RESPONSE TO WHITE PAPER QUESTIONS

<u>Complete the below fields. The white paper will not be accepted without all fields completed.</u>

- Principal Author (*Doe, John J.*): Drory, Niv
- Today's Date *(mm/dd/yy)*: 9/30/14
- Institution of Principal Author: McDonald Observatory, The University of Texas at Austin
- Co-Authors (*Doe, Jane J., Institution, etc.*): Shetrone, Matt, McDonald Observatory; Gaffney, Niall, Texas Advanced Computing Center, UT Austin;
- E-mail Address: drory@astro.as.utexas.edu

- Paper Title (if applicable): Software and the US OIR System

WHITE PAPER QUESTIONS OIR_Study@nas.edu

The National Academies' OIR System Study Committee is charged with identifying the principal federal and non-federal capabilities in the U.S. OIR System and making strategic recommendations to optimize the System for the best science return. It is vital for the committee to receive community input, so we welcome brief White Papers on any related topics of interest. Questions that might be addressed include, **but are not limited to**, the following:

1. What O/IR capabilities are you using, are you planning to use, and will you need through the LSST era?

2. Do you have access to the O/IR capabilities you currently need to conduct your research, e.g. through a proposal process, through collaborators, or via data archives? If not, what is missing?

3. Comment on the need for the U.S. community's access to non−federal O/IR facilities up to 30 meters in size.

4. What could be done, outside of increased funding, that would enable the U.S. astronomical community to realize the goals of the decadal surveys at OIR wavelengths?

5. What is the role that a national observatory should have in an effective ground-based OIR system?

6. What are the U.S. long-term data management, archiving, and mining needs in ground-based O/IR astronomy, including those for LSST?

7. Given the increasing complexity of astronomical instrumentation, where should new major instruments be built (e.g. universities, national labs, collaborations)? How much instrument duplication is desirable or sustainable across different facilities of similar aperture?

8. How can the community ensure that future generations of astronomers have relevant instrumentation, observing, and software skills for the frontier science of tomorrow?

9. Comment on any needed evolution in human astronomical infrastructure, that is, the efficiency of sustaining instrumentation, data, or software teams in centers of excellence relative to assembling the needed skill sets from across the community.

10. What types of scientific and observing coordination among the various NSF telescopes (including Gemini and LSST) and non−federal facilities are the most important for making scientific progress in the next 10-15 years? How can such coordination best be facilitated?

In Response to Question 6 and 11, and on software development within the O/IR community at large:

The computer has become the most important tool for astronomers, apart from the telescope. In the age of large O/IR facilities, many young astronomers no longer use telescopes themselves, so the computer becomes the only tool they use to find, collect, and analyze observational data. Small PI instruments typically have pipelines written by one or very few talented grad students specifically to fulfill their science, and often thesis, needs. Large-facility instruments and data archives cannot be developed that way. They require sophisticated, optimal, and efficient large-scale data processing supporting a large variety of observing modes and science product outputs. While the community trains instrumentalists to design and build the opto-mechanical components for the current and future O/IR facilities, and institutions and university departments exist that specialize and excel in instrumentation, the community offers no formal training in software and we do not recognize specialization in that field. We know, however, of the critical necessity of astronomy-specific domain knowledge for the success of astronomical data processing software projects, and we know that people with only computer science background do not suffice to fulfill the growing software needs of the O/IR community.

The astronomical community has benefitted from governmental support for large observatories and their associated instrumentation but without support for large computing facilities and their associated software products, we are in danger of not taking maximum advantage of the LSST era. NOAO used to support centralized development of data analysis packages (IRAF). The development capability has been largely lost, despite support for IRAF being maintained at a minimal level. IRAF no longer meets the needs of many, despite still being very useful. No community effort to redefine the scope to accommodate todays image and (3D) spectral analysis, statistics, catalog manipulation, and big-data algorithms exists or has been proposed, despite many projects facing similar problems ranging from processing algorithms to storage and data retrieval. To fulfill future O/IR-system needs for facilities such as the LSST, the GMT, the TMT, and their instruments suites, we need a generation of astronomers with at least basic training in computer science and software engineering.

Development of such training could be modeled after the efforts in instrumentation. Many programs offer classes in basic optics and instrument design. In addition, they offer a more hands on experience, i.e. young astronomers are getting involved in instrumentation projects that often become part of their theses. The same should be applied in the field of astronomical software.

Facilities such as McDonald Observatory/UT Austin, which offer telescope access, strong instrumentation programs, and a long history of software development are ideal training environments for astronomers with emphasis on software. Besides the pipelines for instruments, current projects at McDonald include observing planning tools and the queue scheduling system for the upgraded HET, data analysis software for HETDEX, telescope and instrument control systems, and in collaboration with the Texas Advanced Computing Center (TACC), data storage and distribution systems. The TACC is a world leading center for computationally driven research. Its many world-class systems include two different systems in the top 10 International SuperComputing list over the past 6 years and several supporting systems specializing in data visualization and analysis, provides support for a wide array of data research who's research needs demand such powerful systems. TACC provides collaborative support and training to enable development and adoption of new techniques and technologies in a wide array of data research communities. Collaborations like the one between the TACC and the UT Astronomy Department need to be encouraged.

It is of great importance for the future of the US O/IR community to recognize the synergy of large observational projects, ambitious telescopes and instruments and, crucially, breakthrough data analysis software to maintain the leading role in observational O/IR astronomy. While specialization in software skills would be an excellent draw for a graduate program, the goal will be to ultimately combine the software, instrumentation, and telescope infrastructure. The University of Texas at Austin and McDonald Observatory hope to build upon this synergy between the software, instrument development, and telescope access. However, further benefit could be realized by bringing the large surveys efforts under a common supported architecture, e.g. a common set of interface definitions, headers, and metadata. The Virtual Observatory effort is a very good start, but should be extended to include more observatories, instruments, pipeline outputs, and include a wider part of the observational O/IR community. Similarly, integration of data archive search interfaces to a central hub would be useful to minimize duplication of data and effort in an era with budget pressure and danger of closure of smaller facilities. The TACC at The University of Texas at Austin would be one example of a facility, which could support the wider astronomical community in this regard if supported by traditional funding sources.

Lastly, questions relating to career paths for astronomers specialized in software development should be openly discussed in the community to ensure the retention of skills within the astronomical community and to provide the teachers that will train the next generation of software-savvy astronomers.