

Rubin Observatory

Rubin Data Policy

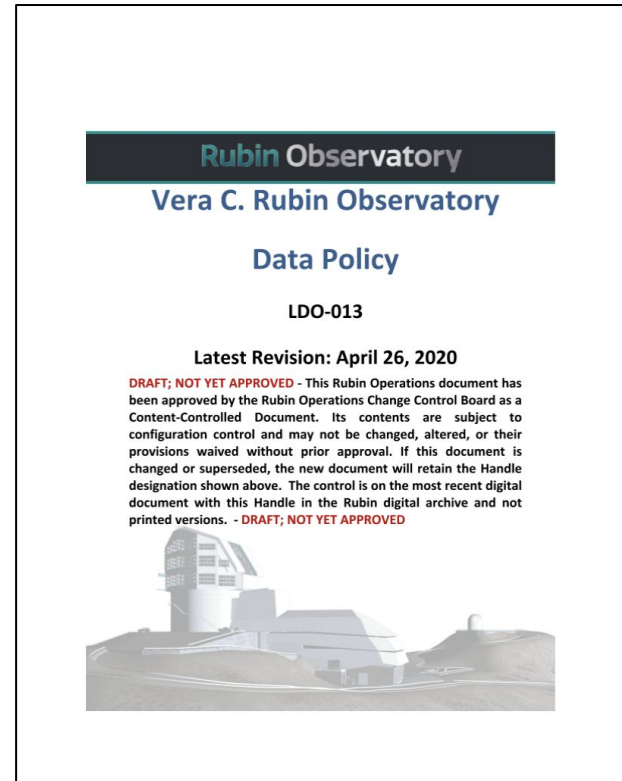
Robert Blum, Acting Director for Operations



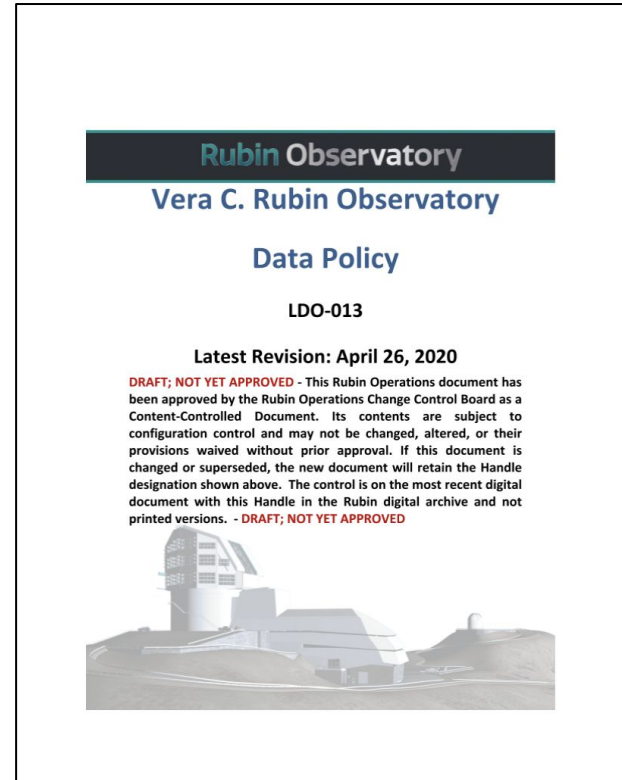
U.S. DEPARTMENT OF
ENERGY



- Approved version by NSF and DOE (November, 2019)
 - Now known as LDO-013 (<http://ls.st/LDO-013>)
 - To be placed under Operations Change Control
 - Public Draft, in force
 - Working with Science Collaborations on several concerns, feedback welcome anytime going forward
- Summary of LSST Data Management Principles, LPM-151 also updated (<http://ls.st/LPM-151>)



- Many updates leading into November approval by agencies
- Some requests for wording or DPOL modifications not yet made (i.e. under consideration).
- SAC input has some outstanding edits (11/24/2019)
- Only DESC has detailed policy/DPOL requests in to Ops team for consideration.
- No updates since November approval except to change to Rubin Data Policy
- **NSF and DOE approval needed**



Proposed: Builder status / permanent data rights for science collaboration "builders" (people who have made such large contributions to an SC infrastructure and community that the SC considers them eligible to co-author all collab papers)

- Incentivizes people to work hard, contribute, be rewarded
- Allows long term planning for young researchers and SCs
- Fair: Ops team supportive of this change
- We believe this needs approval from NSF and DOE depending on where data rights holders move to.

Observer Status

Proposed: Observer (i.e. temporary member) status for any international person needing access to SC information/communications while developing an in kind proposal.

- Rubin can make this happen on our side: Phil and Fed to work with Iain G
- What do SC's need to do?

Grace Period

- Why is the grace period in DPOL 404 specific to US and Chile?
- Data rights for international contributors have always been intended to be covered in the MOA between Rubin (or agencies) and the international group.
- NSF and DOE have signaled that all extensions to data rights for people moving out of the US or Chile may need agency level approval.

- Objection to language below DPOL-302

Beyond scientific use of the LSST data as it is obtained, the benefits of data rights include:

- The right to apply for membership in LSST-sponsored Scientific Collaborations (membership in Scientific Collaborations is entirely up to the Collaboration);
- Access to the LSST help desk and engagement with community science support groups;
- Consideration for membership in LSST community based committees (e.g., Science Advisory Committee, Users sub-committee, Cadence sub-committee)
- Participation of LSST staff at User meetings, similar to participation during Construction (on a best-effort basis).
- Access to the LSST Science Platform on an authorized LSST DAC (at present these include US and Chilean DACs);
- User account management with compute and database allocations.

- Objection to language below DPOL-302

Beyond scientific use of the LSST data as it is obtained, the benefits of data rights include:

- ~~The right to apply for membership in LSST sponsored Scientific Collaborations (membership in Scientific Collaborations is entirely up to the Collaboration);~~
- Access to the LSST help desk and engagement with community science support groups;
- Consideration for membership in LSST community based committees (e.g., Science Advisory Committee, Users sub-committee, Cadence sub-committee)
- Participation of LSST staff at User meetings, similar to participation during Construction (on a best-effort basis).
- Access to the LSST Science Platform on an authorized LSST DAC (at present these include US and Chilean DACs);
- User account management with compute and database allocations.
- Agree, will simply remove any reference to Science Collaborations
- SCs can work with non Data Rights holders

Public Data

- What does it mean to say the data are public?
- What is the project doing to enable public data serving?
- Can the SCs serve public data?

Reproducibility of Results

- What if some group wants to reproduce a Rubin LSST result? They must have data rights or at least access to public data (to complete an analysis once the data set based on a certain DR used for a (published) result has become public.
- This is a powerful argument for the project to enable access to the public data.
- Such requirements may also fall on Derived Data Products hosted by Rubin (use case not covered in current design).
- SAC had a related concern (availability of past DR for reproducibility)