

Princeton University

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

Telephone: (609) 258-3808

Email: strauss@astro.princeton.edu

January 15, 2014

Dr. Joseph Mazzearella, NED Project Scientist
Infrared Processing and Analysis Center
Caltech/JPL, MS 100-22
770 South Wilson Avenue
Pasadena, California 91125

Dear Joe and members of the NED team,

This is a report of the face-to-face meeting of the NED User's Committee we held in Pasadena on December 18, 2013. Here we respond to the questions in our charge, with specific emphasis on the new web interface you have developed, as well as directions that NED is taking in the future.

Let us start by reiterating a familiar refrain: we continue to be enormously impressed by what the NED team has accomplished with quite limited funds. NED has been a core resource for the extragalactic astronomy community for two decades, and has grown and adapted tremendously with the growth of the literature, the advent of enormous surveys, and the changing needs of the community. We are proud to continue our association with NED, and are happy to help make it even better with our advice. We are particularly heartened to see that you are taking the recommendations we made at our last face-to-face meeting to heart, and are steering NED in the directions we suggested. Most of our notes here will be suggestions for further improvement, but let us lead off by stating that we are very enthusiastic indeed about the new Galaxy Environment tool.

The User Interface

The most important topic we discussed at our meeting was the improved user interface for the NED website, building on suggestions that we gave in our face-to-face meeting two years ago. We were in general quite pleased with the new interface, and your vision for its further development. We are concerned, however, by the mixture of the new look on the front page, and the old look one level down. We outline several suggestions here, and give detailed comments in the Appendix to this report. In particular:

- We are concerned about the mixture of old and new on the webpages; the front page looks completely different, but one drills down to familiar pages, and the user is confused about what has and hasn't changed. The words you've added to the front page do help with this.

You should consider making the look and feel of each page on the NED site similar. In particular, it will be valuable to have the top set of pull-down menus available on every page that one navigates to, so one doesn't have to return to the home page to go anywhere else on the site. That feature is currently available at the "With Unprocessed Catalog Sources" page, and no other.

- In the same spirit, it is important to explain that the interface will continue to evolve, and will not remain static. Users will understand this and be comfortable with it, if you state it clearly on the website. This is all about managing the expectations of the users, who will get upset with change that is not explained to them.
- There are a lot of changes under the hood, some of which the user will need to know about. The most important one of these is the distinction between "Sources" and "Objects", a subtle distinction that is difficult to parse. Indeed, we remain concerned that at the moment, the only information available for a source is a position. That is, all the user learns is that, say, a GALEX source is found at this position, with no further information for it (or even how to *get* more information, e.g., via a link to the GALEX archive). This seems to be of limited utility.

In any case, we're happy to see that you've clarified the 'search by position' and 'search by position with sources' distinction in the pull-down menus. Maybe the two should be integrated to one search, in which the default is not to show Sources, but to allow the user to click a box to include them.

- In many cases there are multiple paths that serve only to provide the user with subtly different ways to get to the same end point. Some of these approaches may be redundant ("Coordinate calculator" and "Extinction calculator" both go to the same website!) or obsolete. These different options add clutter, confusion, and developmental complexity for only very small utility gains for a small set of users. We strongly recommend deprecating obscure or seldom used entry points. Quoting from our previous report: "As a general remark, we encourage the NED team to avoid the temptation to add features to the high-level interface that may have minimal or obscure use. Simplicity is a strong goal." We make some specific recommendations in the appendix of this report (see also Tod Lauer's "informal polemic" on the subject dated 1/27/11).

In deciding how to streamline the possibilities, it will be useful to think through a series of "use cases" of the website. Imagine the novice NED user who wants, e.g., to get photometry for their favorite galaxy, or redshifts for a sample of 10,000 galaxies, or the nearest group from a list of 500 positions. In each case, how exactly should they dive into the website to most efficiently get the information they need? Thinking this through (and writing it down) will help decide the best way to streamline the links. As the NUC, we are happy to help you produce a list of such use cases. This effort may well merge with your idea of including on-line video tutorials about how to use NED in various contexts.

- We are enthusiastic about the idea of a “smart box” on the home page, which could accept either positions and object names. Once this is available throughout the website, it may be possible to streamline further. In this same theme, boxes that accept coordinates should take both RA and Dec in the same box, which will make it much easier to upload lists (this ability is currently there in some of the windows, but not all).
- It is unclear to the casual user what the full NED holdings are. A list of the major (> 10,000 objects?) surveys included in NED, together with their basic relevant attributes (see below), would be very useful. It is also worth listing those major catalogs that NED does not (yet) include. In addition, you should make sure that the user knows that the spirit of NED has broadened to cover all objects (even those unresolved!) at high Galactic latitudes.
- NED needs to clearly draw the distinction on its website between search capabilities and a description of their holdings; Hyperleda and IRSA do this in a more intuitive way than does NED. For example, we were unable to discover in a few clicks how current NED’s literature compilation is.

We are writing these words during the week of the AAS meeting, and we see that <http://ned.ipac.caltech.edu> still defaults to the old pages. This is fine; we think that with a bit more work, the new pages will be far more user-friendly and less confusing with some of the changes we suggest here and below. We understand that the implementation of our suggestions will take time, and that this is an evolutionary process. Indeed, we were gratified to see that you had anticipated many of our recommendations in your outline of tasks over the next year or two, and would urge you to put the continued redesign of the look and feel of the site at high priority.

Incorporating New Catalogs

NED has correctly recognized that Very Large Catalogs are a key component of its future. We support your further automating of the process of merging large catalogs into the NED holdings; the old model whereby all discrepancies in overlapping objects were fixed by hand was becoming increasingly untenable. The user of these catalogs will want to know the details of this merging; we urge you to document this fully.

In the same spirit, each large catalog has its individual quirks that you need to take into account; it is very important to document what you learn for each catalog and make that information available. A given catalog may have multiple measurements of photometry for each object (for example, the CFHTLS has roughly 30 different magnitude estimates for each object); a clear statement of which one(s) were used in NED, together with a description of the logic behind this, is key.

The documentation also should include information on the boundaries, footprint, and angular selection function of each catalog; the user will need to know if the fact that there is no detection at a given position in a given catalog is due to the absence of a source there, or the fact that that survey does not cover that region of sky. Incorporating the footprint information in detail will be a big task (perhaps justifiable as an overguide request in the future), and may require a separate data product, but we could imagine coverage maps of major surveys centered on any user-supplied or object location. Vizier already has a crude version of this.

Your plans for the next large catalogs you are going to incorporate make sense. We will of course be happy to work with you to share any relevant expertise that we might have. A few comments on specific catalogs:

- SDSS: SDSS-III will go public with its last data release (DR12) in December 2014. The imaging data in SDSS are now complete and all public, and DR12 will include roughly 2 million redshifts that are not yet in NED. The two Michaels (Strauss and Blanton) will be happy to help with the specifics, as well as to tell you about SDSS-IV, which will contain even more spectra.
- There are major new catalogs of stars in nearby galaxies that could go into NED. The PHAT survey of M31, in particular, will catalog several hundred million stars, and ground-based surveys (PANDAS, for example) extend over enormous solid angles. There are similarly major catalogs of stars in the SMC and LMC. The NED team will have to decide whether they wish to incorporate these catalogs.
- HST is spending an increasing fraction of its time on extensive high-latitude surveys. The catalogs from the CANDELS and CLASH surveys (both of which will have extensive ancillary data in other wavebands) are not yet public, but should be incorporated into NED when they are.
- Would it make sense to start conversations with the principals of some of the major surveys in the future (Gaia, LSST, etc) to have a clear idea of what is coming? It may be in some cases that you can suggest simple modifications of their database schema that will considerably ease your eventual incorporation of their catalogs.

NED's Future Plans

NED is being pulled in a variety of different directions, and one thing that is worthy of further discussion is the “big picture” vision of NED’s goals to prioritize the next steps. Is NED building a general purpose toolkit for research? Or is it providing a sensible matching among multiple data sources with astronomically well-motivated choices for basic parameters? The latter is what NED is very good at; the former is quite a bit more challenging, and taking it on will require a mandate from the NASA Senior Review. But deciding on this will help clarify the next steps for NED.

We were impressed by the list you presented to us of future planned improvements for NED. We won't discuss them all here, but here make a few comments on some of the specific items you gave us.

- We are very enthusiastic about the planned ability to upload long tables to NED in large batch queries. Similarly, it will be important to download any table that one sees on the screen into an easily parsable machine-readable form, in a format and with columns that the user specifies.
- We urge you to continue exploring the possibility of giving users the ability to do more sophisticated and complex queries from the NED database than is possible with GUI interfaces. We understand that it is not as simple as allowing people to write SQL queries, but people often want to explore data in complex ways. This will become increasingly important if NED continues to grow as a platform to do scientific analyses across large datasets.
- We discussed some high-level desiderata in considering what new capabilities and data to incorporate in future releases. An increasingly major theme of large surveys (PTF, DES, Pan-STARRS, LSST) is the time domain. Is this an area where NED can contribute? The survey teams themselves are developing tools to exploit the time domain. NED does not need to duplicate that effort, but rather should look to its principal strength, i.e., bringing together from multiple surveys: does it make sense to combine, e.g., data from SDSS, Pan-STARRS and eventually LSST to look for AGN variability on 30-year timescales?
- In the spirit of combining data from multiple surveys, SEDs and colors across broad wavelength ranges are something that (especially with the demise of the VAO) NED is uniquely positioned to make available. Doing so effectively will require significant care; because of the need to do consistent aperture photometry between surveys, the problem is trickier than simply matching on position. It is also important that whenever possible, each photometry point be flagged with the epoch at which the observations were done.
- We are supportive of the idea that NED become increasingly a facility for statistical analyses of large datasets. The footprint information mentioned above will be a key component of this. However, a guiding principle should be that NED should not attempt to replicate the facilities of individual surveys; for example, the LSST team will always have better control and knowledge of the relevant details of their survey alone. Rather, NED should look to deliver capabilities that draw upon analyses of multiple surveys at once. In this context, NED is increasingly growing into the role that the Virtual Observatory once saw for itself, and is doing so in an organic and natural way. This is a good thing; NED should make sure to limit its ambitions to what makes most scientific and technical sense for NED alone.

- NED has built up a variety of tools over the years, and some of them feel somewhat overdone. For example, for every object, there are eight different velocities and six distances determined from the redshift, using various peculiar velocity corrections. Some of these are more relevant than others, and it would be appropriate for NED to highlight the preferred one in each case (for example, at high redshift, the only one which makes sense is that in the rest frame of the CMB, while one would make a very different choice for a galaxy at 3000 km/s).
- The SED plots given for a galaxy often include quite different quantities, e.g., nuclear vs. isophotal vs. total magnitudes, which really cannot be compared directly. The distinction between them needs to be made clear (e.g., with different colors), and the user should be given the option to, e.g., just plot the total magnitudes.
- We've discussed in the past the idea of giving users the option of creating personal accounts on NED, e.g., to create a NED-style wikipedia on galaxies. Another idea is to imagine a citizen scientist project in which we ask members of the broader community to do some of the time-consuming step of going through the literature on specific objects to sort out errors in NED. Exactly how this would be implemented in practice remains a bit unclear to us, but these ideas are worth exploring.

Miscellaneous

We had only a brief time to discuss the User Survey. We are supportive of your plan to send it out. As we discussed, we recommend asking users to prioritize a list of the various initiatives you have in mind, rather than simply asking for each one separately whether they would like this capability. We also recommend that you ask each user what their level of NED expertise is: are they a power user or one who uses it only casually. One question of yours is about "data mining"; this term is rather vague, and should be defined. Finally, you might consider some open-ended questions about whether people would like NED to include more general purpose tools for research.

We suggest that we meet again on the phone in roughly six months' time, after further progress has been made on the user interface. We would love to have the opportunity to try out the smart box that you are planning, and to be given enough time before the meeting to explore the web site's capabilities extensively to give you more detailed feedback. We would of course be happy to respond to you on shorter timescales as well: feel free to send us e-mail saying, "Hey, we are working on this thing this week; take a look by Friday if you can, and send us your thoughts."

The “Best Practices” document looks very good. This should be distributed to the community, both by linking prominently off the NED page (and perhaps the webpages of various journals) and by posting to astro-ph. Our strongest recommendation is that in the introductory paragraph, you strengthen the statement that following these practices is in the scientist’s, not just NED’s, best interests. There apparently is a way to mark object names in a LaTeX file, the moral equivalent of `\cite` to make references, but the BP document does not seem to make reference to this, and the community is not widely aware of it. We are eager to help you work with the AAS Publication Board to make following these best practices formally part of the editing and refereeing process of the journals.

We look forward to the next face-to-face meeting, perhaps a year from now (as the Senior Review approaches!). It may make sense to hold the meeting over 1.5 days, to allow somewhat more leisure discussion. And we are available to meet by phone before then to talk about specific issues, ideally with the opportunity to look at presentations or check out websites well in advance. Congratulations on your accomplishments thus far; we are always eager to help in any way that we can.

NED User’s Committee:

Dr. Michael Blanton, New York University
Dr. James Condon, National Radio Astronomy Observatory
Dr. Tod Lauer, National Optical Astronomy Observatories
Dr. Janice Lee, Space Telescope Science Institute
Dr. Andrea Prestwich, Center for Astrophysics, Harvard University
Dr. Anna Sajina, Tufts University
Dr. Ohad Shemmer, University of North Texas
Dr. Mark Seibert, Carnegie Observatories
Dr. Michael Strauss, Princeton University (Chair)

Appendix: Detailed Comments and Suggestions

In this section, we include detailed comments and suggestions for the NED interface and holdings from various of the NUC members that we didn't have time to vet as a group. They are a bit haphazard, and may occasionally contradict one another or represent our own misunderstandings. We would also recommend re-reading our report of June 6, 2012, after we first discussed your ideas for the new user interface (we can send you a copy if you would like). There you will see some of these same themes reflected.

The User Interface

- Don't make the classic panel the first option in the pull-down menu on the front page. Also, add a note to this classic panel page explaining that while it has all the buttons of the old page, it does not have the same look; otherwise, this will throw a lot of people off. Indeed, the classic panel has a somewhat unfinished and unprofessional look, and could use some work.
- The default number of objects per page in NED outputs should be a user-defined variable, with a default that is significantly larger than it currently is.
- It is not clear immediately to the user that positions can be entered either in sexagesimal format, or in decimal degrees. This should be clarified. In this same context, it would be useful to make outputs available in both formats.
- Under 'Search Objects', the distinction between searches "by name" and "by IAU name" is unclear. It is not all clear why the distinction is important! This is an example where some streamlining would really help.
- Might the "By Classification" search be incorporated among the choices in the "By Parameters" search?
- We continue to believe that refcode searches are rarely used (we made this comment in our previous report), and could be integrated into a link to the ADS.
- The 'Source Nomenclature' page is unclear; some documentation is needed.
- When doing 'by parameter' searches, there should be a single fillable box when using, say, 'brighter than', which switches to two boxes when using 'between'.
- The menu item called 'Information' is pretty vague. Perhaps it should be called 'Documentation'? Lots of valuable things are hidden there, and the user will not necessarily know to look there. And the mix of material between the 'Information' and 'Tools' menus seems rather arbitrary, with extinction calculators in both places.
- Is there any web interface whereby one can enter a batch job? As it is, it can only be done via e-mail.
- Some of us liked the look and feel of the new front page, with its white background; others of us felt it was a bit dull.
- When you mouse over an item on the pulldown menu, it turns dark blue on a black background. That can be difficult to see.

- When you click on the main menus (rather than just doing a simple mouseover) you are taken to a help page. We found ourselves doing this accidentally many times and getting frustrated. This also means that it is not very “tablet-friendly” where the action is always touch (click), not hover.
- It would be nice to freeze a pull-down menu when you click on it; it would make navigation easier.
- We continue to want the ability to feed the outputs of NED (especially the paper lists) into ADS.
- A useful reference for managing expectations in web design is <http://www.joelonsoftware.com/articles/fog0000000356.html>. Another useful website as you think about refining the website is <http://www.usability.gov/what-and-why/usability-evaluation.html>.

Outputs

- It would be useful to get outputs in the form of FITS binary tables.
- Consider including a positional uncertainty in the output of ‘near position’ searches at <https://ned.ipac.caltech.edu/ngi/?q=nearposn>.
- The user would like to specify exactly which columns of output they receive. The “table-build” tool is not yet integrated in all the pages. In particular, it would be very useful to make the “parameter search” and the “build data table” feature consistent so that one could do a parameter search and use the output directly in the table build tool.
- The current default on ‘Near Position’ searches is to return all objects within 2 arcminutes. Now that most positions are as good as they are, and the number of objects in NED is exploding, this default should be substantially smaller. 10"? 30"?
- Each search returns a “page” that is about 7 screens in length. This could be cut down by more than a factor of two, by reducing the formatting and amount of redundant information. In particular:
 1. The “cross identification” table has only two columns and is much narrower than most of the other tables returned. Doubling the number of columns would halve its length. The list should be ordered alphabetically or by wavelength.
 2. One of the “redshift-independent” distances given is the “metric distance”, which we are not familiar with. “Comoving tangential distance” is another name for the more common “proper motion distance”, There should be a “?” by each cosmological distance that links to an explanation.
 3. The “foreground extinction” tables consume half a screen. Perhaps you should list only the preferred Schlafly & Finkbeiner (2011) values, and have a link labeled “?” or “detailed explanation” with alternative extinctions and all of the text, references, etc. for those who want more information.
 4. Likewise, half of the “basic data” table is used by the caveat “note” and links to additional detailed measurements. The user who wants more should have to click on a “?”, “note”, or “details” button.

5. The “magnitude and filter” line doesn’t always give the filter. For example, on NGC 4645, it is just “12.88”, but I don’t know in what filter or on what magnitude scale, and I don’t know if it is corrected for the “foreground extinction” above.

6. The “quantities derived from redshift” give eight velocities, six Hubble distances, six distance moduli, six angular scales, and four cosmology-corrected distances and distance moduli. But NED has a preferred distance, which it should either mark as preferred or, even better, display instead of the list. For example NGC 4565 has Hubble flow distances ranging from 32.4 Mpc to 60.0 Mpc. To save space and reduce confusion, the user should have to click on a button to get the alternative distances, distance moduli, and angular scales. Each nearby distance should have a “?” button linking to an explanation of why that distance is appropriate or not depending on redshift, sky position, etc.

The cosmology-corrected distance moduli for all but the luminosity distance are formally meaningless and should be deleted.

The “quick-look photometry and luminosities” results break down for cosmologically distant objects. For example, 3C 279 at $z = 0.5362$ is listed with luminosity distance of 2996 Mpc and distance modulus 42.38 mag, but the quick-look luminosities were calculated from the badly incorrect “Average NED-D value” of 1620 Mpc “metric distance” and distance modulus 41.05 mag. 3C 454.3 is lucky enough not to have a redshift-independent distance, and NED is smart enough to use the luminosity distance for it.

Other

- We did not talk about Level5 in this meeting, but we are concerned that its holdings are becoming quite out of date. For example, there are no Planck or recent WMAP articles in the “Cosmological Parameters” or “Background Radiation” subsections of the “Cosmology” main heading. Planck and WMAP are not in the Glossary. The references to “Cosmic Background Radiation”, “CMB”, “CMBR” are all from 2002 or earlier. “Dark Energy” is represented by only one paper from 2001. Under “Fundamental Data”, the most recent article on “Best cosmological parameters” is from 2003. “Atlases and Catalogs” contains nothing newer than 2004. “Spectroscopy” claims to be “a relatively new section in Level5” but was last updated in 2001. Surely the 1999 paper “Optical and UV astronomy in the internet age” by Daniel Golombek is obsolete by now and could be deleted from Level5 “fundamental data”. The “Literature/Thesis abstracts” seems to stop around the year 2004. Most just say “ABSTRACT NOT AVAILABLE in NED”. It sorts only by years, not by author. The same info is on ADS with search by author. Is this worth keeping in NED? In Level5, buried in the “Fundamental Data” section are calculators that convert between Julian date and calendar date. Both are dead links. There is a separate “USNO Julian Date Calculator” that seems to work. To be useful for synoptic data, it should appear on the home page under “Tools”. Another dead link in “fundamental data” is “fundamental constants”. The “extinction law calculator” by Cardelli et al. (1989) is not only obsolete but is a dead link.

- You seem to be moving to a system (like SDSS has recently) where the web interface and the services are somewhat decoupled. You should think about introducing some API access services, that aren't dependent on the web interface. While SDSS has yet managed that full separation, it definitely is the right direction to head.