Astronomical Data Analysis Software and Systems XII
ASP Conference Series, Vol. 295, 2003
H. E. Payne, R. I. Jedrzejewski, and R. N. Hook, eds.

Federating Catalogs and Interfacing Them with Archives: A VO Prototype

Douglas J. Mink, Michael J. Kurtz

Smithsonian Astrophysical Observatory

Abstract. A common scientific requirement is to perform a joint query on two or more remote catalogs, then use the resulting combined catalog as input to query an archive or catalog. We have developed techniques which enable the routine federation of several of the largest astrometric and photometric catalogs from either in-house or remote copies, and use this federated output to query the several archives of spectral and imaging data which we either manage or maintain local copies of.

Allowing the federation of arbitrary sections of large catalogs, with user defined match criteria; and then allowing this result to be used to query several large archives of spectral and imaging data (also subject to user constraints) is a key goal of all VO projects. The problems we have solved in developing our methods will also have to be addressed by any VO project which delivers similar capabilities.

1. Federating Astronomical Catalogs

Catalogs are flat databases keyed on object positions, including other information such as magnitude or flux, spectral type, epoch, redshift, etc. Combining information from different catalogs requires knowledge about the parameters of each catalog. The VOTable standard will eventually be able to describe the contents of a catalog, but for now, since astronomers actually use relatively few large catalogs, this information can be built into software.

Iterated searching can be done through a command line catalog portal, WCSTools' **scat** (Mink 2002), which was written to access the largest astrometric and photometric catalogs, such as the HST Guide Star Catalog and the USNO-A2.0 Catalog. Originally it worked only from in-house copies of catalogs, but over the past two years, the ability to use remote copies accessed over the Internet using the HTTP protocol has been added. **scat** can be used on either end of such a connection as it understands HTTP queries and can return search results as the same tab-separates files it expects to receive over the net.

Input search lists can also be submitted as tab-separated tables, so queries can be chained. **scat** output can be formatted so that the output "catalog" includes information from both the search list and the searched catalog in each entry.

© Copyright 2003 Astronomical Society of the Pacific. All rights reserved.

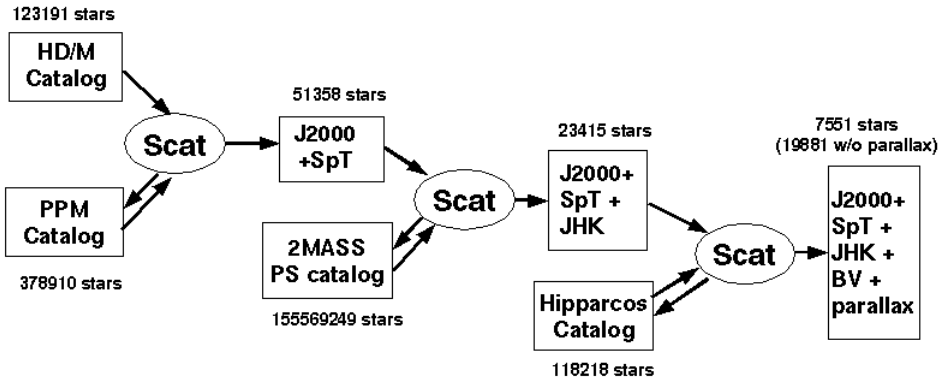

Figure 1. Matching the HD/Houk catalog (spectral class) with PPM, Hipparcos, and 2MASS catalogs gives current positions, parallax, and V and K magnitudes.

2. Combining Spectral Classes and Infrared Colors

The best star catalog for spectral types and classes is the Henry Draper catalog as updated by Houk et al. (1975-1999), but the original HD B1900 positions are kept, and the best photometric catalogs, such as the 2MASS PSC, the GSC II, and the USNO-A2.0 are in J2000 coordinates. WCSTools **scat** can convert between coordinate systems, epochs, and equinoxes, producing a final catalog with visible and IR magnitudes and spectral types. Figure 1 shows how **scat** was used iteratively by searching the PPM catalog to get 2000 positions, the 2MASS Point Source Catalog to get JHK infrared magnitudes, and the Hipparcos Catalog (ESA 1997) to get distance and V magnitudes. Figure 2 shows where stars of luminosity class V show up among the stars in the merged catalog.

3. Serving the SAO Spectral Archives

Archive access can be implemented as a web service using legacy software and providing catalog cross-references After archiving uniformly processed spectra taken on our ground-based telescopes for over 20 years, SAO now has over 215,000 high dispersion and 110,000 low dispersion spectra archived. The Smithsonian Institution is chartered to promote the "increase and diffusion of knowledge," and that has been interpreted to mean that our data should be made as public as possible. 35,000 public spectra from two instruments are currently online, covering many interesting objects in the northern sky, and more are coming. Figure 3 shows how a user interface to the Updated Zwicky Catalog (Falco et al. 1999) archive allows searching a catalog of spectra (using the ubiquitous **scat** program through a Perl script). Once the desired spectra are located, selected spectra can be displayed and downloaded as either FITS or ASCII files.

The Virtual Observatory has been looked at as a hierarchy, with archives at the bottom and portals at the top, but that is only one way to view it. Archives can link to each other and back to the portals as well as acting as browsers into

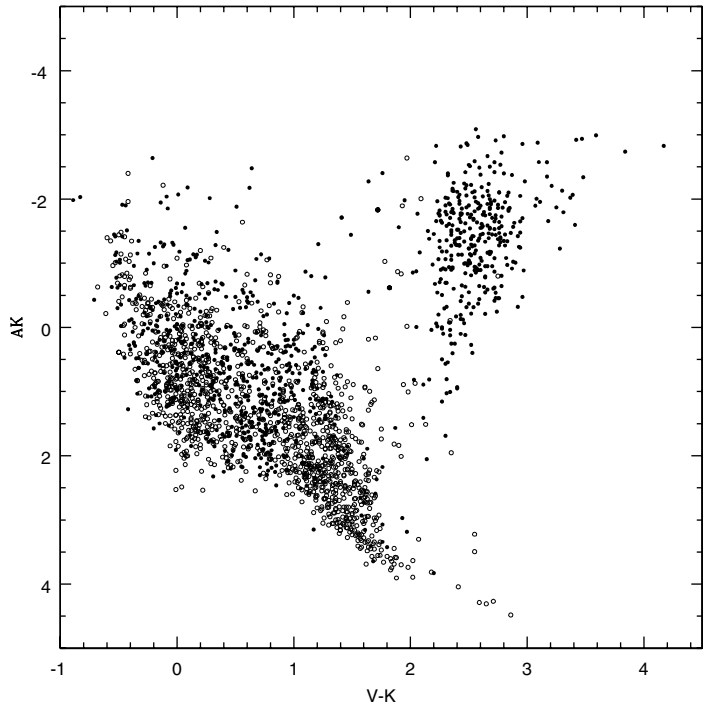

Figure 2. Absolute K magnitude of merged catalogs plotted against (V-K) color. Circles are luminosity class V stars; filled dots are other classes.

their own data. Our spectral portal uses SIMBAD and NED to resolve named astronomical objects into coordinates, and they in turn each link from their entries for the Updated Zwicky Catalog to our spectral archive. Our displays for individual spectra act as single object portals into the virtual observatory, with links back to NED and SIMBAD, as well as the Digitized Sky Survey, and other spectra in our archives.

4. Conclusions

Allowing input and output catalogs to be identically formatted makes merging and comparison easy. Archives should have user-friendly interfaces for browsing their data, and catalog-centric methods work well. The Virtual Observatory is a web of information; connections to and from existing portals are important. Access to the entire WCSTools package is at

`http://tdc-www.harvard.edu/software/wcstools/`.

Acknowledgments. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

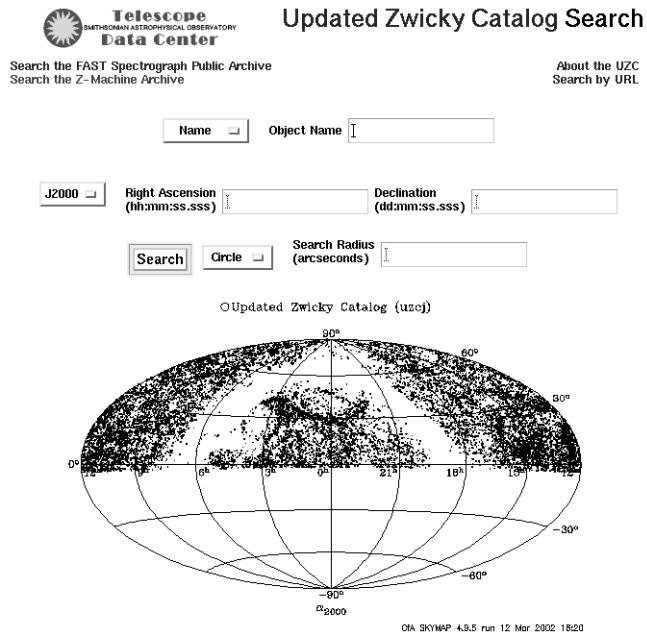

Figure 3. User interface to the Updated Zwicky Catalog spectra.

References

European Space Agency 1997, The Hipparcos and Tycho Catalogues, ESA SP-1200

Falco, E. E., Kurtz, M. J., Geller, M. J., Huchra, J. P., Peters, J., Berlind, P., Mink, D. J., Tokarz, S. P., Elwell, B. 1999, PASP, 111, 438

Houk, N. & Cowley, A. P. 1975, Michigan Catalogue of two-dimensional spectral types for the HD stars, Ann Arbor: Dept. of Astronomy, Univ. Michigan

Houk, N. 1978, Michigan Catalogue of two-dimensional spectral types for the HD stars, Vol. 2, Ann Arbor: Dept. of Astronomy, Univ. Michigan

Houk, N. 1982, Catalogue of two-dimensional spectral types for the HD stars, Vol. 3, Michigan Spectral Survey, Ann Arbor: Dept. of Astronomy, Univ. Michigan

Houk, N. & Smith-Moore, M. 1988, Michigan catalogue of Two-Dimensional spectral types for the HD stars, Vol. 4, Michigan Spectral Survey, Ann Arbor: Dept. of Astronomy, Univ. Michigan

Houk, N. & Swift, C. 1999, Michigan catalogue of two-dimensional spectral types for the HD Stars, Vol. 5, Ann Arbor: Dept. of Astronomy, Univ. Michigan

Mink, D. J. 2002, in ASP Conf. Ser., Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley (San Francisco: ASP), 169