Astronomical Surveys and Big Data
ASP Conference Series, Vol. 505
A. M. Mickaelian, A. Lawrence, & T. Yu. Magakian, eds.

Armenian Virtual Observatory: Services and Data Sharing

Aram V. Knyazyan,[1] Hrachya V. Astsatryan,[1] and Areg M. Mickaelian[2]

[1]*NAS RA Institute for Informatics and Automation Problems (IIAP), Yerevan, Armenia;* `aram.knyazyan@ipia.sci.am,hrach@sci.am`

[2]*NAS RA V. Ambartsumian Byurakan Astrophysical Observatory (BAO), Byurakan 0213, Aragatzotn Province, Armenia;* `aregmick@yahoo.com`

Abstract. The main aim of this article is to introduce the data management and services of the Armenian Virtual Observatory (ArVO), which consists of user friendly data management mechanisms, a new and productive cross-correlation service, and data sharing API based on international standards and protocols.

1. Introduction

The generated data volumes in modern astronomical sky surveys can range from several terabytes to petabytes. In some cases such surveys every day may generate terabytes of data. Astronomical data volume and complexity require a joint effort both of astronomers and IT specialists to develop state-of-the-art infrastructures satisfying their needs, such as Astronomical Virtual Observatories (VO) (Hearnshaw et al. 2006; Knyazyan 2011a).

VO is a platform for launching astronomical investigations: it provides access to huge data banks, software systems with user-friendly interfaces for data processing, analysis and visualization, and even access to computers on which the work can be carried out. VOs over the world are seamlessly networked, and their resources can be accessed over the internet by astronomers regardless of their location, expertise and the level of access to their own advanced computing facilities. Due to their nature, VOs make an immense impact on the way astronomy is done in the developing world. There are several VO projects in operation in different countries in the world. The biggest projects include the National Virtual Observatory of the USA, the European Virtual Observatory, which brings together many European countries, and AstroGRID, which is a VO project based in the UK.

The Armenian Virtual Observatory project (ArVO; http://arvo.sci.am) is being developed since 2005. It is a project of the Byurakan Astrophysical Observatory (BAO; http://www/bao.am) and the Institute for Informatics and Automation Problems (IIAP; http://iiap.sci.am) aimed at the deployment of a modern virtual environment to tackle data management challenges.

The Digitized First Byurakan Survey (DFBS; http://byurakan.phys.uniroma1.it) is the one of the key contribution of ArVO into the International Virtual Observatory Alliance (IVOA; http://www.ivoa.net). The cooperation between BAO and IIAP enabling to develops the ArVO portal, which in case of necessity may use the computational resources of the Armenian National Grid Infrastructure (ArmGrid) (Astsatryan et al. 2015a).

The remaining content of the paper is organized as follows. The introduction of data management component is presented in Section 2 and the cross-correlation service in details can be found in Section 3. Finally, the conclusion and directives for future research are drawn in Section 4.

2. Data Management

ArVO data management service is providing a way to manipulate tables of astronomical data in a uniform way. During previous years Byurakan astronomers have done a huge work to digitalize the plates of past years. These FBS data is the core of the ArVO database. A service has been developed allowing authorized users to upload the files remotely via Web interface or via special scripts from access node. All real data is stored in the dedicated storage element of the ArmGrid infrastructure and the metadata in the database.

ArVO Portal provides a data search engine to find images by coordinates, search-area radius and EQUINOX. The result of search engine includes some basic parameters of a requested image like observation date, coordinates, image size, etc. (Fig. 1). Then image can be viewed through the web more detailed, or user can download Flexible Image Transport System (FITS) file of image for more detailed research. User can make a gallery of featured images and download it as one zipped file (Astsatryan et al. 2010).

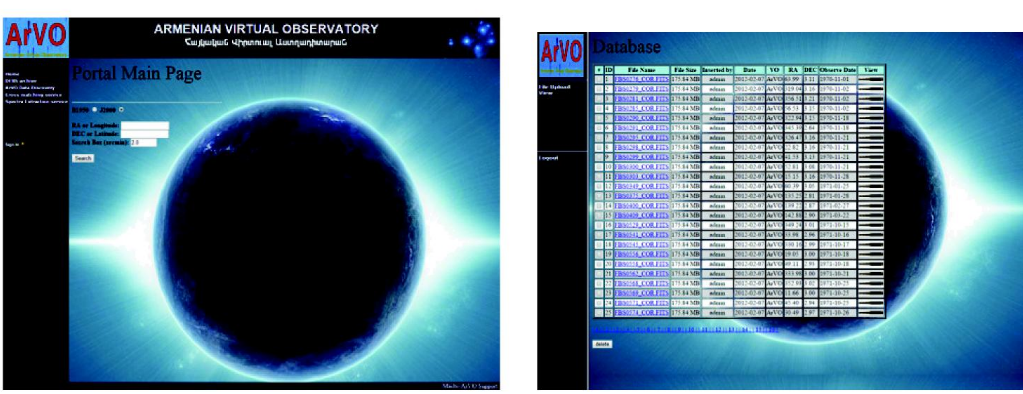

Figure 1. ArVO web portal *Left:*data search engine. *Right:* data search result.

API service for astronomical data sharing has been developed based on IVOA Simple Image Access Protocol (SIAP; http://www.ivoa.net/documents/SIA/20150730/PR-SIA-2.0-20150730.pdf) standard. It gives possibility to make a search request to ArVO portal using standard http, and as a response get astronomical metadata in VOTable format.

Different required and optional parameters for search criteria can be passed through http request, such as POS (RA and DEC coordinates), SIZE (search-area radius), image format and size, naxis, etc. The result of SIAP search is in VOTable format, and includes all astronomical metadata of found images, their physical address URLs and so on. SIAP protocol is used by lots of astronomical search engines, and makes the integration and data sharing simple and easy. An example of SIAP request to ArVO database is http://arvo.sci.am/api/search.php?pos=322.3,3&size=1 (see on Fig. 2 the result of this request).

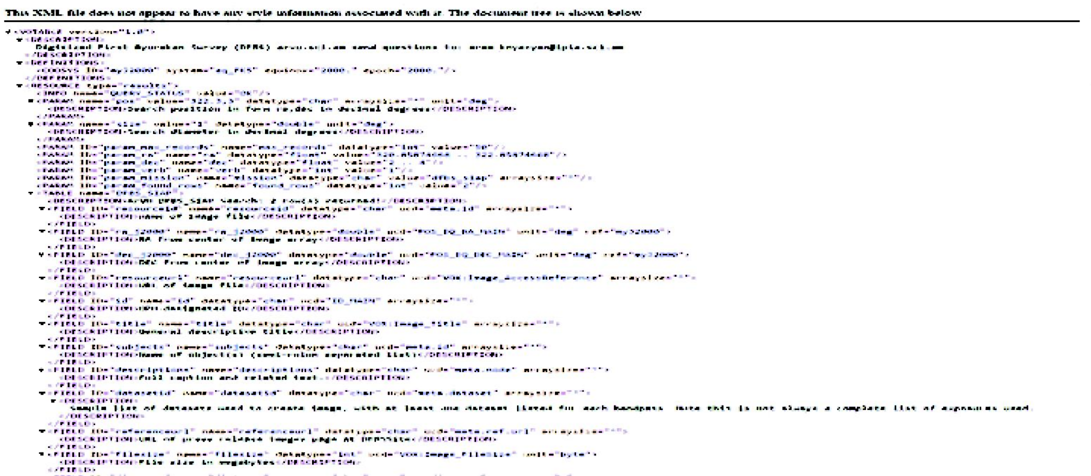

Figure 2. Result of SIAP request to ArVO.

3. Cross-correlation Service

Discovery of variable and proper motion stars (Knyazyan et al. 2011b), as well as other space objects is one of the most important objectives in astronomy. It is possible through comparison, i.e. cross-correlation of the astronomical catalogs resulting from different observations of the same part of the space.

Cross-correlation of two catalogs is a comparison of coordinates line by line. The catalog's each line coordinate is correlated to the second catalog?s all lines coordinates and if the difference of the values of the coordinates is smaller than a prior determined certain value, then it can be concluded that the same object is present in both the first and the second catalogs. As for those lines, which don't have conjunction, deserve serious exploration. In case there is no conjunction, then that object is a new object, a proper motion star or an asteroid. However, it isn't always applicable, since there is possibility that the mentioned object emerges just as a ?random signal?, which can be an error of the extraction making program or an error arising from the digitization of the plate, etc. The selection of the value of comparison of the radius is also of a very high significance.

In any catalog the coordinates of objects are in some way deviated from their real value. This deviation in the catalogs is presented by the parameter of the coordinate root-mean-square (RMS) error. Its value is different for various objects. Moreover, the RMS has a value both for right ascension and declination, which means that we can specify both RMS_x and RMS_y errors. On Fig. 3 the RMS of a star is presented by the example of a DSS plate.

On the right, there is the real position of the star and on the left is the position of the star on the plate and hence, as a result of the extraction, the coordinate of that position is specified in the catalog, which is evicted from the real one by sizes RMS_x and RMS_y accordingly.

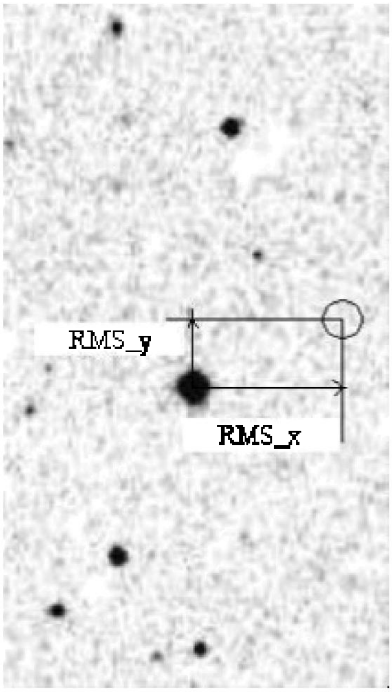

Figure 3. RMS of a star presented by the example of a DSS plate.

Thus, the coordinates of objects in the catalogs, created in the result of reviewing by the observatories of the same scope of the space are expressly different. Furthermore, these deviations are different for various objects. For this reason the selection of the tracing radius value during the cross-correlation is an important issue, since it should be as big to exceed the aggregate deviation of two objects with the maximal RMS. But in this case a new problem is appearing. When we select big search radius, in the result of cross-correlation for some objects the number of objects in the dense fields will be more than one.

Hence it can be concluded that during the cross-correlation it isn?t efficient to select the same search radius for all the objects, since it brings to miscalculations. The cross-correlation will be more efficient, if for each object the search radius is comparable to its RMS. In that case the above described problems will not rise, which means that the probability of the exact conjunction of objects will not reduce and for one object the number of conjunct objects from the other catalog will not be such significant as during the cross-correlation by the same search radius and it will be reasonably optimal.

A new algorithm has been suggested for astronomical catalogs cross-correlation, when search radius for each pair of objects is comparable to RMS. In order to cross-correlate two catalogs it is required to count mean values of RMS errors for each two correlated objects from different catalogs (Knyazyan et al. 2011c).

A new software package has been developed, which gives possibility to correlate two astronomical catalogs by using as the search radius for each two objects their mean RMS values multiplied with some "k" constant. The experiments have shown that optimal value for constant is k=3: That is, search radius value equals 3*RMS, but it is possible to use other values for "k" depended on current objective.

A new software package has been developed, which gives possibility to correlate two astronomical catalogs by using as the search radius for each two objects their mean

RMS values multiplied with some "k" constant. The experiments have shown that optimal value for constant is k=3: That is, search radius value equals 3*RMS, but it is possible to use other values for "k" depended on current objective.

The computational resources of ArmNGI are used for cross-correlation by dividing the process into several processes, and each process runs on one CPU/Core, which decreases the correlation time.

According to the algorithm the requirements for catalogs are the inputs for the software input. Catalogs have to contain objects RA and DEC coordinates in degrees, as well as their RMS_x and RMS_y error values in milli-arcseconds. The program output includes two catalogs. One of them contains rows of matched objects, and the other the rows of not matched objects. Each row contains matched objects information from their root catalogs and in addition distance between matched objects (by x, y and total), percent of distance value from RMS, etc. If for an object there are more than one matching, it is possible by using an option to keep only one matching, in which case the distance from two objects will be the smallest, so the optimal matching will be found.

Byurakan astronomers test new cross-correlation algorithm, and the results show that it is more productive than classic cross-correlation algorithm (Abrahamyan et al. 2012, 2015). The cross-correlation service is available in ArVO Portal (see Fig. 4).

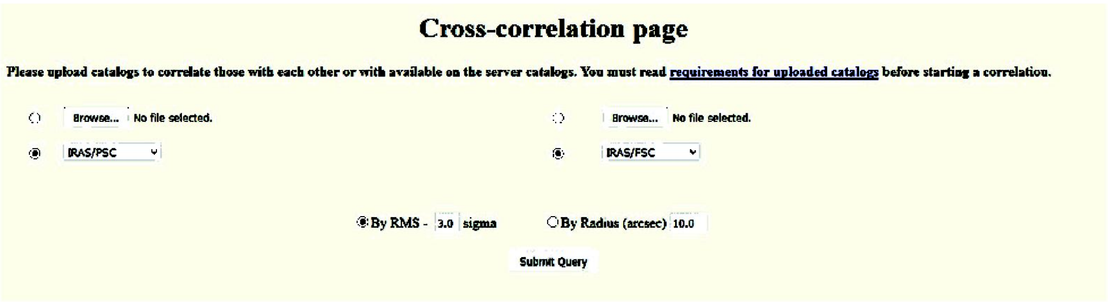

Figure 4. ArVO cross-correlation webpage.

4. Conclusion

The usage of computational and storage resources of ArmNGI and international standards based on SIAP allow ArVO to provide users optimized services. Particularly, the ArmNGI data storage handles the huge amount of astronomical data and high performance computational resources (Astsatryan et al. 2004) decreases the calculation time.

In the future we aim at finishing the development of the portal, digitalizing more astronomical local data, finalizing ArVO SIAP API and integrating it in known astronomical search engines, applying IVOA standards in advance and increasing the number of researchers making use of the Grid infrastructure, and to support them. A federated identity in the Earth Science domain will be developed to provide single sign on services for earth scientists and stakeholders by combining all the available services, such as (Astsatryan et al. 2015b,c).

References

Abrahamyan H., Mickaelian A., Knyazyan A., Harutyunyan G. 2012, Joint Catalogue of IRAS PSC/FSC – Proc. of Young Astronomers Conference of CIS Countries, Yerevan, Armenia, p. 223-229

Abrahamyan H., Mickaelian A., Knyazyan A. 2015, The IRAS PSC/FSC Combined Catalogue, Astronomy and Computing 10, p. 223-229

Astsatryan H., Shoukourian Yu., Sahakyan V. 2004, The ArmCluster Project: Brief Introduction, Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA '2004), Volume III, ISBN: 1-932415-25-4, Las Vegas, Nevada, USA, CSREA Press, June 21-24, 2004, p. 1291-1295

Astsatryan H., Knyazyan A., Mickaelian A., Sargsyan L. 2010, Web portal for the Armenian virtual observatory based on Armenian national grid infrastructure - Parallel computing and control problems, Moscow, p. 109-114

Astsatryan H., Sahakyan V., Shoukourian Yu., Cros P.-H., Dayde M., Dongarra J., Oster P. 2015a, Strengthening Compute and Data intensive Capacities of Armenia, IEEE Proceedings of 14th RoEduNet International Conference - Networking in Education and Research (NER'2015), Craiova, Romania, pp. 28-33, September 24-26 2015

Astsatryan H., Hayrapetyan A., Nersisian W., Asmaryan S., Muradyan V., Guigoz Y., Giuliani G., Ray N. 2015b, An Interoperable Cloud-based Scientific Gateway for NDVI Time Series Analysis, Elsevier Computer Standards & Interfaces, 31(40), doi: 10.1016/j.csi., p. 79-84

Astsatryan H., Petrosyan Z., Abrahamyan R., Shahnazaryan A., Sahakyan V., Shoukourian Y., Melkonyan H., Kotroni V. 2015c, WRF-ARW Model for the Prediction of High Temperatures in South and South East Regions of Armenia, 11th IEEE eScience Conference, Munich, Germany, 31 August - 04 September, 2015, p. 207-213

Hearnshaw J., Martinez P. 2006, Astronomy for the developing world - IAU Special Session, p. 309-319

Knyazyan A. 2011a, Astrophysical Virtual Observatories and Armenian Virtual Observatory Development - Vestnik, vol. 14, Yerevan, Armenia, p. 111-119

Knyazyan A., Mickaelian A., Astsatryan H. 2011b, Astronomical plates spectra extraction objectives and possible solutions implemented on Digitized First Byurakan Survey (DFBS) images - Information Theories and Applications, p. 243-247

Knyazyan A., Mickaelian A., Astsatryan H. 2011c, Astronomical Catalogs Cross-Correlation Objectives and Illustration of a New Correlation Algorithm - Proceedings of CSIT2011 Conference, Yerevan, Armenia, p. 395-397