

Princeton University

Department of Astrophysical Sciences  
Peyton Hall  
Princeton, New Jersey 08544-1001

Telephone: (609) 258-3808

Email: [strauss@astro.princeton.edu](mailto:strauss@astro.princeton.edu)

February 27, 2011

Dr. George Helou, IPAC Executive Director  
Dr. Joseph Mazzearella, NED Project Scientist  
and Task Lead  
Dr. Barry Madore, NED Senior Scientist  
Infrared Processing and Analysis Center  
Caltech/JPL, MS 100-22  
770 South Wilson Avenue  
Pasadena, California 91125

Dear George, Joe, and Barry,

This is the report of the NASA/IPAC Extragalactic Database (NED) User's Committee, following our face-to-face meeting on January 27-28, 2011 in Pasadena. We start by summarizing our main findings. Our most important conclusion is one that I believe you are quite aware of: as NED enters its third decade, the role it plays in world-wide extragalactic astronomy research and support of major NASA missions is more crucial than ever. We were tremendously impressed by the ever-growing number of web hits to the NED database, and the huge number of refereed papers that explicitly acknowledge NED. This is a direct demonstration that NED is enabling a huge range of important science for an enormous world-wide community of astronomers, and it is particularly impressive that this tremendous accomplishment has been, and continues to be, carried out by such a small team.

Our principal recommendations to NED are as follows:

- NED funding should be maintained or increased to allow at least the current complement of PhD-level scientists to carry out its twin core missions of archiving data from the extragalactic literature, and incorporating major surveys.
- We are concerned that NED's contributions to the VAO are larger than what VAO is paying for, and urge that the two be made commensurate.
- The NED website should be revamped, to clean out items that are rarely used, to make more prominent the most popular links, and to highlight new capabilities.

- The NED team should prepare a ‘best practices’ guide to authors, so that the papers they submit will use proper naming of astronomical objects and will have tables that NED will easily be able to incorporate.
- NED and VAO should work together with the AAS publication board to explore ways to archive the data files used in scientific publications.
- NED should include *all* objects from large surveys above some Galactic latitude cut, as well as galaxies and galaxy candidates at lower latitudes.
- The detailed criteria by which objects from large surveys are included in NED need to be properly documented.
- This committee should plan to meet more often, perhaps once a year, using web-conferencing tools to minimize travel.

Those are our highest priority recommendations. In what follows, we give the details, and add some lower-priority and more detailed recommendations.

NED finds itself with two closely related mandates that define its core mission: archiving the extragalactic data from the ever-growing refereed astronomical literature, and archiving the major extragalactic surveys that are driving so much of modern astronomical research. The number of papers published each year continues to grow, and the size of the surveys that NED will have to incorporate is growing even faster (SDSS now has photometry of several hundred million galaxies, while LSST will have roughly 10 billion galaxies in its final catalog). The work of incorporating this tremendous stream of data into NED, while impressively automated, still requires a great deal of expert human attention at the level of PhD astronomers. The NED model whereby top PhD astronomers make NED their career (e.g., Joe Mazzearella, Marion Schmitz, Harold Corwin) is to be commended and encouraged, and future hiring decisions should look for people of similar talents who are likely to stay with NED for the long term. Our first recommendation, therefore, is that **NED funding be maintained to allow at least the current complement of scientists to continue working on the project. It would be highly desirable to increase the number of PhD scientists working on NED by at least one person.**

The tools that NED has built to fulfill its mandates, from its name resolver to a plotter of spectral energy distributions, are of use beyond NED, and are being used in particular by the Virtual Astronomical Observatory (VAO). As such, NED is playing a key and increasing role in the VAO, a development that we applaud. We have felt over the years, and continue to believe, that NED in many ways is a realization of many of the capabilities that the Virtual Observatory is aiming to achieve. Now, with the VAO, that the US Virtual Observatory effort has the goal of becoming directly useful to astronomical research, we are delighted to see that many of the tools and to some extent, the design philosophy of NED are being incorporated into VAO. We were also very pleased to see that the VAO was taking on some responsibilities that NED previously had, in particular working with the journals about improving their archiving of data.

In this context, it is worth remembering that VAO and NED have quite different approaches to the data they serve, and NED is far more than just a server of content to VAO. In particular, NED provides vetting for all the data it incorporates. Hence, its need for PhD astronomer-level expertise is absolutely essential. Thus, NED has an irreplaceable role in its collection of science-grade data which goes well beyond the services it is supplying the VAO.

However, the work and services that NED is providing to VAO is causing an increasing technical and managerial burden on the NED personnel, and it appeared to us that the manpower that NED is providing the VAO exceeds what the VAO is paying for. Much of what VAO is asking from the NED team is not directly in support of NED's core mission of serving extragalactic data from the literature and from surveys. We therefore recommend that **NED carry out a full accounting of what the manpower it is currently providing to VAO, and request a commensurate monetary support from VAO.**

NED is also playing a key, and growing role in supporting a variety of NASA and ground-based missions which are taking up an ever-larger fraction of NED's bandwidth and computational resources. This is a good thing, of course, and is one of the (many!) justifications for NED's existence. NED's growth, including its computational resources, needs to account for this demand. In particular, it needs to be understood to what extent large missions or projects intend to make routinely heavy use of NED, and what paths are available to NED to support the implied increase in demand. It is not unreasonable to consider that such heavy use come with some support contribution. One solution is for NED to offer mirror sites to the heaviest users, shifting the very specific compute/IO/bandwidth demands off of the specific NED facility.

Our remaining suggestions are more specific and detailed issues related to NED functionality and user interface. We group them in several broad categories:

#### **The NED website and interface**

- The NED website has gotten quite complicated, with 39 separate links on its front page to various tools, paths into the NED database, and so on. We were struck by the statistic that 98% of NED users use the 'By Name' link in the upper left of the page. Many of the links are tools that make more sense to access via the 'by name' interface (e.g., the **Search Spectra** page) or are services that are done better by other websites (such as the 'Abstracts' search, which ADS does far better). We urge the NED team to redesign the whole website, removing or making less prominent those links on the front page that are rarely used, making more prominent the principal ways to access the NED database, and so on. It may be worth consulting with professional website designers for approaches to this. One of us (Tod Lauer) has sent you some quite detailed specific suggestions for the website; this letter is not the place to go into such detail. Of course, it would be wholly appropriate to include a link to 'classic NED' for those users who prefer the current interface (and it might be interesting to see how often people use that link!).

- A related issue is that with the current interface structure, it is difficult to highlight new NED capabilities and tools. There is a “What’s New” link on the current NED website, but it needs to be made more prominent. Similarly, the ‘send comment’ button should also be more prominent.
- Given the VAO’s high profile in the community, and the obvious questions about the relationship between NED and VAO, it would be very appropriate to include information on NED’s role in VAO on the NED front page.
- It is worth investigating enhanced methods to offer help to new NED users, such as videos (YouTube?) showing users how to perform both simple and sophisticated queries.
- In the spirit of increasing NED’s visibility to the world, it would be useful to have a journal article, perhaps to be published by PASP, describing NED’s capabilities. This article should be available as a link from the NED front page. There have been ADASS articles about NED in the past, but they have not been very high-profile in practice.
- Each of NASA’s data archives has its own home-grown interface, which needs to be learned separately by those using them. Is there any possibility of coordinating these interfaces between the archives, both to save development money, and to make them easier to navigate for those of us who use multiple archives? We have no specific suggestions along these lines, but we urge the NED team to bring up this possibility in discussions with other archives.

### **NED and the Journals**

- As we indicated above, we were pleased to hear that the VAO is working with the astronomical journals about publication models to improve the archiving of data: both making more uniform the tabulation of data and the naming of objects within journal articles, and allowing the journals to archive the data that lie behind the results presented in articles (e.g., the actual spectrum that might be shown in a specific figure). NED will want to be part of this process as well, of course, and we encourage continued dialogue between NED, VAO, and the the AAS Publications Board (and other representatives of the major journals) on this subject. This was in fact a major theme of the User’s Committee report from four years ago. Given the ever-increasing size of the astronomical literature, improvements in the way that data are represented in journals is key for NED’s (and the VAO’s) ability to keep up, without a steady (and presumably unsustainable) increase in the size of the NED team.

One specific short-term recommendation is that NED prepare a ‘best practices’ guide to authors of journals, describing such things as the proper naming of objects, the best way to create tables, examples of some of the more common types of errors and ambiguities that astronomers have in their papers, and so on. This should be done in coordination with the AAS Publications Board; ideally, these best practices would be included in the instructions to authors for each of the major journals.

- It has been suggested that NED lead the effort within the VAO to help journals actively archive the datasets behind papers published in those journals. This is an excellent idea, and very much is in support of NED’s core mission, but there is a danger of the task becoming too large for NED’s current team, and it may require additional resources. In addition, the problem of archiving data for publications goes well beyond NED’s extragalactic mandate; it should of course include data on stars, molecular clouds, and asteroids, not to mention theoretical “data” such as the results of computer simulations. So NED cannot be expected to solve the entire problem of archiving data for the journals single-handedly.

### Specific NED Tools and Holdings

- We were quite impressed by the ability to query on specific galaxy attributes (usually related to morphology); this is an important new capability within NED. However, the logic of the query was not always clear, and was made ambiguous by the large number of similar-sounding attributes that exist. In order to do useful science with these, scientists need to know *exactly* what the logic of a given query is. For example, when one asks for objects with the attribute “diffuse envelope”, will they also get the objects labeled “diffuse extended envelope”? Another confusing example was that a query that asked for both spirals and irregulars returned no objects (because the English ‘and’ really means the Boolean ‘OR’). The help available on the attributes page needs work to clarify all these things. A specific suggestion is that when a user enters a query into NED, the resulting SQL or Boolean that will be run on the database be displayed. Even better would be to give the user the ability to write their own SQL, and/or modify the SQL query that they are shown.
- Surveys are a key component of NED’s holdings. While NED is rightly focused on extragalactic objects, it is not always obvious which objects in a given survey are in fact galaxies. For example, many galaxies in 2MASS are unresolved, and thus were not included in NED when the 2MASS data were first loaded unless they happened to match a previously known galaxy. If one of these objects is later found to be a galaxy (for example, from its optical image), there is no mechanism to include the 2MASS match in NED. The approach of simply including *all* objects from these surveys into NED is not workable, given that in many surveys, such as 2MASS, the vast majority of sources are stars at low Galactic latitudes. With this in mind, we recommend that NED should include:
  - *All* objects from surveys (perhaps after some basic quality cuts) above a certain Galactic latitude (which may depend on Galactic longitude, e.g., to match stellar density contours within the Milky Way), and
  - *All galaxies* at lower latitude, where galaxies are identified as such by extent, colors, or matches to existing NED objects. The details of this would of course need to be determined for each individual survey.

- In that spirit, it is important for NED to document in some detail what its holdings of each survey are. That is, there are fairly involved criteria (and the above suggests further criteria in addition) by which objects from a given survey are included in NED: color cuts, morphology cuts, quality flag cuts, and so on. For astronomers to get full use out of the NED listings, they need to understand what these cuts are, and what the relevant attributes of each survey are, including caveats on the quality of its photometry (for example, the fact that photometry of galaxies of large ( $> 2'$ ) diameter tends to be biased low in SDSS because of sky subtraction problems). We therefore recommend that NED provide detailed documentation for each survey. Links from object photometry to this documentation would also be very useful. A detailed documentation of how the cross-matching between surveys was done would also be useful (perhaps in the PASP article we recommended above). Finally, it would be very useful to be able to determine whether any given object falls within the footprint of a given survey (which is needed to interpret how the absence, e.g., of an IRAS flux for that object can be interpreted).
- On a related note, it would be very useful for NED to be able to characterize the completeness of its holdings overall as a function of redshift, position on the sky, and luminosity. If users could determine what fraction of *all* galaxies above some luminosity (in some band) are included in NED in a given area of sky and redshift range, they could use NED for comprehensive statistical studies of all sorts. Determining this completeness function would be a major research project, but has the potential to greatly increase NED's usefulness. We urge the NED team to explore this possibility, and determine what resources would be needed to carry it out.
- For well-known galaxies, NED's list of relevant papers can be enormously long (for example, NGC 1068 has 1794 papers listed), making it difficult to find the one paper one might need. The ADS has powerful literature search tools to allow one to find, e.g., papers by specific authors, or with specific phrases in their abstract. It would therefore be very useful to be able to send a NED-generated list of papers to ADS to do further queries on it to find papers of interest. There are no doubt additional synergies that can be developed between NED and ADS; the goal is to offer the user the best of both sites.
- There is a great deal of lore about individual galaxies, some scattered in the literature, much in various expert's heads. One way to capture that lore would be to have a wikipedia-like system, with a page for each galaxy, where individuals could write their thoughts on objects they are familiar with. We haven't thought through the details of this (how would facts be referenced and supported? How would the lists be moderated?), but we urge that the NED team explore this possibility.
- NED tools are often used to plan observing runs. One useful additional tool, which should not be difficult to implement, is the ability to make an airmass chart (i.e., airmass as a function of time) for a given object, given the observatory coordinates and the time of year.

- The NED batch mode is extremely useful, but is currently limited to only 500 objects at a time. This seems like an artificial limitation, and it would be great to increase this limit considerably.

### **NED User's Committee, NED User's Community**

- This was the first meeting of the NED User's Committee in four years. This is clearly too long: we would like to be in the position to give input into NED more often. This would allow us to spend less time on introductory/background material at this meeting, and rather focus on those aspects of NED that are new, and where our input is most needed. We suggest that we meet roughly once a year, perhaps using web-conferencing tools so that we need not all travel to Pasadena.
- Too much of the meeting was in "receive" mode, i.e., with the committee sitting quietly through powerpoint presentations. We would have liked more opportunity and time for open discussion. The presentations brought up a large number of issues, but it was not always clear where our feedback was most needed.
- The last survey taken of NED users was in 2008. It may be time to consider doing this again.

I hope you find these suggestions useful. As you know, we're all big fans of NED, and we're happy to be able to help by serving on the User's Committee. The job you're doing is becoming ever more daunting with the exponential increase in the size of the literature and surveys, and we continue to be amazed that you are staying on top of it all.

Yours Sincerely,

Michael Strauss, for the User's Committee

NED User's Committee Membership:

Dr. Daniela Calzetti, University of Massachusetts

Dr. James Condon, National Radio Astronomy Observatory

Dr. David Hogg, New York University

Dr. Tod Lauer, National Optical Astronomy Observatories

Dr. Janice Lee, Observatories of the Carnegie Institution of Washington

Dr. Andrea Prestwich, Center for Astrophysics, Harvard University

Dr. Anna Sajina, Tufts University

Dr. David Sanders, University of Hawaii

Dr. Michael Strauss, Princeton University (Chair)