Astronomical Data Analysis Software and Systems XV **P160**
ASP Conference Series, Vol. XXX, 2005
C. Gabriel, C. Arviset, D. Ponz and E. Solano, eds.

Search of distant radio galaxies as a subject mediator example

Zhelenkova Olga P.

SAO RAS, Niznij Arkhyz, KChR, Russia, 369167

Briukhov Dmitrij O.

IPI RAN, Moscow, Russia

Kalinichenko Leonid A.

IPI RAN, Moscow, Russia

Vitkovsky Vladimir V.

SAO RAS, Niznij Arkhyz, KChR, Russia, 369167

Abstract. We try to resove the problem of efficient integrated representation of multiple astronomical resources with help of creation of mediators supporting interaction between a researcher and relevant data sources and services through a subject domain description for a class of tasks. The Big Trio project is carried out in SAO RAS headed by the academician Parijskij Yu. N. The main project goal is an investigation of radio sources in the sky strip obtained in the "Cold" deep survey with the RATAN-600 in 1980 and getting maximal information about them. Distant galaxy candidate choice from radio source lists and catalogs is derived with tested selection methods by certain radio source parameters. An application of subject mediator is considered on the example of searching of distant radio galaxies with selection method applying in the project.

1. Approach to integration of distributed resources

The widening gap between the scientists and the sources of the data requires a major paradigm shift in the way of scientific problems solving over multiple large distributed data sources and services (that are concentrated in specialized centers of data and computational facilities). Various technical infrastructures are intended for this way (such as Web-services, Grid architectures, middleware frameworks). In spite of that, the problem of efficient integrated representation of multiple sources for a researcher remains to be open. Two principally different approaches to this problem exist: moving from multiple sources to a researcher and problems (in this case an integrated representation of multiple sources is created independently on the tasks) and from a problem to sources (in this case a subject domain definition for a class of problems is created, the relevant sources are identified and mapped into this definition). The first approach is not scalable. For example, currently in astronomy the number of existing sources (catalogs, archives) has an order of many thousands. The integrated schema of a set of sources is to be modified each time when a new source appears. Another approach assumes creation of mediators supporting interaction between a researcher and relevant data sources and services through a subject domain description for a class of problems.

Definition of a subject domain in terms usual for researchers should serve as a base (an abstraction) for the mediator specification independent of existing data sources and services. Subject mediators are emphasized that support representation and access to various subject domains in astronomy. They lead to more knowledge-based organization. Based on such abstractions, an intermediary layer is formed by subject mediators providing metainformation uniformly characterizing their subject content. A canonical information model is used for the mediator definitions making possible to query such abstract content and compute the result. Introducing of such mediators frees the researchers from having to identify relevant entities among multiple heterogeneous sources and services defined in non-uniform terms, semantically reconcile and correlate them, formulate research tasks in unusual terms, structures, functions and processes. Each mediator supports the process of systematic discovery and registration of sources uniformly expressing their definitions in terms of the mediator (Briukhov et al. 2001). The mediator's canonical information model should be powerful enough for uniform equivalent and complete representation of various data sources and services.

Two separate phases of the mediator's life cycle are distinguished: consolidation phase and operational phase. Consolidation phase is intended for creating a definition of the mediator. A consensus in the subject community should be reached on the mediator ontology, concepts, data structuring and behaviors to consolidate definitions of the subject. During the operational phase the sources relevant to a subject mediator are dynamically registered in it so thatand queries (Kalinichenko et al. 2004) and tasks can be executed. The source capabilities (ontologies, types, functions, query language capabilities, etc.) are expressed in terms of the subject mediator's metainformation and respective wrappers are developed (if required). Subject mediator architecture is planned as a part of the Russian Virtual Observatory Information Infrastructure (Briukhov et al. 2005).

2. Example of subject mediator for distant radio galaxy search

For an example of a subject mediator implementation we consider a distant radio galaxy search problem. During several years the Big Trio project is carried out in SAO RAS headed by the academician Parijskij Yu. N. The main project goal is a distant radio galaxy search in the sky strip investigated in the Cold deep survey with the RATAN-600 in 1980 and getting maximal information about the objects. Distant galaxy candidate choice from radio source lists and catalogs is derived with tested selection methods by certain radio source parameters. Similarly to the other research groups involved into the same problems, the following parameters of radio sources were used:

- a slope of radio source spectrum;
- flux density level;
- morphological type;
- angular size.

Radio maps, accurate coordinates for optical identification, morphology and angular sizes were obtained by VLA for the USS subsample of the RC catalog. Special observations were performed for the sources or their images were fetched from the FIRST and NVSS radio surveys. All objects were identified by the 6-m telescope CCD-images. The reliability of radio/optical identification is a difficult problem depending on an astrometric precision of two data sets and structure of optical candidates and radio source morphology. Photometric redshift estimations and ages of galaxy stellar population were carried out with Spectral Energy Distribution (SED) models and the 6-m telescope B, V, R, I observations (Verkhodanov et al. 2001). The worflow of searching of radio galaxies with steep spectra consists of the following steps:

1. select radio sources with spectral index in a given range [i1,i2];
   - select radio sources containing flux data for at least 2 frequencies;
   - calculate spectral index i for selected radio sources;
   - get a list of radio sources with spectral index in the range [i1,i2];
2. select radio sources with the angular size less than a given value d1 and flux value for a given frequency f1 in the range [s1,s2];
3. find optical sources matching by coordinates (admissible w.r.t. coordinate and size of sources tolerances) with selected radio sources;
4. get a list of radio sources for which the matching optical sources were found;
5. on a user request, for a given pair a radio source and matching optical source:
6. show optical image of the source with over plotted intensity contours of radio image;
7. request user confirmation of matching of the selected radio and optical sources;
8. calculate stellar magnitudes of the selected candidates at least in U, B, R filters;
9. calculate (U-B) and (B-R) values for each candidates;
10. plot colour - colour diagram to define more exactly the source list;
11. get a list of candidates for distant galaxies.

The subject mediator's data model is worked for this task (Figure 1). Initial set of catalogs and surveys to be registered at the mediator includes: RC catalog of radio sources; FIRST[1] survey and catalog of radio sources; NVSS[2] survey and catalog of radio sources; SDSS[3] survey; SAO RAS archive direct images in U, B, R filters. During the registration a resource class is modeled as a set of instances (objects) of the mediator class instance type and the description of the source in terms of the mediator schema specifies the constraints on the class instances

---

[1] `http://sundog.stsci.edu/top.html`

[2] `http://www.cv.nrao.edu/nvss/`

[3] `http://cas.sdss.org/`

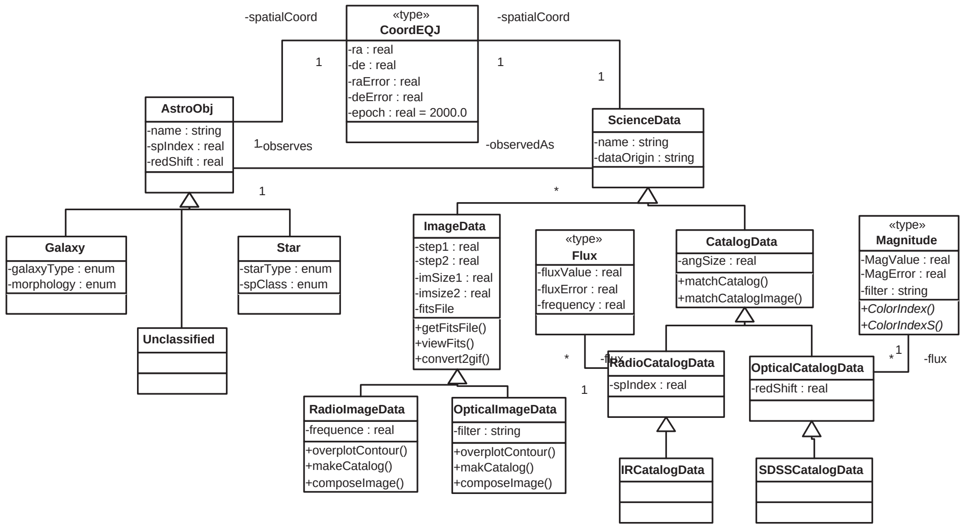

Figure 1. The data model of the subject mediator.

to be admissible for the subject mediator. The registration process includes: ontologically-based (UCD-based) reconciliation of the application contexts of the registered resource and that of the mediator; expression of each resource class in terms of the mediator classes. Using UCDs, we can establish relevance between elements in the mediator and resource schemas. If for a mediator attribute there is no directly relevant a resource attribute, then we construct, if possible, a concretizing (or conflict resolving) function. This function defines how a mediator attribute can be expressed as a function of the resource attributes. We have completed the first iteration of the mediator consolidation phase.

Acknowledgments. The work is supported by the Russian Fund of Basic Researches (grants 03-07-09032, 05-02-27253, 05-07-90413).

References

Briukhov, D.O., Kalinichenko, L.A. & Skvortsov, N.A. 2001, In Proc. of the ADBIS-2001, Vilnius, Lithuania, Springer, 2151, 70

Kalinichenko, L.A., Martynov, D.O. & Stupnikov, S.A. 2004, In Proc. of the ADBIS-2004, Hungary, Budapest, Springer, 3255, 37

Briukhov, D.O., Kalinichenko, L.A., Zakharov, V.N., Panchuk, V.E., Vitkovsky, V.V., Zhelenkova, O.P., Dluzhnevskaya, O.B., Malkov, O.Yu., Kovaleva D.A 2005, Information Infrastructure of the Russian Virtual Observatory (RVO). IPI RAN, Moscow, 173 pp.

Verkhodanov, O. V., Parijskij, Yu. N., Soboleva, N. S., Kopylov, A. I., Temirova, A. V., Zhelenkova, O. P., Goss, W. V. 2001, Bulletin of the SAO RAS, 52, 5