



Dr. Joseph Mazzearella, NED Project Scientist
Dr. George Helou, NED Principal Investigator & IPAC Executive Director
Infrared Processing and Analysis Center
Caltech/JPL, MS 100-22
770 South Wilson Avenue
Pasadena, CA 91125

June 15, 2023

Dear Joe, George, and members of the NED Team,

This is a report of the NED Users Committee (NUC) meeting with the NED team held in person at IPAC in Pasadena and over zoom on May 15, 2023. At this meeting we heard updates from the NED team on software and hardware upgrades, improvements in cross-matching techniques, advancements in the Gravitational Wave Follow-up Service, two publications related to the NED Local Volume Sample and to community best practices, and other work done since our last meeting via zoom in March 2022. We also heard from Anthony Kremin on the status of the DESI redshift catalogs and from the NED team on their initial plans for the upcoming programmatic review.

We continue to be impressed with the multifaceted work the NED team does to provide the community with a robust, adaptable database and to facilitate open access to science. NED remains unique among other archives in providing multi-wavelength, cross-matched, and authenticated data that maximizes the science return from any NASA dataset. NED further continues to play a crucial role in enabling new NASA science as evidenced by the reported doubling of the query rate leading up to the most recent JWST deadline.

Two major strategic goals discussed in past NUC meetings have been to increase NED's role in time domain astronomy and to speed up NED's ingestion rate. The NED team reported on significant progress towards both of these aims. The quick and timely publication of the NED Local Volume Sample will maximize the global usage of NED by teams pursuing GW-EM follow-up efforts. The work with the DESI team in advance of their first data release has put NED in the position to streamline the ingestion of their estimated 40 million redshifts and to provide a leading example for how these collaborations can work in the future. We are happy to see the NED team laying more groundwork for streamlining ingestion from the literature as well, including the publication of their Best Practices document and work to communicate directly with journals. The NED team has met these important goals all while doing the less exciting but still incredibly important work of database restructuring to update hardware/software, improving the cross-match algorithm, and producing three major data releases with over 275k new object links. We are also happy to see the NED team's continued prioritization of open access to science through their effective and well-attended talks and workshops at the annual AAS meetings.



Following what we heard and discussed at this meeting, our recommendations to the NED team are as follows:

1. Continue to prioritize early ingestion of DESI data and to work closely with the DESI team to incorporate information on things like redshift warning flags. Request that the DESI team point to NED in their release announcement(s) if possible.
2. Leverage the NED team's experience working with the DESI team to extend this effort to other large archives. For example, invite someone from Rubin LSST and eventually Roman and Euclid to join the NUC.
3. Continue the effort to improve the MatchEx source matching algorithm. The NUC is impressed with the work to improve and streamline MatchEx and encourages the NED team to publish their results and/or document the algorithm. The robust capability for associations, especially in regards to GW follow-up and large surveys of resolved sources like those with Roman, will only become more important. This is a high impact opportunity for NED to be the leading resource on this methodology.
4. Generalize the NED time domain efforts beyond GW science to maximize the impact. For example, develop a NED service to connect the most probable host galaxy information directly to the transient data stream either through TNS discoveries or through re-streaming ZTF/Rubin alerts. NED could state the most probable redshift and incorporate priors based on source classes (e.g., supernova vs kilonova vs neutrino counterpart).
5. Continue efforts to decrease ingestion time from the literature. The NUC applauds the NED team's creative solutions to spread the word about best practices at the author level through the publication of their Best Practices document and at the journal level through the trial with PASA. Because the PASA connection has not yet yielded a significant number of use cases, consider reaching out to the AAS journals editor about conducting a trial with a different journal, particularly one that is based in a closer time zone. For the more immediate future, create a Jupyter notebook (preferably one that is prominently displayed on the NED webpage) with instructions for table formatting. Consider prioritizing messaging around table entries/formatting that make ingestion more time consuming.
6. Write a paper discussing the fiducial redshift tree. This is a carryover suggestion from the previous NUC report. Possible venues for this publication could be the AAS bulletin, an ADAS paper, or a research note on the NED website.
7. Set aside time as a team to construct mission and vision statements that articulate a clear plan for the future given the upcoming programmatic review and the 2020 Decadal Survey recommendation to increase interoperability between the different archives. Full integration of NED capabilities into the planned NASA archives science platform is critical to the long-term success of NED. We strongly encourage NASA HQ to involve the NED team in the archives science platform development process, and the NED leadership to insist on such representation.



Given the upcoming programmatic review, we further offer our advice on the priorities for that review.

Priority #1: Enhancing extragalactic time domain and multi-messenger astronomy: The NUC agrees that time domain efforts are a high impact opportunity for NED and encourages the team to focus on the areas where NED is unique, specifically the high quality of verified information about host galaxies and associations. An example use case where NED brings unmatched value could be users conducting a search for a host for a particular event by starting with priors, i.e. a limit on the host redshift or the fact that it is expected to be star forming. The NUC agrees a yes/no flag to note whether there is any kind of time variability associated with the galaxy would be helpful but advises against going too far into reproducing the work of databases devoted to time domain efforts (e.g., properties of lightcurves).

Priority #2: Data capture tool to expand open science: The NUC fully supports this creative approach to crowd sourcing the formatting of data to speed up the ingestion rate into NED. The fact that 50% of papers would be neglected if the NED team did not incorporate papers with <10 sources makes it clear that the continued ingestion beyond only large survey papers is important. Such a cut would also exclude scientifically important objects like the highest redshift sources being discovered with JWST. Perhaps rewards could be given to those authors who do a particularly good job of following the guidelines. For example, authors can be given an “open access science star” recognition that they can put in their data management plan when applying for grants or their science could be highlighted on the NED page.

Priority #3: Enrich NASA science platform with NED data and analysis tools: The NUC agrees that a platform that enables sophisticated queries and the derivation of physical quantities would be a valuable asset to the community and is a natural progression for NED. We encourage the team to continue to consider how such a platform could be integrated into and share infrastructure with those already being built by IRSA, etc.

Priority #4: AI-assisted data fusion to increase open science yield from NASA data: The NUC supports efforts to include AI as much as possible to help with ingestion so that the person power required for the vetting of high quality data can be maintained. We encourage the NED team to identify algorithms that they consider critical in the big data era (i.e., one that aids with making associations in high-resolution large area imaging like that from Roman). Given the NED team’s unique knowledge, they would be well-suited to curate benchmark samples that such algorithms could be tested against.

Priority #5: Open Science Repository following the Exoplanet Follow-up Observing Program (ExoFOP) model: The NUC is intrigued by an extragalactic version of ExoFOP as there is currently no equivalent and a curated image repository would be a valuable resource for the community. However, the NUC has some concerns that the time required for this effort could balloon very quickly. We suggest the NED team first allow the community to define itself and then prioritize NED team contributions according to the aspects that are most used by the community and most relevant to NASA science.



We are, as usual, incredibly impressed with the clever and resourceful work of the small NED team. This update was very responsive to the suggestions from the last NUC report and we appreciate the careful consideration of our past advice.

Sincerely,

The NED User's Committee

Dr Pauline Barmby, University of Western Ontario

Dr Rachael Beaton, Space Telescope Science Institute

Dr Brad Cenko, NASA Goddard Space Flight Center and University of Maryland

Dr Mansi Kasliwal, California Institute of Technology

Dr Mark Lacy, National Radio Astronomy Observatory

Dr Preethi Nair, University of Alabama

Dr David Schlegel, Lawrence Berkeley National Laboratory

Dr Ohad Shemmer, University of North Texas

Dr Sabrina Stierwalt, Occidental College (Chair)