Telescope Science Institute (STScI); and the University of Maryland. We shall show how a positional query for astrophysical data in a region of arbitrary geometrical boundary can be carried out using these basic components. We shall also describe a scheme by which user software can be deployed to a data center to extend its services, and how the system will return to the researcher only the desired scientific results. This capability is very important for multispectral studies using the large all-sky surveys that reside in distributed data archives.

116.02

XDF - an XML-based Scientific Interchange Language

E. Shaya, B. Thomas, J. Gass, J. Blackwell (Astronomical Data Center/ GSFC-NASA, RITSS), C. Cheung (Astronomical Data Center/GSFC-NASA)

We present the basics of the eXtensible Data Format (XDF), an XML-based scientific data interchange language developed at the Astronomical Data Center (NASA). XDF addresses common needs of scientific data, and provides a generic wrapper around most existing data files (see related poster paper "An XML Representation of FITS").

The goal of XDF is to serve as an interchangeable nucleus for various sub-field specific scientific languages (as may be used by NVO). The features of XDF include: n-dimensional tables, complex hierarchical structures, a data model that fuses axis information with the data structure, referencing mechanism for tiling together large arrays or equating axes in separate arrays, automated unit conversion (see related poster paper "Generating Basic Units with XML"), and browser readiness. Both Java and Perl software packages are available for ease of integration with existing software.

116.03

An XML Representation of FITS - Introducing FITSML

B. Thomas, E. Shaya, J. Gass, J. Blackwell (Astronomical Data Center/ GSFC-NASA, RITSS), C. Cheung (Astronomical Data Center/GSFC-NASA)

We describe a means by which FITS data may be encapsulated in XML using the eXtensible Data Format (XDF) - an XML based language for encapsulating n-dimensional scientific data. Our goal is to re-map (rather than to redefine) the FITS standards into an XML format. The advantages include: greater interoperability, parsing by XML aware browsers and applications, hierarchical structure for improved searchability, automatic validation, default values for header descriptions, extensibility for specialized usage and future development, and piggybacking on industry applications. We present our current work including a DTD and examples and discuss complexities, choices and future work.

116.04

Generating Basic Units with XML

J. Gass, E. Shaya, B. Thomas, J. Blackwell (Astronomical Data Center/ GSFC-NASA, RITSS), C. Cheung (Astronomical Data Center/GSFC-NASA)

A fundamental characteristic of any scientific data is its physical units. Unit comparisons and conversions are required in order to understand the meaning of the data and whether or not certain operations on the data are meaningful. For example, only data with equivalent units should be merged.

Intelligent query of distributed systems, such as is envisioned for the NVO, will need to integrate or merge data with mixed conventions of units. It would be desirable to provide a standard process to enable machine understanding of units. We present a solution to this problem using XML entities. Our solution features a flexibility to encompass current unit systems, ease of human use, and interoperability (XML/MathML based) between heterogeneous operating systems.

116.05

Current Status of the Astrophysics Source Code Library

P. Helbig (Kapteyn Instituut), R. J. Nemiroff (Michigan Tech)

The Astrophysics Source Code Library at http://ascl.net currently lists over 40 codes of value to astronomers and astrophysicists. These codes are all available for free download over the internet. Founded in 1999, the editors continue to invite scientists that write source code in support of small and large research projects to submit their source code – or a link to their source code – to ACSL.net.

116.06

Accessing, Mining, and Archiving an On-line Database – The APS Catalog of the POSS I

R. M. Humphreys (University of Minnesota), J. E. Cabanela (St. Cloud State University), J. Kriessler (University of Minnesota)

The APS Catalog of the POSS I is an on-line database of over 100 million stars and galaxies (http://aps.umn.edu). A unique subset of this database with over 218,000 galaxies within 30 degrees of the North Galactic Pole, the MAPS-NGP, is now available at our web site. This diameter–selected catalog ($\geqslant$ 10 arcsec) is the deepest galaxy catalog constructed over such a large area of the sky (3000 sq. degrees). The MAPS-NGP includes many additional parameters for the galaxy images not available in the APS Catalog.

Working with members of our computer science department, we have developed a morphological classifier for galaxies that divides our galaxy type into three classes – early, intermediate, and late. We have applied data mining techniques to identify the most useful image parameters for input into a neural network and decision–tree based classifier pipeline.

We are also archiving the APS Catalog for distribution to astronomical data centers including NASA's ADC and SIMBAD at CDS. The extragalactic subset will be integrated into the NASA/IPAC extragalactic database(NED). The MAPS-NGP has already been provided to NED.

The APS is supported by NASA's Applied Information Systems Research Program.

116.07

Color-Space Outliers in DPOSS: Quasars and Peculiar Objects

S.G. Djorgovski, R.R. Gal, A. Mahabal, R. Brunner, S.M. Castro (Caltech), S.C. Odewahn (Caltech/ASU), R.R. de Carvalho (ON Brasil), DPOSS Team

The processing of DPOSS, a digital version of the POSS-II sky atlas, is now nearly complete. The resulting Palomar–Norris Sky Catalog (PNSC) is expected to contain $> 5 \times 10^7$ galaxies and $> 10^9$ stars, including large numbers of quasars and other unresolved sources. For objects morphologically classified as stellar (i.e., PSF-like), colors and magnitudes provide the only additional source of discriminating information. We investigate the distribution of objects in the parameter space of $(g - r)$ and $(r - i)$ colors as a function of magnitude. Normal stars form a well-defined (temperature) sequence in this parameter space, and we explore the nature of the objects which deviate significantly from this stellar locus. The causes of the deviations include: non-thermal or peculiar spectra, interagalactic absorption (for high-z quasars), presence of strong emission lines in one or more of the bandpasses, or strong variability (because the plates are taken at widely separated epochs). In addition to minor contamination by misclassified compact galaxies, we find the following: (1) Quasars at $z > 4$; to date, $\sim$ 100 of these objects have been found, and used for a variety of follow-up studies. They are made publicly available immediately after discovery, through `http://astro.caltech.edu/~ george/z4.qsos`. (2) Type-2 quasars in the redshift interval $z \sim 0.31 - 0.38$. (3) Other quasars, starburst and emission-line galaxies, and emission-line stars. (4) Objects with highly peculiar spectra, some or all of which may be rare subtypes of BAL QSOs. (5) Highly variable stars and optical transients, some of which may be GRB "orphan afterglows." To date, systematic searches have been made only for