White Paper ID Number W015

Title of White Paper Canadian Participation in the LSST

ID of Associated Expression of Interest E040

Topic Area of White Paper science programs, science topics and science themes

Executive Summary of White Paper (5000 character limit)

The Large Synoptic Survey Telescope (LSST) will revolutionize optical wide-field astronomy while opening the window to short time-domain astronomy. Many astronomy communities, like Canada's, largely do not have membership in the LSST. As a consequence, they have no proprietary data access but could benefit from LSST data products. We have conducted a survey of the Canadian community to gauge their interests with LSST, and what data products are most desirable to them. We find that the astronomy research of Canadians would greatly benefit from access to LSST data products. Canadian astronomers fall broadly into two categories: those interested in the time domain and LSST's alerts and those interested in more classical wide-field astronomy of the stationary sky. We have identified four data products in particular that the majority of responders desire. These data products represent a relatively low data volume compared to the full LSST pixel-to-catalog products. We present the idea of a Canadian LSST-light Science Platform that will host the four most desirable products upon their public release, serving access to all Canadians. This platform would act as a national facility that would largely satisfy the LSST-related data needs of the majority of Canadian astronomers at a fraction of the overall cost of full national membership. Hosting the public release imagery for worldwide access has been recognized as a highly valuable in-kind contribution to LSST operations, and could facilitate access to proprietary data for Canadian astronomers. Recommendation: we recommend that Canada establish an LSST Science Platform to enable Canadian astronomers to make full use of the LSST data set.

Lead author and affiliation Wesley C. Fraser, Herzberg Astronomy and Astrophysics Research Centre, National Research Council of Canada, 5071 W. Saanich Rd. Victoria, BC, V9E 2E7

Email address of lead author wesley.fraser@nrc-cnrc.gc.ca

Other authors and affiliations

Melissa L. Graham, Department of Astronomy, University of Washington
JJ Kavelaars, Herzberg Astronomy and Astrophysics Research Centre, National Research Council of Canada
James Di Francesco Kavelaars, Herzberg Astronomy and Astrophysics Research Centre, National Research Council of Canada
Renée Hložek, Dunlap Institute and the Department of Astronomy & Astrophysics, University of Toronto
John P. Blakeslee, Gemini Observatory
Patrick B Hall, Department of Physics and Astronomy, York University
Brian R. McNamara, University of Waterloo & Waterloo Centre fo Astrophysics

1 Introduction

The Large Synoptic Survey Telescope (LSST) aims to make major advances in our understanding of dark energy and dark matter, small bodies in the Solar System, transient phenomena, and the Milky Way by surveying $\sim$18,000 deg^2 of the southern sky to $r \sim 24.5$ mag (single image) and $r \sim 27.5$ mag (final depth), over ten years starting in 2022 (Ivezić et al., 2019). The LSST will acquire $\sim$20 terabytes of raw data each night and process it in real time, distributing alerts on objects that vary in brightness or position within 60 seconds, and generating calibrated images and updated source and object catalogs within 24 hours. After the first half-year of operations, and on an annual basis thereafter, LSST will release a yearly data set that includes deep image stacks and source and object catalogs. A full description of the LSST data products is provided in Jurić et al. (2018). With the exception of the public alerts, LSST data products are subject to a two-year proprietary period, after which the data can be shared with anyone worldwide. Only scientists in the US or Chile, or individuals named in membership agreements (referred to collectively as LSST Members in this document), however, may access, analyse, and publish proprietary LSST data.

To enable science with the massive LSST data set, the LSST Data Management System includes the Science Platform: a web-based service for data access, analysis, and processing which includes software tools and computational resources that will be accessible by LSST Members (Jurić et al., 2017). Although the stream of $\sim$10 million real-time alerts per night is public, it is only guaranteed to be delivered in full to five Alert Brokers (Bellm & co authors, 2018) due to the very high bandwidth required. These brokers will serve the alerts to their communities (some brokers plan to provide public access, e.g., Narayan et al. 2018; Smith et al. 2019). Neither brokers nor their users need to be LSST Members. An additional advantage of LSST membership includes participation in the Science Collaborations, which are developing software for scientific analyses of the LSST data.

1.1 Memberships

The following represents the state of LSST membership costs as of May 2019, before the results of the Open Data Framework discussions were made publicly available. Membership fees through direct contribution to operations costs are now being replaced with in-kind contributions to LSST operations and possibly effort towards the development of science products. At time of writing (Sept. 2019), however, the specifics of the return in memberships for amounts of in-kind contribution remain to be negotiated. See the Data Rights and Funding for LSST Operations webpage[1] for more information.

1.1.1 The Old Model for LSST International Memberships: Funding for Operations in Exchange for Data Rights

This section provides guidance on the level of financial contribution that we anticipate LSST will require for in-kind membership contribution. At this time, no detailed in-kind contribution scoping is available but NRC (through the CADC) is in close communication with LSST, NSF, and DoE on this issue.

There were two routes for non-US astronomers to gain LSST membership. The first option was a national level buy-in, as the UK did. The UK Science and Technologies Facilities Council (STFC) award was sufficient for 100 senior memberships (faculty) and 400 junior associates (PhDs and post-docs). The funds include £15 million ($20 million USD in 2015) as a contribution to LSST operations. An additional £3.4 million (£2.7 million award plus matching funds from the STFC) was awarded for the development of software, a national data centre (the DAC) that will host a copy of the LSST data products, and the LASAIR alert broker (Smith 2019). An additional $\sim$£4 million is expected for Phase B, which is intended to cover additional software and data product development up to the start of LSST operations. An additional $\sim$£11.5 million will be requested during Phases C and D, which will cover DAC operations, and Education and Public outreach activities. The expected total for the national UK LSST buy-in is roughly £34 million, or roughly £340,000 per faculty membership (<£300,000 when excluding the UK DAC expenses). For those astronomers not in countries with national membership, individual or institutional-level memberships could be purchased. A full list of member institutional member facilities can be found on the LSST webpages.[2] [3]

Canadian astronomers had the option to buy individual team memberships. Each membership would cost

[1] https://www.lsst.org/news/data-rights-and-funding-lsst-operations
[2] https://www.lsstcorporation.org/node/3
[3] https://www.lsstcorporation.org/international-contributors

$240,000 USD, which provides data rights for the lead astronomer and three junior teammates such as postdocs and PhD students. To have access to NOAO's compute centres where the LSST data will be hosted and manipulated by users costs another $6,800 USD per team. Including typical institutional overheads and exchange rates, the estimated cost per membership was $517,000 CAD. For Canadians, individual team memberships with full data rights and access are roughly 85% of the equivalent price paid by the UK. Membership funds, however, are very difficult to come by in the Canadian funding climate. There are a few notable exceptions.

The University of Waterloo has committed to institutional membership, purchasing full membership for five faculty members, supported by contributions by the National Science and Engineering Research Council, the Waterloo Faculty of Science, the university provost, and a half membership fund provided by the Dunlap Institute (see below). Additional memberships for other Waterloo faculty may be purchased. As part of the preparations for the LSST, a Canadian consortium had formed, led by Dunlap Institute. Dunlap Institute had planned to provide funding to a limited number of Canadian investigators. The nominal plan being 50% funding for 10 Canadian participants with those participants providing the other 50%, plus additional data access fees. Development activities from this consortium include a CFI proposal (PI: Renée Hlozek) to develop software infrastructure aimed at utilizing the LSST alert stream. Additional development activities inside the Canadian Astronomy Data Centre (CADC) are likely, and will depend on the scope of Canadian LSST users' plans.

1.1.2 Assessing Canadian Need for LSST Membership, Going Forward

While visibility of desire for national LSST access has appeared low, anecdotally there has been clearly visible demand for membership by some. To that end, we surveyed the Canadian astronomy community, seeking to gauge the specific interests Canadians have with aspects of LSST science and the availability of funding for them. In this white paper, we discuss this survey and the responses we received. It is clear that there is significant interest in some particular aspects of the Prompt and Data Release products described above.

The results of that survey have made apparent a low-cost route to providing data access to the most desirable data products. This so-called "LSST-light" dataset, would consist mainly of annual release images and source catalogs (including time domain information) made available after the proprietary period, and would take advantage of national hosting and compute facilities in Canada. The LSST-light science platform would host the most popular LSST data products requested by the vast majority of survey respondents. The LSST-light science platform represents an opportunity to create a national level facility that is useful to the majority of Canadian astronomers, but at a cost of roughly only half of the full buy-in costs for memberships for each of the respondents.

2 The Canadian Survey

To gauge quantitatively the Canadian astronomy community's scientific interest in usage of the LSST data products, a survey was sent out to the community, fashioned to assess broadly community scientific interests, LSST data requirements to meet those interests, and plans for access to those data. In the following section, we discuss in detail the responses to the questions posed in the survey. We also discuss a route forward to LSST data access that makes use of the facilities and expertise inside Canada and the CADC that would address the majority of the needs of survey respondents.

2.1 The scientific interests of respondents

Thirty-nine Canadian scientists responded to the survey. Responses arrived from 16 unique Canadian institutions, including both federal research centres and universities. While the response rate might seem low, we emphasize here that the responses to this survey are likely biased towards those Canadians who are directly interested in science fields related to the LSST. It is unlikely that those without interest in transient or wide-field astronomy would have taken the time to respond to this survey.

To gauge scientific interest, two simple questions were asked: "What are your general science interests?" and "What are your LSST-specific science interests?" Multiple responses were possible. We present the summary of responses in Figure 1.

While the interests of Canadians span the breadth of observational astrophysics, the majority interest rests in studies of the local universe and transient sources - both single event transients like supernovae and repeating transients like variable stars.

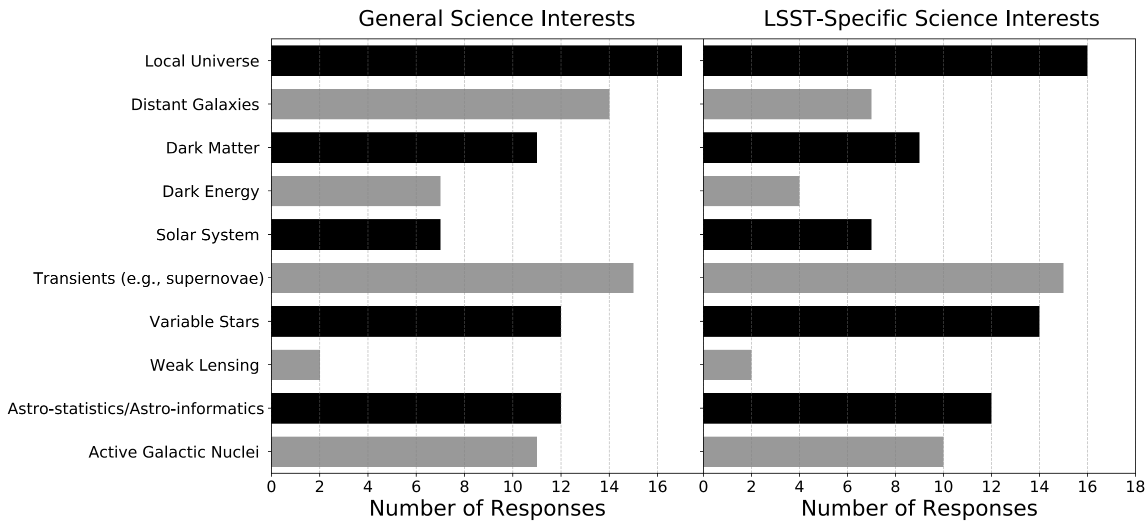

Figure 1: The general and LSST-specific science interests of the survey respondents.

2.2 The Desired Data Products

To assess interest in LSST data products, we presented respondents with a list of prompt and release data products, and asked them to state which products they anticipated using in their LSST-related research. The responses to both are presented in Figures 2 and 3.

Two interesting results are implied by the responses. The first and most notable result is that more than half of respondents are interested in making use of the prompt data products. Most of the respondents with interest in the prompt products wish to make use of a queryable catalog of transient alerts, with 1/3 expressing interest in utilizing the alert stream itself, either from a community broker or from the LSST Science Platform.

The second interesting result is the popularity of the "all-sky" data products. The release products most Canadians are interested in are the reprocessed full source catalogs and deep co-adds. The third most sought after annual product, reprocessed catalogs of transient and variable sources, reflects the most sought-after prompt data product, a queryable database of alerts.

Taken together, these results imply that Canadians fall into two types of users: those interested in fast time-domain astronomy and those more oriented towards classical wide-field astronomy. The first group makes up 1/3 of respondents, which is fractionally higher than those with LSST membership (based on the populations of the various LSST Science Collaborations; Bianco et al. 2019). Clearly, live access to alerts in one form or another is important for Canadian astronomy.

It is also clear that the majority of Canadians interested in use of LSST observations remain interested in classical wide field astronomy of the stationary sky. It seems the majority of that interest can be served with only the four most sought-after of the aforementioned data products. This tendency presents the possibility of serving a light version of the LSST data products to Canadians without the need for many of the expensive products or the compute facilities required for their use. We discuss this scenario in Section 3 below.

Respondents were also asked "Regardless of your membership status and data rights access, how beneficial might access to LSST data, either before or after the two-year proprietary period, be for your personal science goals? Please rank between 1 (least beneficial) and 5 (most beneficial)." The summary of responses is shown in Figure 4. The response is clear: the respondents overwhelmingly agree that LSST data would be greatly beneficial to their science research. The mean response was 3.86, with only three responses of 2 or lower. 70% responded with 4 or higher, demonstrating a clear desire for access to LSST data products.

In addition, respondents were asked if they had access to funding for LSST membership, and if so, from what source (see Figure 4). The large majority, $> 80\%$, have no membership funds of any kind. Only three responded with full funding and four others with partial funding. Those respondents are funded either through DI support (approximately five full memberships), or as faculty at the University of Waterloo.

Finally, respondents were asked if membership were available, would they purchase into the NOAO compute pool ($1700 USD per year per member). The majority (22) responded N/A, i.e., they would not be able to purchase memberships so the question was moot. Of those that did have membership ambitions, however, ten responded negatively, i.e., they would not be able to purchase the needed computing capacity directly, and only four responded positively.

The response to these questions demonstrate a clear desire by the Canadian astronomy community to have some access to LSST data, and, if possible, to participate directly in the alert and static sky science. This interest is present despite the current funding climate which has prevented most Canadian astronomers from securing membership funds.

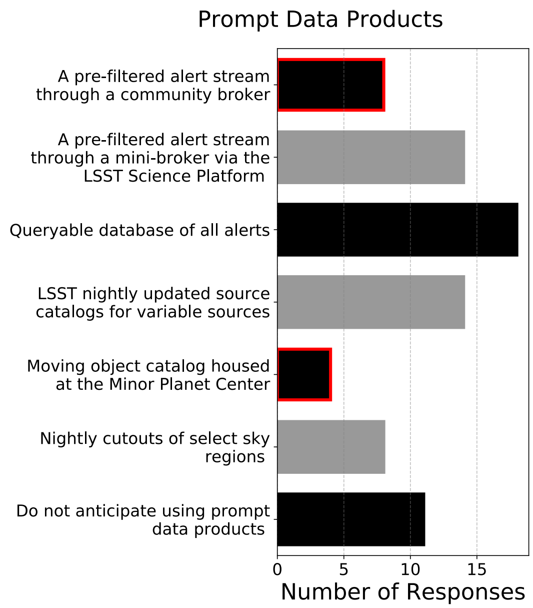

Figure 2: Summary of responses to the question "Which prompt data products do you anticipate using in your LSST-related research?" Publicly available products have been highlighted with a red outline.

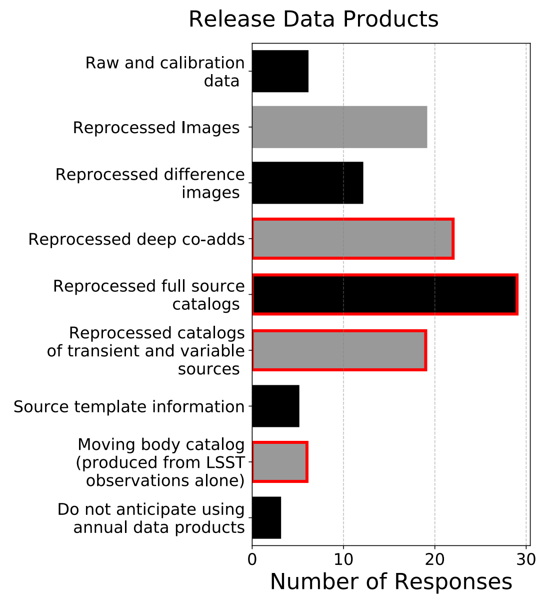

Figure 3: Summary of responses to the question "Which release data products do you anticipate using in your LSST-related research?" Those products that will be publicly available after a two-year proprietary period are highlighted with a red outline.

3 The Canadian "LSST-light" science platform: a low-cost route to LSST data products

From the survey, the most desired data products, in decreasing order, are:

- Annual Data Release full source catalogs
- Annual Data Release multi-band stacks
- Prompt queryable alerts database
- Annual Data Release transient catalog

These 4 products cover the majority of Canadian user's LSST science interests, with 34 of 39 respondents desiring at least one of those products, and 31 of the respondents desiring two or more.

The summary of responses present an opportunity to host a so-called LSST-light science platform: a data centre to host these four most desirable data products combined with interfaces and compute capacity to enable science using those data products. Such a science platform would require only a small fraction of the overall LSST data storage, and provide useful capacity to the majority of Canadian astronomers interested in using LSST data.

The full source and transient catalogs will be roughly 2 petabytes (PB) at first release, and grow to a full 15 PB at survey completion[4]. A single copy of the all-sky co-adds will occupy $\sim$0.6 PB. More precise data volume estimates are not yet available. It is unlikely that more than the most recent two releases need be hosted in the LSST-light science platform. The most complex data product will be the queryable transient catalog. Each

[4] https://www.lsst.org/scientists/keynumbers

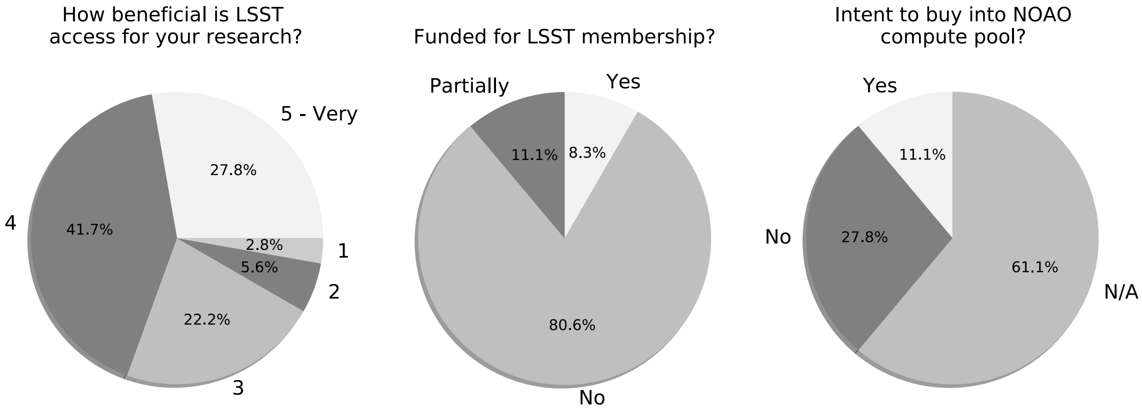

Figure 4: Membership status, compute pool, and benefits to science.

alert packet will be 82 KB or less, with roughly 10,000 alerts produced per visit, on average. As the survey executes, the alerts database will grow linearly to ∼2.2 PB after the full 10 year survey[5]. The total storage requirement of this LSST-light database is roughly 18 PB, or ∼25% of the full LSST data product volume.

Given the unprecedented data volume, it is unlikely that any users will have sufficient personal compute resources with which to process more than minute quantities of the LSST data products. Rather, LSST will operate with large-scale, user-accessible compute resources hosted next to the data at the NCSA, and accessed via the LSST Science Platform (LSP; Jurić et al., 2017; Gregory Dubois-Felsmann & Wu, 2019). The LSP will serve users via the platform, Jupyter notebook, and API aspects, providing the tools to explore, analyse, and reprocess the data (images and catalogs) using the LSST Science Pipelines or user-generated software. For Canadians, a similar setup will be required, even for the LSST-light data volumes. The LSST-light science platform would not only provide the access to the popular data products, but also the compute resources required to make use of those data scientifically.

The LSST catalogs will provide an assortment of measureables of each source, including photometry, morphology, and colour information. In determining what techniques should be implemented in measuring source properties from LSST imagery, input from the US astronomical community was solicited to determine best what techniques are likely to satisfy the majority of astronomers. Other communities, however, may express different science interests. It is for this reason that the UK built their own LSST data centre that will host a complete copy of the entire LSST archive but also provide customization and enhancements to meet its own interests. For example, the LASAIR alert broker has been developed to add additional source data to alerts to enable best follow-ups of supernova transient events, a popular topic amongst UK astronomers.

In a similar fashion, the Canadian LSST-light science platform would provide critical data customization and enhancements of the scientific utility of this already powerful data set. Exactly what customization is most beneficial to Canadians will require additional input from the Canadian community, but could include: additional detailed morphological information of galaxies to improve weak lensing measurements; specialized PSF deconvolution to detect better unresolved binary stars and planetesimals; or alternative source centroiding that would be more appropriate for highly multiplexed, multi-object spectroscopy, just to name a few. The LSST-light science platform would act as a national facility, providing a one-stop shop for all Canadian LSST requirements, providing data for Canadians that have been analyzed and enhanced by Canadians.

3.1 Hosting at the CADC

The Canadian Astronomy Data Centre (CADC) is a national astronomical database facility serving large volumes of astronomical imagery and catalogs to the Canadian and international communities. The CADC provides access to ∼2 PB of astronomical data with discovery via a robust query interface that enables spatial, temporal, and wavelength-based searches and retrieval using astronomy-domain specific approaches. The CADC

[5] `https://dmtn-102.lsst.io/`

also provides specialized compute facilities for the astronomical community, enabling complex astronomical archival analysis and more general astronomical computation. The CADC has a successful history of enhancing the telescopic data products. A recent example is the UNIONS project, which aims to enhance the Panoramic Survey Telescope and Rapid Response System (PANSTAARS) data set with all-sky u-band data acquired at the CFHT. The CADC is a natural place to develop the proposed LSST-light science platform.

Building off the infrastructure that already exists at the CADC, the LSST-light data set should be queryable in the same way as current CADC-hosted data are. These data should also be processable through the CANFAR virtual machine system, which would act as the primary method through which computationally expensive scientific processing will be performed. Furthermore, to provide direct personal access to the LSST-light products, CANFAR should be enhanced to provide a Jupyter notebook-based interaction, following that being developed by the LSST. These three points of access will provide the necessary and sufficient data access and processing capabilities for Canadian users. In addition, bringing the LSST data set into the CADC/CANFAR environment will automatically make that data set exposed to IVOA protocols, enabling tools developed for other projects (e.g., CIRADA) to be easily adapted to the LSST data set.

The Canadian LSST science platform we describe above will require some advanced development. Long term management and operations of the LSST-light science platform will require a modest increase in staffing to support data management, science, and ongoing development, and to maintain the hardware on which the data will be stored. Further, some front-loaded effort will be required to provide roughly two years of initial software development support. If funds were successfully found for such a LSST-light science platform (likely a combination of CFI and NRC funds), a bid for CANFAR compute resources will be submitted to Compute Canada.

There are likely to be additional costs associated with the simple action of transferring the LSST data from NSCA servers to those managed by the CADC. While these costs are still to be determined, we can make a rough estimate of the magnitude of those costs from the current pricing structure of Amazon Web Services (AWS). From the data volume considerations expected, an annual cost of \$220,000 USD is estimated, with an additional ~\$50,000 USD in total transfer costs associated with the alerts database. The annual release products are likely to be transferred in batches two years after their release. How the alerts database is transferred largely depends on the results of a CFI proposal (PI: Renée Hlozek) to provide an alert broker for the few Canadians who will hold membership with the LSST.

3.2 Cost comparison with national buy-in

Under the previous buy-in model approach (see §1.1.1), the full costs for direct membership purchase for the LSST was roughly \$517,000 CAD per senior researcher at time of writing. The majority of survey respondents are either already faculty, or will be at an equivalent level by the time LSST commences operations. If we assume the response represents the majority of Canadian interest in the LSST, then a national buy-in, with 40 membership positions, would cost roughly \$20 million CAD. The benefits to full buy-in are obvious. Members will have proprietary access to all LSST data products, including pixel products, source catalogs, and alerts, and will have access to the NCSA compute resources tied directly to the full LSST database, at an equivalent level to US and other national members. Without a special influx of funds however, full membership of all Canadian scientists is virtually impossible, as no appropriate source exists within the current Canadian funding climate from which a membership can be funded.

An obvious benefit of the LSST-light science platform is a massive reduction in cost. This likely would include the retraction of some of the funds for ten Canadian memberships that have already been committed by the University of Waterloo and the Dunlap Institute. Most notably, the fund recovery would come from those who do not require access to the actual individual images, but are content with the products provided by the LSST-light data set.

Beyond the lower cost, the clear benefit of the LSST-light science platform is the creation of a science platform for the benefit of all Canadian scientists. The platform would provide data access and compute facilities that would satisfy the majority of current Canadian interests in wide-field optical astronomy, albeit after a two-year proprietary period, and with no direct access to live alerts produced by the LSST. Under LSST's new in-kind contribution membership model, however, Canada could negotiate with LSST to allow Canadian national access to the LSST data in exchange for Canada providing worldwide public access to the public data component of a Canadian LSST Science platform.

For comparison, the United Kingdom plan to buy into LSST as a national partner carried a cost of £15 million pounds in operations costs ($25.5 million CAD) plus the additional costs associated with the operating the UK DAC and buying full data access for 100 faculty and up to four junior team members per faculty member. The total award of £34 million pounds includes all costs of a national UK science centre to provide the necessary national compute facilities, and the facilities to operate the LASAIR full stream alert broker[6]. The overall cost of the UK national buy-in provides a lower cost - roughly 10% less - to LSST access than individual memberships.

4 Considerations for current and future Canadian astronomy facilities

There are many current or upcoming Canadian national astronomical facilities that have major synergies with the LSST-light science platform we outline above. In particular, the Manuakea Spectroscopic Explorer, and the Gemini-North and South Telescopes both will need to utilize LSST data products in their daily operations to fulfill the basic science requirements of each facility. We discuss below both of these facilities and their respective LSST data requirements.

4.1 Maunakea Spectroscopic Explorer

The MaunaKea Spectroscopic Explorer[7] (MSE) is a highly developed vision for the future of the Canada-France-Hawaii Telescope. MSE will be an 11-m, highly multiplexed, fibre-fed, optical spectroscopic survey telescope, providing a national astronomical facility for use by all Canadians. MSE will gather up to 4,332 spectra over a 1.5 square degree field in every exposure it acquires. Though MSE's fast fibre placement makes it an excellent tool for LSST follow-up of moderately bright ($<$ 23 mag) sources, most of the planned science activities centre around gathering spectra of the non-transient local universe (Hill et al., 2018).

A critical outstanding item required for the success of MSE is a deep, all-sky source catalog, without which accurate fibre placement will be impossible. The northern portion of this finder chart will likely come from some combination of the *Euclid* survey and the Pan-STARRS 1 and 2 surveys. The LSST all-sky catalogs and image stacks, however, are ideal for MSE's purposes, providing accurate photometry and astrometry of all sources more than a magnitude fainter than realistically observable by MSE. The LSST will be the only deep survey facility to provide coverage for the entire sky South of 0 degrees declination, or roughly 1/3 of the observable sky over MSE.

The minimum data products that would be useful for producing MSE finder charts are a single release of the all-sky source catalogs and multi-band deep co-adds, above -30 degrees declination (half the coverage of the wide-fast-deep survey). At the very minimum, that data set will total at least 1.3 PB (the first annual release). While a relatively small data volume compared to the full LSST-light data set outlined above, even just the repeated handling and manipulation of a PB-scale data set still comes at substantial effort and expense.

With fully integrated compute and storage facilities, a LSST-light science platform would provide an essential service to the MSE. In particular, such an integrated facility would afford the straightforward creation of the catalog manipulation and target selection tools required of the MSE observing planning, and enable the creation of a combined photometric-spectroscopic imaging data set that would be realized by combination of the hosted LSST data products.

It should be noted that the current LSST data rights policy prevents usage of proprietary data by non-members. This policy therefore excludes use of the LSST by the MSE collaboration until after the proprietary period lapses. Under the evolving 'in-kind' data rights policy for LSST, Canada may provide some MSE access to LSST partners in exchange for LSST data access within the proprietary period, i.e., MSE access could be an in-kind contribution to LSST.

4.2 Gemini South as an LSST Follow-up Tool

Canada has been a long term partner in the the pair of 8-m Gemini telescopes, holding at least a 10% share since their inception. The Gemini telescopes act as one of Canada's primary large optical and NIR ground-based telescopic facilities.

A primary recommendation of the 2016 Gemini strategic vision was to modernize Gemini to provide improved support for LSST follow-up:

> Beyond 2021, Gemini should exploit its geographical location and agile operational model in order to be the premier facility for

[6] https://lasair.roe.ac.uk/
[7] https://mse.cfht.hawaii.edu/

> the follow-up investigation of targets identified by the Large Synoptic Survey Telescope.[8]

This strategy is favoured over developing synergies with upcoming 30-m class telescopes like the Thirty Metre Telescope, as it is seen as a more useful specialization that takes better advantage of the special capabilities of the Gemini telescopes. At the time of the last strategic vision survey, a large fraction of Gemini users advocated for a substantial fraction of Gemini being dedicated to following up LSST transients. For example, $\sim$30% of users think that at least 30% of Gemini's time should be spent on LSST follow-up. The reasons for this direction are obvious: LSST standard visit images will have $\sim$30-second exposure times and provide many transients brighter than 23$^{\mathrm{rd}}$ magnitude, which are tractable targets for detailed follow up with a facility like Gemini.

The on-going upgrades and enhancements to both Gemini telescopes are largely guided by its community's desire to enable fast transient follow-up of LSST transients and to maximize science output from observations of non-transient LSST targets. Example enhancements include an upgrade to the Gemini-North laser for improved adaptive optics imaging as well as improvements to the software and hardware back-end to reduce dramatically response times and enable dynamic scheduling responses to transient alerts. In particular, Gemini has become a partner of the Astronomical Event Observation Network (AEON[9]), and will respond to brokered LSST alerts. Full details of the upgrade program are available in the recently released Strategic Plan (Blakeslee et al., 2019).

In addition to strong support for LSST follow-up, equally strong support was given to preserve PI-led science capabilities for normal, non-transient science. Such projects would be member-proposed, standard-queue observations but would also encompass programs triggered in response to LSST alerts and observations of non-trigger targets. This strategy serves both the LSST and the non-LSST Gemini users. Although many Canadian astronomers have interest in transient science, about twice as many fall into the latter category. While many upgrades to the Gemini telescopes are driven by transient follow-up, most of these are generic and will improve most aspects of observing at Gemini.

Canadian astronomers could consider the LSST-light science platform as a necessary upgrade in their observing capabilities. As for MSE, the LSST-light data products will first and foremost provide interesting targets for detailed follow-up with the Gemini telescope. If hosted at the CADC, their compute infrastructure will make it relatively straightforward to mine the transient and stationary source catalogs for targets of a given type or class, e.g., Cepheid variables, high-redshift active galactic nuclei, or even targets that haven't been reliably classed from LSST observations alone.

Gemini provides observing capabilities extending beyond those of LSST, including medium- to high- resolution spectroscopy and high spatial resolution observations in both the optical and NIR. The combination of observations from both facilities will provide opportunity for an extremely rich scientific data set that wouldn't be possible with either telescope alone. Clearly, access to the LSST-light data set (at the very least!) is necessary for Canadians to remain at the forefront of optical-NIR observational astronomy.

Recommendation: This white paper recommends a Canadian LSST Science platform be established to enable Canadian astronomers to make full use of the LSST data set. Worldwide public access to the public LSST data housed in such a science platform will provide a highly valuable in-kind contribution to LSST operations.

[8] http://www.gemini.edu/files/general-announcements/2021beyond_strategic_vision.pdf
[9] http://ast.noao.edu/data/aeon

1: How does the proposed initiative result in fundamental or transformational advances in our understanding of the Universe?

The LSST will sample the southern night sky on cadences of hours to weeks, months, and years over multiple bands. This unique sampling of the visible universe will catalog the contents of our solar system and galaxy and discover thousands of distant supernovae that will enable detailed studies of the formation and content of our universe. No other survey of this scale in area and time sampling has ever been achieved so far.

2: Are the associated scientific risks understood and acceptable?

The major science risk to Canada is that we are not currently participants in this project. Although the 'Alert Stream' will be publicly accessible, the selection function of those alerts will not be known and out of the control of non-members. Access to the pixel data will enable the full use of the LSST data, not just the alert stream. A major risk to Canada, if we have access, is that we do not currently have adequate computing and human resources available for Canadians to exploit their access to these data.

3: Is there the expectation of and capacity for Canadian scientific, technical or strategic leadership?

Already, within Canada, a number of researchers are taking on leadership roles within the LSST project (see the LSST Science White Paper). In addition, the CADC is a world leader in provisioning of data systems for astronomy and direct participation in an LSST Science platform will further solidify that leadership.

4: Is there support from, involvement from, and coordination within the relevant Canadian community and more broadly?

The establishment of an astronomy science digital platform will allow Canadians access to the LSST data and enable their existing plans for LSST data usage. The LSST Canada consortium has submitted a separate white paper outlining their science ambitions. An LSST-Light science platform will be a central component of a Canadian Digital Platform for Astronomy (advocated in the Digital Research Infrastructure in Astronomy white paper).

5: Will this program position Canadian astronomy for future opportunities and returns in 2020-2030 or beyond 2030?

Projects like ALMA, LSST, SKA, and MSE are occurring in an era of digital transformation with the rapid up-take of 'Statistical Learning' (also called Machine Learning) across multiple science domains. An LSST science platform for Canada would form a core component of astronomy's transition into this challenging but exciting new era of data complexity. Canada also has ambitions to build a Canadian SKA Regional Centre (CSRC), this will require Canada to enhance its capacity for storage and computing in astronomy and establish an enhanced science portal/platform for domain specific access to Digital Research Infrastructure (i.e. a more robust CADC and substantially enhanced CANFAR). An LSST-light science platform would form the nucleation of the process to build the systems needed for the CSRC. MSE and *Euclid* will be operating in 2030 and beyond, and will also require an enhanced science platform. Building an LSST science platform that evolves to include also the CSRC will make Canada well prepared for MSE and *Euclid*.

6: Is the cost-benefit ratio, including existing investments and future operating costs, favourable?

The enabling component of an LSST science platform is the drive towards massive data systems within the commercial sector pushing downward the costing of capital. (The retention of high quality personnel within research data centres is a related significant challenge.) LSST data will begin arriving as the costs for storage begin to taper towards levels that make the project highly sustainable. Although there is significant data growth throughout the LSST project, the concomitant decline in storage costs means that running an LSST science platform can be achieved with a reasonably flat budget profile. The challenges and opportunities to work with this world-leading data set will play an important role in Canada maintaining its expertise in data management.

7: Are the main programmatic risks understood and acceptable?

The major risk is achieving the required funding rapidly enough to enable the Canadian LSST science platform to be operational by the start of observing at LSST (2022). Agreement on LSST in-kind contributions is being negotiated now with formal proposals required by **March 2020**. Movement forward on such negotiations should be aligned with recommendations from the LRP2020 process. The LSST schedule is being maintained and slippages in first-light time of more than 6-12 months are not likely.

8: Does the proposed initiative offer specific tangible benefits to Canadians, including but not limited to interdisciplinary research, industry opportunities, HQP training, EDI, outreach or education?

A Canadian LSST science platform, enabled in an environment close to the computing resources needed for advanced Machine/Statistical Learning, will provide fundamental training in the use of such techniques in a complex data universe. In Canada, we will be linking these data sets to existing Canadian astronomy data, thus enhancing further the value of those assets.

References

Bellm, E., & co authors. 2018, Plans and Policies for LSST Alert Distribution

Bianco, F. B., Banerji, M., Bochanski, J., et al. 2019, arXiv e-prints, arXiv:1907.09027

Blakeslee, J. P., Adamson, A., Davis, C., et al. 2019, arXiv e-prints, arXiv:1909.09196

Gregory Dubois-Felsmann, Frossie Economou, K.-T. L. F. M. S. R. P., & Wu, X. 2019, Science Platform Design

Hill, A., Flagey, N., McConnachie, A., & Szeto, K. 2018, MSE Science Book, `https://mse.cfht.hawaii.edu/misc-uploads/MSE_Project_Book_20181017.pdf`

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Jurić, M., Ciardi, D., & Dubois-Felsmann, G. 2017, LSST Science Platform Vision Document

Jurić, M., et al. 2018, LSST Data Products Definition Document

Narayan, G., Zaidi, T., Soraisam, M. D., et al. 2018, ApJS, 236, 9

Smith, K. W., Williams, R. D., Young, D. R., et al. 2019, Research Notes of the AAS, 3, 26