Tools and Services to Support LSST Science

Suvi Gezari (University of Maryland), Abi Saha (NOAO), M. Virginia McSwain (Lehigh University), Federica Bianco (New York University), Will Clarkson (University of Michigan, Dearborn), Melissa Graham (University of Washington), Henry Hsieh (Planetary Science Institute), Bryan Miller (Gemini Obs.), Dara Norman (NOAO), Travis Rector (University of Alaska), Steve Ridgway (NOAO), Marcelle Soares-Santos (Brandeis University), Lou Strolger (STScI) and Kathy Vivas (NOAO)

Following guidance from the NSF, NOAO is developing the LSST Community Science Center (LCSC) to provide support and infrastructure for community science with LSST. The LCSC Working Group (WG) was formed to study the goals, requirements, and aspirations for community science with LSST, and to recommend priorities for the essential functions of the LCSC. Additional information can be found in the October 2017 NOAO Newsletter (https://www.noao.edu/noao/noaonews/oct17/116news.pdf), 'Preparing for Community Science with LSST' on page 11.
This white paper outlines some of the tools and services that NOAO (NCOA) should continue to develop in order to support LSST science and other data intensive research.

NOAO users of LSST data will need computational tools to help access Level 1 (nightly), Level 2 (annual), and Level 3 (user-created) data products. While the LSST Project is developing their own Science Platform with Level 1 and 2 analysis tools and pipelines, there is still a good deal of uncertainty within the community about the nature of the final Science Platform products. Level 3 products will be public codes contributed by the user community to deal with the Level 1 and 2 products. The diversity of Level 3 code origins will require special effort to ensure broad community usefulness. There is opportunity for NOAO to collaborate with the LSST Project to deliver several useful tools and services.

NOAO Data Lab

The NOAO Data Lab, currently very far along in its development, is being designed as a data mining resource for various NOAO survey data. It provides a user-friendly interface for the scientists to build queries and analyze data products using the Jupyter Notebook format. It also runs in “the cloud” with free space on NOAO computers for users who have registered on the site. A limited number of example queries and basic analysis tutorials are already available to help users get started. Visualization tools for the available surveys are under development. The development of more documentation and tutorials would broaden the reach of the Data Lab.

The design development and funding for Data Lab will affect its ability to meet the community’s computing needs (including data storage and bandwidth). Once the LSST data become available, the size of the dataset may lead to an explosion of follow-up

questions that challenge the computational power of the Data Lab and overwhelm the available resources. A competitive proposal process might be necessary to temper the computational demands on the Data Lab. A discussion of the best strategies to maximize science with available resources is needed.

ANTARES

NOAO is also developing ANTARES as a broker service that will be able to classify new, stationary, transient sources almost in real time, using customized filters, and identify unidentified or rare types of objects in the LSST survey. It will then issue alerts to the community to enable rapid follow-up of transient sources. This will be an integral part of the transient source follow-up strategy.

One concern about ANTARES is how flexible it will be at providing or accommodating filters that match the community's needs. If NOAO (NCOA) resources are too limited to accommodate all community filter requests, an external allocation committee may be necessary to select a finite number of customized filters for specific science cases that will interact with the full data stream.

LSST User Support

LSST will revolutionize the way that science is done, and there is a great deal of uncertainty among the community about how to operate in this new paradigm. A robust platform of user support tools must be offered during the planning stages of LSST to help train the community, and the tools must be sustained and developed as needed once data become available.

With these requirements in mind, we recommend the following user support tools and services, with slightly declining but nearly equal priority:

- Level 1 and 2 data access to the NOAO community is essential. The NOAO Data Lab sets a clear path for this requirement, and we strongly endorse its continued development as a tool to access NOAO survey data, including future LSST data. We therefore recommend the establishment of a dedicated LCSC Data Portal to meet these needs, in a manner that complements but does not duplicate the LSST Science Platform that is under development. Furthermore, we recommend that NOAO continue developing more tutorials for the Data Lab to illustrate various science use cases to demonstrate the data mining potential and help educate the community on its use. Hosting the SV data at NOAO will be a tremendous opportunity for the US community.

- The ANTARES alert broker will enable an efficient strategy to follow-up new and interesting transient sources. We strongly recommend that NOAO continue to develop this critical tool.

- The NOAO can provide a curated hub for sharing Level 3 data products sourced from the community. Level 3 data products and pipelines will be the products of

LSST Science Collaborations and individual research groups who wish to make their analyses public. NOAO should encourage the sharing of these Level 3 products and archive them (within the NOAO Data Lab) in order to enhance the science output from the community and reduce redundancy. Guidelines for code sharing protocol and enforce rules for referencing the Level 3 products in new publications should be provided.