

Dr. Harry Teplitz (IRSA lead)
Dr. George Helou (Director of IPAC)
Infrared Processing and Analysis Center
Caltech
Pasadena CA 91125

Dear Harry and George,

This letter reports on the IRSA User Panel meeting held in Pasadena on October 28, 2016. The members of the Panel that were in attendance are Mike Person (MIT), David Tholen (University of Hawaii), Mark Lacy (NRAO), Klaus Pontoppidan (STScI), Alexandra Pope (UMass), Naveen Reddy (UCR). Olivier Dore (JPL) was unavailable. The full day meeting consisted of presentations from Harry Teplitz, Steven Groom, Vandana Desai, Walter Landry, Lisa Storie Lombardi & Luisa Rebull on the various activities of IRSA. The meeting was also attended by George Helou, director of IPAC. The role of the User Panel is to evaluate IRSA's content and services from the user's perspective, to advise on the priorities for ongoing/planned activities (in light of current funding level), and to suggest improvements or new services, not covered in the current work plan.

In this report, prepared by the Panel, we first overview the presentations delivered by members of the IRSA team and then make recommendations on the current IRSA activities and their priorities. Before we do that though, we would like to commend the IRSA team for their responsiveness to the recommendations of previous user Panels.

IRSA is NASA's primary archive for infrared and sub-mm astrophysics missions, containing 14 distinct ground-based, airborne, and space-based IR/sub-mm missions. IRSA provides access to more than 100 billion measurements of astronomical sources, including all-sky coverage in 20 bands, from 1.2 microns to 1 cm. IRSA implements the Spitzer Heritage Archive for the Spitzer Space Telescope, one of NASA's four great observatories, and it currently maintains much of the technical expertise from that mission. It is also the data archive for the WISE and NEOWISE surveys and the US archive for the Planck mission. Finally, IRSA hosts contributed data products, that can be searched in a way consistent with the primary IRSA data.

Overall, the User Panel believes that **IRSA is providing a crucial service to the astronomical community**. Its role in supporting astronomical research is bound to become more and more important with new NASA IR missions coming online. IRSA's impact is demonstrated by the constant increase in the number of refereed astrophysics journal articles that use IRSA holdings (60% growth since 2008), in the growing number of data queries (over 34 million in the first 8 months of 2016) and in the impressive volume (and growth rate) of its holding (currently amounting to more than 1 PB of data). About 40% of all approved NASA ADAP programs involve the analysis of IRSA data sets. IRSA is and will remain a crucial resource for new science and future mission planning. The IR/sub-mm wavelength regime allows addressing most of the scientific priorities in NASA's science plan, from the origin of the first stars and QSOs, to discovery of asteroids & NEOs. The IRSA archive will be the foundation of new strategic initiatives, providing targets for SOFIA, ALMA and JWST, complementing JWST, WFIRST and LSST science, and aiding in the observational planning and calibration of future missions.

IRSA is continuing to ingest new data from active NASA missions, as well as newly processed data from completed missions:

- *Spitzer* will continue to acquire data until early 2019; observations are already planned and DDT time can be awarded at this point. After the end of the mission, all data, documentation and analysis tools will be completely transferred to IRSA, initially to the Spitzer Heritage Archive. Given the complexity of the data and tools, the User Panel believes that IRSA should begin working on facilitating the transfer of the tools not already at IRSA well before the closeout.
- IRSA is preparing for the 2017 NEOWISE-reactivation mission data release, which will include two complete sky coverages, bringing the total to six coverages since the start of reactivation. The Panel felt that it would be important for the archive users to have the capability of co-adding multiple epochs in a simple way, and strongly support the IRSA's plan of making all the NEOWISE single exposures available to the existing WISE coadder tool.
- IRSA should continue to serve the Planck dataset to the astronomical community, and include the ESA's Planck Public Release 3, when available. The Panel also believes that IRSA should deliver the alternate processed dataset by JPL, should it become available, within the limited funding allocated by the Senior Review.

The User Panel believes that IRSA should work on maintaining strong expertise on Spitzer, WISE, Herschel and Planck, despite the likely reduction of available funding. The archival data from these missions will be mined by the community for many years to come.

The User Panel endorses IRSA's plan to implement a streamlined archive for the NASA Infra-Red Telescope Facility (IRTF) and to develop new features for solar system use cases. We recommend that the release of IRTF data be limited to raw data, and that IRSA seeks funding to enhance the archive to allow more sophisticated visualization and analysis within IRSA. Similarly, IRSA is getting ready to import SOFIA's archive, currently hosted at the SOFIA Science Center. The Panel believes that the addition of the SOFIA data benefit the visibility and usage of these data. Archiving the SOFIA data is important, and having them in a form that is useful for the users should also be a priority, as long as resources are not taken away from other high priority activities. We recommend that IRSA get sufficient funding to fully support archival research with these complex data.

IRSA should continue its work to standardize representations for data and metadata, which will optimize multi-mission searches and visibility of IRSA holdings to third party archives. The choice of Common Archive Observation Model (CAOM, already in use by e.g. MAST and CADC) will facilitate exchange of mission metadata, cross-mission data searches, and implementation of ObsCore VO standard.

Future developments planned at IRSA are:

- Data volumes and database tables are increasing exponentially. At IRSA, the "big data" problem specifically translates into a rapidly growing database content with associated increase in the query load (e.g. through VO protocols) and into an increasing need to

support online statistical analysis of database content. We support IRSA's plan to explore how analysis steps can be brought to the archive, without the need to download large amounts of data. The Panel liked the idea of exploring pilot projects that could be done with existing resources, and we are eager to participate. However, small projects feasible within the current budget will likely not solve the problem of how to optimize archival research in the new era of big data. This is a priority for archival research and IRSA should seek funding to support a bigger effort, commensurate to the problem at hand, that cannot be solved within the current funding situation.

- We believe that personal work spaces, where users could temporarily store the results of data queries and run simple data discovery tools, will be extremely useful, and the idea should be developed further.
- The User Panel endorses the effort to optimize interoperability between applications. A set of independent specialized tools for discovery, search, visualization, and analysis will allow a modular development of new projects and will provide advanced functionalities without a fixed workflow dictated by a single tool.
- We support the continued development of the data visualization capabilities of different datasets hosted within IRSA. Lower priority was given to the development of the spectral-cube visualization tool. The Panel was concerned that the development of such a tool could potentially result in a huge amount of work. Keep it small.
- Active communication with the community of users is important. We previously encouraged IRSA to engage in more outreach at science meetings and provide more tutorials and case studies to aid newer researchers in coming to the archive. Over the last year, they engaged in numerous simultaneous outreach efforts. They should continue in pursuing multiple channels (social media, webpage news, e-mails), in order to reach a broader community. We support the idea of a newsletter twice a year.
- We also encourage IRSA to continue with workshops IRSA has held at astronomical meetings, such as the AAS - DPS. We suggest that attendance can be improved by improving the communication with meeting organizers, in order to minimize conflicts with other meeting activities.
- We support IRSA's effort in the activities concerning the Virtual Observatory. In particular, the work performed to guarantee uniform interfaces, data registries and archive stability will allow the Virtual Observatory to easily access the IR datasets hosted at IRSA, guaranteeing the use by JWST tools and future IR surveys.

To conclude, we would like to reiterate the vital role that IRSA is playing in allowing original astronomical research and supporting upcoming NASA missions. We believe that the current workforce is working very effectively, but we are concerned that any additional requirements (IRTF, SOFIA, technology) be adequately funded to avoid exceeding available resources.

Sincerely yours,

Claudia Scarlata, on behalf of the IRSA User Panel members: Mike Person, David Tholen, Mark Lacy, Klaus Pontoppidan, Alexandra Pope, Naveen Reddy, Olivier Dore

1. To what extent are you aware of the Virtual Observatory (VO) interfaces to the services of IRSA, NED, MAST, HEASARC, and ADS?

The User Panel was split regarding this question. Some of us, having served in the User Panel for some time, are fairly well aware of on-going efforts to integrate the Virtual Observatory tools into the IRSA and the ADS, and have experimented with them on their own. The majority was not aware of these interfaces and did not know that tools they used were related to the VO.

2. Do VO activities further the science goals of IRSA?

The Panel felt strongly that yes, additional interfaces to archival data is absolutely necessary to preserving the utility of archives, and in the case of IRSA science is directly impacted by VO activities pointed at IRSA. Additionally, integrating the VO tools would help in greatly reducing the overhead involved during the discovery phase and the retrieving of the data.

3. Are you aware of using these interfaces directly or indirectly?

This question is very similar to the first one. Some of us, being in a privileged position of having been briefed on VO interfaces several times by IRSA staff, are aware of these interfaces, while others were not.

4. Are there additional priorities in providing machine based APIs to our archive services that we should be looking at?

The two major responsibilities of any archive are preservation and accessibility. Providing useful APIs to access the data stores is certainly a necessary step in maintaining the usefulness of the archives. The Panel believes that Integrating the API functionality into the day-to-day programming packages (e.g., AstroPy) is important. There are already on-going community-based efforts towards this goal, but it would be good to have at least one officially supported package with all the various API intractability included.