operations trouble shooting at USDF.

3.8 Emails and email lists

Email is inherently insecure (few of us use encryption) and list servers are fairly open.

3.8.1 Proposed policy for Email

Do not attach Rubin pixel images or parts of image to emails. Use authenticated links to images where needed.

A Implementation details

We shall add details here on availability of

1. easy links to images
2. the project Slack workspace and timeline for moving over channels
3. Notebook tools
    - Render or export to pdf
    - Timesquare enhancement to render any notebook in an authenticated manner.
4. Closed google drive for presentations/docs

B References

[SITCOMTN-076], Bechtol, K., on behalf of the Rubin Observatory Project Science Team, S.R., 2024, Information Sharing during Commissioning, URL `https://sitcomtn-076.lsst.io/`, Vera C. Rubin Observatory Commissioning Technical Note SITCOMTN-076

[DMTN-199], O'Mullane, W., Allbery, R., AlSayyad, Y., et al., 2023, Rubin Observatory Data Security Standards Implementation, URL `https://dmtn-199.lsst.io/`, Vera C. Rubin Observatory Data Management Technical Note DMTN-199

C Acronyms

Acronym	Description
CCD	Charge-Coupled Device
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.
DESC	Dark Energy Science Collaboration

DM	Data Management
DMTN	DM Technical Note
LSST	Legacy Survey of Space and Time (formerly Large Synoptic Survey Telescope)
LSSTC	LSST Corporation
QA	Quality Assurance
S3	(Amazon) Simple Storage Service
SITCOM	System Integration, Test and Commissioning
SP	Story Point
SQuaRE	Science Quality and Reliability Engineering
URL	Universal Resource Locator
USDF	United States Data Facility
VPN	virtual private network