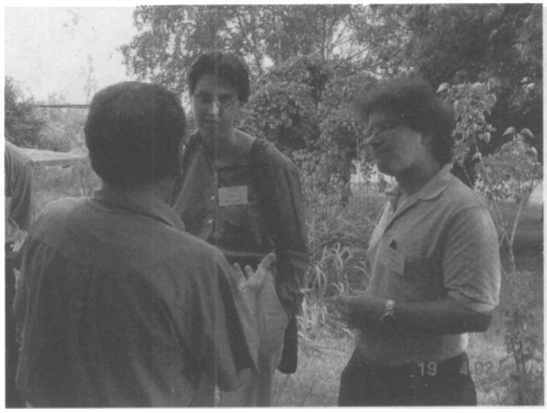
Part 6
Future Projects

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press



Ed Khachikian, Caryl Gronwall, Joe Mazzarella

The Conference Banquet: Areg Mickaelian, Wolfgang Voges

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press

AGN SURVEYS
ASP Conference Series, Vol. 284, 2002
R.F. Green, E.Ye. Khachikian, D.B. Sanders

Using the NASA/IPAC Extragalactic Database (NED) and Federated Virtual Observatory Archives for Multiwavelength Studies of AGNs

Joseph M. Mazzarella and the NED Team

California Institute of Technology, Jet Propulsion Laboratory, MS 100-22, Pasadena, CA 91125

Abstract.

We live in an exciting era that offers increasing opportunities for people all over the world to make discoveries about the Universe using interconnected archives on the Internet as a primary research tool. We review how NED (*http://ned.ipac.caltech.edu*) can be used in concert with globally distributed online archives to perform multi-wavelength, cross-correlated studies of AGNs and other galaxy types. The present status and planned evolution of NED capabilities are discussed.

1. Introduction

The NASA/IPAC Extragalactic Database (NED) is an online research facility designed to support scientists, educators, space missions and observatories in the planning, execution and publication of research on extragalactic objects. The foundation of NED is a growing database of galaxies, quasars and all types of extragalactic objects that can be searched by positions, redshifts, object types, references, authors, and multi-wavelength cross-identifications produced from thousands of catalogs and journal articles. The primary goal of NED is to maintain an up-to-date panchromatic fusion of basic data for all known (cataloged and published) extragalactic objects, including pointers to the astrophysical literature and to relevant distributed archive resources. Scientists working in observational extragalactic astronomy use NED in their research at nearly every step, from proposal planning, through data collection, analysis, and publication.

2. Cross-Identifications and Data Integration

2.1. Multi-wavelength Cross-Identifcations and Associations

Cross-identification refer to the process of establishing which observation in one survey catalog (e.g., FIRST) corresponds to the same astrophysical source in surveys at other wavelengths (e.g., SDSS and 2MASS). The process is much more difficult than it may first appear, because observations taken with different telescopes and at various wavelengths often differ in substantial ways including different positional uncertainties, systematic errors in astrometry (some catalogs), different survey resolutions (e.g., IRAS and APM compared to ROSAT), and

problems such as matching double-lobe radio sources with their parent galaxies. Objects also populate a hierarchical Universe: AGNs, supernovae, and HII regions reside in their host galaxies; galaxies are members of pairs and groups; pairs and groups are typically members of clusters; and galaxy clusters reside in superclusters separated by vast voids. Therefore, complex relationships between objects are needed (one-to-one, one-to-many, many-to-one, many-to-many), in addition to *statistical associations* for cases in which firm cross-identifications cannot be established. NED activities revolve around a systematic process of constructing and revising cross-identifications and statistical associations between millions of entries in multi-wavelength catalogs and publications.

2.2. Database Contents

The database content of NED is updated periodically on the home page (Figure 1). To date NED contains: 5.9 million cross-identifications in thousands of multi-wavelength surveys and journal articles; 4.7 million unique extragalactic objects; 4.5 million photometric measurements covering gamma-rays through radio wavelengths; 2.0 million detailed position measurements with uncertainties; 1.7 million bibliographic references to 50,000 articles; 291,000 redshifts; 748,000 science-grade FITS images and remote links; 54,000 detailed notes from catalogs and journal publications; and 28,000 abstracts of journal articles and Ph.D. theses.

The essential data for sources in NED include positions, redshifts, morphological types, nuclear spectral types, panchromatic photometry, and images. When available, uncertainties in the measurements are also stored and provided in the query reports. Photometric data are stored in original units and converted to common frequency (Hz) units and flux density units ($W\ m^{-2}\ Hz^{-1}$) for construction of Spectral Energy Distributions (see Figure 5); the data are also tagged with their aperture sizes or status as a "total flux" measurement.

3. Query Services

3.1. Web Interface

There is insufficient space here for a comprehensive history and complete technical review of NED. A discussion of the motivation, initial design, and early history of NED was given by Helou et al. (1991). NED is accessible on the World Wide Web at *http://ned.ipac.caltech.edu.* Figure 1 shows the NED Web interface main menu (home page). Following is a review of the primary services.

NED can be searched for extragalactic objects using menus designed for searches 'By Name', 'Near Name', 'Near Position', 'IAU Format', or 'By Reference Code' (a 19 digit code developed jointly by NED, ADS, and CDS[1] to uniquely identify publications in astronomy). Figure 2 illustrates an example of the Essential Data presented after a query of the galaxy NGC 4151.

Figure 3 shows hyperlinks to 'External Archives and Services', easily found when the user scrolls down below the Essential Data that come directly from

[1] *http://adswww.harvard.edu, http://cdsweb.u-strasbg.fr*

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press

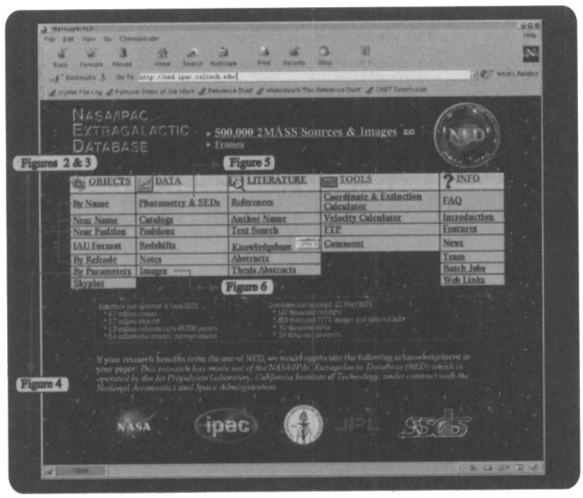


Figure 1. The NED Web interface main menu available at *http://ned.ipac.caltech.edu*. The annotations refer to example query results displayed in Figures 2-6.

the NED database. Links here allow the user to retrieve images and query original catalog data or observation log entries. The first set of hyperlinks are to data related directly to an object name; the second set of links are to services that can be queried at the object's coordinates. A summary of the available resources includes: original catalog record entries in VizieR at CDS; the NVSS, and FIRST catalog and image servers and the Observation Log of the VLA telescope from the *National Radio Astronomy Observatories (NRAO)*; infrared mission archives at *IRSA/IPAC* (2MASS, IRAS, etc.); visual and UV mission archives at *MAST/STScI* (HST, IUE, etc.); and high energy mission archives at *HEASARC/GSFC* (ASCA, CGRO, Einstein, etc.) New archive services are being added as they become available. This tool makes it very easy for researchers to locate existing observations in one or more of the major surveys or observatories, providing a major step toward federating distributed archives.

The 'By Parameters' (recently renamed 'Advanced All-Sky') menu allows the user to query NED using joint constraints on redshift range, sky area, object types, or survey/catalog name prefixes. With this feature one can dynamically regenerate a classic sample that contains not simply the original measurements (available elsewhere, such as VizieR), but rather the most precise and currently available source positions and redshifts, with links to up-to-date references, multi-wavelength photometry, images, etc. For example, today anyone can use NED to generate an up-to-date compilation analogous to the '*Catalog of Markarian Galaxies*' (Mazzarella & Balzano 1986), or generate a current data set for the entire *Third Cambridge (3C)* radio galaxy sample, with the click of a mouse. Since entries in NED are continuously updated through a synthesis

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press

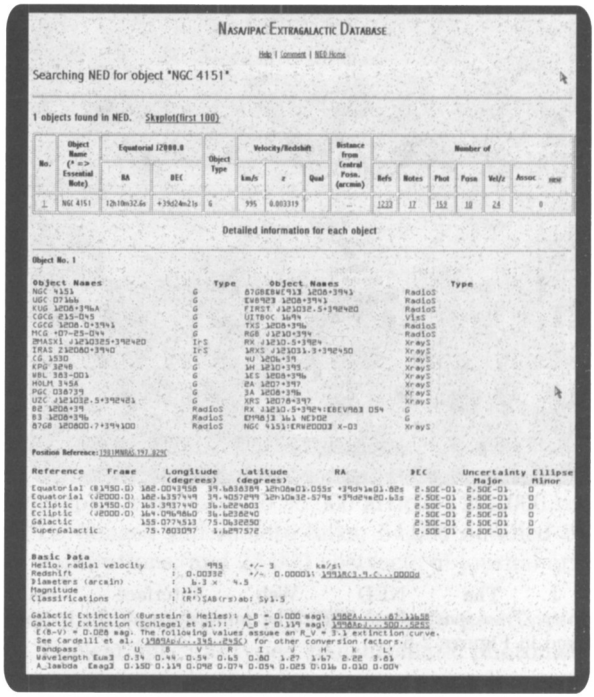


Figure 2. Essential Data returned by NED from a query of NGC 4151: includes coordinates and redshift (with uncertainties), multi-wavelength survey cross-identifications and object types, size, magnitude, classifications, Galactic extinction along the line of site, as well as links to query references, notes, photometry, positions, velocities, and images. This information is followed by links to **External Archives** (Figure 3).

of the literature and large surveys, errors in original catalogs are often found and documented. Therefore, *for many studies it is more efficient and effective to cross-correlate new observations against the data fusion in NED rather than against catalogs in their original published forms.* A visualization of one example of this powerful feature is shown in Figure 4. Other queries that can be performed using this tool, including the recent introduction of photometric constraints, are demonstrated in a new 'Tutorial Examples' feature on the interface menu.

The *DATA* column allows the user to enter an object name (e.g., 'NGC 4151', 'SN 1993G', '2MASXi J1132350+582422', 'SDSS J1044-0125', 'Antennae') and query NED for 'Photometry & SEDs', 'Catalogs', 'Positions', 'Redshifts', 'Notes', or 'Images'. Figure 5 illustrates an example of multi-wavelength photometric data and Spectral Energy Distribution (SED) plots.

Figure 6 shows an example of the rich variety of multi-wavelength FITS images available for immediate download and visualization using the Aladin Java

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press

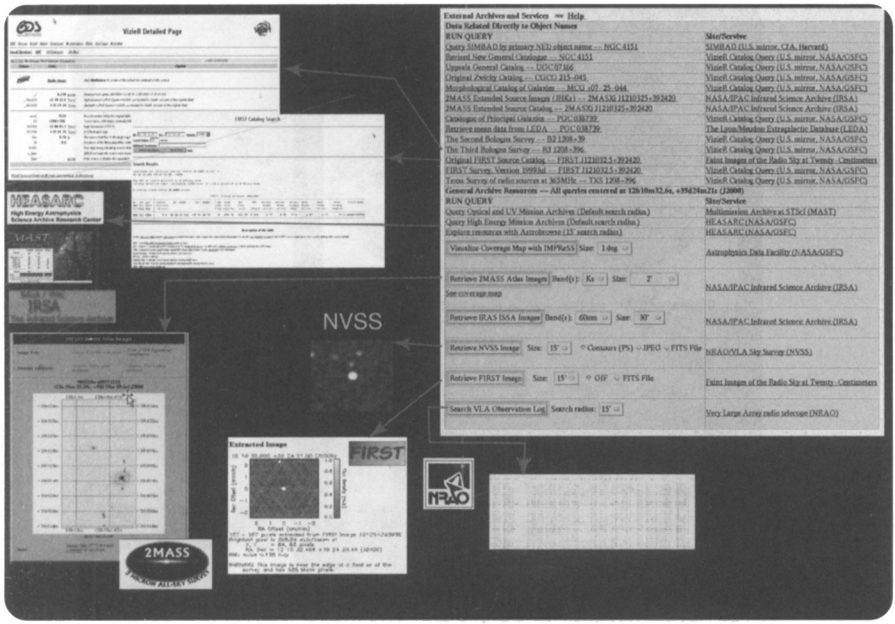


Figure 3. Queries to globally distributed archive data are available with a single mouse click. The first set of links are anchored to data related directly to an object name; the second set of links query services at the coordinates of the NED object. This example is a continuation of the report from a query of NGC 4151 (Figure 2).

applet. Aladin's interoperability with NED and other distributed data services provides a visual summary of the multi-wavelength sky. Aladin was developed at the CDS and configured with NED through a cooperative agreement.

In the *LITERATURE* column the user can: (1) enter an object name and access the 'References' related to that object; (2) enter an 'Author Name' and retrieve a list of references; (3) search journal article 'Abstracts'; (4) search 'Thesis Abstracts'; (5) use the 'Text Search' tool to perform keyword searches on the journal and thesis abstracts in NED or the full text content of LEVEL5; and (6) access the LEVEL5 'Knowledgebase'. LEVEL5 provides hyperlinked review articles and documents of current and lasting interest to cosmologists and extragalactic astronomers. Contents include a glossary of terms, essays, research papers, and reviews. Cited extragalactic objects are cross-linked to NED Basic Data queries, and all available citations are hyperlinked to NASA's Abstract Data Service (ADS), to on-line NED abstracts, or to preprints on astro-ph.

The *TOOLS* column of the main NED menu contains a 'Coordinate & Extinction Calculator' that performs coordinate conversions and precession and displays line-of-sight Galactic extinction estimates using two techniques (Schlegel et al. 1998; Burstein & Heiles 1982). The 'Velocity Calculator' computes conversions between redshifts for extragalactic objects in different reference frames: heliocentric, Local Group, Galactic Standard of Rest, and 3K Microwave Back-

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press

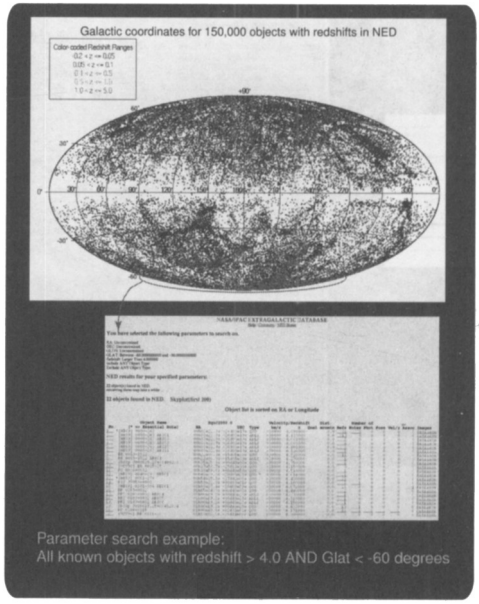


Figure 4. Objects that satisfy the criteria [$Redshift > 4.0\ AND\ Glat < -60°$] returned after submitting a query using the 'By Parameters' menu (see Figure 1).

ground. There is also a link to NED's public FTP site, primarily used by users to pick up their output from the NED batch mode.

The *INFO* column of the main menu contains links to a Frequently Asked Questions ('FAQ') document, an 'Introduction', summary of 'Features', information about the NED development 'Team', forms and documentation for submitting NED 'Batch jobs', and finally a page of useful 'Web Links' relevant to extragalactic astronomy. Information about new content and capabilities is available at the top of the main interface.

3.2. Batch Mode

NED can process requests for large amounts of data as a 'Batch Job.' Using this mode simply involves submitting to NED via email a "batch form" containing a list of objects or positions, or other constraint parameters (e.g., redshift, object type, or survey/catalog name prefix).[2] After the request has been processed, NED sends the user a notice by return email indicating where the resulting data files may be copied via FTP.

[2]The forms are available from the 'Batch Jobs' link on the main NED menu, and they should be emailed to *nedbatch@ipac.caltech.edu* for processing.

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press

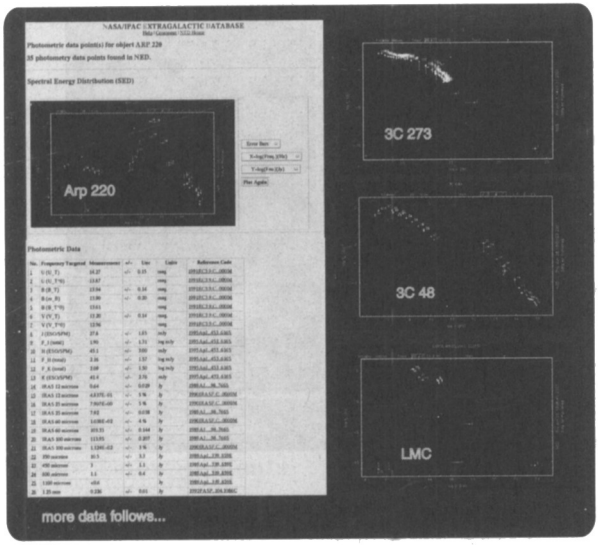


Figure 5. Multi-wavelength photometry and spectral energy distributions (SEDs) covering gamma-ray through radio wavelengths. The data are available in original units as published, and also in common units ($Hz, Wm^{-2}Hz^{-1}$) used to construct SEDs. The data include uncertainties and references. The SED plots are dynamic, with configurable axis units (e.g., wavelength or frequency for the abscissa and f_ν, νf_ν or f_λ for the ordinate) and optional error bars.

3.3. Client/Server Mode Connectivity

For many years NED has provided a 'server mode' with custom client (C) software [3] that has been used by computer programmers all over the world to build applications that issue queries and retrieve data from the NED database in a format that can be integrated into their services. Astrophysics archive centers and observatories use NED's server mode extensively to resolve extragalactic object names into celestial coordinates, and to retrieve lists of objects by specifying an input position and search radius. A number of sky visualization tools also make use of NED's client/server capabilities.

4. NED in the Era of a Global Virtual Observatory

The vision of a 'virtual observatory' (VO) involves interconnected, globally distributed archives from observatories and large-scale sky surveys which are fed-

[3]The NED client C code is available at *ftp://ned.ipac.caltech.edu/pub/ned/client.3/*.

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press

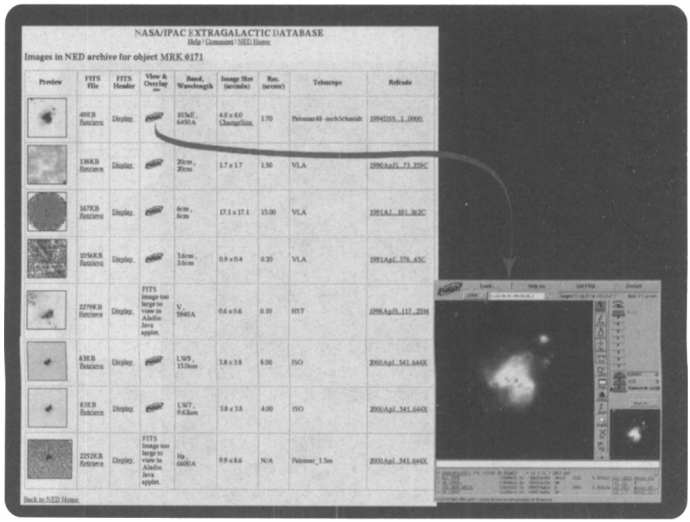


Figure 6. Multi-wavelength galaxy images, including previews (GIF) and science-grade data (FITS). Image overlays and graphical inter-activity between sky coordinates (from information in the FITS image header) and object markers (from NED and other databases) are available using the Aladin Java applet (CDS).

erated using common database query standards and data interchange protocols, combined with user interfaces and data mining tools that integrate and analyze the fused, multi-wavelength data sets to facilitate making new discoveries about the Universe (regardless of the location of the data or the investigators). A popular level description of the concept is given in an article by Cowen (2001). Helou et al. (1991) pointed out that there is a dual challenge presented by the explosive growth in astronomical data: "dealing with the sheer volume, but also inter-connecting intelligently the huge variety of information available." The NED team shares a common vision regarding what a VO can enable for all fields of astrophysics, and we are actively involved in collaborations designed to lay the foundation of the VO and extend its functionality[4]. There is much work to be done on all fronts by a broad community of astronomers teamed up with computer scientists and programmers. It is useful to summarize the role of NED in the emerging global VO in the context of current capabilities and future enhancements that will inter-operate with related projects.

[4]The U.S. National Virtual Observatory (NVO) development Web site is *http://www.us-vo.org*.

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press

4.1. Current Capabilities

NED will continue to establish and improve high fidelity relationships between multi-wavelength data with anchors to the literature for millions of extragalactic objects, using a combination of computer software that utilizes positional uncertainties and astronomical source properties, followed by human inspection to resolve important, complex cases that cannot be fully automated. In addition, the NED team is participating in the collaborative development, testing, and deployment of the next generation of catalog cross-ID software tool-kits. We will also continue to enhance NED's object-based portal into distributed data sets (Figure 3). This work will involve keeping up with evolving technology for interoperability between archive query services, primarily XML and the associated family of protocols for implementing modern Web Services[5] as they are adopted by the community. The galaxy spectral energy distributions will also continue to grow to include data from large-scale surveys (2MASS, SDSS, etc.), space missions (SIRTF, GALAX, Chandra, etc.) and of course literature articles.

4.2. Future Enhancements

In addition to staying current with the literature and extragalactic source observations in new large-scale sky surveys, the NED team is committed to providing new functionality to extend the usefulness of NED as a research tool for astronomers. The newer NED capabilities with direct relevance to the VO concept were reviewed above. Over the next few years NED plans to provide the following enhanced capabilities as resources permit: (1) development of a spectral archive for extragalactic sources, including data reports and queries that involve nuclear spectral classifications (Seyfert 1 & 2, LINER, HII/starburst, etc.) and spectroscopic line measurements; (2) enhancements to the 'By Parameters' (Advanced All-Sky) tool to support queries based on multi-wavelength flux and color criteria; (3) upgrades to the 'Batch Mode' to support larger result sets with output content and formats that can be configured by users for easy input into data analysis applications; (4) production of an XML server mode to support people who want to write software to analyze NED data or integrate query results into new VO services and tools.

5. Summary

NED provides data and relationships between multi-wavelength observations of millions of extragalactic objects. The goal is a comprehensive panchromatic census of objects in the extragalactic sky. Project activities revolve around an ongoing fusion of data from sky survey catalogs and the literature, focusing on established and candidate extragalactic sources, and maintaining cross-identifications, statistical associations, and anchors to online references and pointers. NED serves as an interface into its own database and now also as a portal for the extragalactic research community into an emerging federation of astrophysics data centers and service providers with queries indexed by object names and

[5]Namespaces, DTD, Schema, XSL, SOAP, WSDL, etc. See *http://www.w3.org.*

https://doi.org/10.1017/S0252921100031110 Published online by Cambridge University Press

coordinates synthesized by NED. As a key participant in the global Virtual Observatory (VO), NED will continue evolving the core functions that it provides today, while supporting new initiatives through use of XML standards to establish higher degrees of interoperability with other archives, collaboration in the development and application of new tools for bulk dynamic catalog cross-identifications, serving large multidimensional data streams to interface with data mining and visualization applications, and in general enabling new opportunities for discovery by leveraging information technologies.

Acknowledgments. The NED Team consists of Kay Baker, Judy Bennett, Harold Corwin, Cren Frayer, George Helou, Anne Kelly, Cheryl Lague, Joseph Mazzarella, Barry Madore, Olga Pevunova, Marion Schmitz, Brian Skiff, and Dianna Schettini. The Apache Web servers are maintained by Rick Ebert, and Informix database administration is performed by Nian-Ming Chiu. We also thank the IPAC Systems Group for their fine computer and network infrastructure support. This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA). NED is funded by the Mission Operations and Data Analysis program of NASA's Office of Space Science.

References

Helou, G. et al. 1991, in Databases and Online Astronomy, eds. M. A. Albrecht & D. Egret (Kluwer), 89

Cowen, R. 2001, Science News, 159, 124

Burstein, D. & Heiles, C. 1982, AJ, 87, 1165

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Mazzarella, J. M., & Balzano, V. A. 1986, ApJS, 62, 751