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 is the report of the IRSA User Panel meeting on December 3, 2012. The members who attended and contributed to this report were Mark Dickinson (NOAO), Aaron Evans (Virginia), Oliver P. Doré (JPL), Robert Jedike (Hawaii), Janice Lee (STScI), Matt Malkan (UCLA), Tom Megeath (Toledo) and Sachin Shenoy (ARC). Stellar Kafka (CIW) participated remotely. The full day meeting consisted of presentations from Harry, Steven Groom, Vandana Desai, Peter Capak, Trey Roby, Lee Armus and Lisa Storie Lombardi. The meeting was also attended by George Helou, director of IPAC. The various activities of IRSA were overviewed at this meeting as well as the results of the IRSA user survey.

IRSA is the primary archive for infrared and sub-mm data from NASA's astrophysics missions. It curates data from 12 separate ground-based, airborne, and space-based IR/sub-mm missions. With the upcoming release of the Planck database, IRSA will provide the astronomical community unprecedented access to the entire sky at 20 wavelength bands spanning 1 micron to 1 cm. IRSA also curates the Spitzer Heritage Archive for the Spitzer space telescope, one of NASA's four great observatories, and it maintains much of the technical expertise from that mission. The recent release of the Spitzer Enhanced Imaging Products, providing the community with supermosaics and point source catalogs of Spitzer data combined from multiple observations and programs, is a major achievement for IRSA.

The combination of IRSA's unmatched archive from NASA IR/sub-mm missions, its expertise with infrared data, and its all sky coverage across 13 octaves of wavelength make it a unique and powerful resource for the astronomical community. The data in IRSA impacts every area of astrophysics: exoplanets and planet formation, stellar astrophysics and studies of brown dwarfs, galactic and extragalactic star formation, and surveys of the high-Z universe. The data also provides a key resource for other NASA missions, such as Kepler and HST, and should play an essential role in optimizing returns from JWST, EUCLID and WFIRST. The data also supports observations with ALMA and other ground based observatories. The data archives curated by IRSA will not be surpassed for the foreseeable future: comparable data cannot be obtained from ground based telescopes and there are no plans for spaced based observatories with the wide field coverage, sensitivity and wavelength coverage achieved by Spitzer, WISE or Planck. Sustaining IRSA and maintaining its scientific and technical expertise is of key importance for the NASA astrophysics program and the astronomical community as a whole.

Over the last four years, IRSA has deftly met the challenge of balancing three major activities. The first is the archiving of the Spitzer, WISE and Planck data. This activity continues with new

data from the Spitzer Warm mission, the release of the Spitzer enhanced products, the release of the WISE 3 band data, and the upcoming releases of the Neo-WISE and Planck data. The IRSA archive has increased in size by more than an order of magnitude since 2008, to a current size of 500 TB, and another 30% increase is expected in the next two years. The second activity is the ongoing migration of the database to a new database management software to save money by taking advantage of a CALTECH site license. The final challenge is the development of a unified interface for finding, accessing and displaying data. IRSA has achieved this through the deployment of a software infrastructure flexible enough to easily accommodate future growth in both the archive size and the complexity of the data services.

The panel is highly impressed by the enormous progress that IRSA has made in all three of these activities, particularly given the very modest staffing at IRSA. Clearly, the continuation of these activities should remain the highest priority for IRSA. The importance of the archives maintained by IRSA is demonstrated by downloads rates which typically exceeded 10 TB per month in 2012.

In addition to continuing with the above three activities, the panel feels that it is important for IRSA to develop a strategy for maintaining its leadership in IR data archives and all sky databases. There are three approaches to maintaining this leadership. The first approach is to continue to actively seek out new missions and datasets for IRSA to archive. The second is to continue improving the data products offered by IRSA's existing holdings. The third is to improve tools for discovering and analyzing IRSA data, as well as data from surveys currently outside of IRSA such as the SDSS, with the goal of enabling new kinds of investigations.

The panel makes the following specific recommendations:

The continued improvement of tools to find, access, and examine data across multiple missions should remain a high priority. The IRSA data discovery tool could be expanded to interface with a wider number of surveys, including Spitzer data, WISE images and perhaps SDSS images. It may also be of value to combine the data discovery service with a *finder chart v2* like interface, so that the data can be examined before download. Selecting sources not only by their positions, but also by their fluxes and colors would also be useful. A particular challenge for the infrared data curated by IRSA is dealing with photometry of extended objects across many wavelength bands with different angular resolutions. Strategies for providing photometry in multiple, adjustable apertures should be considered. An SED plotter, which uses both photometric data and IRS spectra, may also be of great use for rapidly examining individual sources. More generally, IRSA should work to break down the barriers between images, spectra and SEDs, and plots of quantities derived from these data, so that the user can seamlessly exploit all three of these data modalities in exploring rich data sets. IRSA should also consider adapting software methods used in other popular data archives such as Topcat to avoid duplicating effort.

Continuing the development of capabilities for multi-epoch and moving target studies are of great value. IRSA already has developed a moving target tool for the WISE mission. It would be

of value to expand this ability to obtain multi-epochs using multiple missions. For example, the ability to track proper motions or variability between 2MASS, Spitzer and WISE could enable the studies of the proper motions and long term variability of cool stars and brown dwarfs.

Expanding the areal coverage of image products is also of importance, particularly since the boundaries of the data tiles can be a major inconvenience when dealing with IRSA surveys. The upcoming release of larger 2MASS mosaics will be a great benefit. Furthermore, the existing capability of the web image server on Montage to build larger mosaics of 2MASS, DSS, SDSS and WISE should be better advertised. The Montage capabilities are currently hard to discover (they would not be apparent to a casual user who doesn't read through the sidebar on the IRSA homepage). Perhaps, this capability could be integrated into IRSA's data discovery tool, giving the users the chance to generate a mosaic centered on the desired region (at a cost of a longer download time). IRSA's capabilities for examining, enhancing and serving out Planck data are of great importance and should widely advertised, as it is widely superior to what is available at ESA. If IRSA has extra resources, they could also try to implement some a celestial sphere visualization for Planck, and perhaps other all sky databases.

Continuing to cultivate a loyal user base should be a priority. Given the rapid pace of change within IRSA, it is key that IRSA aggressively advertise new capabilities and track usage. Although the new holdings such as WISE are bringing new users, many IRSA tools may remain underutilized, and IRSA needs to determine if underutilization is due to lack of awareness, lack of utility, or the presence of a better capability elsewhere. Many users may have bookmarked old tools, thus advertisements may be associated with older tools that they are replacing (for example, *finder chart2* advertised on the *finding chart* page). To assess each tool and IRSA as a whole, perhaps the users should be asked if a few tailored questions when they register, and when they use a specific tool or archive (instead of asking the user to fill out an entire survey). As recommended before, an opt-out email to IRSA Users that regularly informs them of new holdings and tools may also be valuable.

The panel also recommends that IRSA seek external funding for creating enhanced data product. Such external funding may be led by IRSA members, or done in collaboration with other teams. Not only would such funding produce new data products, but it may also help maintain the scientific vitality of IRSA by supporting postdocs or students. One worry is that seeking external funding could weaken the case for direct funding from NASA. However, proposals for external funding would tend to focus on a specific scientific goal. In contrast, direct funding from NASA may be best used for creating enhanced data products that have broad appeal and are not optimized for a specific scientific goal (such as the Spitzer Enhanced Image Products). We also continue to recommend that IRSA solicit the community for enhanced data products and adopt strategies for efficiently incorporating community data sets.

Public outreach is important. The committee is impressed by NITARP, an IPAC led program to engage high school teachers in research utilizing primarily IRSA data. We endorse IRSA's participation in this program as long as there is adequate funding.

With funding for the unified Virtual Astronomical Observatory coming to an end, IRSA has the opportunity to establish leadership in developing tools for virtual observing. Tools for advanced statistical exploration using the positions, sizes, colors, fluxes, etc in the archive would be of great value. Given the large size of the IRSA archives, and in particular the size of the WISE archive, providing public access to high power computers near the archive is certainly worth exploring for the long term. This could be similar to services provided to Herschel users by the NHSC, in which users can have access to powerful machines capable of operating efficiently on the entire database. In its long term planning, IRSA could explore charging a nominal amount for access to high power computing since users would save money that they would otherwise spend on purchasing their own computers. With cloud computing becoming commonplace, this novel approach may have broad appeal.

Sincerely,

Tom Megeath Chair