Report of the
Astrophysics Archives Senior Review
For the Astrophysics Division, Science Mission Directorate

17-19 May 2011

Senior Review Panelists
Andrew Connolly, Eileen Friel (Chair), Arne Henden, Gregory Laughlin, Pamela Marcum, Alice Monet, Donald Petravick, Christopher Reynolds, James Schombert, Thomas Vestrand

Introduction
Recent National Research Council reports ("*Portals to the Universe*", the Decadal Survey "*New Worlds, New Horizons*") stress the central role and the growing importance of data archives to astronomy today. Astrophysics data centers have moved beyond simple archives that served as the final deposits of the raw data collected by a mission to become data centers where the data are curated and high-level data products are distributed to the science community. During the last decade, NASA's Great Observatory missions have entered a new era of legacy datasets that are enabling innovative approaches to research and data-mining. Large scale sky surveys are becoming available in all regions of the electromagnetic spectrum and even more extensive ones are in the development stages. Leveraging NASA's astrophysics data centers, the NSF/NASA-funded Virtual Astronomical Observatory (VAO) is entering into its operational phase with the goal of enhancing public access to the vast array of both space-based and ground-based archived data.

In response to the 2007 NRC report on the NASA astrophysics data centers, NASA Science Mission Directorate (SMD) reassessed how the Astrophysics Archives were reviewed and initiated an Archives Senior Review process. Operating missions are reviewed in a separate Senior Review process on a two-year cycle. In addition, as part of the FY2009 President's Budget request, the funding for all of NASA's Astrophysics Archives and their associated data centers were integrated into a single budget line, and the Astrophysics Data Curation and ARchiving (ADCAR) program was introduced.

Purpose and Charter of the review
To maximize the scientific return from its missions, NASA regularly seeks expert evaluation from the scientific community. The Astrophysics Archives Senior Review, held every three years, conducts an independent, comparative evaluation of the activities of NASA's Astrophysics Archives and their associated data centers. The last Senior Review of the archives was carried out in 2008.

The Senior Review evaluates proposals from active archive centers for continued and augmented funding. Results of the Senior Review are used to define implementation strategies and give programmatic direction. The review evaluates the proposals in an absolute as well as a relative sense, expressing priorities to consider under circumstances of constrained funding.

NASA has directed the Senior Review to carry out the following:

1) In the context of the science goals, objectives, and research focus areas described in the Science Mission Directorate's Science and Strategic Plans, rank the scientific merit on a "science per dollar" basis – based upon the expected returns from the various archives and their associated data centers reviewed during the three-year period from FY12 through FY14.

2) Assess the cost efficiency, technology development and dissemination, data collection, and archiving and distribution, as secondary evaluation criteria, after science merit/ usefulness.

3) Based on (1) and (2), provide findings to assist with an implementation strategy for Astrophysics Division data curation and archiving for the period FY12 - FY14, including an appropriate mix of:

- *continuation of projects as currently baselined,*
- *continuation of projects with either enhancements or reductions to the current baseline;*
- *consolidation of projects and activities to enhance efficient management of limited budgetary resources.*

4) Make similar preliminary assessments for 2015 and 2016.

The final report is provided to Dr. Jon Morse, Director, Astrophysics Division, Science Mission Directorate, NASA Headquarters.

The Senior Review evaluated the scientific productivity, technical status, data dissemination, future plans and budget of the six active archive and data centers in the ADCAR program:

- The High Energy Astrophysics Science Archive Research Center (HEASARC) at NASA Goddard Space Flight Center

- The Multi-Mission Archive (MAST) at Space Telescope Science Institute

- The NASA Astrophysics Data System (ADS) at the Smithsonian Astrophysical Observatory

- The NASA/Infrared Processing and Analysis Center (IPAC) archives/centers :
  - The Extragalactic Database (NED)
  - The Infrared Science Archive (IRSA)
  - The NASA Star and Exoplanet Database (NStED)

Review Procedure

Each of the archive centers above was instructed by NASA to prepare proposals for continued and augmented funding for the period FY2012-2016 and given guidelines for proposal content and budget presentation. Each proposal described the centers' current status, including its holdings, services and tools provided, metrics on usage, scientific contributions, and relation to NASA strategic goals, objectives and research focus as articulated in the 2010 NASA Science

Plan. Proposals also presented descriptions of current projects and activities, as well as plans or possibilities for future development over the next 5 years. Budgets and FTE requirements were presented for both in-guide and over-guide budget requests.

The review was held 17-19 May 2011 in Arlington, VA. The review began with an overview by Dr. Jon Morse, Director of the Astrophysics Division, outlining the current Astrophysics Division portfolio and budget, and highlighting recent changes in the status of NASA missions relevant to the Astrophysics Archives activities. Dr. Thierry Lanz, Program Executive for the ADCAR Program, followed with an in-depth review of the charge to the Senior Review, a discussion of the ADCAR budget, a brief review of the highlights of the proposals under review, and some clarifying information on the ongoing joint NSF/NASA Virtual Astronomical Observatory activities of relevance to the Astrophysics Archives.

Each of the archival centers gave presentations to the Senior Review panel over the first two days of the meeting. Each center was represented by 3-4 people who met for a 90-minute period scheduled with the panel. Prepared presentations were given which addressed the highlights of each proposal, provided updates since the proposal submission when appropriate, and responded to several questions the Senior Review panel had requested of all centers following an initial organizational telecon two weeks prior to the face-to-face meeting. During this period, there was time for questions and discussion with the center personnel. The Senior Review panel wishes to thank all center staff for their cooperation both in preparing very informative presentations and for their responsiveness during the discussion.

Following each presentation, the Senior Review panel met in a brief executive session (including NASA personnel), to discuss the presentation and identify any further questions it wished to ask of the center personnel. In several cases, the panel asked to talk again briefly with center staff to clarify points or to ask additional questions.

Following the presentations from all centers on the second day of the meeting, the Senior Review panel discussed each of the proposals in turn, identifying both its strengths and weaknesses, and reviewing its over-guide budget requests. At the end of the day, a preliminary scoring was made. On the final day of the meeting, a final ranking was adopted and a set of recommendations was made, including priorities on over-guide requests.

The panel also discussed a number of issues that were applicable to all of the archive centers. Those of note are called out below.

Throughout the meeting, NASA officials were helpful in providing background information and guidance on the process of the review and were very responsive to questions from the panel.

Outcome of the Review

The final ranking of the six centers took into consideration all the criteria established in the charge to the Senior Review, in both an absolute and relative sense. The panel found the first three centers (ADS, HEASARC, MAST) to be at essentially the same, excellent level of performance and merit, so ranked them identically. These were very strong proposals that fully responded to the Senior Review Call, and contained no major weaknesses. The next two centers

(IRSA and NED) were performing very well but the panel identified some weaknesses in their proposals that led them to rank them slightly lower, but found little distinction between the two in overall level. Finally, the last of the centers, NStED, was found by the panel to be substantially weaker than the others; the proposal contained serious flaws. Our evaluations and recommendations for the individual centers follow in individual sections below, in the order of their appearance in the table below.

Archive center	Ranking
ADS	Excellent
HEASARC	Excellent
MAST	Excellent
IRSA	Very Good
NED	Very Good
NStED	Fair

The panel also discussed a number of overarching issues of relevance to all of the centers.

- The panel appreciates the established, although at times complex, processes by which decisions are made about the long-term disposition of archival data sets. However, the panel noted that these decisions have led at times to having archives located at one center, while expertise and/or programmatic focus resides at another. These situations, some of which are called out in the individual center evaluations, may present a confusing set of alternatives for the user, and more importantly, have the potential to limit the effective and optimal use of the archives. Recognizing that these institutional arrangements are often there for good reasons, the panel nevertheless urges NASA to keep foremost in their consideration the ultimate needs of the user in making these decisions in the future. Both NASA and the scientific community are best served by a system of archival holdings that actively facilitates the opportunity to do science and does not create impediments to the effective, creative, and maximal use of its data holdings. Identifying, clarifying, and advertising each center's specialization, and establishing guidelines in Calls for Proposals for new missions that recommends the proposer use the archival center appropriate to the nature of the proposed data, might be one strategy that would insure that a mission's data is archived where it will be optimally maintained and disseminated, rather than leaving the question of data archiving entirely up to the proposer of the new mission project.

- The archival centers currently rely on user interfaces that focus on individual, human interactions. We are, however, entering an era of machine-to-machine and robotic communications, which are necessitated by the vast amount of information in datasets and databases. These forms of communication come with an array of complex issues. The panel

recommends that all the centers carefully consider how they could benefit by taking advantage of these new modes of robotic communication and how to incorporate them in their future planning.

- The panel was pleased to see the attention all data centers gave to the adoption and use of Virtual Observatory (VO) protocols and standards, issues of interoperability and their modes of engagement with the VAO effort. It was clear to the panel that the NASA data centers, the VAO, and ultimately the scientific community, stand to benefit enormously by the full incorporation of standards and effective coordination of efforts. While the panel was glad to see evidence of improved mechanisms for communication between the NASA centers and the VAO, as the VAO establishes itself and responsibilities become better defined, the panel felt there was ample room for further development. The panel urges all the centers, the VAO, and NASA (and NSF) to work toward continued clarification and definition of the interfaces, relative roles, and responsibilities that aim at a full integration of efforts. These activities should be a high priority for the archival centers and the necessary resources from their core budgets should be directed toward VO-compatibility and interoperability.

- The panel saw some extremely effective instances of collaboration between individual archival centers, several of which are called out in the reports on the centers that follow. These examples demonstrate the importance of sharing expertise, experience, development efforts, and even established tools and services. These collaborations are cost-effective, efficient, and benefit the user in promoting familiar interfaces and capabilities across different archive centers. The panel felt, however, that this was an area that deserved continued attention and effort by both NASA HQ and the centers. Even while recognizing that there is, in certain circumstances, reason for multiple approaches to a given problem, the panel was concerned that there was still potential for unnecessary duplicative effort, and that opportunities for mutual benefit were being missed. The panel urges the centers to continue looking for methods to promote communication between centers, beyond the useful activities of the ADEC, and urges NASA HQ to provide the necessary oversight to ensure that the centers work effectively and collaboratively.

- Each of the archival centers demonstrated a strong connection to their user communities, with mechanisms that are tailored appropriately to their users and respective community needs. We saw evidence that the centers are responsive to their communities, rely on their advice and turn to them frequently for input both on existing and on potential future services. There is enthusiastic and widespread support for most of the center activities among the user communities, and increasing community reliance on the data and bibliographic resources they provide as judged by publication and citation rates based on archival data. The panel was also impressed with how center management handles the day-to-day and short-term balancing and prioritization needed to maintain their effectiveness, particularly under the sometimes uncertain and changing circumstances during mission lifetimes and within strict budgetary and resource constraints.

The panel noted, however, much less thought being given to the development of a vision and long term goals that would enable strategic planning within the decision-making process. We feel it is important for NASA and all the centers to undertake longer-term, strategic

planning that will provide a coordinated, coherent vision for the 10-15 year horizon. In spite of the challenges and uncertainties inherent in considerations of the future, including rapidly changing technologies and variable funding profiles, it is nevertheless important that the centers have a vision of their future and longer term goals that can help guide the critical decisions of where to place development efforts, make strategic resource allocation, and position themselves for the future. We thus recommend that NASA, in conjunction with the archival centers, and with input from the community, undertake long-term planning across the centers within the ADCAR portfolio. We note that this effort will have the benefit of facilitating and increasing coordination among the centers, addressing some of the concerns raised above.

Individual Reports for each Center

SAO/NASA Astrophysics Data System (ADS) (Rank: t-1)

The NASA Astrophysics Data System (ADS) provides free access to articles and preprints from an extensive range of astronomy and physics journals and other sources, and is publically available to a U.S. and international community which makes extensive use of this facility. The heart of this archive is a search engine that identifies publications on any astronomical subject and provides links to abstracts and articles, as well as related information such as number of citations for a particular article, and links to published datasets. The user interface provides for customization through the myADS portal, which can be set up to provide weekly notices on articles satisfying a set of individually programmed criteria.

Relevancy:
Given that the ADS focus is not one of archiving and distributing NASA mission data, this archive is in somewhat of a different category compared to the other NASA data archives that were also evaluated in this review. Nevertheless, the NASA relevancy of the ADS archive comes from its inherent ability to enable science that is closely aligned to NASA SMD's Science and Strategic plans. The ADS provides important information that contributes to all stages of scientific inquiry, beginning with the preparation of proposals that lead to data acquisition and/or analysis in support of NASA-related science investigations and ending with the final publication and dissemination of results.

Science strengths:
The ADS serves as an effective tool that enables science by streamlining the process of acquiring relevant information from the literature. The proposal presented several measures by which the ADS services have facilitated astronomical research and demonstrated that its impact and usage continues to grow.

The ADS team has done an exceptional job in responding to the criticisms of the previous Senior Review. A completely revamped user interface, with improved search capabilities, appears to have addressed one of the major concerns from the previous review regarding separating the "wheat from the chaff" in search results. The interface has been greatly simplified, such that both complicated or very simple search inquiries can be accommodated with a straightforward

and easy-to-use format within the new ADS Labs environment. The new name-resolver that enables robust distinction between authors having similar names is a significant improvement, as is the new support for author names in native languages. The plan to do full-text searches within all of the archival holdings is also a considerable enhancement.

The responsiveness of the ADS to the community is commendable: the ADS team has formulated their new design around inputs they have received through emailed criticisms and requests from the community, tracked metrics derived from actual community usage of the ADS website, and by maintaining a presence at AAS meetings. This new interface is to be launched at the summer 2011 AAS Meeting.

The team has recognized the need for long-term strategic planning, and is encouraged to initiate the development of a 10-year plan that addresses, among other factors, sustainability of in-house expertise on the staff and maintaining a competitive edge in a rapidly-evolving world of electronic information and new methods of media distribution.

Proposal weaknesses:
Some of the new tools being made available with the ADS upgrade, such as the graphic that illustrates hierarchical relationships between papers or authors, received mixed reactions from the review panel. The ADS team is encouraged to track usage of these additional tools and consider removing or revising those which do not appear to be contributing in a useful manner. For example, some panel members found the graphic showing hierarchical relationships between authors and journal articles to be quite busy, but recognizes that it holds the promise of conveying useful information if adequate filters or other ways to "drill down" into the graphic, to eliminate undesirable clutter, could be applied.

Overall assessment and recommendations:

The new ADS interface and functionality is a significant improvement to the old system, and should be able to meet the standards and expectations of the most web-savvy end-user for at least the next five years. The obviously strong relationship that the ADS team has established with the user community is commendable; the panel encourages the ADS to continue to allow user feedback to help shape and direct the ADS design in the future, particularly with regards to the new tools being made available in this Summer 2011 release. The panel recommends that NASA continue to fund the ADS at the full in-guide budget.

The High Energy Astrophysics Science Archive Research Center (HEASARC) (Rank: t-1)

HEASARC (High Energy Astrophysics Archive Research Center) is the original multi-mission science archive for NASA, and has developed into the dominant provider of archival data and data analysis tools for the high-energy astrophysics community. HEASARC continues to be extremely effective at supporting the high-energy astrophysics user community, and continues to work with missions currently in formulation to ensure a smooth data archiving and distribution pipeline.

LAMBDA (Legacy Archive for Microwave Background Data Analysis) serves as the archival data center for NASA's Cosmic Microwave Background (CMB) missions COBE and WMAP and is planning to ingest ground-based and sub-orbital CMB data. While it is apparent that the integration of LAMBDA and HEASARC has produced institutional/organizational efficiencies, the overall scientific merits of the merger remain unclear and should be assessed by the next Senior Review.

Science Strengths:
HEASARC remains a crucial service to the high-energy astrophysics community, and offers services/support that reach well beyond this core community. The standardization and continued development of data formats and analysis tools across missions is important both to the user community and to mission teams. In turn, HEASARC remains very responsive to the needs of its communities.

The previous Senior Review recommended that HEASARC step up its innovation efforts. HEASARC was very responsive to this recommendation and, in a very timely and useful update, is preparing to deploy its new interface (Xanim). The panel also applauds HEASARC's efforts to engage and, where appropriate, implement new technologies such as virtualization, cloud computing, and tablet/smart-phone access. Hera, HEASARC's remote analysis interface, is another example of innovation and has already proven to be of use to the education community; the panel agrees that Hera should be further developed, although we urge HEASARC to remain receptive to its user community in judging the level of effort expended on this project.

The panel appreciates that all of HEASARC's core activities and innovation plans have been included in the in-guide budget.

Science Weaknesses:
LAMBDA's role as the primary archive for CMB science is weakened by the exclusion of Planck data. LAMBDA should seek to collaborate with other CMB archives in order to best serve the CMB community. To effectively guide LAMBDA's mission given the different cultures of traditional high-energy astrophysics and CMB studies, HEASARC/LAMBDA should maintain a CMB-focused Users Group (or a CMB sub-group of the HEASARC Users Group).

Also, as a minor concern, the formulation and the goals of the GRB thematic database were not articulated clearly. Again, HEASARC should be guided by the user community when considering further developments of this database.

Relevancy strengths:
HEASARC serves a large number of past and current NASA-supported missions and already has an impressive set of future missions for which it will act as the data archive. Thus, HEASARC remains extremely relevant to NASA's goals.

Organizational strengths:

The management and organizational structure of HEASARC is clearly very effective at implementing its primary mission. They have close connections to the communities that they serve.

Overall assessment and recommendations:

HEASARC serves a valuable role in the portfolio of NASA data archives, curating a large number of past and current missions, and with plans to act as data archive for future missions. The panel recommends that NASA continue to support HEASARC at the full in-guide budget.

The panel reiterates the recommendation from the previous Senior Review and encourages HEASARC/LAMBDA to coordinate actively with the other NASA archives holding CMB data to prevent unnecessary duplication and to optimize science use.

Multi-mission Archive at the Space Telescope Science Institute (MAST) (Rank: t-1)

The Multi-mission Archive at the Space Telescope Science Institute (MAST) is the Optical/UV data center for NASA's space-based missions. Its current archive comprises 160TB of data from 15 separate missions, four of which (HST, Kepler, GALEX, and XMM-OM) are in operation today. The role of MAST is to curate and serve data from these missions, and to provide ancillary information and tools that will enhance the scientific impact of the NASA archive.

Strengths:
The data sets served by the MAST archive are highly relevant to NASA and its core mission. MAST's focus on NASA's optical and UV telescopes, ingesting ground-based data where it enhances the scientific utility of the space-based data, is entirely appropriate. MAST is a lean operation that delivers substantial scientific returns for a relatively small investment. The panel noted that the newly released Kepler data are of great importance to the community and we encourage the archive to continue to develop and support the functionality of these data. Where possible, MAST is encouraged to develop tools to assist users in making the best possible use of the data, especially those requiring time series analysis.

The interfaces to the MAST archives have undergone significant development since the last Senior Review with the deployment of the Hubble Legacy Archive viewer, the GALEX viewer and a number of enhancements to the visualization of catalog and imagery data. These improvements, including CasJobs for the GALEX catalogs, enable access to the data at many different levels of expertise (from browsing individual sources to sophisticated SQL-based queries that can join multiple data sets). The continued ingestion of all of the MAST data sets into CasJobs and the development of a common visual interface to these data using the HLA viewer was considered a great strength of this proposal.

The usage of MAST, the number of papers generated from MAST data, and the number of citations to these papers continues to grow. That there are now more papers produced from

archival searches than from non-archival data is indicative of the impact of MAST on the community. The archive has been extremely responsive to input from its user community, improving access to the data through an increase in network bandwidth and by developing a common interface to the MAST data.

The MAST developers have demonstrated good use of the institutional resources available to them. This includes: access to the HST instrument scientists, use of in-house IT support, access to database experts resident within STScI and at Johns Hopkins University, and collaborations with CHANDRA and HEASARC science and data and international archives in Canada and Europe. This sharing of resources and technologies has led to the development of common tools (e.g. the CHANDRA footprint service), increased functionality, and has lessened the impact of the closure of the ESO/ST-ECF on the HST community. Continuation of this collaborative approach is strongly encouraged.

MAST continues to implement and support Virtual Observatory protocols that will enable access to NASA data across multiple data centers. It is leading in the development of a VAO Portal that will unify the interfaces to all archived data. Cross-archive access will be critical to the success of the VAO as well and to enhancing the scientific utility of NASA's archives. The panel's view is that VO interfaces are a core component of the archive's mission.

Weaknesses:
The proposal did not address the question of how the institutional knowledge associated with the NASA missions that are, or will be, in close-out can be preserved by the MAST team. We encourage the team to develop support plans for data such as GALEX that incorporate preserving not just the bits but also the knowledge associated with these data (including how the data were taken and processed, and the properties of the instruments themselves). In particular, the team needs to consider long term planning for Kepler information and instrument expertise.

While the team demonstrated that they were very responsive to the community and to the needs of the missions they support (with a well defined series of short term goals), it was felt by the panel that the proposal would have been improved by a discussion of a long-term strategic plan for the archive including the long-term preservation and distribution of data and the expected evolution in technological resources.

Overall assessment and recommendations:

The panel recommends continued support for MAST through its in-guide budget. We consider MAST and its services an indispensable part of the NASA data archives. While we did not feel that MAST made a strong scientific case for its over-guide requests, the panel believes that the potential loss of the GALEX photon data is of sufficient importance that the request for over-guide funds to support preservation of the photon data was warranted. We do, however, consider that the over-guide request for support for high-resolution images to be of low priority.

NASA/IPAC Infrared Science Archive (Rank: t-2)

The NASA/IPAC Infrared Science Archive (IRSA) curates and serves scientific data products from the NASA infrared and sub-millimeter projects and missions. Along with many historical datasets such as IRAS and 2MASS, IRSA is now serving the Spitzer Heritage Archive and will serve the WISE and NASA Planck Archives as they become available during the review period.

Relevancy:
The infrared is a key wavelength regime for many current scientific investigations. Spitzer was the first instrument to detect the actual radiation from an exoplanet; WISE has detected over 150,000 asteroids and comets, including some 500 near Earth objects; Planck will continue our exploration of the Cosmic Microwave Background (CMB) and provide tests for dark energy models. These are all topics highlighted in the Decadal Survey and are major initiatives by NASA. IRSA provides access to the results from these missions, along with software tools to investigate the products and cross-correlate them with other wavelengths. The IRSA archives will be primary sources for generating ALMA and SOFIA targets.

Strengths:
IRSA continues to provide access to a wealth of data from past (IRAS, 2MASS, MSX, etc.) and existing missions (Spitzer, WISE) as well as planning the integration of new missions (Planck, Herschel, SOFIA public products). IRSA's co-location on the Caltech campus with IPAC facilities provides cross-fertilization between team members. Their user groups and accessibility at AAS meetings have given them many avenues for advice and support.

IRSA provides imaging, photometry, spectroscopy and some time-domain data as well as a suite of analysis tools, both imported from the original spacecraft teams and those they have written themselves. They now serve 20 all-sky IR surveys from 1 micron to 10 millimeters. They hope to provide the WISE point-source data for time domain studies. The panel encourages a strong emphasis on the time domain, as this is where much of the new science is headed.

IRSA is now the Spitzer cryogenic-mission data archive, with the major Spitzer Heritage Archive (SHA) incorporated in their holdings. The warm mission will continue for another year, with its data archived at IRSA. IRSA is the only route to any Spitzer data, as Spitzer has turned off their access tools. The panel was in unanimous agreement that proper support of the SHA should be a high priority of IRSA over the next funding cycle, including proper staffing to handle user requests through expertise and knowledge of the full Spitzer suite of instruments.

IRSA also holds the WISE Archive, currently containing the Preliminary Data Release, with final data transfer occurring during 2012. This is the largest IRSA dataset and will grow by a factor of two by mission end.

IRSA provides access to their data holdings through the usual web portals. The web portal was recently redesigned, adding a basic query interface that is a much nicer way for new users to access the site using a single-line "object" request. The original Gator interface can still be used if desired. Power users can use the CatQuery page, which includes a full description of their html "wget" Applications Programming Interface (API), which will return results in ASCII, VO

Table, or HTML format. SQL queries are also supported.

The Spitzer, WISE and Planck archives already use a new web interface that will give a unified user experience across their holdings. It includes such features as finding all images that contain a target or those images that contain the entire search region, and image previews before download.

IRSA has been very responsive to the recommendations of the 2008 Senior Review. They incorporated the Spitzer archive and expertise into their system, and expanded the suite of tools available. They are moving all of their holdings over to a unified data access method, as evidenced by the single portal for Spitzer and WISE data.

Weaknesses:
The current IRSA system and the in-guide request seem completely reasonable to the panel. Transitioning from Informix to Oracle (using CalTech's site license) is a good idea but with potential problems in understanding the consequences on the internal data integrity; it will also involve some effort in rewriting tools. The panel encourages IRSA to work more closely with the other NASA centers to plan strategically for consistent wide-wavelength access for the user community.

The very large over-guide request needs to be considered carefully. Integration of external value-added databases comes with a cost; perhaps some of this cost can be placed on the group providing the external database (e.g., if they want to be included, they have to format the database in a specific way to make it easier for IRSA to import). The panel was surprised that bulk download was not included in the in-guide request, since such user requests must happen often. Likewise, the in-guide request mentions VO-compliance, but the over-guide request also includes VAO enhanced program interfaces; given IRSA's key role in the VAO, the panel feels that these activities should be a high priority within the core budget. The panel was not convinced that the WISE over-guide requests to add extended structure maps are necessary, and questions whether the WISE source extractor does anything significantly different than other available software such as SExtractor. IRSA requests three FTE years for the conversion of this software, which seems excessive.

The over-guide requests significant increased staffing to cover the Spitzer, WISE, Planck, and Hershel missions. The panel realizes that proper instrument support enhances the scientific value of the datasets from these missions, but feels that the proposal did not provide sufficient justification for the 1.5 FTE/instrument staffing request. We would assume that each mission would provide the best quality final product for archiving, so an initial "help desk" might be necessary, but how long it is required and the scientific return by having expertise several years after a mission end is hard to quantify.

Overall assessment and recommendations:

IRSA curates and serves archives that play an important role in astrophysics and offer the potential for continued and expanded discovery. The panel fully supports IRSA's in-guide request to continue and enhance these products.

However, the panel found many portions of the large over-guide request to be less well justified. Of highest priority the panel found the request for exploration science enhanced products for the Spitzer warm mission and the WISE archive, the request for bulk download preparation, and the contributed science products and it supports funding augmentations to make these possible. The panel realizes that proper instrument support enhances the scientific value of IRSA's new and planned datasets, but feels that the proposal did not provide sufficient justification for the specific staffing request for science user support. The panel recommends that NASA examine these staffing and support requirements critically before providing the necessary funding. The panel did not find the requests compelling and does not recommend funding for the WISE data analysis tools, enhancing the Planck/CMB archive, and the enhanced extended structure maps.

NASA/IPAC Extragalactic Database (Rank: t-2)

NASA/IPAC Extragalactic Database (NED) is NASA's premier center for the service of extragalactic metadata, as a clearinghouse for imaging and spectrophotometric data. NED's dataset is multi-wavelength, providing a user not only with critical resources enabling science, but also the ability to do realtime pure science within their archive using online tools through NED's website. NED provides meta-data links to data and resources at the other NASA data centers as well as links to the literature through ADS; thus providing a comprehensive tool for the typical user to go from data analysis to publication.

Science and Technical Strengths:
NED has an extremely efficient operation, capable of ingesting a steady stream of literature data as well as coordinating large datasets and verifying complicated lists. Its holdings have increased thirty-fold since the last review with no increase in release time to the community.

NED has access to numerous science and technical resources by leveraging IPAC science and engineering staff as well as the research environment offered by Caltech/JPL. They have responded to previous concerns about cheaper storage and maintenance contracts with a series of upgrades and cost sharing with other NASA resources at IPAC.

NED operates in close consultation with its community and user group. The future plans for NED have not been defined in isolation; NED has excellent staff and receives superior advice from its users and the immediate Caltech/JPL community.

The time is approaching for NED to consider long-range plans to maintain staff and technical excellence in the face of possible future retirements. Institutional memory is core to NED's excellence; IPAC/JPL may not have resources in the future to maintain a seamless change in NED staff.

The panel singles out the Level 5 knowledge base as a unique resource for the community due to its special release permissions and overall knowledge content. Its access numbers have increased dramatically, and the panel encourages further development of this resource.

The panel agrees with NED's user group in that NED's interface is in need of an overhaul as it has become cumbersome in the current browser environment. However, the discussion of what improvements would be made with over-guide funding were less well defined than other portions of the proposal. The panel recommends NED work to better understand the style of its users in terms of number of queries (i.e., power users, mid-range users, single query users) and to use these statistics in formulating their short range plans for interface design and long range plans in terms of new complicated datasets. Given its importance for the user, the panel supports over-guide funding for interface improvements.

The Physics within Galaxies aspect of the over-guide is highly recommended as it is long overdue for NED to become involved in kinematic data and resolved components in an era of high resolution space imaging. However, this type of data is extremely complicated and must be organized in a coherent fashion to avoid unintended misuse by the community. Dealing with cubes of data is a new avenue for the NED technical team. The panel supports this project, but recommends that a more detailed plan for how to organize and display data supporting Physics Within Galaxies be completed before this task is undertaken.

Proposal Weaknesses:
NED's plan to absorb large future datasets was not well defined in the proposal. Other than a clear division into space-based versus ground-based datasets, there was little discussion on the science value compared to cost for datasets requested by the over-guide portion of the proposal. We note that this weakness was also noted in NED's previous Senior Review. We assumed the NED community was consulted on the priorities for each new dataset, but again the proposal was less than informative of this process.

The panel agrees that SED data is an exciting component to NED's meta-data/VO capacity. However, a non-trivial fraction of galaxies (but with high interest, i.e. AGN's) have variable luminosities. The temporal nature of such data limits the quality and extent of the new aspect to NED's holdings until the time domain problem is solved (this issue is particularly salient in the high-energy regime). Properly time-tagging such data would require a major effort by the NED staff. The panel placed other activities at a higher priority than this effort.

Data mining, statistical work and footprint information services proposed by the project are problematic. The core difficulty lies in the fact that this information can only be extracted by knowledge of the original footprint and quality of the data. Almost any problem imagined by the panel would be better served by attacking the data at the original center that holds the level 1 information. The argument for adding these new services to NED was not compelling.

VAO efforts were requested as over-guide. However, the panel believes that NED, as a mini-VO project, should take the lead in VAO interoperability as part of their core mission. Such prioritization would send a strong message to the community that VAO efforts are important and central to meta-data projects like NED. We understand that resources are limited, but defining VAO activities as outside NED's main function by over-guide funding seemed inappropriate.

Overall assessment and recommendations:

NED plays an important role for the extragalactic community. The panel was impressed by the efficiency and growth of NED since the last review. We fully support their in-guide activities as critical to NED functionality.

The panel recommends support for over-guide tasks focused on new interface design and incorporating kinematic and resolved components information. These tasks were given high priority by their user group and the community, and demonstrate that NED has future plans aligned with a solid broader vision. The panel does not recommend funding for other over-guide requests. NED needs to apply more effort in defining which future datasets to ingest, developing footprint information, and its VAO activities. These concerns are not due to limits to the current NED expertise, but rather a better roadmap for achieving its goals.

NASA/IPAC Star and Extrasolar Planet Database (NStED) (Rank: 3)

The detection and characterization of extrasolar planets is one of the fastest moving and most exciting fields in astronomy. With the early and immediate successes of the Kepler mission, and with the iconic post-discovery follow-up characterizations of exoplanets by HST and Spitzer, NASA now has its strongest claim to date to a clear and unambiguous global leadership role in this area. It therefore makes natural programmatic sense for NASA to serve and maintain the definitive database that covers both extrasolar planets and their host stars.

NStED has been tasked to assume this role, and during the past several years, has greatly increased its holdings of stellar time-series photometry from transit surveys, and has ingested a variety of exoplanet related data, including planet orbital and physical properties, as well as stellar and radial velocities and other host star properties.

The panel unanimously recognized, however, that despite the clear overall desirability of a unified and authoritative exoplanet database, the NStED project has been placed by circumstance in a very difficult, effectively untenable position. NASA's primary exoplanet-related data are already being hosted elsewhere: The HST data are served by MAST, Spitzer's legacy data sets are hosted at IRSA, and, most critically, the Kepler light curves (which are currently in great demand, both by the astronomical community and even by members of the public) are served by MAST. NStED, therefore, is not the go-to source for NASA's highest-profile exoplanet-related

data sets. This difficulty is greatly compounded by the cancellations of the SIM and TPF missions, for which NStED was slated to be the archival host site.

In an attempt to offset this structural disadvantage, the NStED team has made a strong and commendable effort to accumulate as much exoplanet-related data as possible. This endeavor is complicated, however, by an atmosphere in which the discovery of planets orbiting nearby stars remains largely a high-profile PI-driven effort. Outside of the NASA missions, many of the most productive exoplanet-related data sets (such as the California and Geneva groups' radial velocities) remain largely proprietary, and indeed, in most cases, the data that are available are of somewhat secondary importance. For example, while there is some value in having 20 million super-Wasp photometric light curves, these data are of somewhat limited utility for planet hunting without the supporting spectroscopic follow-up. The Kepler light curves, furthermore, are so accurate and numerous that they largely supersede ground-based light curves (and even to a significant extent the CoRoT light curves) for the purposes of study of stellar variability. The most uniquely valuable current holding of the NStED archive is likely the TrES photometry from the Kepler field, yet the archive provides no immediate or convenient method for cross-linking the TrES targets with their corresponding KIC identifiers.

An additional issue facing the team is competition from other sites. The Paris Observatory-hosted exoplanet encyclopedia has been in operation for over 15 years, and remains the default destination for many workers seeking physical and orbital data for planets in the current census. As another example, exoplanets.org skims the informational cream (and likely a significant fraction of the potential user base) from NStED while serving a full-featured, fully explorable catalog of planets.

In short, NStED faces a difficult climate for operating effectively, and the panel feels strongly that if the effort is to go forward, then NASA needs to supply a clearer mandate to the project. NStED needs to offer a compelling product that is in strong and demonstrable demand by the community, and which is not easily or readily obtainable elsewhere. Exoplanet characterization is likely to be a key scientific driver for the JWST mission, and so it is important that NASA shape the narrative and the scientific agenda for exoplanet-related research. NStED can play a vital role in this regard, but fundamental changes need to be made if the effort is to succeed.

Moving to the specific details of the proposal, a number of concerns were voiced by the panel:

(1) It was recognized that while the NStED is significantly younger than the other organizations under review, the usage and citation statistics for the past three years are quite low in comparison with the peer group, indicating that community support for NStED has been slow in materializing. The panel was puzzled by the apparent disconnect between the sparse usage statistics and the report that the "NStED Users' Panel has been extremely supportive and excited about the content and direction of NStED."

(2) There were significant concerns raised by panel members regarding the utility of some of NStED's core services. It was not clear, for example, whether injection of summary statistics into the Kepler light curve FITS files provides significant value. It appears that these statistics are calculated in the absence of trend filtering. The panel was concerned that photometric transients

and offsets associated with the peculiarities of the satellite's observing sequence (which are often very significant features in a given light curve) may be reflected in, and may perhaps even dominate, the summary statistics.

(3) The panel voiced significant reservations regarding the utility of the periodogram services, which featured strongly in both the proposal and in the presentation. There was concern that in the Lomb-Scargle implementation, aliases will generate very problematic ambiguities in the interpretation of periodogram results. A dramatic illustration of the problem is provided by the recent re-assignment of 55 Cnc's period from 2.7 to 0.7 days. Uncritical use of periodograms likely delayed the transit detection of this key planet by five years. With respect to the BLS service, the panel felt that application of this algorithm to the Kepler light curves prior to any trend filtering may generate potentially misleading conclusions, especially where very subtle planetary signals are concerned. The panel felt that the periodogram plots in the proposal were somewhat misleading. They depict eclipsing binary stars that present signals that are orders of magnitude larger than those presented by transiting or Doppler-detected planets.

(4) Regarding the proposed plans for utilizing over-guide funding, the panel felt that the plan to support direct imaging work was not well motivated. Specifically, it was not clear that an image service for published science-quality products would be of significant value to the community. Given the current scientific landscape, it is not clear that NStED will have the wherewithal to negotiate a value-added hosting service of what are likely to be highly proprietary multi-epoch direct imaging data.

Overall assessment and recommendations:

The panel recognized the leadership position NASA holds with respect to exoplanet discovery and characterization and the overall desirability of a unified and authoritative exoplanet database. NStED, however, faces a difficult climate for operating effectively, since it holds none of the primary exoplanet datasets, and competes with other well established and frequently used sites providing exoplanet data.

The panel feels strongly that if the NStED effort is to go forward, NASA needs to supply a clearer and achievable mandate to the project. To be successful, NStED needs to offer a compelling product that is in strong and demonstrable demand by the community, and which is not easily or readily obtainable elsewhere. It currently does not meet these expectations. NASA should examine its goals and strategies for serving the community in exoplanet research and if NStED is to continue to play a role, it should be provided with a clear mission and the environment and resources to achieve it.