UK Data Facility Update

Rubin Data Facilities Meeting
11th–13th February 2025, IN2P3, Lyon

LSST:UK Computational In-kinds

In-kind contributor to Vera C. Rubin Operations
- Data Release Processing of 25% of annual survey data
- Full IDAC serving 20% of international Rubin community
  - Assumed to be 1,000–1,500 data-rights holders
- Community Alert Broker (Lasair)
- Derived data products focused on UK science priorities (DEV)
  - Crowded-field crossmatch; Near-IR data fusion; TiDES/ 4MOST follow-up; DESC

Various other contributions that are less computational

LSST:UK Consortium

- Representation from every (that is, 36) UK astronomy groups

LSST:UK Consortium

- Coordination of UK involvement in Rubin
  - Responsible for securing and managing funding for UK Programmme
    - Called *LSST:UK Science Centre Programme (LUSC)*
  - Delegated responsibility for administration of Rubin data-rights in UK
- Roles
  - Project Leader – Bob Mann (Edinburgh)
  - Project Scientist – Graham Smith (Birmingham)
  - Communications Officer – Eleanor O'Kane (Edinburgh)
  - Commissioning Coordinator – Graham Smith (Birmingham)
  - Project Managers – George Beckett and Terry Sloan (Edinburgh)
  - Education and Public Outreach Coordinator – Chris Lintott (Oxford)
  - Consortium Board Chair – Mike Watson (Leicester)

LSST:UK Science Centre Programme

- Five-phase programme
  - Phase A (Jul'15—Jun'19)   –   Development
  - Phase B (Jul'19—Mar'23)   –   Commissioning
  - **Phase C (Apr'23—Mar'27)**   –   **Early Operations**
  - Phase D (Apr'27—Mar'31*)   –   Standard Operations
  - Phase E (Apr'31--~2036*)   –   Standard Operations (cont'd)
- Forecast budget of £72.6M (including £26.4M capital)
  - UK Gov't (BEIS) approved long-term STFC funding via business case in 2022
- Infrastructure element funded through IRIS Programme
  - Core funding for DRP and DAC (inc. Lasair)
  - Additional (DEV) contributions via IRIS Resource Scrutiny and Allocation Panel
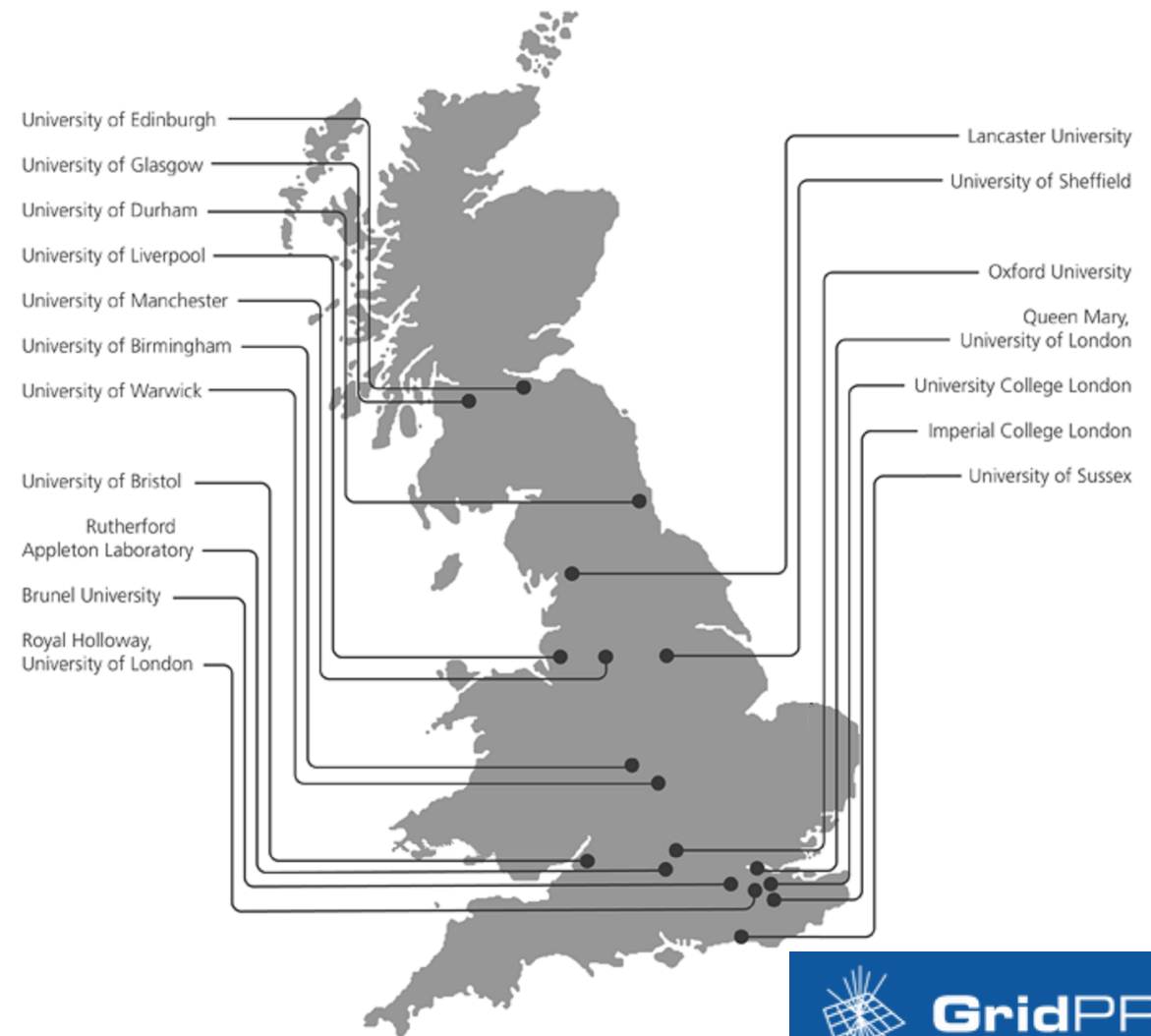
IRIS

- Cooperative federated community creating digital research infrastructure for PPAN in UK
- Range of capabilities:
  - HPC computing (mostly CPU/ some GPU)
  - Grid computing (LHC-like compute and storage)
  - Cloud computing (mostly, Scientific OpenStack / Kubernetes)
  - Storage – Object Storage / tape-backed HSM
  - Links into UK NREN – JANET
  - Shared services – accounting, directory, IAM
- LSST:UK computing infrastructure is provided by IRIS (www.iris.ac.uk)
  - Through partner sites (inc. GridPP, DiRAC, STFC RAL)
- Resource Scrutiny and Allocation Panel
  - Additional resources via annual application (from shared capacity)

GridPP (UK DRP Host)

- Community of particle physics and research infrastructure engineers throughout UK
- UK's contribution to WLCG
- 18 sites inc. UK's WLCG Tier 1 at RAL
- Significant compute and storage
- IRIS member and resource provider
  - Also used by Dune, LIGO and Euclid

Somerville (Primary IDAC Host)

- IRIS research cloud hosted at UoE Advanced Computing Facility
- Software stack based on Scientific OpenStack
- As of March 2025
  - 5 Petabytes of Ceph storage (scientific-data archives)
  - 100 Terabytes of fast, NVMe-based Ceph storage (low-latency databases)
  - 1,984 (virtualized) CPU cores along with 9 Terabytes of memory
    - 100 TB local disk (NVMe/ HDD) targeted at distributed data analysis framework
  - 100 Gbps internal data network
  - 2×100 Gbps uplink into JANET (UK NREN)
  - 15-node Qserv (1 PB NVMe)
- 1.5 FTE of on-site support (UoE)
  - Level 3 support from StackHPC

SOMERVILLE
CLOUD

Data Release Processing

UK DRP Team

Name (Inst.)	Role	Effort (FTE)
Matt Doidge (Lancs)	Workflow Manager	0.5
Peter Love (Lancs)	Workflow Manager	0.5
Dave McKay (Edin)	Processing Scientist	0.5
James Mullaney (Sheff)	Production Scientist	0.5
Tim Noble (RAL)	Data Wranger	1.0
Steven Simpson (Lancs)	DRP Ops Support	0.5
Mathew Sims (RAL)	DRP Ops Support	0.5
System Admin (Lancs and RAL)		2×0.5 FTE

Lancaster Data Facility Details

- Colocated with the "UKI-NORTHGRID-LANCS-HEP" WLCG site (the other half of Lancaster FTEs)
  - Mostly providing resources for ATLAS.
- Compute provided from a portion of the University's main HTC cluster.
  - The WLCG gridsite and Lancaster DF are the largest stakeholders, Matt is Co-Admin.
  - "Grid" interface provided by NorduGrid ARC CE
  - Backed by Slurm batch system
- Storage is shared with the WLCG site
  - CephFS system (Running Ceph Reef. 8+3 Erasure Coded)
  - 11PB total volume
  - Fronted by an XRootD cluster, consisting of a Redirector and 8 Gateways.
  - Access to data both internally and externally via davs/https to the gateways
- Connecting to the NREN over a dedicated 100Gb link

RAL Data Facility Details

- Co-located with the "RAL-LCG2" WLCG site (used by many other experiments)
  - Mostly providing resources for ATLAS, CMS, LHCb, Alice, DUNE, SKA.
- FTS deployment for data movement across experiments
  - 8 servers, but dropping to 4, due to funding
- Compute provided from 10% non-LHC resources
  - Job interface provided by NorduGrid ARC CE
  - HTCondor batch system
- Storage is shared with WLCG site
  - Ceph object store (Running Ceph Pacific (Soon to be Quincy). 8+3 Erasure Coded)
  - 100PB total volume (4PB for LSST)
  - Fronted by XRootD cluster – Redirector and 26 Gateways (each with 25Gb/s NIC)
  - Access to data both internally and externally via davs/https to the gateways
- Connectivity to site is 2×100 Gb/s LHCOPN (upgrading to 400 Gb/s total)
  - 4×100Gb/s external network (includes LSST traffic)

DRP Status

- Processing jobs are running on both UK sites
  - Ctrl-ingestd (registers data into Butler once replicated by Rucio) tested at Lancs
  - Pipeline Rucio-register (registers Pipeline outputs for Rucio) also tested
- Test DRP workflows (from HSC and DESC DC2) routinely running in UK
  - Confirms capability for end-to-end processing, at small scale
  - Moving LSST Pipeline CI to UK (and Fr) DF, to free up capacity at USDF for Commissioning
- File transfer tests continue
  - Addition of Zip support to Pipeline should allow small files to be collated for more efficient transfer
- Operational support model to be finalised

UK IDAC

UK IDAC Team

Name (Inst.)	Role	Effort (FTE)
Greg Blow (Edin)	Qserv Admin	0.5
Gareth Francis (Edin)	Science Platform + Lasair	0.9
Dave McKay (Edin)	Service Manager	0.5
Daniel Pizarro	System Admin	1.0
Mike Read (Edin)	Science Support	0.5
Eckhard Sutorius	Data Wrangler	0.5
TBD	*Feb'25 recruitment*	2.0

IDAC Plans

- LSST:UK preparing for Full IDAC
  - Running for duration of survey
  - Current and most recent previous DR
  - Serving 20% of anticipated Community
  - Running Rubin Science Platform
- Co-located with Lasair Broker
- Hosting User-generated Products
  - Fused LSST-VISTA data products
  - Crossmatch catalogues for selected ancillary surveys
  - Popular ancillary surveys
- Running on IRIS cloud
  - Hosting early datasets (see later)
  - Deployed on Scientific OpenStack
  - Running at Edinburgh (and RAL)
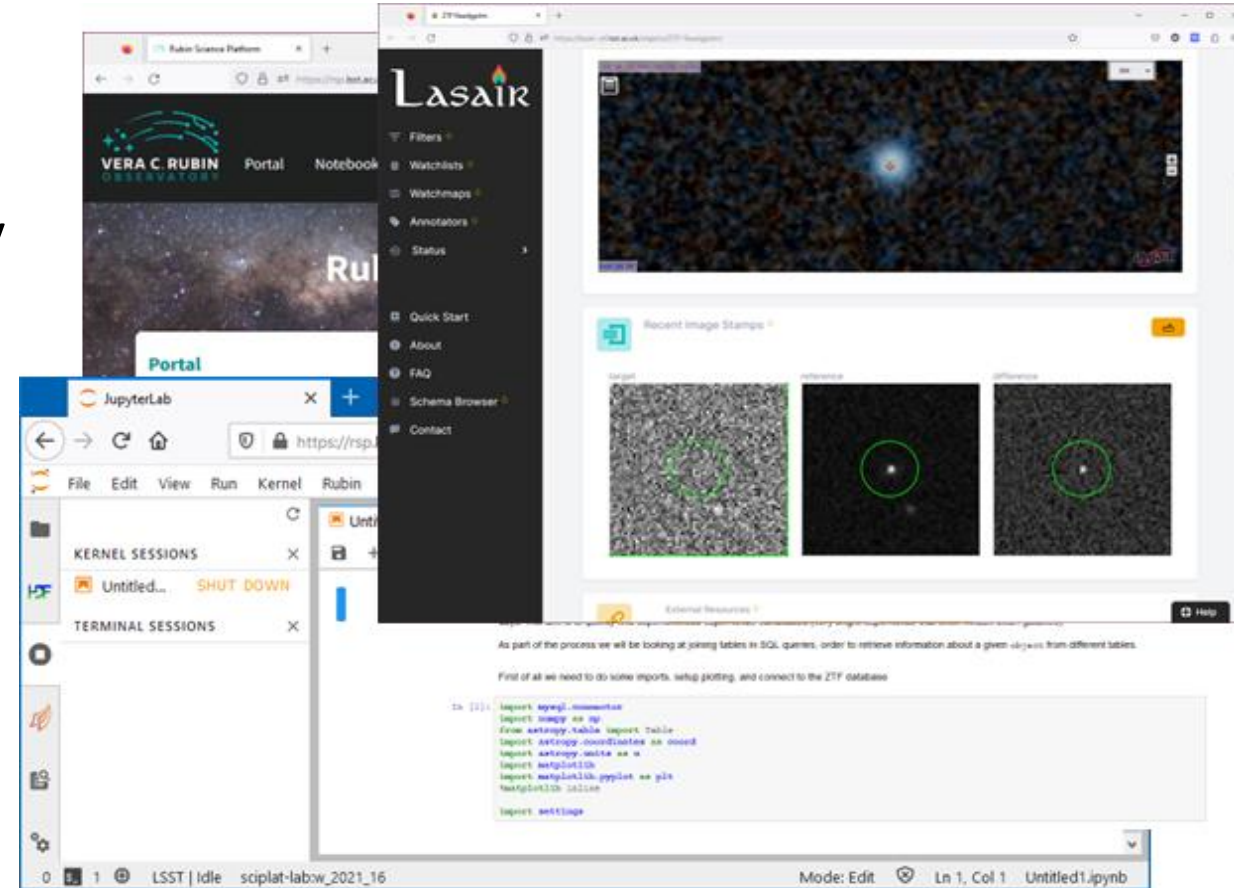
UK DAC Architecture
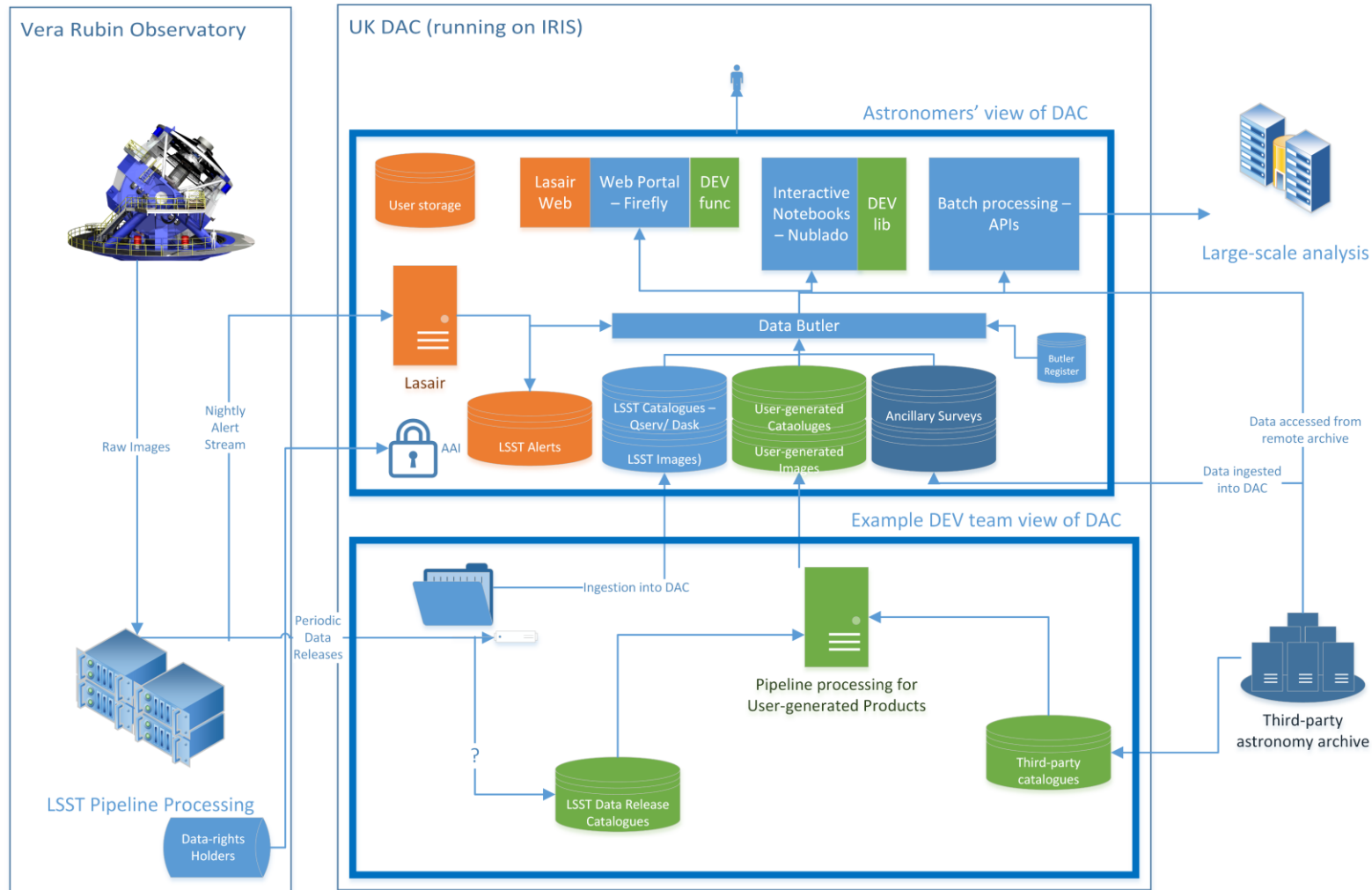
Roadmap

	UKDAC3	UKDAC4	UKDAC5	UKDAC6	UKDAC7
Delivery Date	2023-03-31	2024-03-31	2025-06-30	2026-03-31	2027-03-31
Platform	IRIS	IRIS	IRIS	IRIS	IRIS
Functionality	•RSP components: all	RSP components: all - Inc. Nublado Ver. 3	RSP components: all	RSP components: all	RSP components: all
Data	UKDAC2 holdings, plus WP3.11 (Gaia-CatWISE catalogues)	UKDAC3 holdings plus: • DP0.2 (cats and images) • DP0.3 (cats) • VISTA-HSC (im+cats) •ZTF DR20 (sample)	UKDAC4 holdings plus: •DP1 (Apr—Jun'25)	UKDAC5 holdings plus: •DP2 (May—Jun'26)	UKDAC6 holdings plus: •DR1
Access	•Authentication: GitHub credentials •Accounts: admin, invited users involved in DEV WPs (e.g., science teams related to WP3.5) and other activities (e.g., Lasair usage).	Auth: GitHub credentials (tbc) Accounts: Small number (~20) interested users	Auth: OpenID Connect (as per project) Accounts: ~100, depending on demand	Auth: OpenID Connect (as per project) Accounts: ~400 (UK/ Intl)), depending on demand	Auth: OpenID Connect (as per project) Accounts: ~1,000 (UK plus Intl), depending on demand
Documentation	Project-provided, supplemented by UK-specific online documentation to support authorised users	Project-provided, supplemented by UK-specific online documentation to support authorised users	Project-provided, supplemented by UK-specific online documentation to support authorised users	Project-provided, supplemented by UK-specific online documentation to support authorised users	Project-provided, supplemented by UK-specific online documentation to support authorised users
Helpdesk	Community forum	Community forum	Community forum	Community forum w/ active engagement from UK DAC staff	Community forum w/ active engagement from UK DAC staff

UK Sizing Model

Infrastructure Sizing

- LSST:UK maintains long-term sizing model, for duration of survey
  - Mix of dedicated, apparatus-like infrastructure and shared services
- Identifies four categories of requirement
  - DRP – computing, working storage, long-term storage
  - DAC – data hosting, science platform, DEV products, user compute, storage
  - Lasair – stream processing, data hosting, user compute, downstream users
  - DEV – compute, working storage (long-term storage in DAC)
- Guided by Project Model (summarized in DMTN-135)
- Used to determine infrastructure costs for STFC business case

Sizing Model (High-level Summary)

DRP

	Preop	Survey Operations										Post-ops	
Capability	FY24	LOY1	LOY2	LOY3	LOY4	LOY5	LOY6	LOY7	LOY8	LOY9	LOY10	FY35	FY36
CPU (M core hrs)	11	11	21	30	40	50	63	73	83	93	103	100	50
Normal/ Object (PB)	8.0	9.0	16.0	23.0	30.0	37.0	44.0	51.0	58.0	65.0	72.0	61.4	0

DAC

	Preop	Survey Operations										Post-ops	
Capability	FY24	LOY1	LOY2	LOY3	LOY4	LOY5	LOY6	LOY7	LOY8	LOY9	LOY10	FY35	FY36
CPU (M core hrs)	0.53	0.88	2.10	2.45	2.63	3.94	5.26	5.26	6.57	7.88	7.88	7.88	7.9
Normal/ Object (PB)	2.2	21.9	50.0	72.1	94.9	117.6	140.2	162.9	185.6	208.3	231.0	231.0	231.0

Risks, Issues, and Opportunities

Potential Opportunities/ Open Questions

- PerfSonar
  - Nodes in Edinburgh, RAL and Lancs
- Inter-DF Qserv experiments
  - Successfully access UK IDAC Qserv instance from RSP/ TAP at IN2P3 with no significant performance degradation
- Use of LHC-One network
  - Potential option for DRP-related traffic
  - Trusted network of sites meaning lower overheads?
- CERN Data Challenge in February
  - Including transfer experiments between SLAC and UK / European grid sites

Thank you

Risks and Potential Issues

- Aspects of DRP operation
  - Important to test multi-site DRP at scale
  - Robust access control and science validation
- Capital budget for UK contributions
  - Costed based on information in sizing model: risk from changes to DRP implementation and hardware-price fluctuations
- Operational support model
  - Still unclear on process for at-site and inter-site support during Operations
- Dealing with user demand immediately after publication of DR
  - IDAC demand management and capacity management elements not yet tested
- Navigating DF site-specific requirements and interfaces
  - SLAC is single point of failure for key services

Data Holdings

- ## Current holdings
  - ### DP0.2 (only subset of Butler repository, due to capacity limits)
  - ### DP0.3
  - ### Gaia-CatWISE crossmatch catalogues (UKD-S9)
  - ### VISTA-HSC fused datasets (UKD-S5)
  - ### ZTF DR20 catalogue (small subset) – for progressing Lasair-RSP integration activity
- ## Next step
  - ### Ingest further collections from DP0.2 repository – scaling up Butler ingest
  - ### Test access to Butler repository in remote S3