LSST:UK Consortium

LSST:UK Science Centre

UK involvement in the Large Synoptic Survey Telescope: Phase B - Resubmission

(Credits: LSST; M. Read)

Contents

Abstract

In 2015, STFC awarded the LSST:UK Consortium £17.7M, securing UK involvement in the Large Synoptic Survey Telescope (LSST) with a £15M contribution to LSST operations, and providing £2.7M for the first phase of UK preparation for LSST science. The Phase A award enabled UK astronomers to start contributing to the LSST Science Collaborations, and supported the successful prototyping of a UK LSST Data Access Centre (DAC), plus initial development of the analysis software that the UK community will require to exploit the huge scientific potential of the unprecedented LSST dataset. In this proposal, we seek funding for the next phase of UK involvement in LSST, which will allow us to participate in LSST commissioning and build on the successes of Phase A by turning the proto-DAC into a production facility and deploying within it the analysis software we need for the start of LSST survey operations.

1. Introduction

Currently under construction[1] on Cerro Pachón in Chile, LSST will conduct the most ambitious optical sky survey yet planned, imaging the whole visible portion of the southern sky twice a week for 10 years from late 2022. It will visit each field more than 800 times during that decade, with exposures taken through six filters (*ugrizy*) spanning the visible portion of the spectrum, from 320 to 1050nm. The integrated depth of the final stacked images (r~27.5 mag, 5σ point-source) will be five magnitudes deeper than the Sloan Digital Sky Survey, and the survey dataset will constitute an unprecedented multi-colour movie of the Universe that will be used to study astrophysical phenomena all the way from the closest asteroids to the nature of the dark energy driving the accelerating expansion of the Universe.

Figure 1.1 LSST in construction on Cerro Pachón in February 2018 (Credit: LSST)

[1] See https://www.lsst.org/news/see-whats-happening-cerro-pachon for the latest summit webcam photos.

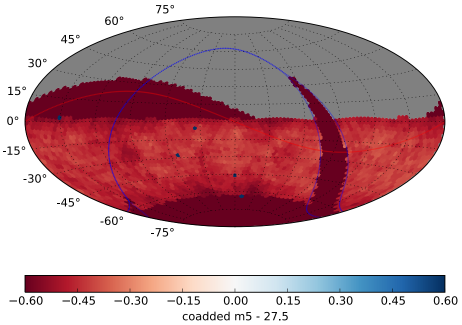

Figure 1.2 The final r band depth in one realization of the baseline survey strategy. (Credit: LSST)

The strength and breadth of LSST science led to its being the highest ranked large-scale ground-based initiative in the 2010 US National Academy of Sciences Decadal Survey of Astronomy and Astrophysics. Within the UK, the STFC Astronomy Advisory Panel listed LSST as a high-priority future facility in its submission to the last Programmatic Review, reflecting not only the importance of LSST data to the UK astronomical community, but also its powerful synergies with the UK's southern hemisphere facilities. From ESO (VISTA, VLT, ELT and ALMA) through to SKA and Euclid, the UK astronomy roadmap for the 2020s presents an exciting array of multi-wavelength possibilities, and having LSST in the list will allow the community to maximise the scientific return on these earlier investments by enabling new synergies between them.

About 85% of LSST's time will be spent on its ~18,000 sq. deg. *Wide-Fast-Deep* survey, which will be supplemented by coverage of other areas of especial interest (e.g. in the Galactic and Ecliptic Planes), plus *Deep Drilling Fields*, where greater depth is required to take full advantage of existing and planned multi-wavelength datasets. LSST may also implement schedule over-rides to localise newly-discovered gravitational wave sources: the 9.6 sq. deg. LSST field of view will be well matched to their positional uncertainties once the next generation of detectors comes online. For many analyses – e.g. determining orbits of solar system bodies and classifying variable stars or supernovae – the cadence of exposures in different bands is crucial, and a top priority for the first half of Phase B is assessing the impact on key metrics from different science areas resulting from varying the survey strategy: e.g. changing from the baseline, in which the integrated exposure time is built up evenly over the sky to a *rolling cadence*, in which, at any one time, a certain region of the sky receives denser time sampling, with that focus region moving over time to attain uniform depth.

1.1 LSST Data Processing

At the end of each exposure, the resulting image will be shipped to NCSA in Illinois, for comparison with a reference image of that field in the same filter, to detect sources that are new, have moved or have changed in brightness. Statistically significant detections will be issued as alerts within 60s of the end of the exposure, while those associated with known solar system objects will be used to update orbits during the following day. Every night this difference imaging pipeline will process about 20 TB of images and generate up to 10 million alerts. The transient alerts and the solar system orbits comprise the *Prompt* (or, ***Level 1***) *Products*. Once a year (twice in the first year of survey operations), all extant data will be run through a direct imaging pipeline to generate the *Data Release* (or ***Level 2***) *Products*, which are standard images and object catalogues, used for static-sky analyses and non-time-critical analysis of variable sources. This pipeline will be run at NCSA and at CC-IN2P3 in Lyon, and the resultant data products shipped to Data Access Centres (DACs) through which registered users may access them. The final data release (DR11, in 2033) is estimated to comprise a 15 PB database and ~400 PB (compressed) of image files. These Project-generated data products will not be optimal for all analyses, so, for full scientific exploitation, they will be manipulated further to yield *User Generated* (or ***Level 3***) *Products*, which may also contain the result of the integration of non-LSST data to produce multi-wavelength datasets (a UK priority and strength).

1.2 The Funding and Management of LSST

The construction of LSST is primarily being funded by US agencies: the Department of Energy is funding the camera ($165M) with the NSF providing $473M for the telescope and the data management system. In addition, $40M of private funding was used for items with long lead times, such as the mirror fabrication and initial site preparation, while IN2P3, the French particle physics agency, has contributed to building the camera, as part of a long-term relationship with DoE labs. The LSST Project is led by its Director, Steve Kahn, from SLAC, his Deputy is Zeljko Ivezic, from the University of Washington, and Bob Blum (NOAO) is the Interim Director of Operations.

LSST operations will be funded by three sources. In addition to further contributions (totalling about $270M) from the NSF and DoE, channelled through AURA and SLAC, respectively, an additional ~$100M is being sought from International Contributors, who obtain data rights for named researchers in return. The international funds are managed by the LSST Corporation (LSSTC), a not-for-profit body that initiated the LSST project prior to the start of federal funding and which will partner with AURA and SLAC to manage LSST operations.

Separate from the LSST Project, which is building, and will operate, the LSST facility, there are a set of *Science Collaborations* responsible for coordinating, by science area, preparation for, and analysis of, LSST data by rights holders within the international community. At present there are eight of these: *Active Galactic Nuclei*; *Dark Energy*; *Galaxies*; *Informatics and Statistics*; *Solar System*; *Stars, Milky Way and Local Volume*; *Strong Lensing*; and *Transients and Variable Stars*. Currently, these vary greatly in size, organisational structure and activity, reflecting differences between the level of preparation and the scale of operations needed for different types of analysis.

1.3 UK Participation in LSST

The UK is not part of the LSST Project, so we focus solely on science, to be performed both within UK groups and in the international Science Collaborations, and to do that effectively requires the coordination provided by our PPRP-funded programme. As detailed in Section 4, the UK's involvement in LSST proceeds through two entities. The *LSST:UK Consortium* comprises all those institutions with a scientific interest in LSST: it now has 35 members, comprising all UK astronomy groups. The *LSST:UK Science Centre* (LUSC) is a distributed team undertaking funded project work on behalf of the UK astronomical community, as represented by the LSST:UK Consortium.

The LUSC programme comprises four strands of activity: (i) **LUSC-DAC** will cover preparation for, and, later, operation of, a UK LSST Data Access Centre, curating data for the LSST:UK Consortium and supporting, through technical advice and computing hardware, their Level 3 analyses; (ii) **LUSC-DEV** will develop the Level 3 software required for the UK to secure leadership of the community's top priority science topics; (iii) **LUSC-TRN** will train young researchers in the data science techniques that will be required to analyse the multi-PB LSST dataset; and (iv) **LUSC-EPO** will develop an Education and Public Outreach programme in collaboration with the well-funded LSST EPO team in the US. We will seek support for these activities from different sources: funding for LUSC-DAC and LUSC-DEV is sought through PPRP; LUSC-TRN will seek funding from the EU and coordinate activities with the STFC CDTs in Data Intensive Science; and, on the advice of Science Board, LUSC-EPO will solely target STFC Public Engagement funding schemes.

1.4 The LUSC Programme

The LUSC programme has been defined to run in four phases, whose boundaries have been set by two Project milestones[2]: the installation of the Commissioning camera (ComCam) in October 2019 and the start of survey operations in October 2022. Those four phases are as follows:

- **Phase A: Development** (July 2015 – March 2019). The principal goal of this phase is to achieve full technical and scientific engagement with the LSST Project and Science Collaborations, enabling the UK community to participate fully in Commissioning and to prepare detailed plans and prototype DAC services in readiness for the start of survey operations.
- **Phase B: Commissioning.** (April 2019[3] – March 2023). This spans LSST Commissioning and the start of survey operations, so its focus will be on determining detailed knowledge of what to expect from the Level 1 & 2 products and, hence, in preparing the Level 3 software required to create the additional user-generated data products needed to meet UK science goals.
- **Phase C: Early Operations** (April 2023 – March 2027). The first few data releases will be used to shake down operations of the UK DAC and to drive to completion the development of the bulk of the Level 3 software.

[2] The LSST Project schedule may be found at https://www.lsst.org/about/timeline. At the moment, it still retains 8.5 of the original 13 months of contingency, so all dates remain correspondingly uncertain.

[3] As discussed in Section 6.1, the Phase B funding we are requesting here will start in 1 July 2019, with bridging funds from April-June 2019.

- **Phase D: Standard Operations** (April 2027 – ~March 2033[4]). Level 3 development work should cease early in Phase D, replaced by science exploitation funding through the Astronomy grants line, while the DAC should be settling down into routine operations.

LUSC-DAC and LUSC-DEV started during Phase A and we expect LUSC-TRN and LUSC-EPO to be launched during Phase B.

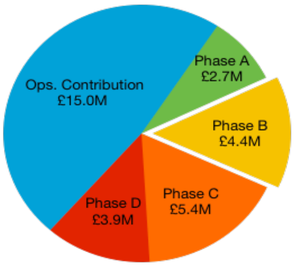

Figure 1.3 The budget for the full 18-year LUSC programme, divided into the four Phases, plus the Operations Contribution. The total is £31.5M, averaging to ~£1.75M per year. The Operations Contribution and Phase A funding have already been awarded.

Table 1 The Phase B Work Packages and their DI staff requirements

WP	Title	DI S.M. Original	DI S.M. Final
1	Management	48	45
2.1	DAC Management	12	11.5
2.2	Data Ingestion and Publication	42	37.5
2.3	Alert Handling Infrastructure	18	17
2.4	Provision of DAC Platform	24	22.5
2.5	Science Support	60	47
3.1	UK Solar System Science Server Software	48	0
3.2	LASAIR: UK transient broker for LSST	48	36
3.3	Transient Classification & Spectroscopic Follow-up	48	24
3.4	UK Variability Broker for LSST	48	0
3.5	LSST & Near-Infrared Data Fusion	72	50
3.6	3D LSST: Photometric Redshifts	48	0
3.7	Morphology & Low-surface-brightness Science	54	30
3.8	Strong Lens Discovery System	23	0
3.9	PSF and Sensor Characterisation & Modelling	48	36
3.10	UK's Contribution to DESC Operations	23	20
3.11	Cross-Matching Catalogues at LSST Depths	24	24
	Total	688	400.5

In Figure 1.3 we show the baseline profile for STFC funding across the full LUSC programme. This yields a total cost to STFC over the 18-year period of £32M, on the assumption that the hardware cost of the DAC from Phase B onwards is met through a coordinated computing infrastructure for STFC science like that being planned by the IRIS (formerly UKT0) consortium, rather than from the astronomy budget. In the remainder of this proposal we detail our costed plan for Phase B of the LUSC Programme, whose Work Packages are listed in Table 1, in both original and descoped form.

2. Objectives

This proposal seeks funding for the second of four phases of UK involvement in LSST.

The objectives of the full programme are:

1. To obtain for the UK community the data access required for full scientific participation in the LSST survey programme and for enhancing the scientific return from other facilities in the UK astronomy programme through incorporation of LSST data;
2. To secure intellectual leadership of the UK community's top priority LSST science areas, by targeting investment in the software and Data Access Centre services needed for their success.

The principal objective of the four-year Phase B programme for which we are seeking funding here is to build on our successful Phase A R&D programme, and the lessons to be learnt from LSST Commissioning, to prepare the UK community for the start of main survey operations. We intend to:

1. Maintain and increase UK leadership positions in the LSST Science Collaborations;
2. Deploy a UK LSST Data Access Centre providing the computational infrastructure needed to support exploitation of LSST data by the UK community. This will be undertaken in collaboration with LSST colleagues in the US & France, and with research computing experts from other disciplines through the IRIS consortium;

[4] The schedule for preparing a long-term archive for the complete LSST dataset after the survey ends remains TBD.

3. Participate in the LSST Commissioning programme, using the Commissioning data to aid development of data analysis software targeted to securing leadership in the UK's top priority LSST science areas; and
4. Develop the services needed to integrate LSST data with existing and future astronomical data resources to facilitate multi-wavelength and multi-epoch analyses.

3. Scientific Justification

LSST's temporal resolution, unprecedented depth and uniform photometry over an entire hemisphere combine to produce a compelling science case, and the LSST:UK Consortium possesses the expertise to secure leadership positions across a broad range of astrophysics. The UK's strengths and heritage in leading astronomical surveys, data processing and analysis combined with our access to ESO facilities provides an exciting platform for scientific leadership and impact in the 2020s. The UK community is now embedded within the LSST Science Collaborations, thanks to the Phase A funding, and we are already shaping our own direction and designing the tools and data products required to ensure we are at the forefront of LSST data analysis.

The descope required for this resubmission resulted in the loss of 4 of the 11 DEV work packages that LSST:UK had prioritised as delivering essential data products needed for the realisation of the science case for UK involvement in LSST. We present that full science case below, including comments on the impact of the descope.

3.1 Cosmology

Cosmology is one of the main LSST science drivers. The STFC Science Priorities identify the major science goals for cosmology as discovering the nature of the dark Universe, and improving our understanding of the earliest moments of time (STFC Science Challenges A1, A3, C4, C5). The AAP Report for the STFC Programmatic Review outlines a multi-probe approach to test the underlying fundamental assumptions of General Relativity, using weak gravitational lensing, galaxy clustering, Type Ia supernovae, and galaxy clusters, as well as cross-correlation with the Cosmic Microwave Background. The AAP report highlights the step-change from DES to LSST and Euclid due to the order of magnitude increase in galaxy numbers, as well as the importance of LSST for finding large numbers of SNe because of its time domain capability.

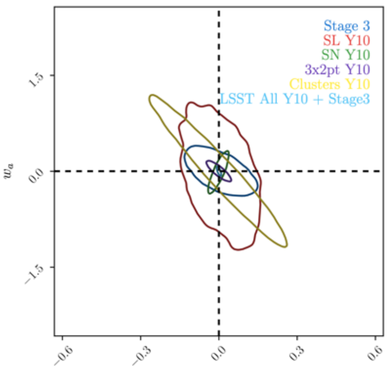

Figure 3.1 Predicted LSST 10-year constraints on the dark energy parameters w_0 and w_a.

The Dark Energy Science Collaboration (DESC) is the most organised and developed of the LSST Science Collaborations and the UK is now in key leadership roles within DESC. We envisage that the high profile cosmological results will come through analysis pipelines developed by DESC. LSST will measure the equation of state of dark energy to unprecedented percent-level accuracy, and will make similarly powerful measurements of the laws of gravity and the evolution of dark matter and cosmic structure. This accuracy will come from the combination of multiple probes tracing different aspects of the behaviour of light and matter (Fig 3.1, from the DESC Science Requirements Group). Weak lensing, the coherent distortion of galaxy shapes over wide fields caused by intervening gravitational fields, causes galaxy images to align around dark matter structure. Measuring it probes dark matter structure, dark energy's history, and the behaviour of relativistic objects under gravity. Galaxy number density measurements trace dark matter with a precise but biased relation, and in combination with lensing have provided the most powerful low-redshift measurements of cosmic structure to date (DES Collaboration et al 2017; van Uitert et al. 2018). Counting galaxy clusters probes the high-density tail of the dark matter distribution, while supernovae trace the relationship between redshift and distance out to large distances, again letting us probe dark energy dynamics, and, finally, strongly lensed quasars (and supernovae) and double source-plane lenses probe ratios of distances with complementary sensitivity to dark energy (Bonvin et al. 2017). When rigorously tested and combined, these measurements will provide the most powerful ever constraints on the cosmological density parameters and equation of state of dark energy, as well as new tests of the ΛCDM cosmology itself. Our next major step forward in understanding the dark universe relies on control of systematic errors through the combination of data from LSST, plus Euclid and the Square Kilometre Array, both already significant UK projects.

The UK has international leadership in these areas, and we are already making a major impact to ensure LSST realises its potential in cosmology. This task is so demanding that it requires the full power of the large DESC collaboration and this proposal directly addresses these challenges with UK leadership in the DESC simulations and operations (see WP 3.10), combining LSST with our near-infrared legacy data (WP 3.5), providing photometric redshifts and error models (WP 3.6), and discovering and classifying type Ia supernovae (WP 3.2+3.3) and strongly lensed galaxies (WP 3.8). *The descoped resubmission loses the work on photometric redshifts and strong lens discovery system entirely, but this will still need to be completed for DESC to achieve its full scientific potential.*

3.2 Transients and variables

LSST will provide an unprecedented sampling of the time-domain universe. No other time-domain experiment can compete with its Wide-Fast-Deep survey for science. The survey will deliver about 10 million transient and variable object alerts per night. This rich photometric data, with exquisite calibration and control of systematics, will be further enhanced by spectroscopic follow-up of ~100,000 selected objects. Spectra will reveal the energetics, distances, luminosities and nuclear physics of explosive transients, and, ultimately, the physical nature of the transient Universe. A combination of selecting the extra-galactic transients and exploiting the UK's strengths in massive spectroscopic follow-up will ensure our leadership in this field. The 4MOST project at ESO will place a multi-object spectrometer on the VISTA 4m telescope and our UK collaboration have secured a leadership role. The future of the ESO NTT is secured and it will be a full-time transient follow-up telescope with a new spectrometer (SOXS, with UK involvement), while, ESO's VLT and, ultimately, ELT are perfectly matched to LSST follow-up of faint and exotic transient or variable sources, as well as faint and distant sources in the early Universe.

Deep colour information will reveal exotic super-luminous supernovae out to redshift z~6 and potentially beyond, while there are immense opportunities to link with the SKA radio surveys, GRB and high energy missions such as Swift and SVOM. Multi-messenger astronomy has finally arrived with the LIGO-Virgo detection of its first neutron star–neutron star merger and the spectacular confirmation of a kilonova powered by the radioactive decay of heavy r-process elements (Abbott et al 2017, Smartt et al. 2017, Tanvir et al. 2017,). LSST will be the optimal survey to target the uncharted parameter space of faint, fast extragalactic transients – i.e. gravitational wave sources, failed supernovae, orphan afterglows, and mysterious fast radio bursts.

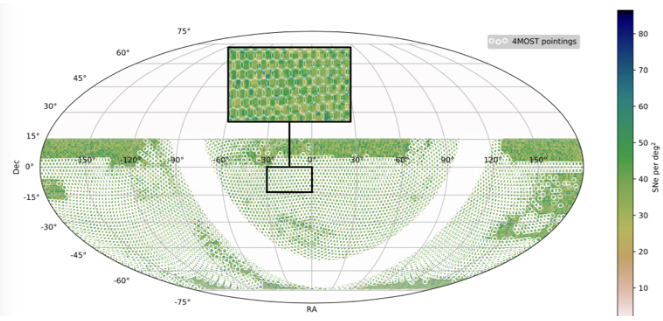

Figure 3.2 The impressive synergy between LSST and 4MOST: green denotes 40-50 LSST SNe per deg², while, in the zoom, each circle shows a 4MOST field.

Galactic plane coverage opens up a huge range of science. Episodic accretion in young stellar objects on timescales of years will dominate the Galactic population of high amplitude variables in the red optical bands (Contreras Pena et al. 2017, Lucas et al. 2017). Thousands of multi-colour light-curves will enable study of the poorly-understood processes in protoplanetary discs from the stellar surface out to 5 au. Accretion bursts in X-ray binaries will also be a major application, thanks to the availability of pre-outburst light curves. At lower amplitudes, the precision and duration of LSST is ideally suited to the study of recently discovered classes of variables such as brown dwarfs (probing cloud stratification and auroral activity), the very numerous OSARGs (OGLE Small Amplitude Red Giants, a new type of standard candle) and, of course, the decadal magnetic cycles of normal main sequence stars. The size of the dataset will lead to detection of rarely-seen variables that trace brief but important events, e.g. giant ring systems around young planets that reveal the formation of satellite systems (Kenworthy et al. 2015), the irregular circumstellar matter distributions around sun-like stars that may trace planet engulfment (Boyajian's star) and the new class of stars known as BLAPs (blue large amplitude pulsators). The survey depth allows monitoring of optical counterparts of ultra-luminous x-ray sources in nearby galaxies. The precision, sampling and depth combination for supernova cosmology will define the field in the next decade and, undoubtedly, LSST will make serendipitous discoveries by pushing the time domain to new sensitivities and unprecedented survey volume.

The UK has substantial heritage and leadership in this science and the WP 3.2-3.4 are targeted to enhance this role. This is further underpinned by the proposed UK work on cross-matching of sources across multi-wavelength data sets in WPs 3.5 and 3.11. *In the Phase B descope, the WP on stellar and AGN variability has been removed and the classification and spectroscopic follow-up WP has been cut back. The main WP for transient detection and providing user interfaces and tools for the science community has also been reduced.*

3.3 Solar System

LSST will directly address STFC's SSAP goals of understanding the dynamical and physical evolution of bodies, and the transport of volatile ices throughout the solar system. While the primary LSST goal to find >90% of Potentially Hazardous Asteroids 140m across will be led by US funding, the survey will also record all solar system objects via MOPS (moving object pipeline system). LSST will discover, and obtain precise photometry of, inner belt asteroids >100m to TNOs >70km in the Kuiper belt. It will be the first comprehensive survey to cover the transition from rubble-pile to coherent asteroids in the main belt, and characterise non-hydrostatic equilibrium bodies in the Kuiper Belt. Sparse light-curve observations will increase known spin periods and pole positions in all populations by a factor of 10 over those from Pan-STARS/PTF, allowing investigation into dynamical and physical evolution throughout the asteroid and Kuiper belts. LSST will provide high-cadence nuclear and coma magnitudes for comets on a systematic basis, similar to the exceptional major campaigns previously performed on a handful of individual objects such as comets 67P (Rosetta target) and ISON. LSST will also provide the first realistic chance of the systematic detection and observation of rare events, exemplified by the discovery of the first Interstellar Object in 2017. It should detect interstellar objects at the rate of ~1 per year, see pre-collision impactors at Jupiter and Mars, collisions in the Kuiper belt, and probe the rate of out-bursting of distant comets and Centaurs.

The UK has expertise in the follow-up of solar-system targets using ESO and other facilities (e.g. Fitzsimmons et al. 2018), and leadership in LSST will follow from our co-leadership of the Pan-STARRS Solar System survey, plus leadership of the ESO Large Programmes on Near Earth Object rotations and Rosetta. WP 3.1 is designed to ensure we maintain that leadership (linked with the cross-matching astrometric challenges addressed in WP 3.11). *The Phase B descope resulted in this WP being removed from the funding envelope entirely. This does leave a significant section of the UK science community with no funded infrastructure to meet their science requirements which are often quite different to static object analysis.*

3.4 Galaxies and the early Universe

LSST will identify billions of galaxies and hundreds of thousands of galaxy groups and clusters out to z=2, and beyond, thanks to the unprecedented solid angle, sensitivity, and uniformity of its survey. It will revolutionize our view of the extragalactic universe, a topic in which the UK has a broad, deep, and rich track record of leadership. Recent highlights include: UK leadership of infrared surveys, including the UKIDSS and VISTA public surveys; high-redshift galaxy-evolution and morphological studies using the HST (CANDELS, GOODS); studies of intracluster light at $z > 1$ (XCS); characterisation of low-redshift galaxy clusters as cosmological probes (LoCuSS); optical/IR/mm studies of strongly-lensed high-redshift galaxies; exploration of the fossil record of galaxy formation at low-redshift (e.g. SAURON, ATLAS-3D); and UK leadership in the Herschel Key Programmes. The long-term scope for UK exploration of the extragalactic universe with LSST data is vast and directly relevant to a large fraction of the UK community, but our proposed Phase B work packages are focused on areas of specific UK strength, and where we can add significant value to the Level 2 LSST data products to the benefit of the whole UK galaxy community.

LSST's impact will be greatly enhanced by the addition of information at other wavelengths to its optical data. In particular, by adding infrared (IR) wavelengths to the photometry in the six LSST optical bands (see WP 3.5), the UK will create a unique, legacy dataset that extend the parameter space of stars, galaxies and quasars explored by LSST, enabling a broad range of science.

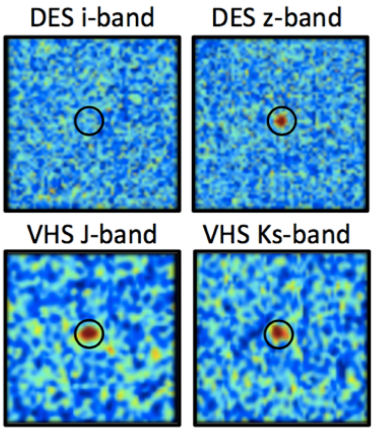

Figure 3.3 Cutouts of a recently discovered z~6.8 quasar, combining optical DES and near-infrared VISTA Hemisphere Survey data. The deep "i-band drop out" provides the redshift identification.

The most distant galaxies and quasars, at redshifts z>7, corresponding to the Epoch of Reionization and beyond, can only be discovered by combining information from LSST (where the galaxies are "drop-outs") and IR surveys (where the galaxies are robustly detected, Fig. 3.3). Identification of large numbers of z~7 galaxies and quasars (e.g. Mortlock et al. 2011, Bowler et al. 2015, Reed et al. 2017) is an important goal for both the LSST AGN and Galaxies Science Collaborations (Robertson et al. 2017). It is an area where LSST will have early impact: e.g. we expect about 10 new z~7 quasars and thousands of the most luminous z~7 galaxies in the LSST Commissioning and Science Verification data alone. At redshifts of z~1-3, corresponding to the main epoch of galaxy formation, IR photometry is essential to obtain robust stellar masses and star-formation histories for galaxies (e.g. Muzzin et al. 2013) and, therefore, to build up a complete picture of the mass build-up of galaxies over cosmic time. The addition of IR photometry to optical surveys can significantly improve photo-z performance, reducing the scatter and catastrophic outlier fraction, as key spectral indicators for old and evolved stellar populations shift into the IR at redshifts above 1 (e.g. Banerji et al. 2008, 2015, Jarvis et al. 2013, Bezanson et al. 2016). More accurate photo-zs will, in turn, enhance LSST's ability to use photometric samples to constrain the dark energy equation of state (e.g. Rhodes et al. 2017).

To exploit the galaxy images that LSST is capable of producing will require solutions to the coupled problem of detecting low-surface-brightness structures surrounding/between galaxies (e.g. tidal features and intra-cluster light) and robust de-blended photometry of galaxies along crowded lines of sight. Overcoming this challenge will unlock the merger history of galaxies (Figure 3.4; Kaviraj 2010, 2014 a,b), and create exciting synergies with future SKA diffuse HI detections. Overcoming this challenge at z > 0.5 is critical to secure identification, and mass calibration, of intermediate and high-redshift galaxy clusters that will deliver the strongest constraints cosmology through the growth of large-scale structure (Kelvin et al. 2012, Williams et al. 2016). Low-surface-brightness science using deep surveys is a UK area of focus and strength (see WP 3.7).

Figure 3.4 Deep CFHT MegaCam image of two local galaxies (Duc et al. 2011), which reaches a 1σ surface brightness limit of mag 29.5 arcsec^{-2}, similar to the expected LSST depth. The elliptical galaxy in the centre of the frame shows an abundance of low-surface-brightness tidal features from a recent merger. Right: an SDSS image of the same field, showing that these low-surface-brightness features are not visible in current wide-area surveys like the SDSS.

Strong gravitational lensing directly traces the total mass distributions of lensing galaxies, groups and clusters. It probes new regimes of sensitivity and resolution in high-redshift lensed galaxies, tests detailed predictions of CDM on sub-galaxy scales, and can deliver competitive constraints on cosmological parameters (Suyu et al. 2013). A large number of strong lenses could be used as accurate probes of background shear, thereby providing significant support for weak lensing (Birrer et al. 2017). LSST's sensitivity, image quality, survey volume and time-resolved measurements will transform our use of strong gravitational lenses as astrophysical and cosmological probes, enlarging samples of galaxy-scale lenses from $<10^3$ to 10^5 (comparable with, and complementary to, Euclid), discovering 10^4 strongly lensed quasars and 500 lensed type Ia SN (Goldstein & Nugent 2017); time delays can be determined for hundreds of these, plus thousands of group- and cluster-scale lenses.

The two problems of morphological classification and finding strong gravitational lenses are linked through the common technique of machine learning image analysis techniques. We now have significant machine learning and AI-assisted tools across the UK astronomy community, and this Phase B proposal will link this expertise and initiate a hub for code swapping, expertise sharing and novel development. Five of our original Phase B work packages had machine learning components and this growing area is set to dominate and influence LSST and other UK big data projects, with the STFC-funded CDTs providing a forum for sharing large-scale data analysis techniques and algorithms. *The impact of the descope is to remove the machine learning aspect of the lens discovery system and to significantly reduce the machine learning aspects of the galaxy morphology work. The case for a UK-wide effort to tackle these problems in large datasets remains strong.*

3.5 AGN and supermassive black holes

LSST will provide a step-change in determining the host properties - e.g. star formation, stellar populations - of active galaxies identified at gamma, X-ray and radio wavelengths as well as opening new parameter space to chart newly-triggered black hole activity in real time. The ubiquity of supermassive black holes in all bulge-dominated galaxies and their driving role in the formation and evolution of galaxies throughout cosmic time are now well accepted. Spending around 10% of their lives actively accreting matter, they are the most energetic radiation sources in the Universe. Consequently, the physics of accretion and black hole growth, the nature of galaxy-black-hole coevolution, and the relation to the host galaxy environment are forefront research questions. Accretion is a genuinely dynamic process with the underlying physics working on time scales of minutes to decades and with amplitudes of variability of up to three orders of magnitude. However, our current understanding of AGN originates predominantly from static data, limiting the ability to put the various phenomena and incarnations of AGN activities into a coherent picture. Only recently have we begun to glimpse the diagnostic and discovery power of the time-domain window, e.g. with the discovery of changing-look quasars/AGN or switch-off accretion.

LSST will be transformational for AGN. Its revolutionary combination of wide-area coverage with redshift depth, temporal sampling and a long time baseline will provide the first census of black hole accretion from quiescence to major outburst. LSST will allow real-time follow-up of accreting and quiescent black holes, specifically probing the triggering of black hole activity and jet formation via detection and multi-wavelength follow-up of gamma-ray, X-ray and optical flares produced by tidal disruption events around quiescent black holes. The UK has a uniquely strong international research pedigree in the study of Active Galactic Nuclei (AGN) across the electromagnetic spectrum and their role in galaxy formation and evolution, with leadership of major AGN programmes on eMERLIN, ALMA, JVLA+VLBA, Liverpool Telescope, VLT, Chandra, Swift, & Fermi, ELT, VISTA and DES. *The Phase B descope has mean that the AGN science within the UK Variability Broker has been removed, but the challenge remains as to the UK DAC will support analysis of these objects.*

3.6 UK leadership roles in LSST structures

Science from LSST will be done both within the LSST Science Collaborations and independently by teams with access to LSST data products and value-added data, and the UK community will adopt both approaches, as appropriate. We have the expertise and strength to add enormous scientific value to LSST data within a UK DAC: for example, our expertise in solar system, transient time domain survey operations and combining data over wavelengths (the near infra-red in particular) will provide UK scientists with combined data products that are world leading. Equally, we will play leading roles in the large collaborative teams needed to undertake some analyses. During Phase A, UK scientists have started to gain leadership roles in a number of the LSST Science Collaborations:

- **D. Alonso**: co-convenor of the DESC Large Scale Structure WG and member of DESC Council;
- **M. Banerji**: co-chair of the LSST Galaxies Science Collaboration;

- **B**. **Burningham**: co-chair of solar neighbourhood WG in Stars, Milky Way and Local Volume Science Collaboration (from June 2017);
- **E. Chisari**: member of LSST DESC Collaboration Council (Nov 16-Nov 18);
- **T. Collett**: co-convenor of the LSST DESC Strong Lensing Working Group;
- **W. Hartley** and **O. Lahav**: co-conveners of the DESC photo-z working group;
- **B. Joachimi**: DESC Operations Committee, DESC Membership Committee (until 2017);
- **S. Kaviraj**: co-chair of the LSST Galaxies Science Collaboration;
- **C. Lintott**: leads the LSST EPO development of Zooniverse as a citizen science platform;
- **M**. **Lochner**: co-leader of DESC Observing Strategy Task Force;
- **J**. **McEwen:** member of DESC Membership Committee;
- **H**. **Peiris**: member of DESC Collaboration Council (from 2015) and Publications Board;
- **I. Shipsey**: chair of LSST DESC Advisory Board (2016-17); DESC Spokesperson Nominations Committee (2016-17);
- **S. Smartt** and **K. Smith**: lead the development of LASAIR as a community LSST broker for the TVS Science Collaboration;
- **G. Smith**: chair of clusters Working Group in the LSST Galaxies Science Collaboration;
- **M**. **Sullivan**: co-chair of DESC Follow-up Task Force;
- **A. Verma**: co-chair of the LSST Strong Lensing Science Collaboration & leading the Strong Lensing Group in the Galaxies Science Collaboration;
- **J. Zuntz**: outgoing co-lead of the DESC Weak Lensing working group, Leader of Lensing/Large-scale structure cross-correlation project.

Most of these positions are within DESC, because it is the most mature and coherently-functioning Science Collaboration, with membership and publication policies, and a very detailed Science Roadmap, while others are currently looser associations of scientists developing infrastructure and ideas together. We expect UK astronomers to take leadership roles within the other Science Collaborations as they take shape during Phase B.

4. Project organisation and Project Management Plan

One of the goals of Phase A was to establish the organisational structures needed to support UK involvement in LSST. That has been achieved, with many processes defined and tested in practice, although further organisational evolution will take place during Phase B, to reflect the changing nature of the project and as the various international components of the LSST community become more closely integrated; one of our guiding principles is not to create parallel structures in the UK, but to enable UK astronomers to take leadership roles within the international LSST community. In this Section we outline our project organisation, while the Project Management Plan provided with our resubmission gives a more detailed view of how Phase B will work.

4.1 Internal Organisation

UK involvement in LSST proceeds through two organisations. The *LSST:UK Consortium* comprises all UK astronomy groups with a scientific interest in LSST – which is *every* astronomy group in the UK – and the *LSST:UK Science Centre* (LUSC) is a distributed organisation that conducts funded work on behalf of the LSST:UK Consortium. Figure 4.1 below, taken from *LSST:UK Governance*[5] document adopted by the Consortium Board prior to the start of Phase A and followed since then, shows the principal bodies set up by the Consortium.

The Executive Group comprises the Project Leader and five members elected by the Consortium Board (but not members of that Board). The Project Scientist and Project Manager normally attend Executive Group telecons, as does the Consortium Board Chair. The Consortium Board is the sovereign entity in LSST:UK, but it has delegated to the Executive Group the power to act promptly when needed between Board meetings, and the presence of the Board Chair at the Executive Group meetings helps advise when that is appropriate. The Executive Group also acts as an advisory body to the Board, since almost all Board papers have significant input from the Executive, along with a recommendation for action to be taken by the Board.

[5] https://lsst-uk.atlassian.net/wiki/download/attachments/1146928/governance.pdf

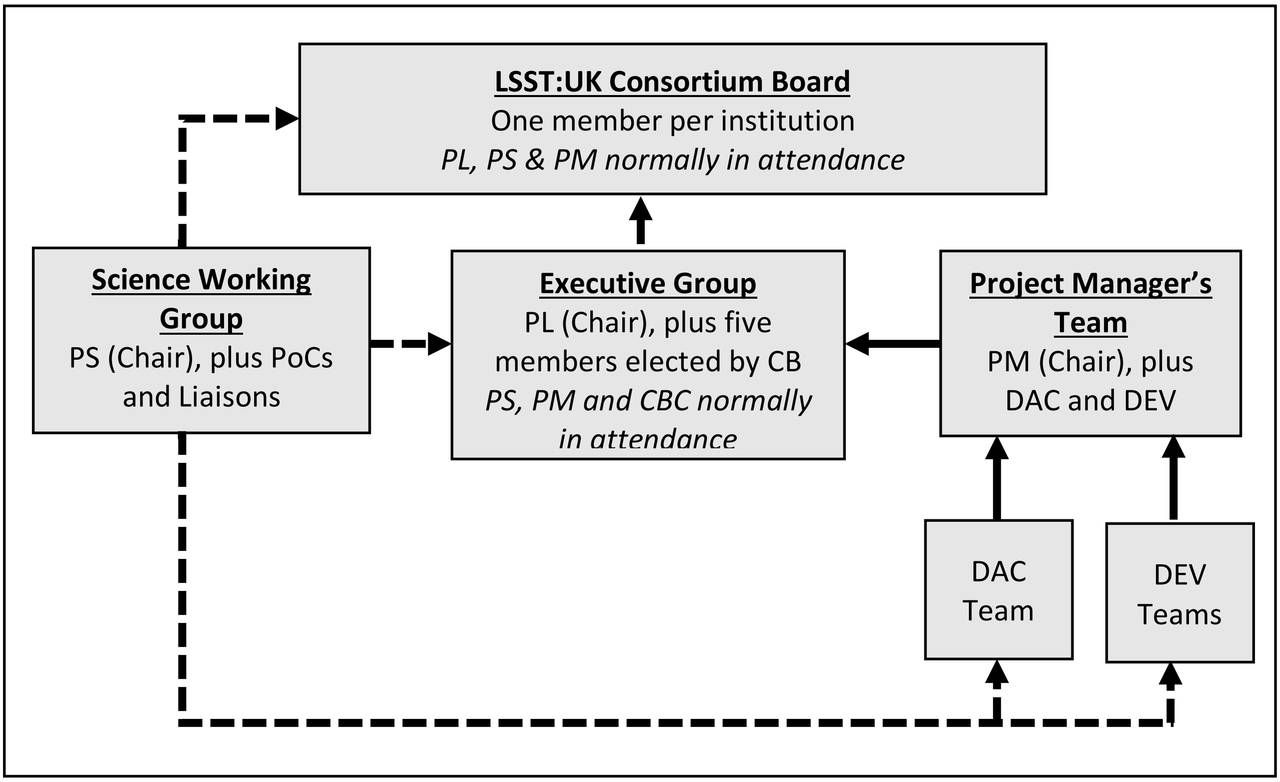

Figure 4.1 The LSST:UK management structure. PL, PS, PM and CBC denote Project Leader, Project Scientist, Project Manager and Consortium Board Chair, respectively. A solid arrow indicates reporting and a dashed one provision of scientific guidance.

Currently, Bob Mann (Edinburgh) is the Project Leader, Stephen Smartt (QUB) is the Project Scientist, George Beckett (Edinburgh) is the Project Manager and Tim Naylor (Exeter) is the Consortium Board Chair, with Roger Davies (Oxford) as the Deputy Chair. The elected members of the Executive are Alastair Edge (Durham), Richard McMahon (Cambridge), Victoria Scowcroft (Bath), Graham Smith (Birmingham) and Aprajita Verma (Oxford); Smith is also the LSST:UK Commissioning Coordinator. Not shown in this diagram is the LSST:UK PI Selection Committee, which is chaired by Vivienne Wild (St Andrews) and which advises the Board on PI selection.

During Phase A, the Consortium has developed a number of internal policies and exercised the procedures to implement them. An important milestone was the adoption of a set of *Professional Conduct Policies*[6], adapted from those of the LSST Dark Energy Science Collaboration and including a *Meeting Code of Conduct* policy that is now used for all Consortium-organised meetings; signing up to these policies is a requirement for retention of PI status and for attendance of LSST:UK meetings, respectively. There also exists a pair of Ombudspersons – currently Andreea Font (LJMU) and Martin Hendry (Glasgow) – with whom concerns over inappropriate behaviour can be discussed at any time. The Board has approved an *Equality and Diversity Statement* and has charged itself with seeking applicants for all Consortium positions who will expand the diversity of the Consortium leadership. In particular, gender balance is considered in the composition of all short-term committees, and gender statistics are monitored in all selection processes, but further work remains to be done to attain appropriate gender representation across all of LSST:UK.

A crucial task for the Consortium is to manage the lists of Affiliate PIs (i.e. faculty members) and Junior Associates (all others) who have LSST data rights; in recognition of the UK's operations contribution, it has 100 AP and 400 JA positions, which are recorded in an annual submission to LSSTC that enables those listed to join Science Collaborations and will, in due course, form the basis of access control lists for the UK DAC. The *LSST:UK Policy for Selection of Affiliate PIs and Junior Associates*[7] enshrines the principles under which these lists are populated, and has been

---

[6] https://lsst-uk.atlassian.net/wiki/download/attachments/50924278/LSSTUK_ProfessionalConductPolicies_v1p0.pdf
[7] https://lsst-uk.atlassian.net/wiki/download/attachments/6946926/PISelectionPolicy20170930accepted.pdf

applied (with slight tweaking in the light of experience) in selection rounds for four annual lists to date. The guiding principle is that PI status is awarded on scientific merit and in three-year terms, to ensure that the PI lists continue to reflect those within the UK community who are most active in LSST; applicants explain why their plans require PI status and those seeking status renewal describe what use they have made of it in the previous three years. In this way, we have managed the growth of the AP list, using it to motivate and reward engagement with LSST science planning, while monitoring its composition by science area and institution, to ensure that it remains representative of the Consortium as a whole. The 2019 selection round is currently underway, but the 2018 list has 90 out of the 100 places filled, and we expect the additional interest engendered by the arrival of Commissioning data to lead to enough strong applications to fill the remainder within the next couple of years. There are currently 98 Junior Associates, and the PI policy gives preference in AP selection to applications from existing JAs who have just taken on permanent academic positions, to encourage researchers to engage actively with LSST early in their careers, as is crucial for such a long-term project. Several of the AP places are filled by particle physicists, who are welcome to engage as fully as they wish with the scientific programme of LSST:UK, while others, with a less direct interest in astrophysics, have sought funding on particle physics consolidated grants to undertake technical work related to the LSST camera, mirroring the situation in the US, where DoE funding has drawn in a large number of particle physicists into LSST, especially into DESC.

The Consortium has also successfully implemented the procedures it has agreed for the periodic renewal of its fixed-term leadership positions. The Board undertook a review of the Project Leader and Project Scientist positions in December 2016, as a result of which Bob Mann's term was renewed to a point 18 months into Phase B. At the same time, Sarah Bridle stepped down as Project Scientist to take over leadership of *STFC Food Network+*, and the resultant search exercise concluded with the appointment of Stephen Smartt from a pool of strong applicants. Tim Naylor has been re-elected as Consortium Board Chair during Phase A, and the election process for the Deputy Chair position is currently underway. Elections for renewal of three of the five Executive Group posts have also taken place, with the (intentionally staggered) elections to re-fill the remaining two to follow once the Consortium Board Deputy Chair election has taken place. The initial appointment of the Project Manager runs until the end of Phase A, as do the terms of the Education & Public Outreach Coordinator (Andrew Norton, Open University) and the Training Coordinator (Nic Walton, Cambridge) who lead, respectively, the LUSC-EPO and LUSC-TRN activities.

4.2 LUSC reporting and planning procedures

As indicated in Figure 4.1, the funded LUSC DAC and DEV teams report progress through their local work package leaders (i.e. typically the PI of the particular institutional grant) to the Project Manager. From this the Project Manager creates a six-monthly report to the Consortium Board and the annual Project Assurance Report for STFC. In addition to the progress report from the Project Manager each six-monthly Consortium Board meeting also receives updates on two living documents: the Project Leader maintains a *Long-Term Plan*, which defines the strategy for LSST:UK and LUSC, while the Project Scientist updates a *Science Requirements Document* that translates the Long-Term Plan into deliverables expressed in scientific terms. This structure has proven to work well during Phase A, with clear reporting lines from the grant-funded LUSC staff all the way up to the Consortium Board which represents the UK community on behalf of whom those staff work, while the Science Working Group is able to inject scientific guidance at whatever level it is needed.

For Phase B, a slightly different approach is required, as the nature of the programme is changing, from the more R&D-focussed activities of Phase A to the production of an operational DAC and DEV software that can be used in earnest from Commissioning onwards. To that end, we have created a Project Management Plan (included as an attachment in this resubmitted application) which will form the basis of the Project Plan against which progress will be monitored during Phase B. Amongst other things, this details the process whereby the Executive Group will perform the acceptance of deliverables, as well as its role in overseeing change management through consideration of change requests (e.g. proposing variations to the scope or timescale of deliverables) submitted by the Project Manager on behalf of a Work Package Leader. If an approved change request has significant resource implications, the Project Leader will be actioned by the Executive Group to seek permission from the STFC Oversight Committee for the drawing down of appropriate Working Allowance funds

4.3 External interactions

LSST:UK and LUSC have a number of important relationship with external entities, in the UK and beyond. For Phase A, the LUSC has had light-touch oversight from STFC - with Chris Woolford (Astronomy Facilities Programme Manager) attending Consortium Board meetings, and provided before each with the current versions of the Long-Term Plan, Science Requirements Document and Six-Monthly Report – but we expect that STFC will establish an Oversight Committee for Phase B.

An increasingly important interaction is with the IRIS Consortium, which is planning a coordinated computing infrastructure for STFC science: Mann and Beckett sit on its Delivery Board and Beckett is also a member of its Technical Working Group.

Scientific interactions with the international LSST community are mediated by the Science Working group, chaired by the Project Scientist and comprising *Points of Contact* with the different LSST Science Collaborations (and working groups thereof) and *Liaisons* with major external projects (e.g. SKA, Euclid). There are currently three UK Science Collaboration Co-chairs – Aprajita Verma in the Strong Lensing Science Collaboration, plus Manda Banerji (Cambridge) and Sugata Kaviraj (Herts) in the Galaxies Science Collaboration – and LSST:UK members are making an increasing impact as the Science Collaborations are internationalising, as noted in Section 3.6.

As noted in Section 1, the International Contributors to LSST operations are represented within the operations management structure by a representative from the LSST Corporation (LSSTC), who is advised in that role by the LSSTC Corporate Operations Committee (COC). Mann was invited onto the LSSTC Operations Taskforce that defined the terms of reference and membership of the COC and he is now the UK's representative on it: the UK and France are the only two International Contributors with a permanent seat on the COC. Membership of the COC also entitles Mann to attend meetings of the LSSTC Board of Corporate Directors on behalf of LSST:UK. Beckett is currently a member of an important, fixed-lifetime working group set up by LSSTC and the LSST Project to consider the roles and responsibilities of international DACs.

5. Summary of Phase A Outcome

The LUSC Phase A programme comprised Work Packages for LUSC Management, LUSC-DAC and LUSC-DEV, with the DEV WP divided into five sub-Work Packages conducting R&D in four different science areas and on one technical topic that will influence LSST science. At the time of writing, all WPs are on track for successful completion before the end of Phase A, although a few deliverables have had their deadlines extended, due to the time taken to replace project staff who have left.

5.1. WP1: Management

The Phase A Management WP comprises five Tasks: development of the Long-Term Plan and of the Science Requirements Document; management of DAC and DEV staff; external liaison; and reporting to STFC oversight bodies. The two living documents continue to be updated, and Project Assurance Reports are being supplied to STFC annually. The Project Manager chairs monthly meetings of the DAC team (with more regular meetings for the teams undertaking each current DAC activity), while the more research-oriented nature of the Phase A DEV R&D WPs has led the PM to monitor the progress of the DEV staff mainly through reports from their local supervisors. The liaison activities have progressed well, with regular contact maintained with the LSST Director, and with the Project Leader joining the LSSTC Corporate Operations Committee. The Project Leader and Project Manager have regular telecons with the technical management of the LSST archive sites at NCSA and CC-IN2P3 and are members of the IRIS Delivery Board and its Technical Working Group.

5.2. WP2: LUSC-DAC

The Phase A DAC WP had four Tasks, which we report on briefly in turn below.

Task 2.1: Testbed DAC system

This Task has both software and hardware aspects. Given its modest funding, the UK DAC must deploy (possibly with tailoring) open source software developed by the Project, rather than develop its own user access software from scratch. The Project's DAC centres on the LSST Science Platform (see Fig. 5.1) which provides support for three different modes of data analysis: simple visualisation and basic interactive data exploration will be performed in a **web portal**; **Jupyter Notebooks** will allow a higher level of interactive analysis through the scripted manipulation of data; and the use of **web APIs**, such as International Virtual Observatory Alliance (IVOA) standards, will enable the use of existing tools (e.g. TOPCAT, Aladin) and the extraction of data for analysis outside the DAC.

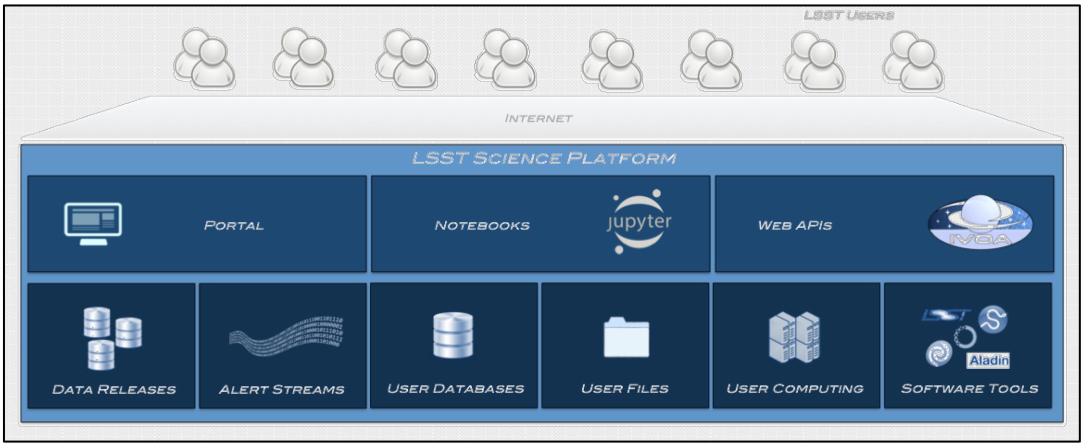

Figure 5.1 The LSST Science Platform (from Juric et al 2017)

The DAC team in Edinburgh has studied all three aspects of the Science Platform during Phase A, gaining experience with management of the relevant technologies and prototyping extensions needed to support particular UK science priorities. For example:

- **Web Portal**: IPAC's Firefly system will provide the main LSST web portal, but a crucial addition in the UK DAC will be portal-based support for the exploration of transient alerts. This will extend the portal developed by the QUB team, a version of which is currently deployed on DAC hardware in Edinburgh as part of LASAIR, our alert handling testbed. A first version of LASAIR (lasair.roe.ac.uk) has been deployed to handle the alert stream from the Zwicky Transit Facility (www.ztf.caltech.edu), an LSST precursor whose event-generation software is being developed by the same team as will produce the LSST Prompt Products.
- **Jupyter Notebooks**: The UK DAC must support users running shared Notebooks within an environment that makes effective use of compute and storage resources while enforcing the authorisation required to respect LSST data access rights. To probe the technical challenges involved in that, we have collaborated with Edinburgh colleagues running the Noteable (https://noteable.edina.ac.uk) service and have deployed a JupyterHub serving Notebooks that run on a Kubernetes cluster on the University's OpenStack cloud, to be followed by deployment on IRIS, once an authorisation mechanism for IRIS has been agreed.
- **Web APIs**: DAC staff are working to ensure that IVOA standards will scale to the data rate of LSST alerts. Roy Williams is one of two editors of the VOEvent standard, while Dave Morris is vice chair of the IVOA Time Domain Interest Group, and, with funding from ASTERICS (asterics2020.eu), they are driving the evolution of the relevant IVOA standards in the light of LSST Level 1 requirements: e.g. providing VOEvent packets in the Avro data serialisation format used by the Apache Kafka framework adopted by the Project. This work will continue under ESCAPE, the recently-funded successor to ASTERICS in the H2020 programme.

Following Science Board guidance at the time of the Phase A award, the hardware aspect of the DAC testbed is being addressed through IRIS. It was awarded the full £1.5M on offer in STFC's 2017 *Computing Capital Enabling E-Infrastructure* call, to provide cloud compute resources at RAL and Manchester, a ~1 PB database cluster to be installed in the University of Edinburgh's Advanced Computing Facility (ACF), plus capitalised software development at Cambridge and RAL. Figure 5.2 shows the design for the database cluster, which includes a modest local cloud compute resource, funded through the Phase A DAC grant, to provide a DAC testbed.

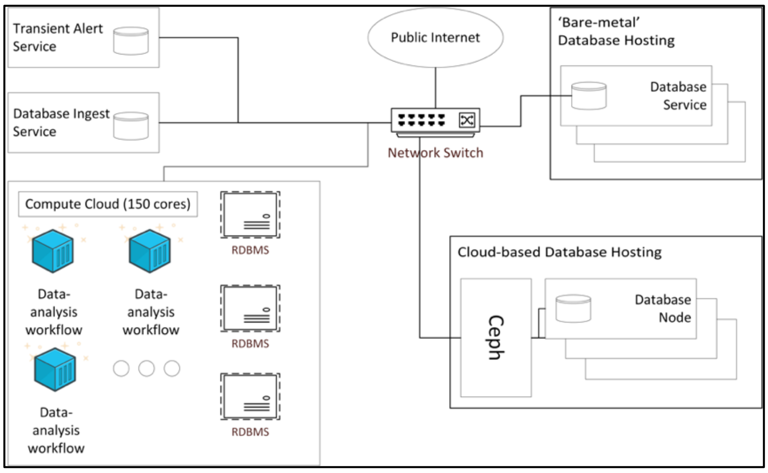

Figure 5.2 The design for the LSST:UK/UKT0 database testbed.

This design was developed in consultation with colleagues from the French LSST DAC team at CC-IN2P3 in Lyon, in the light of the experiments they have performed assessing the scalability of qserv, the LSST database system. It features storage nodes in two configurations - "bare-metal" nodes, similar to those used for WFAU's current sky survey databases, and a cloud-based hosting system using the Ceph distributed file system – to probe the trade-offs between performance and ease of administration that must be balanced in a multi-PB production database system like that which will run in the UK DAC. The co-located compute cloud to run analysis tasks (e.g. from Jupyter Notebooks) against LSST qserv databases was included in the design to allow the comparison of the performance of local and remote (e.g. in the IRIS clouds at Manchester or RAL), providing important information for deciding how to provision the UK DAC portion of the future IRIS system.

Task 2.2: Participation in LSST Data Challenges

The data challenges originally planned by the Project Data Management team have not taken place, but much of the same functionality is being exercised in the Data Challenges being undertaken by the Dark Energy Science Collaboration. The DAC team has been participating in DESC DC2, which will involve running the Level 2 pipeline at CC-IN2P3 on simulated images, providing very valuable experience in the transfer and storage of Level 2 products before the arrival of the first data.

Task 2.3: Support of Phase A DEV work

As noted below, the DAC team has provided computational support to WP3.1 and 3.3.

Task 2.4: Implementation Plan for UK DAC

Several iterations of DAC operations plans have been made. The software plan centres on implementation of the LSST Science Platform, but details of where tailoring is required are dependent on the Phase B DEV programme. The same goes, to an extent, for the storage requirements, but estimated (file and database) storage and compute requirements for the first three years of Phase B were included in the "*eInfastructure for the exploitation of the UK National Facilities and STFC Frontier Science programmes*" business case that resulted in £16M being awarded by BEIS to the IRIS consortium. A more concrete implementation plan is being developed, as the technical details of the IRIS system – e.g. its common authentication, authorisation & auditing mechanisms – are taking shape, with DAC staff active in the relevant IRIS technical working groups.

5.2. WP3: LUSC-DEV

The Phase A LUSC-DEV programme is divided into five Work Packages:

WP3.1: Weak Lensing

The weak lensing WP was initially led by Sarah Bridle (Manchester) and then by Catherine Heymans (Edinburgh). Joe Zuntz was the funded postdoc in Manchester, but, following his move to a tenure-track post in Edinburgh, the role was taken over by Robert Schuhmann (who had previously worked on WP3.3 at UCL, as described below). The team have assessed the performance of the Level 2 data products for weak lensing through analysis of image simulations and images from the pre-cursor Hyper Suprime-Cam (HSC) survey reduced using an early version of the LSST pipeline. They have also built a suite of diagnostic tools to verify data quality, ready for application to Commissioning data, and worked with the DAC team to prepare for running large scale simulations on GridPP. The

WP3.1 team have been very active within DESC: Zuntz is convener of the Weak Lensing WG and is one of three Weak Lensing “Pipeline Scientists”, leading development of mission-critical software.

WP3.2: Milky Way

This WP, undertaken under the supervision of Vasily Belukurov (Cambridge), is developing better algorithms for star/galaxy separation and for detecting tidal streams in multiband photometric catalogues. Good progress has been made, although the deadlines for some deliverables have been moved into the final year of Phase A, due to a delay in finding a suitable replacement for the grant-funded postdoc originally working on this project. The team has been making detailed comparisons of star/galaxy separation in Sloan Digital Sky Survey and Dark Energy Survey datasets, along with a re-reduction of SDSS data using the LSST pipeline. They have developed a star/galaxy classifier based on Gaussian Mixture Models that outperforms existing tools, and can be improved further with the inclusion of multi-colour information. Detecting tidal streams in the large LSST dataset relies on having an accurate, smooth model sky background across the whole survey region, and the team have been developing a new approach using piecewise radial basis function interpolation on the sphere. This is proving effective, and should allow the automated detection of much fainter streams than currently known, shedding important light on the Milky Way merger history.

WP3.3: Transient Server and Supernova Photometric Classification

This WP comprises three activities. Hiranya Peiris and Jason McEwen led work at UCL developing machine learning techniques for supernova classification (Lochner et al 2016) that was adopted by the Dark Energy Science Collaboration and used extensively (Marshall et al 2017) in the international initiative to optimise the LSST survey strategy. Mark Sullivan (Southampton) has been leading work to address one of the most significant gaps in the survey strategy simulations, namely the lack of realistic light-curve templates for stripped envelope core collapse supernovae. They have developed templates using the extant data for this class of sources, and made them publicly available. The DAC team has helped the Southampton team prepare for large-scale simulation runs using these templates on GridPP, and this work is playing an important role in the PLAsTiCC transient classification data challenge currently being performed by the Dark Energy and Transients & Variables Science Collaborations, and which Peiris and Sullivan are helping to coordinate. Stephen Smartt’s team at QUB have been working with the DAC team in Edinburgh to extend their existing alert handling and transient classification system so that it will scale to LSST data rates. A prototype of the resulting broker, LASAIR, is taking shape, built upon the same technology (the Apache Kafka stream processor) chosen by the LSST Project Data Management team for their own alert handling work, and retaining the user-friendly scientific capabilities of the existing QUB Sherlock system; its development is being driven by experience handling the Zwicky Transient Facility alert stream. This is already public, with further functionality enhancements to come before the end of 2018[8].

WP3.5: Solar System

This WP, led by Alan Fitzsimmons (QUB) is developing software to enable automated stacking and image analysis of solar system objects, and provide filter-dependent light curves for phase function and spare light-curve fitting. This project was scheduled for the latter half of Phase A, so, is less well advanced than the others, but good progress has been made, with full access being granted to the QUB team of the images and photometric catalogues from the ATLAS transient survey, which will serve as an excellent testbed for the software development. Their software uses ephemeris snapshots for all known moving objects in the solar-system to predict where they will lie in ATLAS exposures and thereby extract measurements of them from pre-extracted source catalogues. This is yielding data on thousands of asteroids and is a very useful first step in preparation for LSST.

WP3.6: Sensor Characterisation

The CCD detectors in the LSST camera are unusual in several respects: they are thick, back-illuminated, back-biased, have 16 outputs and run at 550kHz. Ian Shipsey (Oxford) has used local funds to commission a state-of-the-art optical test stand (shown in Section 9.5 below) that allows characterisation of several aspects of CCD performance complementary to the acceptance testing

[8] See http://lasair.roe.ac.uk/

being performed elsewhere in the LSST Project and important for understanding image quality systematics. Initial results from this programme have been published (Weatherill et al 2017), and good relationships have been developed with (UK-based) e2v, one of the two detector vendors, with the camera team in the Project and with the DESC Sensor Anomalies Working Group, which will prove invaluable during Phase B, as Commissioning images becomes available to complement the lab data and allow a more detailed modelling of how detector effects translate into systematic errors in cosmological parameters derived from galaxy shape analysis through weak gravitational lensing.

6. Introduction to the Phase B Work packages and UK prioritisation

6.1 Phase B WP selection and prioritisation

The LUSC programme is designed to meet the scientific requirements of the whole LSST:UK Consortium, and a careful, science-driven process has been followed to select the Phase B work packages for inclusion in this proposal. During the first half of 2017 the Consortium held a series of workshops, involving US-based Project staff (both the Data Management and Commissioning teams) and the Science Collaborations, to initiate discussions amongst UK researchers active in all areas of LSST science regarding the desired programme for Phase B. This led, under a procedure defined by the Project Scientist, with guidance from the Executive Group, and endorsement from the Board, to the issuing, in October 2017, of calls for groups to develop LUSC-DEV work package proposals and for institutions to express interest in undertaking a set of specified LUSC-DAC work packages; the different approach for DAC and DEV reflected the differing nature of the two activities.

Three DAC expressions of interest were received, along with 24 DEV proposals, covering the whole spectrum of LSST science, and featuring applicants from a wide range of Consortium institutions. These were assessed by a panel composed of one member each of the Board (Michelle Collins, Surrey) and the Executive Group (Alastair Edge, Durham), plus two representatives from the wider UK LSST science community (Alan Heavens, Imperial; Jacco van Loon, Keele); these individuals were chosen to provide expertise across the necessary range of science areas, while not being conflicted through personal involvement in any of the cases under consideration. To ensure consistency of the Phase B programme with the detailed plans of the Project and of the Science Collaborations, the panel also included two leading members of the LSST community in the US: Melissa Graham (U. Washington) from the Data Management team; and Phil Marshall (SLAC), the Spokesperson of the Dark Energy Science Collaboration, whose coordination of the study of the LSST survey strategy has exposed him to the goals of all the Science Collaborations. The Project Leader and Project Scientist were in attendance, *ex officio*, to provide background and strategic guidance, but were not voting members of the panel, as each was PI of a case under consideration.

The proposals were assessed on the basis of their scientific excellence, their exploitation of particular UK expertise and their ability to help secure UK leadership in our community's high-priority science topics. Proposals that were too close to immediate science exploitation (and, hence, fundable through the Astronomy Grants Panel) were rejected, with serious consideration being given only to proposals for the development of scientific infrastructure, in the form of software and data products required for Level 3 analyses, with realistic objectives and led by scientists with a strong track record. This yielded a series of recommendations that were then debated by the Board, resulting in approval of the Work Package programme presented (together with an Education and Public Outreach Work Package that we have omitted at the request of Science Board) in the original April submission on behalf of the Consortium as a whole, and as representative of its priorities for Phase B.

6.2 Descoping the Phase B programme

The feedback from the July PPRP meeting, included a cost envelope for Phase B, necessitating a significant descope of the original programme. We reconvened the Phase B selection panel to identify a prioritisation amongst the original programme, guided by instructions from the Consortium Board that we should maximise the DI staff effort that could be supported within that cost envelope. To that end, the Consortium Board decided that we should strip the pooled staff and DI support staff from all institutional requests and that we should also halve the rate of DA investigator effort to 0.1 staff year per staff year of DI effort. The Board decided not to allow the delay introduced to the start of Phase B to lead to a knock-on delay to the start of Phase C, so that Phase B should be reduced to 45 months' duration, now running from 2019-07-01 to 2023-03-31.

After some iteration between the Board, the panel and the Work Package leaders the following WP-specific descopes to the other Phase B programme were agreed:

- **WP2**: **LUSC-DAC**. reduced by 1 DI s.y.
- **WP3**: **LUSC-DEV**:
  - WP3.1 (UK solar system server software): cut completely
  - WP3.2 (LASAIR: UK transient broker): descoped from 48 to 36 DI staff months.
  - WP3.3 (Transient classification & spectroscopic follow-up): cut from 48 to 24 DI s.m.
  - WP3.4 (UK variability broker for LSST): cut completely
  - WP3.5 (LSST & near-IR data fusion): descoped from 72 to 50 DI staff months
  - WP3.6 (3D LSST: photometric redshifts): cut completely
  - WP3.7 (Morphology and Low Surface Brightness science): only LSB funded (30 s.m.)
  - WP3.8 (Strong lens discovery): cut completely
  - WP 3.9 (PSF & sensor characterisation/modelling): reduced from 48 to 36 DI s.m.

This yields the revised programme listed in Table 1 above: N.B. the Management Work Package was protected, as were the two smallest DEV WPs (namely, WP3.10 and WP3.11), which had no more than 24 s.m. each to start with.

The scientific loss from having no Phase B funding to support DEV activities in solar system science, stellar & AGN variability, photometric redshift estimation and strong lens discovery is obvious, but the Consortium Board decided that it was better to have a programme featuring a smaller number of sufficiently well-resourced Work Packages than to "salami-slice" every Work Package and risk having a number that would be incapable of reaching their goals. Where the resources have been reduced for a Work Package, a corresponding descope to the work programme has been identified, but there remain risks internal to these Work Packages, which are presented in the Risk Register and which drive the estimate of the Working Allowance required for this revised Phase B programme.

6.3 Costings

We have adopted a standard set of costings. As noted above, we request 0.1 FTE of supervisor effort for each staff year of DI researcher/developer effort, with justification given in each case as to the appropriateness of the supervisors for each activity. We request £4k per (researcher+supervisor) staff year for T&S, since attending LSST meetings in the US will be crucial to the success of all proposed WPs, plus £1k per (researcher+supervisor) staff year for personal computing, etc.

6.4 Phase B Work Package Planning and Progress Monitoring

As agreed with the PPRP Chair for the LSST proposal and STFC Office staff, the detailed planning information for the Phase B Work Packages is presented in the *Project Management Plan* (PMP) included with this resubmission, not in this Case. The PMP identifies Milestones for Phase B, as well as specifying the Deliverables from each Work Package and describing the process by which they will be assessed for acceptance. In the following three Sections, we provide a narrative description for Work Packages 1, 2 and 3, respectively, including Gantt charts to summarise each at the sub-WP level. As requested, we provide brief biographies of the key staff driving the Phase B programme, presented in the description of the first Work Package with which each is associated.

All WP1 and WP2 staff are currently in place, as are the DI staff for WPs 3.2, 3.7, 3.9, 3.10 and one of the two DI staff for WP3.5; recruitments will be required for the second DI staff member working in Cambridge for WP3.5, and one DI staff member each at Southampton and Exeter for WPs 3.3 and 3.11, respectively. The WP2 staff are long-term members of the Wide-Field Astronomy Unit in Edinburgh, and, similarly, several of the WP3 staff are well established in their host institutions, but we include in the Risk Register an item for the potential loss of project staff during Phase B. To reduce the impact of losing crucial project staff, we shall ensure that key project information is well documented on the LSST:UK wiki, and, where possible, experience is shared amongst several staff, but this will be difficult in WP3, where most DEV teams comprise only one member.

7. Work Package 1: Management

Phase B WP1 aims to ensure that all resources are effectively deployed towards the principal Phase B objective of maximising the lessons that can be learnt from Commissioning to prepare the LSST:UK Consortium for the start of survey operations. This will involve liaison with the organisations outlined in Section 4, as well as the project management of the grant-funded LUSC staff.

WP1.1 Maintenance of the Long-Term Plan

The Project Leader is responsible for a living document that encapsulates the Consortium's plan for the full 18-year LUSC lifetime, with increasing detail as it approaches the current date. This will involve a periodic revision of science priorities from the Science Working Group and resource estimates from the Project Manager, together with strategic guidance from the Consortium Board.

WP1.2 Maintenance of the Science Requirements Document

The Project Scientist, aided by the Science Working Group, is responsible for the production and periodic revision of a *Science Requirements Document* that translates the *Long-Term Look* into clearly defined deliverables. This will undergo a major revision before the start of Phase B, based on the outcome of the DEV work package selection, which now defines the UK priorities very clearly.

WP1.3 Management of the DAC and DEV staff effort

As detailed in the Project Management Plan, the Project Manager is responsible for delivering the Phase B objectives, within the constraints of project resources and the time constraints of the funding. The day-to-day management of the DEV teams is devolved to the respective Work Package Leaders, who report exceptions and highlights to the Project Manager on a regular basis. The Project Manager oversees the process by which Deliverables are assessed for acceptance, as well as handling the escalation of any significant exceptions through Change Requests to a possible request for release of Working Allowance funds. A key role for the Project Manager during Phase B is ensuring that the development of the UK DAC proceeds in sync with both that of the international DAC network and the IRIS computing infrastructure.

WP1.4 Coordination of LSST:UK contributions to Commissioning

The LSST Project will invite suitably qualified researchers to participate in Commissioning, some visiting Chile to take part in Commissioning observations, and a larger number visiting SLAC, where the reduction and analysis of the data will be centred. These visits will be an excellent opportunity for UK researchers to get early experience of LSST data products, and provide feedback to optimise the data reduction pipelines. To support this important opportunity fully, we have appointed a Commissioning Coordinator (Graham Smith, Birmingham) to liaise with the LSST Systems Scientist, Chuck Claver, who leads the LSST Commissioning Team. He will identify suitable UK researchers to participate in Commissioning, coordinate UK input to the selection of Commissioning fields, and represent the UK's interests in planning for commissioning of both physical infrastructure and software pipelines developed by UK researchers. We request a travel fund to support these visits.

WP1.5 External liaison

A crucial task in Phase B for the Project Leader and Project Scientist is to maintain a close collaborative interaction with a number of external bodies, most notably the LSST Project and Science Collaborations and the national LSST teams that are being established in other countries as LSST internationalises. To ensure that the UK voice is heard within LSST, the Project Leader will continue to participate in the LSSTC COC and we will seek UK representation on the LSST Science Advisory Committee. We request additional T&S for the Project Leader and Project Scientist to effect this external liaison, plus a central travel fund to enable Consortium members not in receipt of Phase B funding to attend Science Collaboration meetings in the US; our Phase A travel fund proved very effective in facilitating integration of UK researchers into the Science Collaborations.

WP1.6 Reporting to STFC

The Project Manager, Project Leader and Project Scientist will, together, prepare the annual Project Assurance Report, and other reports as directed by the STFC oversight committee.

WP1 Summary Gantt Chart

WP	Task Name	Start	Finish
1.1	Maintenance of Long-term Plan	01/07/2019	31/03/2023
1.2	Maintenance of Science Requirements document	01/07/2019	31/03/2023
1.3	Maintenance of DAC and DEV staff effort	01/07/2019	31/03/2023
1.4	Coordination of LSST:UK contributions to Commissioning	01/07/2019	31/03/2023
1.5	External Liaison	01/07/2019	31/03/2023
1.6	Reporting to STFC	01/07/2019	31/03/2023

Figure 7.1 Summary Gantt chart for WP1.

WP1 Resources

For the 45 months of Phase B, we request DA staff effort for the Project Leader (Mann) and Project Scientist (Smartt), at a constant 0.2 FTE each across Phase B, and for the Commissioning Coordinator (Smith), at a rate of 0.1 FTE, plus DI staff effort for the Project Manager (Beckett) at 1 FTE. To support UK participation in the three-year Commissioning we request a £40k T&S budget, to cover 4 trips to Chile (£1.6k per trip for travel, with local costs met by the Project) and 8 trips to SLAC (£3.8k per two-week trip, staying at the SLAC Guesthouse), plus associated travel by the Commissioning Coordinator. We request £25k per annum for a central travel fund to support attendance of Science Collaboration meetings by Consortium members without Phase B funding, and an additional £2.5k per annum T&S for each of the Project Leader, Project Scientist and Project Manager to support their external liaison roles. We also request £1,383 for videoconferencing and £34,080 for wiki software licencing, respectively, to retain the systems used during Phase A.

WP1 Key Staff

Bob Mann has been advocating for UK involvement in LSST since 2007 and has been LSST:UK Project Leader since the Consortium formed in 2013. He also leads the development of the UK DAC, which builds on his two decades of sky survey data management experience as a member of the Wide-Field Astronomy Unit, of which he has been Director since 2012. Mann is the UK representative on the LSSTC Corporate Operations Committee.
Stephen Smartt has a long track record in running wide-field sky surveys for transient objects. With Ken Smith (and Dave Young) he built and managed the Pan-STARRS Transient Server at Queen's, which is now processing the ATLAS data stream. Smartt is recognised as a scientific leader in this field with a series of major discovery papers. He was PI and Survey Director of PESSTO, ESO's public survey for transient objects.
Graham Smith co-chairs the Galaxy Clusters Working Group of the LSST Galaxies Science Collaboration, is a member of the LSST:UK Executive Committee, is the LSST:UK Galaxy Clusters Point of Contact, and has recently been appointed as the UK's Commissioning Coordinator. He is also a member of the Dark Energy and Strong-lensing Science Collaborations (DESC and SLSC).
George Beckett is a PRINCE2-accredited project manager, supported by a Diploma in Project Management from Stevenson College of Edinburgh, and with 14 years of experience of managing software-engineering projects. During 2012-2015, he was seconded to the Pawsey Supercomputing Centre in Perth, Western Australia, where he held the position of Deputy Director and Head of Supercomputing. His role focused on developing supercomputing facilities for, amongst others, the astronomy community (most notably the Square Kilometre Array telescope and its precursors).

8. WP2: LUSC-DAC

8.1 Background

Under the terms of the UK Memorandum of Agreement with LSSTC, the UK may choose to operate a Data Access Centre or pay a *per capita* data access fee for UK data rights holders to use the

Project-provided DACs in the US or Chile. The LSST:UK Consortium decided at the outset that maximising UK scientific return from LSST strongly motivates operating its own DAC. Doing so will provide greater control over the scientific analyses that we can run than would be the case if the UK data rights holders were simply a small fraction of the user community of a DAC primarily intended to serve the many thousands of US astronomers. It will also enable us to integrate with the LSST data releases the other datasets that we require for much of our priority science. The size and complexity of these – together with proprietary restrictions, in the case of ESO data, for example – preclude their inclusion in the US DAC, which is designed solely to host LSST data. The UK DAC must, therefore, be sufficient to enable the UK community to take control of its science direction and to avoid liability for the data access fees associated with use of the US DAC.

In practice, the international nature of the Science Collaborations argues against DACs for each national community, and savings could be made by coordination to share development and operations resources and, perhaps, to have different DACs specialising in different kinds of analysis that requiring particular hardware, ancillary data or scientific expertise. The practicalities of operating a coordinated DAC network was initially discussed in a working group set up by the Project and in a sub-committee of the LSSTC COC, while a more detailed assessment is currently being undertaken by a working group jointly established by the Project and LSSTC, on which George Beckett represents the UK. Operating a UK DAC within a coordinated international network will allow LSST:UK to retain control where required over aspects of its exploitation of LSST data, while sharing resources and expertise that will reduce the overall cost of doing that in the longer run. The ultimate form of the DAC network for production (i.e. LUSC Phases C and D) will take shape during Phase B, so the objective for the Phase B DAC work package is to perform the deployment work needed to meet the immediate science requirements of the UK community during Commissioning, while undertaking development work that retains the flexibility needed to continue satisfying those demands whether they end up being delivered through a UK-only DAC or a UK DAC contributing its "fair share" of resources to an international network.

Flexibility will also be needed to take advantage of the possible pooling of computing resources across STFC science areas through IRIS. From the UKT0 Collaboration Meeting in March 2018, it is clear that many of the computational requirements of LSST:UK can be met through the same cloud compute and storage services envisaged for the next generation of particle physics computing, but the requirement for multi-PB databases distinguishes the survey astronomy community from the rest of PPAN science, so that aspect of IRIS operations must be led by astronomy. The Edinburgh DAC team's ability to respond flexibly has been demonstrated during Phase A. The Phase A DAC award provides an average of 1.6 FTE, but has funded four staff to date, as we have exploited the pool of experienced staff in the Wide-Field Astronomy Unit (WFAU) and EPCC (www.epcc.ed.ac.uk) to pull in the right effort for each activity, while additional contributions have been made by staff supported by other WFAU projects with overlapping objectives.

8.2 The Phase B DAC Work Packages

We described in Section 5 the plan to have a prototype UK DAC running on the LSST:UK/IRIS database cluster and its co-located compute cloud by the end of Phase A. This includes the technologies (e.g. qserv, Jupyter Notebooks) that we expect will form the core of the production DAC and run within a distributed cloud infrastructure like the eventual IRIS system. The task for the Phase B DAC programme, therefore, will be to develop this prototype into a robust production system, ready for the arrival for the first LSST data, from Commissioning onwards. The total DI effort requested for DAC work in the original Phase B proposal was 13 staff years - 3 FTE in the first three years, and 4 FTE in Year 4, once operations have started, with these resource requirements estimated on the basis of the experience of the Wide-Field Astronomy Unit, which currently deploys ~2.75 FTE to operate its archives for current ESO public surveys, undertaking activities analogous to WPs 2.2, 2.4, 2.5 and parts of 2.1, albeit for datasets that are two orders of magnitude smaller and served to a much smaller user community. *In the descoped Phase B programme, a total of 135.5 sm is available for the DAC, at an average rate of 3 FTE. WP2.1 (DAC Management) and WP2.4 (Provision of the DAC platform) are, essentially, fixed costs, so their FTE levels have been retained from the original Phase B submission, while we have also protected the FTE level allocated to WP 2.3 (Alert Handling Infrastructure), since that is likely to be the most technically-challenging aspect* *of DAC work during Phase B, and is the one area for which WFAU staff have no operational experience from the ESO public surveys, although they are developing relevant expertise through prototyping the handling of the ZTF alert stream in Phase A. That leaves the plans for WP2.2 and WP2.5 to be descoped, as discussed below.*

Work Package 2.1: DAC Management

WP2.1 will liaise with IRIS over DAC resource provision and with LSSTC over implementation of PI data access rights. It will allocate resources to users, plus Virtual Organisation management, management of the DAC team and planning for DAC operations in Phase C. These tasks will be undertaken by a Phase B DAC Management team - comprising the DAC Manager (Roy Williams), the three DAC Co-Is (Peter Clarke, Andy Lawrence and Bob Mann), and the LUSC Project Manager (George Beckett) - which will report to the LSST:UK Executive Group. The team's interactions with the LSST COC, the IRIS Delivery Board and with the leaders of the US and French DAC teams at NCSA in Illinois and at CC-IN2P3 in Lyon, respectively, will ensure the technical cooperation necessary for effective development of the UK DAC.

WP2.1 Resources

We request 0.25 FTE DI effort for a DAC Manager (Roy Williams). Lawrence will lead WPs 2.1 and 2.3, while 2.2 and 2.5 will be led by Mann, and 2.4 by Clarke, each funded at the level of 0.1 FTE per FTE of effort of the DI staff they supervise.

WP2.1 Key Staff

Andy Lawrence is an international leader in all of quasar science, survey astronomy, and internet software for astronomy. On quasar science, his most recent work is in extreme variability - changing look quasars and tidal disruption events - which will be of key interest for LSST. He has led or participated in many large scale surveys, and most notably is the PI of the UKIDSS survey. He created the Edinburgh Wide-Field Astronomy Unit, and was a founding member of the International Virtual Observatory Alliance.

Peter Clarke is a particle physicist with a long-standing role in computational support for the LHC programme, and, more recently, for a wider range of STFC projects as Scientific Director of IRIS. He is Deputy Project Leader of GridPP, providing the UK Grid for the LHC, was formerly Director of the National e-Science Centre, and was the first Chair of CAP. He has worked for several years with Edinburgh colleagues on planning for the UK DAC, most recently focussing on how UK DAC services can be implemented within the common infrastructure to be provided by IRIS.

Work Package 2.2: Data Ingestion and Publication

This WP focuses on the ingestion of data – LSST Level 1 and 2 data products provided by the Project, Level 3 data products derived from it and ancillary data needed for analysis of them – into the UK DAC. LSST Data Release Production will be shared equally between NCSA and CC-IN2P3. A copy of all LSST data will reside in Lyon, so we shall probably transfer Level 2 data from there, rather than NCSA. We already have operational experience of data transfer from IN2P3 to the UK as it is an LHC Tier1 centre. Connectivity from CC-IN2P3 to Edinburgh's Advanced Computing Facility over the production network (JANET, GEANT) appears to be more than adequate and we will involve JISC-JANET in operations planning from the start of Phase B. The LHC virtual data network (LHCONE) may also be used if needed (e.g. to receive the Level 1 alert stream from NCSA).

Images will dominate LSST data by volume, but loading Level 2 catalogues into a database is the challenging ingestion task. That data must be divided into shards by sky region, which are then distributed across the many nodes of a qserv cluster. An additional task for the UK DAC is inclusion (in the same spatial sharding, so on-sky proximity is reflected on disk) of the external catalogues required to support LSST:UK's multiwavelength analysis goals. Many of the largest of these (e.g. UKIDSS, VISTA, WISE, Gaia DR1, etc) are already present in WFAU archives, while others (e.g. PS1, Gaia DR2) will be ingested before the end of Phase A, and more can be added, as requested by LSST:UK researchers (e.g. in support of LASAIR).

WFAU has more than a decade's experience in receiving, preparing, ingesting and publishing data products generated both by remote survey pipelines (e.g. in Cambridge for UKIRT/WFCAM, VISTA and VST) and from survey teams (e.g. for VVV and the Gaia-ESO Spectroscopic survey) producing the analogues of Level 3 data products. The DAC team can, therefore, provide reliable solutions to the various networking, quality assurance and book-keeping challenges arising in this WP.

WP2.2 Resources

For this WP we request a total of 37.5 staff months of DI staff effort: 22.5 s.m. (constant 0.5 FTE) from Eckhard Sutorius and 15 s.m. (constant 0.33 FTE) from Stelios Voutsinas. Sutorius currently leads VISTA and VST data ingestion into WFAU's data archive, and Voutsinas is one of the lead developers of the Firethorn system for publishing and querying data using the IVOA's Table Access Protocol. *The descope in WP2.2 resources (from 42 to 37.5 sm) will result in a reduction in the number of ancillary data sets that will be available for use during Commissioning.*

Work Package 2.3: Alert Handling Infrastructure

This WP will provide the DAC infrastructure to underpin the development and operation of the UK's transient alert and variability brokers; it is, therefore, the technical complement to the more science-focussed activities of WP3.2. It covers the receipt of the Level 1 alert stream from NCSA to the UK DAC, the provision of the containerised platform within which the DEV teams can deploy the filters that they will design to select the events of interest for their science goals, and the onward publication of those filtered alert streams to the world-wide astronomical community in near real time. We have deployed a testbed for the LASAIR event broker system in Edinburgh, which features virtual machines receiving alerts from the Zwicky Transient Facility (ZTF) using the Apache Kafka stream processing system, adopted by the Level 1 development team in the LSST Project. At the moment this is testing our mirroring of the ZTF stream, and before the end of Phase A, we will replay the historical ZTF alert data through it at an increased rate to test its scalability and robustness.

A key WP2.3 responsibility is to build the extensive external catalogues that will provide the contextual information needed to classify LSST alerts through QUB's probabilistic event classifier (based on boosted decision trees, see WP 3.2 below), which will must execute the necessary proximity queries very quickly, to keep up with the real-time alert stream. That task may be suitable for a column-store database (e.g. MonetDB), and testing possible solutions will be a priority for the first half of Phase B, since LASAIR must be fully functioning as soon as LSST alerts appear: the scientific return from transients is a linear function of time, as missed events can never be recovered.

To succeed, LASAIR must enable the design of alert filters in a user-friendly way, and two PhD students that will aid the development, and exploit the early use, of the system, to trial how users will interact with it. Amanda Ibsen, an STFC Data-Intensive Science CDT student who started her PhD with Lawrence and Mann in April 2018, is studying the detection of Tidal Disruption Events and other extreme AGN variables, while a QUB-funded student will provide the science use examples for all other extra-galactic transients (supernovae, kilonovae, GRBs, orphan afterglows etc).

WP2.3 Resources

We request a total of 17 s.m. of DI staff effort: 12 s.m. from Roy Williams and 5 s.m. from Dave Morris. Williams is working with Ken Smith (QUB) to develop the LASAIR prototype running on the ZTF alert stream, while Morris is performing the tests with the Apache Kafka stream-processor. Morris is vice chair of the IVOA Time Domain Interest Group, and, with funding from ASTERICS (asterics2020.eu) – and, from February 2019, from ESCAPE, its H2020-funded successor - is driving the evolution of the relevant IVOA standards in the light of LSST Level 1 requirements.

Work Package 2.4: Provision of the DAC Platform

This WP covers the procurement, installation and operation of the DAC hardware, which we assume will take place as part of the IRIS system, with sysadmin effort contributed by projects like LSST:UK. It also covers support for the installation and maintenance of software for LSST:UK researchers working within the (probably OpenStack) cloud compute environment of IRIS. The DAC team has been experimenting with building Docker containers for LSST simulation and pipeline code, which seems promising from a user perspective, and the next step is to have containers mounting CernVMFS file systems when they start up, making available to them the more stable amongst the software components they require. This approach is attractive, but it will require experimentation with Commissioning data in Phase B to know whether it provides both the flexibility required by users to secure the compute resources they require with the correct software already installed and the ease-of-administration required for a production system with limited sysadmin resources.

Another important Phase B activity will be to ensure that DAC resource provision scales as LSST Data Releases increase in size. In particular, the qserv database system is a conservative design, providing a low-risk solution for the LSST DR1 database, but the explosion of interest in “big data” in the commercial sector is driving rapid development of “NoSQL” data management technologies. Most of these have limited support for functionality equivalent to joins and indexes in relational databases, which may limit their utility in survey astronomy, but it will be important to monitor their development to ensure that open source alternatives to qserv are not ignored.

WP2.4 Resources

We request 22.5 s.m. of effort from Teng Li (at 0.5 FTE throughout Phase B). Li is currently employed half-time on GridPP and half-time on Phase A DAC work. He is setting up and maintaining the DAC testbed during the final year of LUSC Phase A.

Work Package 2.5: Science Support

Science support is a traditional strength of WFAU, which has been helping users maximise the science they get from its sky survey archives for more than 20 years. This experience will be invaluable in guiding users to make effective use of LSST data products, but it will have to be supplemented by expertise to support LSST:UK researchers in using the UKT0 computing infrastructure, which will be new to most of them. Different kinds of support will be needed by individual researchers with an immediate problem that can be solved through an email helpdesk and by larger science teams who need help planning a large-scale analysis.

The DAC team in Edinburgh have been trialling the latter kind of support through ~6-month mini-projects to get significant Phase A DEV workflows running on GridPP. This format, pairing a DAC team member with the necessary computational expertise with a science team with an immediate analysis need, works well, so we propose to continue that model into Phase B, but, as competition for expert guidance increases, it may be necessary to formalise the deployment of that effort; e.g. by having an “effort allocation committee” to assess brief proposals for supported mini-projects.

In addition to UK DAC helpdesk, further science support can be provided by documentation provided by others within the LSST community. LSSTC is currently considering its potential role in providing documentation and training material for the LSST community, and Mann will use his place on the Corporate Operations Committee to further that, perhaps through LSSTC providing funding for people to develop materials that could be shared within the international LSST community.

WP2.5 Resources

For this major activity we request a total of 47 s.m. of DI effort: 22.5 s.m. from Mike Read, 11.25 s.m. from Dave Morris, 10.25 s.m. from Stelios Voutsinas and, in Year 4, 3 s.m. from Nigel Hambly. Read leads science support for WFAU’s existing sky survey archives while Morris and Voutsinas pioneered (Morris et al. 2017) the use of containers for astronomical software deployment, and are currently working with Hambly on STFC- and EU-funded projects to aid user access to Gaia catalogue data. *The resources for this sub-WP have been reduced from 60 to 47 s.m., which will reduce the amount of support that LSST:UK researchers can receive during Commissioning.*

WP2 Summary Gantt Chart

A Gantt chart for Work Package 2 is presented below. The DAC activities underpin much of the work that is undertaken within Work Package 3 and, as such, there are a number of critical paths that need to be monitored carefully. Specifically, the work to define, deploy and operate infrastructure for the LASAIR service is time-critical and must respond to the timeline of the LSST project to ensure that LSST:UK presents a credible case for running an LSST Community Broker on behalf of the international astronomy community. Similarly, Work Package activities have been carefully aligned to the LSST Project schedule, to ensure that DAC services are ready to receive the data from Commissioning and Operations.

WP	Task Name	Start	Finish	Duration
2.1	**DAC Management**	**01/07/2019**	**31/03/2023**	**196w**
2.1.1	Preparation and submission of periodic (IRIS) resource requests	01/07/2019	31/03/2023	196w
2.1.2	Preparation and submission of PI and JA lists to LSST Corporation	01/07/2019	31/03/2023	196w
2.1.3	Operational planning for DEV activities	01/07/2019	30/06/2022	156.8w
2.1.4	Operational planning for Commissioning activities	01/11/2019	31/03/2020	21.6w
2.1.5	Operational planning for activities during full-operation of telescope	01/04/2020	31/01/2023	148w
2.2	**Data Ingestion and Publication**	**01/07/2019**	**31/03/2023**	**196w**
2.2.1	Ingest experiments for DESC DC2	01/07/2019	31/12/2019	26.4w
	Finalise list of ancillary surveys to ingest alongside Commissioning data	31/12/2019	31/12/2019	0w
2.2.2	Setup and ingest of Commissioning data	01/01/2020	31/12/2021	104.6w
2.2.3	Preparation for ingest of LSST Data Release 1	01/05/2020	30/11/2022	134.8w
2.2.4	Ingest of ancillary collections for LSST Data Release 1	01/05/2020	30/11/2022	134.8w
	Finalise list of ancillary surveys to ingest alongside LSST Data Release 1	01/12/2022	01/12/2022	0w
2.2.5	Design ingest process for User-generated Products during Phase C	01/12/2022	31/03/2023	17.4w
2.3	**Alert Handling Infrastructure**	**01/07/2019**	**31/03/2023**	**196w**
2.3.1	Set up of ancillary catalogues for cross-matching with ZTF alert stream	01/07/2019	30/10/2020	70w
2.3.2	Set up of ancillary catalogues for cross-matching with LSST alert stream	02/11/2020	28/02/2022	69.2w
2.3.3	Design and implement services for LASAIR	01/07/2019	31/08/2022	165.6w
	Infrastructure for LASAIR set up and operation	01/09/2022	01/09/2022	0w
2.3.4	Operation of services for LASAIR	01/09/2022	31/03/2023	30.4w
2.4	**Provision of the DAC Platform**	**01/07/2019**	**31/03/2023**	**196w**
2.4.1	Define scope of platform for User-generated Data Products	01/07/2019	31/12/2021	131w
	Resources secured on IRIS (or other infrastructure) for DAC during Commissioning and Early Operations	01/04/2020	01/04/2020	0w
2.4.2	Set up and operation of proto-DAC Platform for Commissioning	01/04/2020	31/12/2021	91.6w
2.4.3	Set up of DAC infrastructure for LSST operations	03/01/2022	31/08/2022	34.6w
	Infrastructure for DAC set up and operation	01/09/2022	01/09/2022	0w
2.4.4	Operation of DAC infrastructure for LSST operations (inc. LASAIR and ingestion of Data Release Products)	01/09/2022	31/03/2023	30.4w
2.4.5	Scaling experiments and validation for Date Release sequence	01/07/2019	31/03/2023	196w
2.5	**Science Support**	**01/07/2019**	**31/03/2023**	**196w**
2.5.1	Set up of LSST:UK Helpdesk	01/07/2019	31/03/2020	39.4w
	LSST:UK Helpdesk operational	01/04/2020	01/04/2020	0w
2.5.2	Operation of LSST:UK Helpdesk	01/04/2020	31/03/2023	156.6w
2.5.3	User documentation	01/07/2019	31/08/2022	165.6w
2.5.4	Training of data-rights holders during Commissioning and LSST operations	01/01/2020	31/03/2023	169.6w
2.5.5	Computational support for DEV activities (mini-projects)	01/07/2019	31/03/2023	196w

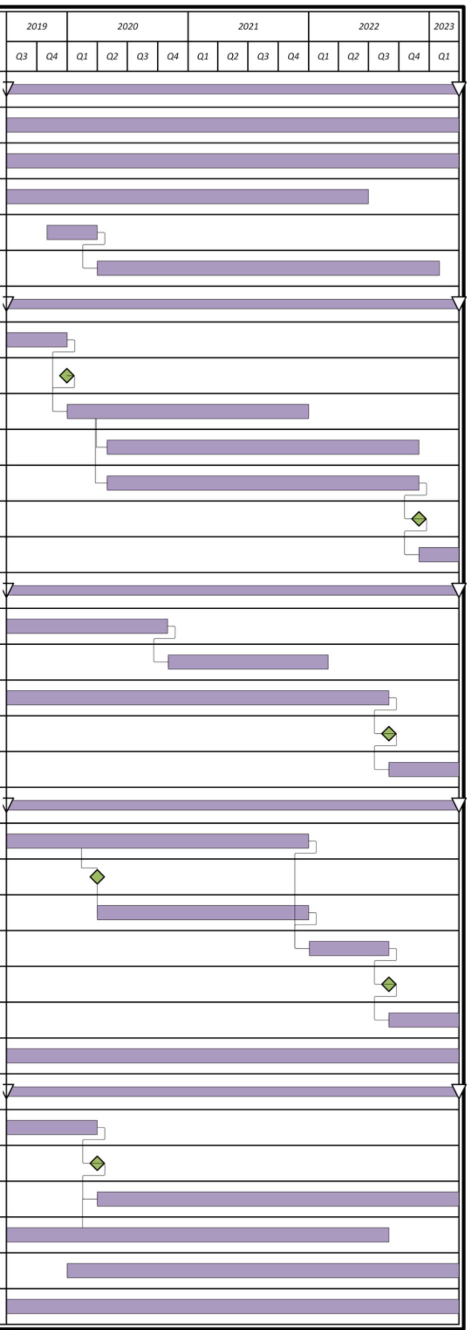

Figure 8.1 Summary Gantt chart for WP2

9. WP3: LUSC-DEV

The objective of the Phase B LUSC-DEV programme is to provide the scientific software and ancillary data products required by the UK community in preparation for the start of LSST operations. Through this coordinated programme we can deliver capabilities required by different subsets amongst the ~500 UK scientists who will have LSST data rights in a much more cost-effective manner than having them develop separately software to run in their own institutions. The scale of the LSST data challenge means that this work must happen now, while the lack of immediate scientific return precludes its being funded through the astronomy grants line.

As described in Section 6.1, 11 DEV Work Packages described were selected as a prioritised list from the 24 submitted proposals to create a coherent programme for Phase B that will build upon the LSST data access delivered by the DAC to incorporate the other UK data holdings (e.g. ESO survey data, near infra-red surveys, high energy and radio catalogues, and other time domain surveys) and provide the software needed to meet our LSST science goals. *Following the descope, we request funding for seven DEV Work Packages, which we describe in turn below, retaining their original numbering for ease of comparison with the original case.*

9.1 WP3.2 LASAIR - the UK transient broker for LSST

WP3.2 goals and deliverables: In the UK DAC, we will build, maintain and run LASAIR, providing a critical and essential service specifically focused on the needs of the UK community. The LSST Project will provide a stream of 10^7 transient or variable sources (total volume 400-600 GB per night).This stream would overwhelm individual scientists, and LASAIR will manage the transient data stream and provide users with a database, web pages, visualisation tools, classification and search queries that will enable LSST transient science to be done by UK and international scientists. We expect LASAIR to be one of the official LSST brokers, receiving the full transient stream.

This is a close collaboration between the UK DAC and QUB. For every alert, we will provide an answer to the question *"what has been detected (resolved in time) at this position in the sky, at every wavelength from x-ray to radio?".* This means in real time we will provide the following:

- **Light-curves**: assimilate all *diaSource* alerts in *diaObjects*: providing interactive webpages (linked to database), plots, ability to select ranges, submit user added points. Previous history from Pan-STARRS, DES, Skymapper, ATLAS, CRTS, PTF/ZTF
- **Postage stamps**: all LSST detections and most recent non-detections. Plus multi-colour images from LSST, near infra-red (VISTA/UKIDSS), H-alpha (VPHAS) and EUCLID, or HST/JWST if space based imaging is available.
- **Massive catalogue cross-match:** with star, galaxy, AGN, x-ray, radio catalogues and provide classification through boosted decision trees (and machine learning) through our already working code "*Sherlock*" (Young et al. 2018).
- **In real-time, cross-match to all other wavelength time-domain surveys:** gamma-ray, x-ray and radio (e.g. MEERKat/Thunderkat through 4pisky.org , Swift, SVOM, eRosita)
- **Cross match to all *previously* known transients:** supernovae, transients, gamma ray-bursts, x-ray and radio burst sources (e.g. searching for currently unknown physical links over time)
- **Spectroscopic and/or photometric redshift:** locate catalogued redshifts of the likely host galaxy and hence absolute mag (we will link to WPs 3.3 for redshifts)
- **Combine all of the above information:** including the first 24hr-48hr lightcurve trend (e.g. rapid rise/decline) to probabilistically classify all transients as : supernova – kilonova – GRB – Tidal Disruption Event – AGN – XRB – CV – eruption star – microlens – orphan
- **Multi-messenger cross-matching:** *GW coincidence tag* based on their 4 dimensional position in space and time compared to LIGO-Virgo gravitational wave events (sky position, distance, and time). All transients will also be *Neutrino coincidence tagged* based on their 3D space time location (sky position and time) with IceCube high energy neutrinos.
- **Provide a stream of transients to 4MOST and SOXS spectroscopic programmes** and ingest the classifications and data from those facilities in return (linked to WP 3.3)
- **Machine learning algorithms for real-bogus classification:** as a final check on real-bogus objects, we will run our own trained ML code to weed out spurious objects (Wright et al. 2016, 2017, Smartt et al. 2016).

LASAIR is a massive database project, assimilating all transient sources together with the all-sky catalogues with Edinburgh (DAC) and Cambridge (WP3.5) and providing users with easy access through user selected web-pages. Users will be able to login to the database and run either *SQL* queries or upload code (through *Jupyter* Notebooks) to run on the whole database. The QUB team now have more than 8 years' experience running this with the Pan-STARRS and ATLAS surveys in real time in Belfast. The combination of this and the UK DAC expertise in Edinburgh is essential to build a robust and larger scale version for LSST. We will be focused on the UK science requirements, but LASAIR will be open to anyone from the LSST community (which is a requirement for official LSST broker status). It will provide the platform for the light-curve fitting and spectroscopic classification work-package (WP 3.3), it will link with 3.7 in further developments in machine learning, and its catalogue cross-matching will use the state-of-the-art algorithms from WP 3.11.

WP3.2 Summary Gantt Chart:

WP	Task Name	Start	Finish
3.2.1	Development and testing of classification algorithms (SHERLOCK)	01/07/2019	31/03/2020
	LASAIR Version 2.0 launched (inc. SHERLOCK Version 2.0 classifier)	*01/04/2020*	*01/04/2020*
3.2.2	Build links to other mult-wavelength, multi-messenger time-domain surveys	01/07/2019	30/06/2021
3.2.3.1	Develop LASAIR web interface and user tools	*01/07/2019*	*30/06/2021*
	Lasair Version 3.0 launched (tech. preview of web interface and user tools)	*30/03/2021*	*30/03/2021*
3.2.3.2	Develop LASAIR web interface and user tools, Pt. 2	01/07/2021	31/03/2023
3.2.4	Integrate spectral information, follow-up, and data access within LASAIR	01/07/2021	30/12/2022
	LASAIR Version 4.0 launched (final web interface and user documentation)	*30/03/2022*	*30/03/2022*
3.2.5	Testing LASAIR during Commissioning	01/04/2020	31/03/2023
	LASAIR Version 4.0 launched (w/ interface for non-proprietary data)	*03/04/2023*	*03/04/2023*

Figure 9.1 Summary Gantt chart for WP3.2

WP3.2 Resources: K. Smith at 1FTE requested. He is currently employed on Phase A and an expert in designing and running massive databases and developing scientific code for exploitation. He has a long track record of database design and management in industry and 9 years of experience in astronomy data management. Retention of Smith's expertise is critical. His experience in Pan-STARRS and ATLAS data processing streams, and constructing billion row database systems make him critical to the success of this project. Smartt (0.1 FTE) is the science lead and local manager and heavily involved in testing, science verification, and managing this within the broader LSST project. QUB will provide a leverage studentship to work on Data Intensive Science to support science development.

9.2 WP3.3 Transient Classification and spectroscopic follow-up

WP3.3 goals and deliverables: We will build the infrastructure for rapid transient follow-up and classification, including interaction with the 4MOST spectroscopic engine, as well as other spectroscopic programs. We will develop the steps required to move from initial photometric prioritisation in LASAIR (WP3.2), to spectral observations of 10^5 transients and their subsequent classification release. This will release spectral classifications back to the LSST:UK community via LASAIR, which we will test during commissioning, allowing the community to do high dimensional searches in photometric and spectroscopic phase space on a massive scale. Beyond Phase B, this will enable us to combine data from the 4MOST multi-object spectroscopic instrument and LSST to produce spectra of ~10^5 transients (or their host galaxies) during LSST operations.

WP3.3-1: Ensure 4MOST/TiDES survey strategy is optimized for LSST transients *(3 s.m.)*. We will ensure the TiDES survey plan is optimised for revisions to the LSST observing strategies. This will optimise TiDES for the evolving LSST field choice, cadence, filters and depth based on the latest LSST OpSim (we are expert in running LSST OpSim, from Phase A). We will superimpose the 4MOST survey on our LSST simulations, and optimise the 4MOST field sequence, observing pattern, and fibres allocated per field. This must be done in late 2019 to allow 4MOST to finalise its overall survey strategy. Although TiDES targets have priority in 4MOST, TiDES must be able to respond to 4MOST requests in order to efficiently observe the LSST survey fields.
WP3.3-2: Develop and apply algorithms to select targets for 4MOST *(9 s.m.).* This task develops the TiDES targeting algorithms from target selection to 4MOST Observing Block (OB) generation. We will use the new public PLAsTiCC simulations of LSST data to show the likely balance of different transient types in the LSST data down to i=22.5 (the 4MOST limit for transient classification in one hour). Our algorithm will prioritise transients according to science case: high priority 'rare' transients (~1-2 per pointing), and lower-priority 'randomly-selected' transients for photometric classification training. In collaboration with LASAIR, we will combine light-curve information with contextual data (for example, host galaxy colours and morphology; transient relative position; existing

photometric/spectroscopic redshifts) for inclusion in our prioritisation. This will allow better prioritisation of exotic physical events such as tidal disruption events from supermassive black holes and superluminous supernovae. We will develop, in collaboration with 4MOST, the procedure to ingest our prioritised LSST targets into the 4MOST OBs for observation. We have an agreement to update OBs up to three days prior to observing with 4MOST, and anticipate reducing this further. Reducing this lag will depend on demonstrating an ability to prioritise candidates quickly and produce error-free catalogues.

***WP3.3-3: Automate classification of resulting spectra, returning classifications to LSST:UK** (6 s.m.).* 4MOST quicklook reductions will be available to TiDES within 3-4 hours via the quicklook reduction server at Paranal. Based on these reductions, this task will ensure automated classification of the transient spectra can be performed either via ML cross-correlation techniques (Muthukrishna et al., in prep) or template fitting. We must determine the level of required human oversight to produce error-free classifications by testing on both simulated 4MOST observations of the different transient types in PLAsTiCC, and using our privileged access to transient spectra from OzDES in the Dark Energy Survey. We will develop software to 'handshake' with LASAIR to return transient classifications to the community.

WP3.3-4: Test and exercise system during LSST commissioning (6 s.m.). By late 2020, LSST should be in commissioning mode for at least part of the camera and running 'mini-surveys'. At least one of these will include cadenced observations of extragalactic fields. We will test our spectroscopic selection algorithms using these data and the AAT/2dF spectrograph (similar principle to 4MOST, but with a smaller field-of-view, fewer lower-throughput fibres, and at a poorer astronomical site). This will ensure that our fibre allocation algorithms are optimal and our classifying routines are working. Again, we will return these classifications to the UK community via LASAIR.

WP3.3 Summary Gantt Chart:

WP	Task Name	Start	Finish	Duration	2019 Q3	2019 Q4	2020 Q1	2020 Q2	2020 Q3	2020 Q4	2021 Q1	2021 Q2	2021 Q3
3.3.1	Ensure 4MOST/ TiDES survey strategy is optimised for LSST transients	01/10/2019	31/12/2019	13.2w									
3.3.2	Develop and apply algorithms to select targets for 4MOST	01/01/2020	30/09/2020	39.2w									
3.3.3	Automate classification of resulting spectra, returning classifications to LSST:UK	01/10/2020	31/03/2021	26w									
3.3.4	Test and exercise system during LSST commissioning	01/04/2021	30/09/2021	26.2w									
	Software for automatic target selection algorithm and classification pipeline for 4MOST is deployed	01/10/2021	01/10/2021	0w									

Figure 9.2 Summary Gantt chart for WP3.

WP3.3 Resources: Our infrastructure must be developed and tested before and during commissioning to ensure we can start immediately when LSST science begins (2022, Fig 9.2). The workload is ambitious and not possible without developer support. WP3.3 requires 2 years (Southampton + 0.1 FTE management for Sullivan for 2 years), starting on October 1st 2019, to allow both input to the TiDES survey planning and overlap with the LSST commissioning phase. *We have descoped this work package to 24 staff months (from 48) to focus on the essential tasks concerned with building the core LSST + 4MOST joint infrastructure. We cannot reduce the effort further and remain viable. Our proposed start date follows the Southampton Phase-A matched-funded work on the LSST observing strategy has completed. This continuity is important to ensure the expertise in LSST simulations and software tools remains in the LSST:UK community. No funding will mean that 4MOST rapid classifications and spectral information will not be returned to the UK community.*

WP3.3 Key Staff

Mark Sullivan has 15 years' experience of leadership in transient sky surveys, including the Palomar Transient Factory, Supernova Legacy Survey, and the Dark Energy Survey. This has included running major spectroscopic follow-up large programmes at the ESO Very Large Telescope, and the Chair of the Science Board of PESSTO. He is a full member of the LSST DESC, PI of the TiDES survey within 4MOST, and co-chair of the DESC Follow-up Task Force.

9.3 WP3.5: LSST and near-infrared data fusion: leveraging the UK's IR data resources

WP3.5 goals and deliverables: The scientific return of LSST will be greatly enhanced by the addition of information at other wavelengths to the LSST optical data. In particular, infrared (IR) photometry for LSST stars, galaxies and quasars will open up the discovery space sampled by LSST to the very distant, obscured and cool Universe. We will build on the UK's leading role in IR surveys with VISTA, to implement a flexible pipeline for ingesting IR data in the UK DAC, and processing the IR pixels jointly with LSST pixels to produce science-ready multi-wavelength catalogues for the community. A critical aspect of combining the LSST data with surveys at other wavelengths is ensuring that the photometry is measured consistently across all wavelengths, taking into account variations in angular resolution, depth and signal-to-noise between the different surveys. Whilst catalogue-level matching between surveys (from WP 3.11) will be crucial to identify matches between wavebands, catalogue-level photometry will not be sufficient for many science applications and joint image-level processing of LSST and IR surveys will enable us to probe considerably deeper than the IR catalogues (e.g. Banerji et al. 2015). We will build a modular, dual-photometry pipeline, based on the LSST stack software, to process pixels from IR surveys jointly with LSST pixels making use of photometric positional and shape priors from one survey to measure fluxes in another. The pipeline will be configurable to new datasets such as *Euclid* as they become available. We will deliver combined multi-wavelength optical+IR catalogues (including sources not detected in the optical e.g. those at high-redshifts) together with source-level metadata, detection and measurement image provenance information and workflow provenance information. All images, weight maps and catalogues will be delivered to the UK DAC and accessible via an SQL-queryable database, and the multi-wavelength information will also be used in transient identification and classification (WP3.2 and 3.3).

Data Management & DAC Delivery: We will ingest and curate the relevant datasets within the UK DAC **(WP 3.5.1)**. Both pixel and catalogue-level data from optical and IR surveys need to be ingested including re-formatting of meta-data to comply with LSST stack processing. In the IR, we will use VISTA as the primary dataset as well as potentially HST data as a test-bed for ingesting upcoming datasets e.g. *Euclid*. Currently available optical pixels from surveys such as DES and HSC will also be ingested together with their associated catalogues, and will serve as training sets for LSST. Finally the outputs from our dual photometry pipeline **(WP3.5.2)** will also be ingested into a queryable database and made easily accessible to the community **(WP 3.5.6)**. This requires 0.33 FTE of effort until month 33 in order to allow new releases of the relevant survey datasets to continue to be ingested after the start of the work package.
Dual Photometry Pipeline: The pipeline will be based on the forced-photometry module within the LSST stack. This module will be configured to perform photometry on VISTA IR images based on LSST detections and vice versa **(WP 3.5.2)**. The combination with the IR is not a Level 2 deliverable and therefore requires L3 DEV effort. Existing photometry pipelines (e.g. SExtractor[9], imcore(list)[10]) used by current surveys such as DES and VISTA, will also be implemented in order to allow benchmarking **(WP 3.5.3)**. The LSST stack is still in developmental stages hence benchmarking is an important step. Direct comparisons between outputs produced using these different approaches will be made. The workflow will also be ported onto generic High Performance Computing (HPC) systems e.g. DiRAC/IRIS **(WP 3.5.4).** Pipeline commissioning and testing will overlap with the commissioning and Science Verification phases of the LSST survey, allowing LSST pixels to be directly processed and combined with the VISTA data during this phase. The key timelines are (i) Commissioning using ComCam (Jan-Sept 2020) (ii) Commissioning using LSSTCam (Oct 2020-Mar 2021) and (iii) Full Science Verification (Apr-Jun 2021). We will ensure fields with good IR coverage are observed in LSST Commissioning and Science Verification. This will be facilitated by the LSST:UK Commissioning Coordinator (G. Smith).
Scientific Validation (WP 3.5.5): Initially the data products will be multi-wavelength catalogues produced by combining current optical photometric surveys such as DES and HSC with VISTA, while LSST pixels will be combined with IR pixels towards the end of Phase B. The key scientific validation steps will be: (i) Assessment of noise properties, depths, seeing FWHM, stellar and galaxy colours

[9] https://www.astromatic.net/software/sextractor

[10] http://casu.ast.cam.ac.uk/surveys-projects/software-release/imcore

produced using different approaches to photometry; (ii) Assessment of the utility of IR data in improving star-galaxy separation, by exploiting the new colour information from the IR; (iii) Other key multi-wavelength datasets: The UK also has strong interests in other multi-wavelength datasets, which can be combined with LSST during science validation. We will provide and validate X-ray (e.g. XMM-Newton, Chandra, NuSTAR, e-ROSITA), Mid-IR (e.g. Spitzer/WISE) Far-IR (e.g. Herschel) and radio (e.g. MeerKAT, ASKAP) data to be accessible to the UK community via the DAC.

WP3.5 Summary Gantt Chart:

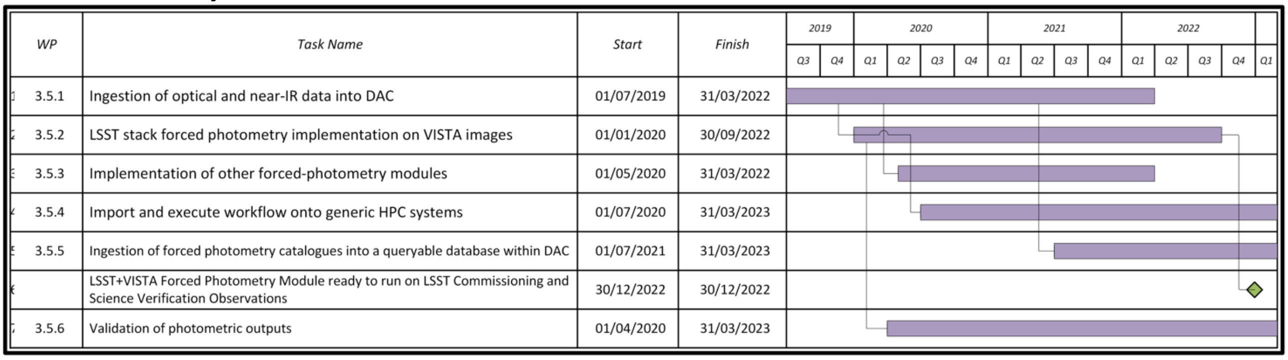

Figure 9.3 Summary Gantt chart for WP3.5

WP3.5 Resources: 1 new PDRA (Cambridge) to develop the dual photometry pipeline, managed by Banerji (0.1 FTE, but no funding requested). The PDRA will devote 0.1 FTE to aiding science validation efforts (WP3.5.5). Additional resources for science validation will come from AGP-funded postdocs. We also request 0.5 FTE (Carlos Gonzalez-Fernandez) for the data management and DAC interfacing effort. He works within the Cambridge Astronomical Survey Unit (CASU), and is responsible for the processing, curation and pipeline development for all VISTA IR surveys. He will reformat VISTA data and metadata to comply with LSST stack re-processing and facilitate ingestion of all relevant data from this work package into the UK DAC. The data management effort will be overseen by McMahon (0.05 FTE) and science validation will be overseen by Jarvis (0.05 FTE).

WP3.5 Key Staff

Manda Banerji has significant experience in wide-field optical and infrared imaging surveys over the last decade. She holds Builder status within the Dark Energy Survey (DES) Collaboration and has led the DES Galaxies and Quasars Science Working Group since 2014. She co-chairs the LSST Galaxies Science Collaboration and is PI of the ESO VISTA VEILS Public Survey, which is completing near infra-red coverage of 3 of the 4 announced LSST deep-drilling fields. Within LSST:UK she is the Infrared Liaison and a member of the Science Working Group.

Richard McMahon has over 25 years of experience in data management of wide field astronomical surveys primarily in the optical and near infra-red. He is PI of the 18,000 deg^2 VISTA Hemisphere Survey, the largest ESO public survey with the VISTA telescope. He is also a member of the Dark Energy Survey (DES) collaboration with 'builder' status.

Matt Jarvis has been the PI of the VISTA Deep Extragalactic Observations (VIDEO) survey for over 10 years, which is one of the key near-infrared data sets to be combined with the LSST data. He is a leader in multi-wavelength extragalactic astronomy, with management experience in surveys spanning optical, near, mid- and far-infrared wavelengths, as well as also being the PI of the new MeerKAT-MIGHTEE radio survey, also being carried out over the LSST deep drilling fields. His research spans, galaxy formation and evolution, the impact of AGN feedback on this evolution, and how we can better use galaxy populations to understand the cosmological model. He also helped develop one of the leader photometric redshift codes in the recent LSST Data Challenge.

9.4 WP3.7: Low-surface-brightness science using LSST

WP3.7 goals and deliverables: With its unique combination of depth and area, statistical low-surface-brightness (LSB) astronomy is one of LSST's niches. LSST will revolutionise galaxy evolution studies, by revealing structures that are invisible in past wide-area surveys, e.g. LSB tidal features and intra-cluster light (ICL). ICL dominates the baryonic content of clusters, which are unique probes of our cosmological model. LSB tidal features encode galaxy assembly histories, making them key tracers of hierarchical structure formation. While LSB science is fundamental to the activities of several science collaborations (e.g. Galaxies, AGN, DESC), the faint/diffuse nature of LSB structures makes them susceptible to two, serious data-processing issues: sky over-subtraction and shredding of galaxies and their tidal features by de-blenders. Since the default sky-subtraction and de-blending pipelines from the Project are optimised for accurate photometry in deep/crowded fields (i.e. smaller spatial scales than LSB structures), preparatory work is essential for any LSB science to be possible using LSST. WP 3.7 will provide the necessary mitigation of these issues by delivering (1) optimised sky-subtraction to preserve LSB structures on any spatial scale and (2) algorithms to mitigate shredding by de-blenders. Together, these are essential for enabling all LSST LSB science and are Roadmap priorities (LSST Galaxies Roadmap; Robertson et al. 2017). They are also unique UK contributions, with no LSST Project work in this area.

Task 1 (leading to Activities 3.7.1 – 3.7.5): LSB-optimised sky subtraction

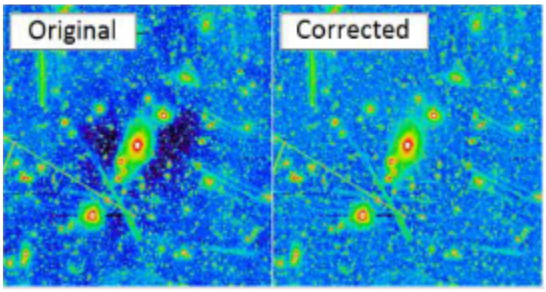

Figure 9.4 Left: XCS 35 in HSC DR1 i-band, reduced through the current LSST pipeline. Sky over-subtraction is visible around bright and extended sources. Right: Test of a non-parametric mixed model flux threshold correction (Kelvin+ in prep), which successfully restores over-subtracted regions.

We will construct new sky-estimation algorithms, tailored for LSST LSB science on a wide range of spatial scales. Fig 9.4 (left panel) shows how the current LSST DM pipeline significantly over-subtracts flux around bright and/or extended sources in HSC imaging, removing LSB signatures. Furthermore, contamination of LSB signals can occur due to imaging artefacts like ghosting and CCD edge effects. We will build on non-parametric mixed-model flux threshold approaches we have developed that show promise in correcting over-subtracted regions (Fig 9.4, right panel). A variety of alternative approaches to traditional sky-estimation techniques will also be considered. E.g. the Fornax Deep Survey (Iodice + 16, ApJ, 820, 42) utilize techniques to combine pseudo-random pointings of multi-wavelength imaging data, an ideal case-study for LSST. Akhlaghi +15 (ApJS, 220, 1) present a new, noise-based non-parametric technique for detecting nebulous objects, which proves extremely successful in the LSB regime. Additional approaches will analyse imaging in Fourier space (which allows low-frequency components, e.g. sky, to be separated from high-frequency counterparts e.g. stars/galaxies), utilizing the wealth of possibilities afforded by the multi-wavelength LSST dataset. Algorithms will be benchmarked using mock images from the Horizon-AGN and Cluster-EAGLE cosmological hydro-dynamical simulations (for the tidal-feature and ICL-related algorithms respectively), inserted into HSC/LSST ComCam frames to ensure realistic representations of noise/background sources/camera effects in the data.

Task 2 (leading to Activities 3.7.6 – 3.7.7): Mitigating shredding of galaxies and tidal features

Any deblender, while typically preserving total flux, will shred galaxies and LSB tidal features into different objects. This task will (1) optimize de-blenders to minimize shredding and (2) re-associate shredded galaxies with their tidal features using a machine-learning morphological analysis code.

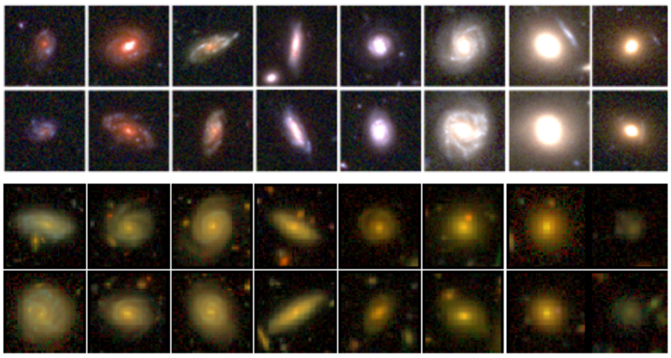

Figure 9.5 Top two rows: H18 implemented on HST data produces clean separation of objects that are composed of pixels with different properties (e.g. colour and texture).Bottom two rows: An implementation on ground-based images, from HSC, that are close to LSST resolution and 10-year depth (Martin, Kaviraj, Geach in prep.)

For this, we will use mock images of merger remnants from Horizon-AGN and the sky-subtraction machinery, tailored for tidal features, developed in Task 1. We will first tune aspects of the Project de-blender (e.g. applying constraints on symmetry/monotonicity or using complex morphologies such as example-based matches or Gaussian-Process approaches) to minimize shredding. Other available de-blenders will also be considered.

We will then use an unsupervised machine-learning morphological-analysis algorithm (Hocking et al. 2018; H18 hereafter), that efficiently groups together similar pixels (and objects built from those pixels, like galaxies) in survey images. The unsupervised nature of the algorithm is particularly well-suited to the unprecedented data volume and short cadence of LSST, since it is unfeasible to repeatedly produce large training sets (as would be required for supervised machine-learning techniques) on relatively short timescales.

Briefly, the algorithm samples pixel patches in real or simulated survey images and maps these patches onto a 'feature' vector, which holds statistical information (e.g. colour/texture) about the patch. Patches are clustered using growing neural gas and hierarchical clustering and connected component analysis is used to group similar patches into galaxy 'detections'. Objects with similar morphologies can then be easily collected into classes (Fig 9.5). Originally developed using HST-CANDELS data (Fig 9.5, top two rows), we have adapted the algorithm to work with images (real or simulated) that have ground-based resolution and depth similar to LSST (see bottom two rows of Fig 9.6 for such an implementation on HSC Ultra-deep; Martin, Kaviraj, Geach in prep). We will use this code to create algorithms for re-associating galaxies and tidal features that have been shredded by deblenders. Since tidal features form from galaxy material, they inherit features like colour and will, therefore, have similar feature vectors. Galaxies can be reconstructed by searching for and collating shredded objects in proximity to them that have similar feature vectors.

WP3.7 Summary Gantt Chart:

WP	Task Name	Start	Finish
3.7.1	Explore methodologies for accurate sky subtraction	01/07/2019	31/03/2020
3.7.2	Testing on simulated images	01/10/2019	30/06/2020
	Low-surface-bright Sky Subtraction Software source decision	*01/07/2020*	*01/07/2020*
3.7.3	Finalize optimization for intra-cluster light	01/07/2020	26/02/2021
3.7.4	Finalize optimization for tidal features	01/07/2020	26/02/2021
3.7.5	Write sky-subtraction package for LSST software stack	01/03/2021	28/06/2021
3.7.6	Optimize deblenders to minimize shredding of tidal features	01/07/2021	30/09/2021
3.7.7	Write software for re-associating galaxies and shredded components	01/10/2021	31/12/2021

Figure 9.6 Summary Gantt chart for WP3.7

WP3.7 Resources: We request 2.5 staff years for Lee Kelvin, who is an expert in sky-subtraction techniques. His time will be split across Herts/LJMU. He will spend 12 months at LJMU followed by 18 months at Herts, so as to interact with staff with specific expertise in the task components and deliver the activities in the sequence proposed in the Project Management Plan. We request 0.1 FTE for Collins (for 12 months), to assist with the analysis of the Cluster-EAGLE simulations and 0.1 FTE for Kaviraj (for 18 months), who will help implement the Horizon-AGN simulation and, together with Geach (0 FTE), will assist with implementing the machine-learning algorithms in Task 2. Recruitment of Kelvin is critical for this WP: he is an expert in source detection, modelling and sky-subtraction techniques, tailored towards low-surface-brightness flux preservation, and already has experience of creating such algorithms for us with a multitude of wide-area multi-wavelength imaging campaigns, such as SDSS, VST KiDS, VISTA VIKING and HSC.

WP3.7 Key Staff:

Sugata Kaviraj has been at the forefront of observational and theoretical work in the low-surface-brightness regime, performing pilot observational studies using small, deep surveys like the Stripe 82 and making testable predictions for the low-surface-brightness regime using cosmological hydro-dynamical simulations. He is recognised as a leader in this field, with several major papers on how low-surface-brightness structures like tidal features can be used to reconstruct galaxy assembly histories and constrain the standard hierarchical paradigm. He is co-Chair of the LSST Galaxies Science Collaboration and served as Chair of the Galaxies Low-Surface-Brightness Working Group.
Chris Collins is a senior researcher in the field of observational galaxy evolution and large-scale structure. He has made major contributions to the understanding of the growth of Brightest Cluster

Galaxies and pioneered measurements of low-surface brightness intra-cluster light emission from clusters approaching z=1. This work demonstrates the importance of the contribution of the ICL to understanding the stellar mass growth at the centres of galaxy clusters.

9.5 WP3.9: LSST Point Spread Function, sensor characterisation and modelling

WP3.9 goals and deliverables: (1) To carry out detailed laboratory characterisation of the LSST sensors, enabling sensors to be quality controlled, calibrated and modelled; (2) To develop methods for measuring and modelling the overall atmosphere+telescope+camera point spread function (PSF), including the lab-measured effects, building on existing techniques for inferring the PSF and its variation across the LSST field from science observations. Work will be carried out in collaboration with the LSST Camera project, the Dark Energy Science Collaboration (DESC) including the Sensor Anomalies Working Group (SAWG) of DESC.

This WP is the only one proposed that works directly on LSST hardware. The lab measurement is recognised as a critical contribution to DESC science roadmap deliverables. The outcomes are essential for the wide range of LSST science and will be directly used by the Dark Energy and Transients collaborations for weak lensing and photometry measurements. A precise determination of positions, fluxes and shapes of celestial sources through an understanding of the PSF is crucial for accurate galaxy shape measurement, used to infer the weak lensing distortion, and for accurate galaxy photometry, needed both for photometric redshift estimation (vital for dark energy science) and for transients detection. Each component of the PSF (atmosphere, optics and sensors) needs to be extremely well understood, modelled and controlled. Modelling of the stochastic atmospheric fluctuations is challenging given the limited number of unresolved stars available to sample the PSF. Variable telescope wavefront perturbations need careful modelling. Dynamic charge collection effects in the sensors introduce correlated noise and asymmetric distortions, which are dependent on the brightness of sources. The LSST commissioning phase is an important opportunity to test our models. ComCam data combines all components of the PSF, while lab testing isolates the sensor PSF.

In addition, all LSST science goals benefit from optimisation of the sensor timing and bias parameters, as they affect the noise, linearity, full well capacity and charge transfer efficiency. Optimising the LSST focal plane (189 sensors, from two different vendors: e2v (UK) and ITL (US), each with differing optimal timing and bias parameters) is not a fully understood process. Our existing working relationship with e2v technologies in the UK, is beneficial to the LSST project and to UK industry involvement in LSST.

Work package description: As a continuation of our Phase A work (partly presented in Weatherill et al 2017b) and from which we were able to make informed input about optimum clock levels to the LSST camera team), this work package will translate the lab characterisation into a component of the overall PSF model, that will be tested against Comcam data and the effect of PSF modelling errors will be propagated through shear measurement to the weak lensing cosmology analysis.

Laboratory measurement and characterisation (Shipsey and Weatherill): In the lab it is comparatively easy to measure the image profile of a point source. The brighter-fatter effect (BFE) also appears in calibration flat fields, as correlations between neighbouring pixels. As flat field calibration data will be readily available, whereas non-linear point source calibrations will not, it is powerful to be able to infer BFE contributions from flat fields (e.g. Gruen et al 2015, Guyonnet et al 2015, Coulton et al 2017), but more accurate methods are needed. The development of BFE corrections is aided by knowledge of device physics encoded in numerical simulations. We will use commercial TCAD (Technology Computer Aided Design) software to verify and augment simulations currently used to model LSST sensors. We propose also to develop and extend the existing LSST simulations to handle near-full saturation (following on from Weatherill et al, 2017a) and partial depletion conditions which are not yet well understood or reproduced in existing simulations. We will assist the LSST Camera Team during commissioning in timing and operation optimisation.

Propagation to telescope PSF modelling, weak lensing measurement and cosmology analysis. The full PSF model comprises components (time-averaged atmospheric wavefront distortions, telescope optics, sensor response and telescope guiding) which are constrained by observations of stars within science images. The accuracy requirements are very stringent – for weak lensing, the accuracy must be comparable to that required for Euclid, despite being a ground-based telescope with the additional atmospheric component. The large aperture of LSST, yielding high signal-to-noise ratio on stars, is key to being able to model the PSF to sufficient accuracy. *An important development is full propagation of PSF modelling errors through to the cosmology analysis, to prioritise modelling of effects and ensure that the science goals may be met.*

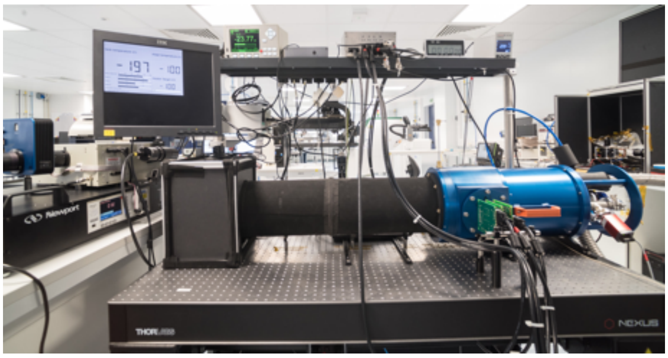

Figure 9.7 LSST Test Stand in Oxford OPMD Lab

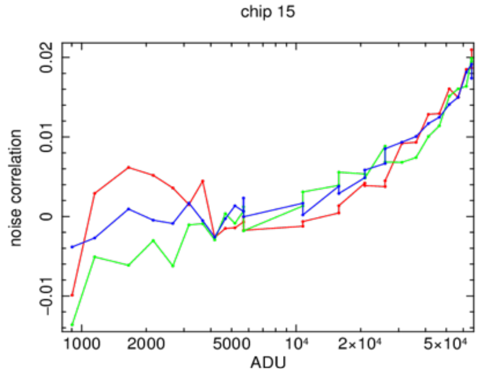

Figure 9.8 Anomalous correlated noise in the KiDS survey

Miller has led PSF modelling and shear measurement in the ground-based weak lensing surveys, CFHTLenS (Miller et al 2013), CS82, RCSLenS (Hildebrandt et al 2016), and KiDS (Kuijken et al 2015). These have non-linear sensor effects (Fig. 9.8 preliminary analysis by Miller, showing correlation in noise between neighbouring pixels in the serial readout direction, for pixels close to (blue curve), mid-range (green curve) or far from readout (red curve). The aim of the Phase B work is to synthesise a complete PSF model. There is no one method adopted for modelling atmospheric wavefront fluctuations – DESC is investigating Gaussian process models and *we will investigate a wavefront-domain Gaussian process framework.* However, to correctly transfer modelling of bright stars to faint galaxies, we need to understand the non-linear sensor effects. Application of models to the Comcam data will be an important test step.

WP3.9 Summary Gantt Chart

WP	Task Name	Start	Finish	Duration
3.9.1	Optimise sensor docking and bias parameters	01/07/2019	31/12/2020	78.8w
	Sensor characterisation complete	01/01/2021	01/01/2021	0w
3.9.2.1	Extension of LSST detector simulations with TCAD	*01/01/2020*	*30/06/2020*	26w
3.9.2.2	Extension of LSST detector simulations with TCAD, cont'd	01/07/2020	30/06/2021	52.2w
3.9.3	Incorporation of sensor effects into LSST PSF modelling	01/07/2020	30/12/2022	130.6w
3.9.4	Assist LSST Camera Team during commissioning in timing/ operation opt.	01/01/2021	30/12/2022	104.2w

Figure 9.9 Summary Gantt chart for WP3.9

WP3.9 Resources

Lance Miller (0.05 FTE), Ian Shipsey (0.05 FTE), current PDRA Dan Weatherill (1 FTE for 3 years = 36 months). Oxford have invested £250k to create a class 10000 clean room with smaller class 100 area for LSST and specialised CCD testing laboratory equipment (see Fig 9.7) at a level of £140k. This work package leverages full access to this valuable laboratory and equipment resource and only requests staff effort time.

WP3.9 Key Staff

Ian Shipsey has been a member of the LSST project camera team since 2008 bringing extensive instrumentation and data analysis experience from particle physics to the project. He interfaced LSST with UK CCD camera vendor Teledyne-e2v. With Weatherill he conducts characterization of LSST CCDs to optimise performance to maximise science reach. He served as Chair, LSST DESC Advisory Board (2015-17), and is an elected LSSTC Director (2018-22). As a member of the LSSTC Operations Taskforce he was instrumental in obtaining a permanent seat for the UK on the LSST Corporate Operations Council. He is Chair of the LSSTC Admissions Committee. His science interest is the measurement of dark energy via weak lensing tomography, a measurement that will depend on an exquisite understanding of the LSST PSF.

Lance Miller: Miller has had a long-standing interest in survey science and developed the PSF modelling and galaxy shear measurement methods used for the CFHTLenS and KiDS ground-based weak lensing surveys. He leads the development of the PSF measurement and modelling for ESA's Euclid mission, as well as working on shear measurement and galaxy modelling for that mission, and has taken an active interest in PSF modelling for LSST DESC. Miller is a full member of DESC.

9.6 WP3.10: The UK contribution to DESC Operations

WP3.10 goals and deliverables: (1) A programme of software infrastructure is vital for handling the large volume of DESC data and running it through reduction, analysis, and testing pipelines. Zuntz will build on his ongoing development work to deliver pipeline infrastructure for running the Weak Lensing and Large-Scale Structure pipelines from images (real and simulated) to cosmology. This consists of: interfaces to workflow management systems to run pipelines at supercomputer scale, including variant pipelines to measure cosmic magnification and 3D (non-tomographic) shear; interfaces to null and regression test suites; components of the DESC parameter estimation and covariance estimation codes. (2) DESC is running two large simulated data challenges, which form a critical part of preparation for cosmological analysis of LSST data. Perry will build on his previous DAC role to deliver suites of image simulations for these challenges. This will consist of contributing to the two codes used to generate image simulations, building frameworks to run them at scale, running them on UK facilities, and using LSST and DESC software pipelines to analyse the output.

Operations: In exchange for full data access the UK DESC community are expected to contribute towards development of pipelines for the reduction and analysis of DM-generated Level 3 data into summary statistics and tests useful for UK and wider science cases. These contributions will both directly enable UK science and provide UK personnel with the right to lead high-profile science cases on early LSST data. DESC has identified and prioritised several key infrastructural ("*operations*") roles. In consultation with them we have identified two roles the UK contribute to meet our requirements, weak lensing pipeline scientist and simulations & data wrangler, as critical in the preparation for LSST data.

WP 3.10.1 The Pipeline Scientist will design and implement the catalogue-to-cosmology pipeline infrastructure for the weak lensing and large-scale structure joint analysis, implementing (at LSST scale) algorithms for data reduction and analysis from DM catalogues into two-point measurements and other summary statistics. They will also write and manage the infrastructure framework for running the collected complete pipeline. This expertise is also key to supporting UK-specific research specialities which require variant pipelines, e.g. magnification (Joachimi group, UCL; Duncan et al 2016), joint catalogue analyses with Euclid or SKA (e.g. Brown group, Manchester; Harrison et al 2016), and 3D lensing (Peiris and McEwen groups, UCL, DESC 3D Data Compression Taskforce). Without further UK funding this position will pass to a US institution, losing the UK's leading role.

WP 3.10.2 The Data & Simulations Wrangler will manage simulation data generation, storage, and analysis as a key part of the data challenge simulation process (see below). Without UK funding in this area we will lose leadership in the generation of these simulations, making UK facilities, such as the STFC Common Cloud Platform, significantly less useful for major DESC usage. There will also be no UK expertise in running the LSST Data Management software stack at scale.

WP3.10 Summary Gantt Chart

WP	Task Name	Start	Finish	2019 Q3	2019 Q4	2020 Q1	2020 Q2	2020 Q3	2020 Q4	2021 Q1	2021 Q2	2021 Q3	2021 Q4	2022 Q1	2022 Q2
3.10.1.1	Development and use of DESC TXPipe and FireCrown software	01/07/2019	30/06/2021												
	Science-grade TXPipe and FireCrown software completed	*01/07/2021*	*01/07/2021*												
3.10.1.2	Further development and use of DESC TXPipe and FireCrown software	01/07/2021	30/06/2022												
3.10.2	*Data generation, reduction and storage for DESC data challenge*	*01/07/2019*	*30/06/2021*												

Figure 9.10 Summary Gantt chart for WP3.10

WP3.10 Resources: This request is for funding for Zuntz and Perry to take on these roles. We request 0.2 FTE for Zuntz as pipeline scientist for years from July 2019-2022, and 0.5 FTE for Perry as Data Wrangler for years 2019-2021 (the period of the data challenges). Zuntz also requests 0.05 FTE management time to supervise Perry during 2019-2021. Perry is a postdoctoral HPC professional and has been funded under the Phase A DAC grant and serving in the role of Data & Simulations Wrangler (see below) during Phase A. During this time he has contributed to the ImSim and PhoSim simulation codes (analyzing and breaking code bottlenecks), tested analysis frameworks against the Data Challenge 1 simulations, managed the deployment of data on UK testbed systems, and built Docker-based analysis environments for image simulation now in use by the working groups. He is tasked with managing the running the LSST image analysis stack on Data Challenge 2 data once it is generated (in the final year of phase A).

WP3.10 Key Staff
Joe Zuntz is a Chancellor's Fellow at the University of Edinburgh, and has recently completed a term as the co-lead of the DESC Weak Lensing working group, where he oversaw the DESC WLPipe project to prototype a weak lensing pipeline. He is now leading the Weak Lensing + Large Scale Structure joint pipeline project for DESC.

9.7 WP3.11: Cross matching and astrometry at LSST depth

WP3.11 goals and deliverables. LSST catalogues will be so crowded (even far from the Galactic Plane) that standard algorithms for cross-matching with other surveys will fail. Hence we will provide (through the DAC) a service which uses state-of-the-art cross-matching algorithms which include the effects of crowding, and partially mitigates them. We will provide matches with a wide range of surveys including VISTA, VPHAS, WISE and Spitzer, with the ability to extend to EUCLID, giving the UK a significant advantage in exploiting LSST data. In addition we will provide algorithms which calculate the effects of crowding on proper motions and parallaxes.

A significant fraction of LSST science will be based on finding reliable counterparts to objects in LSST images. Indeed, a key UK strength is using our legacy wide-field surveys to help interpret LSST data. We have already drawn the links in science cases 3.1 to 3.4, and for cross-matching for transient classification (WP 3.2). Once a cross match has been found WP 3.5 can provide the best fluxes, using the assumption of a common position. The problem though, is that the depth of LSST catalogues means their crowding will be similar to WISE, so that at least a quarter of stars and galaxies will have an unrelated object within a standard 2 arcsecond "matching circle". Worse still, the combination of depth and ground-based seeing means that there will be large numbers of faint, unresolved contaminants within stellar point spread functions. These move centres-of-light such that the true astrometric uncertainties can be far larger than the formal astrometric uncertainties (Wilson & Naylor 2017; Fig. 9.11). This affects not just the cross-match problem, but also parallaxes and proper motions. Thus this WP is also important for Solar System science, and a range of Galactic science problems including defining the properties of the moving groups, which are crucial to planet searches and characterisation.

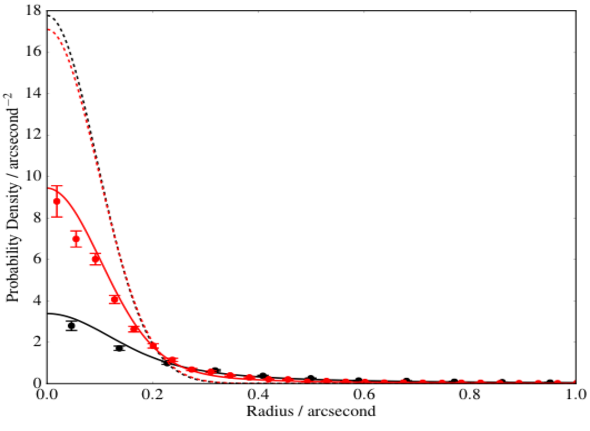

Figure 9.11 The underestimation of the astrometric uncertainty due to crowding. This figure compares the distribution of the separations between objects in WISE (which has similar crowding to LSST) and their Gaia counterparts (black error bars), with the expected distribution using the quoted WISE uncertainties (black dotted line). The solid black line is our model of the astrometric uncertainties which includes the effect of contamination, and produces the long wing to the distribution. The red coloured data are from a somewhat less-crowded region, and demonstrate how the true astrometric uncertainties in less crowded regions tend towards the uncrowded model.

Work package description. For any pair of catalogues there are three distinct tasks to carry out. **Task 1.** Enable the full power of the astrometric uncertainties provided by LSST by implementing a Bayesian match which assumes a Gaussian distribution of counterpart separations (Sutherland & Saunders 1992). This can reduce the number of false matches compared with error circle matching by an order of magnitude or more. **Task 2.** The effects of faint unresolved contaminants on the astrometric uncertainties are important in crowded fields, producing many false negatives if not corrected for. So task 2 is to model and correct for these non-Gaussian effects. **Task 3.** This in turn can have the effect of weakening the ability to decide between counterpart and field stars. So task 3 is to use a newly developed algorithm to include the photometric information from the two catalogues, which typically improves the Bayes factors of true counterparts by a factor of 10 (Wilson & Naylor 2018). The entire package will be a unique UK contribution to LSST, since we know of no US groups working in this area.

WP3.11 Summary Gantt Chart:

ID	WP	Task Name	Start	Finish
1	3.11.1	Model for contamination at LSST depth	01/11/2019	04/05/2020
2	3.11.2	Demonstration software for one example catalogue	05/05/2020	02/11/2020
3	3.11.3	DAC integrated software for one example catalogue	03/11/2020	02/02/2021
4		*Integration of algorithms from WP3.11.2 into DAC*	*03/02/2021*	*03/02/2021*
5	3.11.4	Final catalogue collection available at DAC	03/02/2021	02/11/2021

Figure 9.12 Summary Gantt chart for WP3.11

WP3.11 Resources. We request 1FTE PDRA for two years to carry out the majority of the work, and 0.1FTE for Naylor to manage that. Our strategy is to create prototype software which matches LSST with one other catalogue (WP3.11.1 and 3.11.2) and then deploy production software on the DAC (WP3.11.3). We will then add other catalogues (WP3.11.4). Flexibility in the final number of catalogues in WP3.11.4 gives us protection against overruns.

WP3.11 Key Staff
Tim Naylor led the catalogue cross matching effort for the MYStIX project, the Exeter component of the eSTAR project, and is currently responsible for the scheduler and archive software for the Terra-Hunting Project. He has a track-record of delivering novel astronomical algorithms including his optimal extraction for imaging photometry and Bayesian fitting of colour-magnitude diagrams.

10. Risk Analysis, Working Allowance and Contingency

We present an initial Phase B Risk Register as an appendix to this case, and have used it to estimate the level of Working Allowance and Contingency required for the Phase B programme. We have divided our risks into three classes: (i) those which would cause an additional cost to the project that could be addressed by calling upon Working Allowance funds; (ii) the risks which would result in a loss of value to the project that could not be ameliorated by use of Working Allowance funds; and (iii) one risk which yields a Contingency item. One situation – the UK failing to gain access to Commissioning data – has given rise to one risk in each of classes (i) and (ii).

To estimate the Working Allowance required for the Phase B programme we have considered all risks of class (ii) and computed for each the product of their Residual Risk Likelihood and their Cost. The Working Allowance is then simply the sum of those terms over the class (ii) risks and totals £497.5k (£398k STFC contribution), which is ~9% of our total Phase B budget of £5.5M (£4.4M STFC contribution).

The information contained in this appendix reflects our current knowledge of the likelihood and cost of the risks identified, but there are several cases in which there remains some uncertainty, due to constraints external to LSST:UK, so the Risk Register will have to be updated as these situations become clearer. Some clarification may be achieved before the start of the Phase B grants: for example, the Working Group set up by the LSST Project and LSST Corporation to develop a specification for nodes of the proposed international DAC network is due to report by the end of 2018, and that may include a minimal staffing requirement that could exceed the level of DAC funding included in WP2, necessitating the use of Working Allowance funds if the UK DAC is to qualify as a member of that network.

11. Summary

One of the distinguishing features of the UK involvement in LSST is its breadth. The breadth of the science enabled by LSST's unique – wide, fast, deep, multi-colour, multi-epoch – survey has resulted in engagement by a remarkably broad subset of the UK astronomical community – members of every astronomy group in the UK and from almost all sub-fields within astrophysics – plus, increasingly, a number of particle physicists, mirroring the situation in the US and in France.

The main goal for Phase B is to prepare the UK astronomical community for the start of LSST survey operations, by maximising the lessons to be learnt from LSST Commissioning. To that end, the LSST:UK Consortium has held a series of workshops over the past year to engage all its members in Phase B planning, and this proposal represents the community's collective view of its requirements for Phase B. The work packages included here are the prioritised outcome of assessment by an independent panel: the LUSC-DAV work package builds on the successful technical prototyping of Phase A and participation in the IRIS initiative to develop a coordinated computing infrastructure for STFC science; and the LUSC-DEV work packages are the prioritised selection resulting from evaluation against criteria of scientific excellence, ability to secure UK leadership and capacity to deliver scientific infrastructure to the benefit of the LSST:UK Consortium. The descope required by the specification of a funding envelope by STFC has removed a number of high priority work packages – in solar system science, photometric redshift estimation, strong gravitational lens discovery and the study of AGN variability - which will have an impact on the ability of UK researchers to secure leadership in key science areas, unless alternative sources of funding can be identified, but what remains is a programme of world-class R&D in preparation for one of the most exciting astronomical projects of the coming decade.

In addition to its scientific potential, the scale of its dataset and the duration of its survey make LSST an ideal project upon which to base programmes of Education and Public Outreach and of data science skills training. The LSST:UK Consortium will take advantage of those opportunities during Phase B, by applying for STFC Public Engagement schemes to tailor to UK circumstances resources being developed by the well-funded LSST EPO programme in the US and by coordinating with the STFC Data Intensive Science CDTs, respectively.

The next few years will be very exciting for everyone connected to LSST, as a construction site in the Chilean Andes becomes an operational facility. The Phase B programme proposed here will ensure that the UK community is ideally placed not only to exploit the immense scientific potential of LSST itself, but also to benefit from the enhanced return that incorporation of LSST data will bring from the investments already made in the other southern hemisphere facilities within the UK astronomy roadmap for the 2020s.

References

Abbott B.P. et al. 2017, PRL 119, 1101; Akhlaghi et al. 2015, ApJS, 220, 1; Angus et al. 2018, MNRAS, 474, 2094; Avestruz et al. 2017 arXiv:1704.02322; Banerji et al. 2008, MNRAS, 386, 1219; Banerji et al. 2015, MNRAS, 446, 2523; Bezanson et al. 2016, ApJ, 822, 30; Birrer et al. 2017 arXiv:1710.01303; Bonvin et al 2017 MNRAS 465; Bowler et al. 2015, MNRAS, 452, 1817; Carter & Winn 2009, ApJ, 704, 51; Collett 2015 ApJ 811 20; Collister & Lahav (2004), PASP, 116, 345; Contreras Pena et al.,2017, MNRAS, 465, 3011; Coulton et al 2017, arXiv:1711.06273; DES Collaboration 2017 arXiv:1708.01530; Davies, Serjeant & Bromley 2018; Duncan et al. 2016 MNRAS 457, 764; Duc et al. 2011, MNRAS, 417, 863; Ferreira Lopes & Cross 2016, A&A, 586, 36; Ferreira Lopes & Cross 2017, A&A, 604, 121; Fitzsimmons A., et al. 2018, NatAst, 2, 133; Fraser et al. 2008, Icar 195, 827F; Goldstein & Nugent 2017 ApJ 834 L5; Gregory & Loredo 1992, ApJ, 398, 146; Gruen et al 2015, Jour. of Instr., 10, C05032; Guyonnetel at 2015, A&A, 575, 17; Harrison et al. 2016 MNRAS 463 3674; Hartley et al. (2013), MNRAS, 431, 3045; Hartley et al. 2017 MNRAS 471 3378; Hildebrandt et al. (2017), MNRAS, 465, 1454; Hildebrandt et al 2016, MNRAS 463, 635; Hocking A. et al 2018, MNRAS, 473, 1108; Iodice E. et al. 2016, ApJ, 820, 42; Jacobs et al. 2017 MNRAS 471 167; Jarvis et al. 2013, MNRAS, 428, 1281; Juric et al. 2017, https://ls.st/LSE-319; Kasliwal et al. 2015, MNRAS, 451, 4328; Kaviraj 2014, MNRAS, 440, 2944; Kaviraj 2014, MNRAS, 437, L41; Kaviraj 2010, MNRAS, 406, 382; Kelvin et al. 2012, MNRAS, 421, 1007; Kenworthy, M., et al, 2015, ApJ, 800, 126; Kovacs et al. 2002, A&A, 391, 369; Kuijken et al 2015, MNRAS 454, 3500; LSST Science Collaboration 2017, arXiv:1708.04058; Lang et al. 2016, AJ, 151, 36; Lanusse et al. 2018 MNRAS 473, 3895; Lochner et al. 2016, ApJS 225, 31; Lucas, P., et al., 2017, MNRAS, 472, 2990; Marshall, Verma, More et al. 2016 MNRAS 455 117; Merlin et al. 2016, A&A, 595, 27;Miller et al 2013, MNRAS 429, 2858; More, Verma, Marshall et al. 2016, MNRAS, 455 1191; Morris et al. 2017, A&C, 20, 105; Mortlock et al. 2011, Nature, 474, 616; Muzzin et al. 2013, ApJ, 777, 18; Plavchan et al. 2008, ApJS, 175, 191; Reed et al. 2017, MNRAS, 468, 4702; Rhodes et al. 2017, arXiv:1710.08489; Robertson B. et al. 2017, arXiv:1708.01617; Sadeh et al (2016), PASP, 128, 104502; Schwamb et al. 2018, arXiv: 1802.01783; Smartt S.J. et al. 2016, MNRAS, 462, 4094; Snodgrass et al. 2005, A&A 444, 287; Sokolovsky et al 2017, MNRAS, 464, 274; Soo et al. (2018), MNRAS, 475, 3613; Suyu et al. 2013 ApJ, 766, 70; Tanvir N. et al. 2017, ApJL, 848, 27; Taylor & Joachimi (2014), MNRAS, 442, 2728; van Uitert & KiDS (2018), MNRAS, accepted; Weatherill et al 2017b, Jour.l of Instr., 12, C12019; Weatherill et al, 2017a, Jour. of Instr., 12, C05008; Williams, Baldry, Kelvin, et al. 2016, MNRAS, 463, 2746; Wilson, T.J. & Naylor, T. 2017, MNRAS, 468, 2517; Wilson, T.J. & Naylor, T. 2018, MNRAS, 473 5570; Wright D.E. et al. 2015, MNRAS 449, 451; Wright D. E. et al. 2017, MNRAS, 472, 1315; Zuntz et al. 2015 A&C, 12, 45; Zuntz et al 2017 arXiv:1708.01533.

Acronyms

4MOST – 4-metre Multi-Object Spectrograph Telescope; **AAP** – Astronomy Advisory Panel; **AGN** – Active Galactic Nucleus; **AP** – Affiliate PI; **AURA** – Association of Universities for Research in Astronomy; **BLAPS** – Blue Large Amplitude Pulsators; **CANDELS** – Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey; **CC-IN2P3** – Centre de Calcul de l'IN2P3; **CCD** – Charge-Coupled Device; **COC** – Corporate Operations Committee; **COC-DAS** – Corporate Operations Committee Dark Access Sub-committee; **CRTS** – Catalina Real-time Transient Survey; **CV** – Cataclysmic Variable; **DBMS** – Database Management System; **DES** – Dark Energy Survey; **DESC** – Dark Energy Science Collaboration; **DiRAC** – Distributed Research utilising Advanced Computing; **ELT –** Extremely Large Telescope; **GOODS** – Great Observatories Origins Deep Survey; **GRB** – Gamma Ray Burst; **HST** – Hubble Space Telescope; **HSC** – Hyper Suprime-Cam; **IN2P3** – Institut National de Physique Nucleaire et de Physique des Particules; **ISON** – International Scientific Optical Network; **IVOA** – International Virtual Observatory Alliance; **JA** – Junior Associate; **JVLA** – Jansky Very Large Array; **LoCuSS** – Local Cluster Substructure Survey; **LSB** – Low Surface Brightness; **ML** – machine learning; **MOPS** – Moving Object Processing System; **NCSA** – National Centre for Supercomputing Applications; **NTT** – New Technology Telescope; **OSARGS** – OGLE Small Amplitude Red Giants; **PESSTO** – Public ESO Spectroscopic Survey of Transient Objects; **PLASTiCC** – Photometric LSST Astronomical Time-series Classification Challenge; **PoC** – Point of Contact; **PSF** – Point Spread Function; **PTF** – Palomar Transient Factory; **SKA** – Square Kilometre Array; **SLAC** – Stanford Linear Accelerator Center; **SOXS** – Son of X-Shooter; **SRM** – Science Road Map; **SSAP** – Solar System Advisory Panel; **SVOM** – Space Variable Objects Monitor; **TNOs** – Trans-Neptunian Objects; **TVS** – Transients and Variable Stars; **UKIDSS** – UK Infrared Deep Sky Survey; **VLBA** – Very Long Baseline Array; **VLT** – Very Large Telescope; **VST** – VLT Survey Telescope; **VVV** – VISTA Variables in the Via Lactea; **WISE** – Wide-field Infrared Survey Explorer; **XCS** – XMM Cluster Survey; **XRB** – X-ray Binary; **ZTF** – Zwicky Transient Facility

Note: As is standard for STFC funded projects, we quantify risks in terms of likelihood (L) on a scale of 0-1, and impact (I) on a scale of 0-100. The total (T) risk factor is the product of these.

A risk with total risk factor above 25 is classified as a medium risk; above 50 is a major risk

Risk exposure 'cost' figures in blue cells represent the potential loss of value to the project rather than additional funding that would be required from STFC

Risk exposure 'cost' figures in orange cells represent a potential call on contingency funding

Ref	Risk Description	Potential impact on project	Owner	Inherent Risk			Existing Controls	Current/Proposed mitigation	Residual Risk			Risk Exposure		Action Required	Working allowance (£k)
				L	I	T			L	I	T	description	Cost (£k)		398
R01	Loss of key project staff	Schedule delays; loss of skills/ knowledge; damages to internal morale	**WP1**	0.5	60	30	Local management procedures	Motivate and reward key staff; anticipate any unavoidable issues and plan to replace with minimum overlap; aim to avoid key knowledge residing with one person	0.2	30	6	Interruption to project work. Recruitment costs (assume 20% staff turnover and £1,000 per recruitment) and lost productivity (assume 2 months to train	**293**	**maintain item and monitor**	**59**
R14	Weak currency exchange rate increases Sterling cost of PI/ JA subscriptions	Greater funding required to fulfil costs/ reduced number of PI positions	**WP1**	0.4	70	28	Work with STFC (Colin Vincent) to address early	Identify additional funding/ reduce number of PI positions	0.2	40	8	Increase in project costs—exposure figure based on cost at exchange rate $1.2 USD/ £	**1,334**	**maintain item and monitor**	
R15	Insufficient DAC hardware provided through IRIS initiative	Inpedes downstream UK science goals; UK liable for data access fees	**WP1**	0.5	50	25	Engage with IRIS to influence infrastructure choices to meet LSST:UK requirements	Prioritise LSST:UK science cases that use available resources; keep LSST:UK researchers informed of IRIS opportunities	0.25	30	7.5	Reduced science output. One year of access charges for 100 PIs and 250 JAs to access US DAC, at $1k	**233**	**maintain item and monitor**	**58**
R02	Major delay with LSST construction project	Increased lifetime cost; STFC funding in Phase B is spread more thinly; planning uncertainty	**WP1**	0.3	70	21	Track LSST roadmap locally, and highlight changes	Engage and assist LSST project; anticipate issues as early as possible; rephase UK project as necessary. Understand LSST has eleven-month contingency in plans.	0.2	40	8	Additional cost to maintain continuity of Management and DAC expertise. Expsoure costed based on twelve months of key WP1 and WP2 task costs in	**281**	**maintain item and monitor**	**56**
R13	Leaving EU adversely affects funding of training network	Funding level reduced, training programme de-scoped	**WP1**	0.3	70	21	Monitor EC funding situation	Identify alternative funding mechanisms	0.1	70	7	Reduced capacity to fulfil UK science goals. Lost project income from EU	**800**	**maintain item and monitor**	

R04	Late recruitment of staff	Delays against the schedule	**WP1**	0.5	40	20	Local management procedures	Monitor progress in consortium partners;	0.25	20	5	Some schedule delays. Risk exposure based on average delay of 1 month across all appointments, and assume no extension to Phase B and data	**121**	**maintain item and monitor**	
R12	Failure to join LSST Corporation	Inability to influence operational decisions to benefit of UK science goals	**WP1**	0.6	30	18	Progress w/ support of STFC (Colin Vincent)	Work through Science Collaborations to prioritise UK science goals and within LSSTC Corporate Operations Committee to voice UK concerns that arise during operations.	0.2	30	6	Reduced capacity to fulfil UK science goals. Lost opportunity estimated at value of one DEV work package	**273**	**maintain item and monitor**	
R05	UK does not gain access to commissioning data	Impedes downstream UK science goals	**WP1**	0.3	60	15	Document aim in LTP, and develop strategy in SRD	Engage LSST technical/management staff early; target relevant contributions; have appointed Commissioning Coordinator to liaise with Project team. Identify alternatives to commissioning data—HSC for Imaging; ZTF for transient alerts	0.1	40	4	Some risk to UK science; loss of value of Phase B commissing work	**81**	**maintain item and monitor**	
R21	UK does not gain access to commissioning data	Impedes downstream UK science goals	**WP1**	0.3	60	15	Document aim in LTP, and develop strategy in SRD	Engage LSST technical/management staff early; target relevant contributions; have appointed Commissioning Coordinator to liaise with Project team. Identify alternatives to commissioning data—HSC for Imaging; ZTF for transient alerts	0.1	40	4	Additional effort to complete DEV activities	**382**	**maintain item and monitor**	**38**
R16	LASAIR not selected as community broker	Impedes downstream UK science goals	**WP3**	0.3	40	12	Engage with Project regarding broker selection criteria	Plan for use of another community broker	0.2	20	4	Reduced science output. Loss of impact of Transient activities in Phase B	**495**	**maintain item and monitor**	
R08	Simulation resources insufficient	Impedes downstream UK science goals	**WP3**	0.2	50	10	Learning from US LSST experience	Begin testing early; consider buying more hardware; utilise third-party resources such as provided by GridPP and IRIS	0.1	20	2	Some risk to UK science. Additional computing cost based on topical DESC Data Challenge requirements (20M CPU hours)	**147**	**maintain item and monitor**	**15**
R09	Deep Drilling fields exclude favoured UK regions	Impedes downstream UK science goals	**WP3**	0.2	50	10	Informal relations with LSST staff	Establish UK members on key Science Working Groups; Interact early with key LSST staff	0.1	15	1.5	Reduced impact of some DEV activities	**148**	**maintain item and monitor**	

R17	LSST Project fails to secure sufficient funds for operations	Impedes downstream UK science goals, by shortening survey	**WP1**	0.3	30	9	Monitor status of funding, via Corporate Operations Committee	Revise science plans to take account of reduced survey length/ scope	0.3	15	4.5	Reduced science output. Exposure based on loss of DAC income, if survey reduced to 9 years	**358**	**maintain item and monitor**	
R10	UK unable to influence survey cadence to meet UK requirements	Downstream UK science goals put at risk	**WP3**	0.2	30	6	Liaise with LSST: Science Advisory Group; Informal relations with LSST staff	Establish UK members on key Science Working Groups; Interact early with key LSST staff	0.1	20	2	Reduced impact of some DEV acitivities	**259**	**maintain item and monitor**	
R18	Failure to secure funding for LSST:UK Phases C and D	Significantly reduces ability for UK astronomers to realise science goals. Additional demain on Astronomy grant line for DEV work	**WP1**	0.1	70	3.5	Work w/ STFC to address early	Identify additional funding for LSST DAC access, plus apply to Astronomy grant line for DEV work.	0.02	70	1.4	Lost impact from DAC work in Phases A and B, plus DEV work in Phases C and D	**4,800**	**maintain item and monitor**	
R11	Qserv architecture will not scale to meet UK science requirements	Downstream UK science goals put at risk	**WP2**	0.2	30	6	Early Data Challenges	Begin testing early and assess; consider alernative solutions as necessary; discuss solutions with LSST DM team	0.1	20	2	Risk to UK science goals. Exposure based on need to re-run database evaluation, planning and implementation (3 SY); plus additional 10%	**359**	**maintain item and monitor**	**36**
R19	Insufficient effort to achieve critical objective of DEV work packages	Reduced impact of DEV work	**WP3**	0.3	30	9	Project-management approach, based on experience from Phase A	Employ design-to-schedule approach and track progress on quartlerly basis	0.2	30	6	Additional effort needed to achieve core DEV outputs	**358**	**maintain item and monitor**	**72**
R20	DAC effort is insufficient to meet minimum commitment for International DAC Network	Project liable for LSST Data Access fees	**WP2**	0.5	30	15	Project is contributing to International DAC Network strategy	LSST Project cost estimates for DAC scrutinised, and validated based on UK experience from previous sky surveys	0.25	30	7.5	Additional DAC effort needed in PY3 and PY4 (based on indicative minimum DAC effort of 5 FTE)	**256**	**maintain item and monitor**	**64**