Draft version February 19, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

Rubin Observatory Data Facilities

Adam Bolton,[1] William O'Mullane,[2] Kian-Tat Lim,[1] Fabio Hernandez,[3] Mark G. Beckett,[4] Michelle Butler,[5] Tim Jenness,[6] Orion Eiger,[1,7] Peter Love,[8] Timothy Noble,[9] Stephen Pietrowicz,[5]

Rubin Observatory Data Management Team

[1]*SLAC National Accelerator Laboratory, 2575 Sand Hill Rd., Menlo Park, CA 94025, USA*
[2]*Vera C. Rubin Observatory, Avenida Juan Cisternas #1500, La Serena, Chile*
[3]*CNRS, CC-IN2P3, 21 avenue Pierre de Coubertin, CS70202, F-69627 Villeurbanne cedex, France*
[4]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[5]*NCSA, University of Illinois at Urbana-Champaign, 1205 W. Clark St., Urbana, IL 61801, USA*
[6]*Vera C. Rubin Observatory Project Office, 950 N. Cherry Ave., Tucson, AZ 85719, USA*
[7]*Kavli Institute for Particle Astrophysics and Cosmology, SLAC National Accelerator Laboratory, Stanford University, Stanford, CA 94025, USA*
[8]*Lancaster University, Lancaster, UK*
[9]*Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell, UK*

(Dated: February 19, 2025)

ABSTRACT

As the Commissioning Execution Plan (LSE-390) says, "The project team shall deliver all reports documenting the as-built hardware and software including: drawings, source code, modifications, compliance exceptions, and recommendations for improvement." As a first step towards the delivery of documents that will describe the system at the end of construction, we are assembling teams for producing of the order 40 papers that eventually will be submitted to relevant professional journals. The immediate goal is to accomplish all the writing that can be done without data analysis before the data taking begins, and the team becomes much more busy and stressed.

This document provides the template for these papers.

1. INTRODUCTION

The Data Facilities support Rubin Data Management by providing the underlying infrastructure used for productions and...

(for the layperson, what are we trying to do) (data rights, security requirements)

2. DATA STORAGE

2.1. *Identity Management and Security*

2.2. *Archives*

2.2.1. *Images*

2.2.2. *Engineering/Facilities Database (EFD)*

2.2.3. *Large File Annex (LFA)*

2.2.4. *Summit Database*

2.2.5. *Alert Production Products*

2.2.6. *Prompt Products Database*

2.2.7. *Data Release Products (includes ConsDB snapshot)*

2.3. *Storage Types*

2.3.1. *POSIX*

2.3.2. *S3*

2.3.3. *WebDAV*

2.3.4. *Tape*

2.4. *Databases*

2.4.1. *SQL*

2.4.2. *InfluxDB*

3. DATA MOVEMENT AND MANAGEMENT

3.1. *Long-Haul Networks*

3.1.1. *Summit to USDF*

3.1.2. *USDF to Europe*

3.2. *Rucio and Butler Integration*

3.3. *s3daemon*

3.4. *Disaster Recovery*

4. MULTI-SITE PROCESSING

4.1. *CVMFS Software Distribution*

4.2. *Embargo and Unembargo*
4.3. *Processing Middleware: The Butler and BPS*
4.4. *Campaign Management*
(implications of facility setup on CM)
5. MONITORING
5.1. *Compute*
(utilization, memory, etc.)
5.2. *Networking*
(bandwidth, latency, errors/drops, etc.)
5.3. *Storage*
(size, bandwidth, quotas, etc.)
5.4. *Applications and Service Logs*
6. DEVELOPER RESOURCES
6.1. *USDF RSP*
6.2. *Interactive Nodes*
6.3. *Build and Test System*
7. DATA ACCESS CENTER(S)
7.1. *US DAC and Hybrid Model*
7.2. *Chilean DAC*
7.3. *IDACs*
8. MANAGEMENT AND COORDINATION FOR OPERATIONS
8.1. *Planning*
(sizing models, etc.)
8.2. *Communications Mechanisms*
8.3. *Incident Response*
9. FACILITY SPECIFICS
9.1. *Summit Data Facility*
(instantiation of the rest here)
9.1.1. *Commissioning Cluster*
9.1.2. *Teststands*
9.2. *USDF*
9.3. *French Data Facility*
9.4. *UK Data Facility*
9.4.1. *Lancaster*
9.4.2. *Rutherford Appleton Lab*

REFERENCES