Science Discovery With Diverse Multi-Wavelength Datasets Fused in the NASA/IPAC Extragalactic Database (NED)

Joseph M. Mazzarella

IPAC, California Institute of Technology, Pasadena, CA 91125
mazz@ipac.caltech.edu

Abstract. Data from space-based observatories, large sky surveys, and the astrophysics literature are growing at an unprecedented rate. The ongoing expansion in volume, velocity and variety of astronomical data is creating exciting opportunities for making ground-breaking discoveries from interconnected archives and databases. I discuss how recent advances in joining data across the spectrum from NASA’s GALEX, 2MASS and AllWISE sky surveys with over 118,000 other unique catalogs and journal articles, combined with new capabilities of the user interface, are helping astronomers make discoveries directly from information fused in the NASA/IPAC Extragalactic Database (NED). After a review of recent scientific results that were enabled or facilitated by NED, I briefly consider some challenges and limitations in joining heterogeneous datasets, and conclude with a reminder of the importance of state of the art astronomical archives in expanding the opportunities for new science discoveries from multi-wavelength big data in the 2020s. NED is operated by the California Institute of Technology under contract with NASA.

1. Introduction

The NASA/IPAC Extragalactic Database (NED, https://ned.ipac.caltech.edu) is a thematic, web-based research service used by scientists, educators, and mission operations for observation planning, data analysis and science discovery for objects beyond our Milky Way galaxy. NED provides a systematic fusion of data from hundreds of sky surveys and more than a hundred thousand research publications. The holdings and services cover the observed electromagnetic spectrum from gamma rays through radio frequencies, and are continuously updated with data extracted from new journal articles and astronomical source catalogs. NED has been in operations since 1990, growing in content and capabilities with advances in databases and information technology. As of October 2019, the system contains over 1.1 billion distinct objects and 1.5 billion multi-wavelength cross-IDs, the result of a hundredfold expansion since a decade ago (Mazzarella & NED Team 2007). This growth is not only in size, but also complexity: data from more than 118,000 catalogs and publications are currently combined in NED.

2. Combining heterogeneous, multi-λ measurements to facilitate research

NED simplifies and accelerates scientific research on extragalactic objects by distilling and synthesizing data across the spectrum, and providing value-added derived quan-
Mazzarella

tities and functionality. As new data about extragalactic objects are published in the peer-reviewed literature and in well vetting mission and observatory archives, they are identified, distilled, cross-identified and integrated into a unified database schema designed to simplify queries and data access. The system contains published information including object names, coordinates, redshifts, redshift-independent distance measurements, photometry, sizes, classifications and attributes, references, and object notes. NED also hosts a repository of images and spectra contributed by authors of journal articles. Value-added derived quantities include distances based on the Hubble flow, velocity corrections and cosmological corrections, metric sizes, spectral energy distributions, and luminosity estimates. Queries can be submitted from programs and scripts via APIs, or through a web-based graphical user interface which has recently been modernized and extended with linked interactive tables, graphics and image overlays.

The high rate of usage and citations in the literature indicate NED has become an essential tool for many researchers, and it is an integral part of modern astrophysics research on extragalactic objects. While NED is frequently cited as an important reference for specific data, recently there have been a number of science discoveries made using NED’s data synthesis as a primary resource for the discovery, and others where NED has facilitated the research in a substantial way. We turn now to reviewing some recent scientific results that were directly enabled or facilitated substantially by NED. These include the discovery and characterization of super spiral galaxies, use of data from NED in studies of galaxy distances and measurements of the Hubble constant, analysis of variations of the pitch angles of spiral arms with wavelength to test density wave theory, and rapid identification of the electromagnetic counterpart to the first gravitational wave event detected from a pair of merging neutron stars.

3. Super spiral galaxies

The largest and most luminous spiral galaxies known have luminosities more than 8 times the characteristic knee in the galaxy luminosity function (L*), diameters as large as 130 kpc, steller masses of 30 to 300 times that of our Sun (M⊙), and star formation rates (SFR) ranging from 5 to 65 M⊙/yr (Ogle et al. 2016). Known as "super spirals", this class of objects was discovered while characterizing the completeness of the redshift measurements in NED. Data joined within NED that enabled this work include redshifts, object types, diameters, SEDs and derived quantities: luminosity (SDSS), stellar mass (2MASS), star formation rate (GALEX, Herschel). Figure 1 shows the distribution of colors, star formation rates, and stellar masses for the initial sample of super spirals (Ogle et al. 2016).

Recent follow-up work includes a catalog of 1,525 of the most optically luminous galaxies from the Sloan Digital Sky Survey at z < 0.3 (Ogle et al. 2019b), and discovery that most super spirals rotate substantially faster than predicted by the baryonic Tully-Fisher relation established for lower mass galaxies (Ogle et al. 2019a). These results indicate that super spirals are under-massive in stars considering their huge halo masses, and that there is a fundamental upper limit to a galaxy’s mass in stars of about 10^{11.8} M⊙, corresponding to a critical dark matter halo mass of M_{halo} = 10^{12.7} M⊙. Both of these studies also relied heavily on NED. Figures 2 - 4 show highlights from this work.
Figure 1. (a) SDSS and WISE colors of super spirals (circles) compared to other SDSS galaxies classified as late type galaxy (LTG) or early type galaxy (ETG) by Lintott et al. (2008). The infrared transition zone (IRTZ) is the mid-IR equivalent of the optical green valley (Alatalo et al. 2014). (b) Star formation rates and stellar masses of super spirals compared to the SDSS-WISE sample of Chang et al. (2015). The dashed line indicates the star-forming main sequence at $z \sim 0$ (Elbaz et al. 2007). This is Figure 1 from Ogle et al. (2016).

Figure 2. SDSS images of super spiral and super lenticular galaxies, ordered by decreasing $r$-band luminosity. Each image is 150 kpc on a side. This is Figure 11 from Ogle et al. (2019b).
Figure 3. Fastest rotating super spiral 2MFGC 12344. a) SDSS g, r, i band image. b) 2D spectrum of Hα and [N II] λ6585. c) Rotation curve (solid blue line, red points) with $v_{\text{max}} = 568 \pm 16$ km/s at $r = 41$ kpc. d) Residuals to cubic spline fit, with standard deviation 16 km/s. e) Galaxy continuum profile along the slit. This is Figure 1 from Ogle et al. (2019a).

Figure 4. Baryonic Tully-Fisher relation (BTFR) between baryonic mass and the terminal rotation velocity of a galaxy’s rotation curve. Masses in stars for the super spiral and comparison samples are estimated using WISE W1-band photometry, assuming $M/L = 0.6$. The photometric uncertainty is smaller than the size of the plot symbols (0.01-0.02 dex). The observed BTFR (data points) is compared to the Lelli et al. (2016) power-law fit (solid line). This is Figure 2a from Ogle et al. (2019a).
4. NED is used heavily in studies of the Hubble Constant

Fernández Arenas et al. (2018) recently used a "standard candle" distance indicator provided by the relationship between integrated Hβ line luminosity and the velocity dispersion of the ionized gas of H II galaxies and giant H II regions to study new measurements in 13 galaxies. They derived a best estimate of the Hubble constant of $71.0 \pm 2.8$ (random) $\pm 2.1$ (systematic) km s$^{-1}$ Mpc$^{-1}$. Their study made use of NED as a substantial source of joined data, including Cepheid distances and Hα images used to target their spectroscopic observations. See Figures 5 and 6.

Figure 5. L - $\sigma$ relation for Chávez et al. (2014) and Bordalo & Telles (2011) combined samples. The solid line is the fit to the HII Galaxies (HIIGs). The inset equation is the distance indicator where the slope is obtained from the fit to the HIIGs and the zero point determined as described in Fernández Arenas et al. (2018). This is Figure 18 from Fernández Arenas et al. (2018).

Figure 6. Hα image obtained from NED (left), the high-resolution spectral profile for the giant H II region (center), and the low resolution spectrum (right). This is part of Figure A1 from Fernández Arenas et al. (2018); see the article for more details.
It has been argued that cosmology is in a state of crisis given the significant difference between values of the Hubble constant inferred from the cosmic microwave background as observed by the Planck mission and those measured locally using observations of Cepheid variables and distant supernovae to study the Hubble flow locally. A new value measured by Freedman et al. (2019), using a high precision tip of the red giant branch distance indicator, falls squarely between the two previous values, reducing the tension with Planck. The team based their results on a three-year study using the Hubble Space Telescope, and they acknowledged being heavily reliant on NASA’s on-line archives, including the Astrophysics Data System (ADS), the Mikulski Archive for Space Telescopes (MAST), and NED in the planning, execution and interpretation of the observations. See Figures 7 and 8.

Figure 7. Hubble diagram for galaxies with TRGB (left panel) and Cepheid (right panel) distances. The black filled circles are for galaxies in common with both the TRGB and Cepheid samples. The velocities have been corrected for the presence of nearby mass concentrations using NED, as described in Freedman et al. (2019). A slope of $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is shown in black. This is Figure 7 from Freedman et al. (2019).

Figure 8. Plot of $H_0$ measurements as a function of year of publication. The black points and shaded region are determined from measurements of the CMB, those in blue are Cepheid calibrations of the local value of $H_0$, and those in red are TRGB calibrations. The red star is the best-fit value obtained by Freedman et al. (2019). Error bars are $1\sigma$. This is Figure 17 from Freedman et al. (2019).
5. Variation in pitch angles of spiral arms test spiral density wave theory

An article titled “Investigating the Origins of Spiral Structure in Disk Galaxies through a Multiwavelength Study” by Miller et al. (2019) performed detailed analysis of images of galaxies in NED at various wavelengths representing stellar populations with different mean ages. Each group of stars formed a spiral arm with a slightly different “pitch angle” with respect to the center of the galaxy, matching the prediction of density wave theory. Figure 9 from their article illustrates the variation in pitch angle with the image wavelength. This research was also featured in a science news article “Research on disk galaxies sheds light on movement of stars”.

Figure 9. Illustration of spiral-arm structure based upon pitch angle measurements at multiple wavelengths measured from images in NED. The authors observed lower pitch angle (tighter arms) for 3.6 \( \mu \)m and the B band, higher pitch angle (looser arms) for the u band and even higher pitch angles (looser arms) for 8.0 \( \mu \)m, FUV and H\( \alpha \) images. This means the star-forming arm is upstream from the blue arm, which is in turn upstream from the red arm. This conforms with spiral density-wave theory’s prediction. This is Figure 1 from Miller et al. (2019).

6. NED use in multi-messenger astronomy

Beyond multi-wavelength astronomy is multi-messenger astrophysics, which involves combining measurements of electromagnetic radiation with other energy carriers, most notably gravitational waves. Here too NED is proving to be a useful tool.

On 17 August 2017, the Laser Interferometer Gravitational-Wave Observatory (LIGO) and the Virgo interferometer detected gravitational waves (GWs) from an event known as GW170817. Nearly simultaneously, the Fermi and INTEGRAL telescopes in space detected a gamma-ray burst labeled GRB 170817A. About 11 hours after the GW event, Coulter et al. (2017) discovered a fading optical source coincident with GW170817, Swope Supernova Survey 2017a (SSS17a), located in an S0 galaxy 40 Mpc away known as NGC 4993. This was the first detection of two neutron stars merg-

ing together both in gravitational waves and in EM radiation emitted in a cataclysmic event known as a kilonova, and data in NED played an important role in assisting the Coulter et al. team in narrowing down the candidate galaxies for a rapid identification of the EM counterpart. Figures 10 and 11, from Coulter et al. (2017), show the top candidate galaxies in the GW localization area and the discovery image. A detailed account of the timeline of events, and the role of NED, is given by Steer (2017).2

Figure 10. The 50%, 70%, and 90% probability contours are shown for the localization of GW170817. Gray circles represent galaxies in the galaxy catalog observed by the Swope Telescope on 17 to 18 August 2017 to search for the EM counterpart to GW170817. The size of the circle indicates the probability of a particular galaxy being the host for GW170817. The blue square labeled “9” contains NGC 4993, the location of which is marked by the blue circle, and SSS17a. This is Figure 2 from Coulter et al. (2017).

Figure 11. Optical images of NGC 4993, with north up and east left; the field is 3 arcmin by 3 arcmin. (a) Hubble Space Telescope F606W-band (broad V) image captured 4 months before GW170817. (b) Swope image of SSS17a (i-band) obtained on 17 August 2017 at 23:33 UT. SSS17a is marked with the red arrow. This is Figure 4 from Coulter et al. (2017).

These results inspired development of a new NED gravitational wave follow-up service. Released in March 2019, in time for Observing Run 3 (O3) of LIGO/Virgo, this tool is designed to leverage data in NED to facilitate searches for electromagnetic counterparts of GW events (Cook et al. 2019). Within minutes after the LIGO-Virgo collaboration issues a new GW alert on the NASA Gamma-ray Coordinates Network

---

(GCN), NED responds by posting a list of candidate galaxies in the event’s 90% probability volume. The service currently supports galaxies with distances less than 200 Mpc, and the team is in the process of extending the distance limit and characterizing the completeness of available galaxy distance measurements. See Figure 12.

Figure 12. The NED gravitational wave follow-up service, available at https://ned.ipac.caltech.edu/gwf/overview.

7. Summary

The NASA/IPAC Extragalactic Database (NED) has been undergoing continuous expansion in data content and functionality improvements. The recent fusion of fundamental measurements from NASA’s GALEX, 2MASS and AllWISE sky surveys with attributes joined from over 118,000 other catalogs and journal articles, is vastly increasing opportunities for multi-wavelength research and discoveries. A number of findings published recently in which data fused in NED played a substantial role were reviewed.

I conclude with a quote from a science white paper submitted to the U.S. 2020 Decadal Survey on Astronomy and Astrophysics by Fabbiano et al. (2019): “The history of astronomy shows that paradigm changing discoveries are not driven by well formulated scientific questions, based on the knowledge of the time. They were instead the result of the increase in discovery space fostered by new telescopes and instruments. An additional tool for increasing the discovery space is provided by the analysis and mining of the increasingly larger amount of archival data available to astronomers. Revolutionary observing facilities, and the state of the art astronomy archives needed to support these facilities, will open up the universe to new discovery.” We can all look forward to many exciting discoveries in the 2020s enabled by the worldwide ecosystem of interconnected astronomical archives and emerging science platforms.
Current NED team members are Ben Chan, Tracy Chen, David Cook, Rick Ebert, Cren Frayer, George Helou, Tak Lo, Barry Madore, Joseph Mazzarella, Olga Pevunova, Marion Schmitz, Ian Steer, Scott Terek, Cindy Wang, and Xiuqin Wu.

NED is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References


