# Astronomical Surveys, Catalogues, Archives, Databases and Virtual Observatories

A. M. Mickaelian*, H. V. Abrahamyan, G. M. Paronyan, and G. A. Mikayelyan

NAS RA V. Ambartsumian Byurakan Astrophysical Observatory (BAO), Byurakan 0213, Aragatzotn Province, Armenia

**Abstract**

Astronomical surveys and catalogs are the main sources for the discovery of new objects, both Galactic and extragalactic. Archives and Databases maintain billions of astronomical objects; planets, comets, stars, exoplanets, nebulae, galaxies, and quasars. We will review the current background in astronomy for further all-sky or large-area studies. Modern astronomy is characterized by multiwavelength (MW) studies (from gamma-ray to radio) and Big Data (data acquisition, storage, mining and analysis). Present astronomical databases and archives contain billions of objects observed at various wavelengths, and the vast amount of data on them allows new studies and discoveries. Surveys are the main source also for accumulation of observational data for further analysis, interpretation, and achieving scientific results. We review the main characteristics of astronomical surveys (homogeneity, completeness, sensitivity, etc.), compare photographic and digital eras of astronomical studies (including the development of wide-field observations), and describe the present state of MW surveys. Among others, Fermi-GLAST, INTEGRAL (gamma-ray), ROSAT, Chandra, XMM (X-ray), GALEX (UV), DSS1/2, SDSS, Hubble, Gaia (optical), 2MASS, IRAS, AKARI, WISE, Herschel (IR), NVSS and FIRST (radio) surveys and major astronomical archives and databases will be presented and discussed, as well as surveys and databases for variable and transit objects.

***Keywords:*** *astronomical surveys – multiwavelength astronomy – catalogs – archives – databases – Virtual Observatories – Big Data – Data Science*

## Astronomical Surveys and Catalogs

Astronomical surveys and catalogs are the main source for astronomical data, for discovery of astronomical objects and various parameters related to them. Surveys are quite different and have a number of parameters defining their tasks and needs. For an astronomical survey, it is most important to define the task and corresponding parameters. Main parameters of astronomical surveys:

- **Wavelength range.** You find different objects at various wavelengths (from gamma-rays to longest radio), depending of what astronomical objects radiate at what wavelengths and the strength of their radiation. As this depends on the mechanisms of radiation and for the thermal one, mainly the temperature (as well as some other parameters), we find different objects and physical conditions, hence depending of what we search, corresponding wavelengths should be used.

- **Method.** Methods may be direct imaging, photometry, spectroscopy, variability, polarimetry, etc., also depending on which of them efficiently reveal what we search for.

- **Sensitivity.** The sensitivity of the survey (**limiting magnitude** in optical range) is the parameter defining the deepness of the survey and the number of expected objects, i. e. we may select between relatively surface studies and deep ones.

- **Resolution.** The resolution is another limitation depending on the task of the survey; if we need to have faster collection of well resolved objects or detailed studies of small features. Two types of resolution may be regarded; **spatial** (for resolving of close objects or morphological details) and **spectral** (for resolving close spectral lines and other features).

*aregmick@yahoo.com


doi:https://doi.org/10.52526/25792776-22.69.2-179

- **Sky area.** Ground based surveys have always been limited to the sky area depending on the geographical location of the telescopes, mainly geographical latitude but also very often the longitude. Space telescopes do not have such limitations. Nevertheless, any survey is solving some task and for this, definite sky area is being used.

- **Coverage.** Surveys are being accomplished at various surfaces; from very small deep surveys (ex. 2 arcmin by 2 arcmin) to large area (thousands of sq. degrees) and all-sky ones (mostly from Space observatories). Typically, the coverage is somehow anticorrelated to the sensitivity, as the capabilities of accumulating data are limited; the deeper the survey, the smaller the coverage and vice-versa, the larger the coverage the smaller the sensitivity.

- **Time domain.** The Universe is changing and observations give data for the definite moment (epoch, given the astronomical changes are slower). Time domain astronomy appeared after accumulation of many observations at different epochs during dozens of years. It is important to compare data from different time domains and obtain understanding on their variability. Some objects have very fast variability, so that even seconds are important. In such cases fast photometers are needed.

- **Completeness.** This is one of the most important parameters of the surveys, as the observers should provide an understanding on what can be expected from the point of view of detection and the nature of the discoveries. Thus, completeness of the **detection** and the **classification** are quite different. E. g., the detection limit of Markarian Survey for galaxies is almost $18^m$, however by Markarian criteria objects brighter than $17^m$ can be selected as UV-excess galaxies.

One of the most famous astronomical surveys, Sloan Digital Sky Survey (SDSS) was started in 2000 and is still active (Figure 1). Its main goals are cosmological, detecting and studying numerous galaxies and quasars, however many stars and other objects are also being detected. The total unique area covered is 14,555 sq. deg. In the photometric part, the number of catalog objects is 1,231,051,050, among them unique detections are 932,891,133. The number of unique, primary sources: total 469,053,874; stars 260,562,744, galaxies 208,478,448, unknown 12,682. In spectroscopy, the total number of spectra is 4,851,200, including useful spectra 4,151,126; galaxies 2,541,424, QSOs 680,843, stars 928,859, sky spectra 394,231, standards 88,788, and unknown objects 217,055. We give in Table 1 the list of the most important astronomical

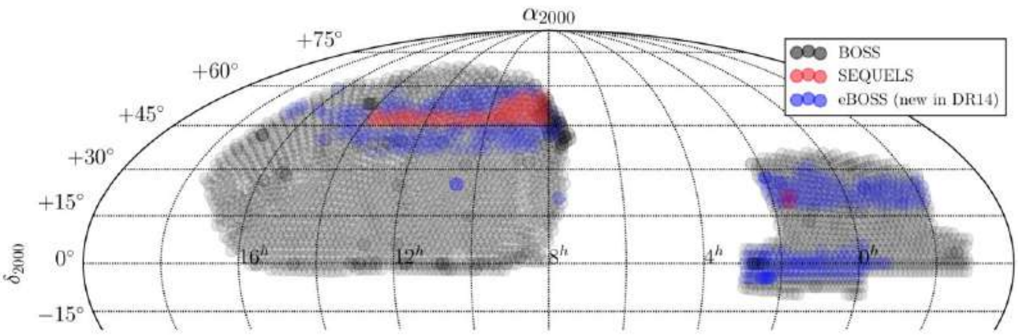


Figure 1. Sloan Digital Sky Survey (SDSS).

surveys in all wavelengths, from gamma-ray to radio.

Some understanding on the distribution of surveys, their data and their importance is given in Figures 2-4. In Figure 4, only extragalactic surveys are taken into consideration.

We give in Figure 5 the numbers of catalogued objects in astronomical catalogues of different wavelength ranges, from gamma-rays to radio. The difference is so big that even at logarithmic scale the numbers for gamma- X-rays, FIR, submm/mm and radio are almost negligible. We give in Table 2 the numbers by wavelength ranges.

## Astronomical Archives and Databases

Astronomical observations, hence data are maintained in astronomical archives, mostly organized by the observatories, where the observations have been carried out. Later on, Space observations appeared and dedicated astronomical archives and databases appeared to maintain and share the data.


doi:https://doi.org/10.52526/25792776-22.69.2-179

Table 1. Most important all-sky or large area astronomical surveys in multiwavelength astronomy.

| Survey | Years | Wavelength | Results | Number of sources |
|---|---|---|---|---|
| **Fermi-GLAST** | 2014 | 10 MeV-100 GeV | Sky survey | 3 033 |
| **ROSAT** | 1999 | 0.07–2.4 keV | Sky survey | 124 730 |
| **Galaxy Evolution Explorer (GALEX)** | 2013 | 1350–2800 Å | Sky survey | 82 992 086 |
| **USNO B1.0** | 2003 | visible | Sky survey | 1 045 913 669 |
| **Guide Star Catalog (GSC 2.3.2)** | 2003 | visible | Sky survey | 945 592 683 |
| **Sloan Digital Sky Survey (SDSS)** | 2000-pres. | visible | Sky survey | 1 231 051 050 |
| **HIgh Precision PARallax COllecting Satellite (HIPPARCOS) (Tycho-2)** | 1993 | visible | Astrometry | 2 539 913 |
| **Global Astrometric Interferom. for Astrophysics (Gaia)** | 2022 | 3200–10000 Å | Astrometry | 1 811 709 771 |
| **DENIS** | 2001 | 0.8-2.4 $\mu$m | Sky survey | 355 220 325 |
| **Two Micron Astronomical Sky Survey (2MASS)** | 2003 | 1.24, 1.66, 2.16 $\mu$m | Sky survey | 470 992 970 |
| **Wide-field Infrared Survey Explorer (WISE), AllWISE catalog** | 2013 | 3–28 $\mu$m | Sky survey | 747 634 026 |
| **Infrared Astronomical Satellite (IRAS)** | 1988-1990 | 8–120 $\mu$m | Sky survey | 405 769 |
| **AKARI** | 2006 | 7–180 $\mu$m | Sky survey | 1 298 044 |
| **NRAO/VLA Sky Survey (NVSS)** | 1998 | 21 cm | Sky survey | 1 773 484 |
| **Faint Images of the Radio Sky at Twenty Centimeters (FIRST)** | 1999 | 21 cm | Sky survey | 946 432 |

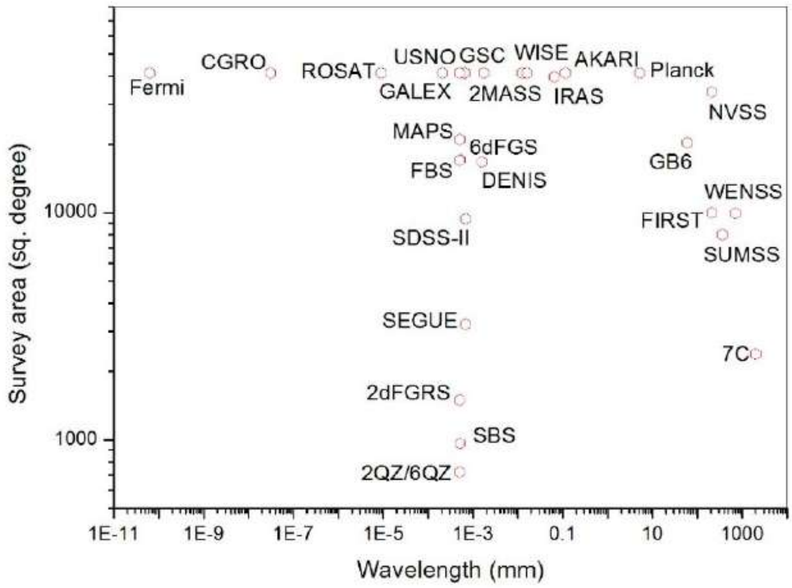


Figure 2. Survey area vs. wavelength (in mm) for most important astronomical surveys.

One of the oldest and most important astronomical databases is the **Wide-Field Plate Data Base** (**WFPDB**, http://www.skyarchive.org) created and maintained by Milcho Tsvetkov (Bulgaria). It contains 414 astronomical archives, in total 2,204,725 photographic plates from 125 observatories obtained between 1879 and 2002. The database includes 2,128,330 direct and 64,095 objective prism plates. The largest archives involved are: Harvard (USA) – 600,000 plates, Sonneberg (Germany) – 270,000 plates, Italian archive – 87,000 plates, Ukrainian archive – 85,000 plates; SAI (Moscow, Russia) – 50,000 plates. Among the objective prims Schmidt telescope observations, some 2500 First and Second Byurakan Surveys (FBS and SBS) plates are listed, plates that provided numerous astronomical discoveries. We give in Fig. 6 images from the digitized version of the FBS (**DFBS**), the software bSpec for extraction and analysis of DFBS spectra and webpage interface for working with DFBS images and spectra.

Among the multimission archives, most famous are:

- **High Energy Astrophysics Science Archive** (**HEASARC**; http://heasarc.gsfc.nasa.gov). It is the primary archive for HEA missions in gamma-rays, X-rays and extreme UV. Available data from ASCA, BeppoSAX, Chandra, EUVE, GLAST, HETE-2, INTEGRAL, ROSAT, RXTE, Astro-E2, Swift, XMM-Newton, etc.

- NASA/IPAC **Infrared Science Archive** (**IRSA**; http://irsa.ipac.caltech.edu). A multi-mission archive for NASA's IR and submm (IR/SM) astronomy data. Available data from 60 source catalogs, 22 image data sets and 7 spectroscopic data sets.

- **Multimission Archive at Space Telescope** (**MAST**; http://archive.stsci.edu). Supports a variety of astronomical data archives with a primary focus on scientifically related data sets in the


doi:https://doi.org/10.52526/25792776-22.69.2-179

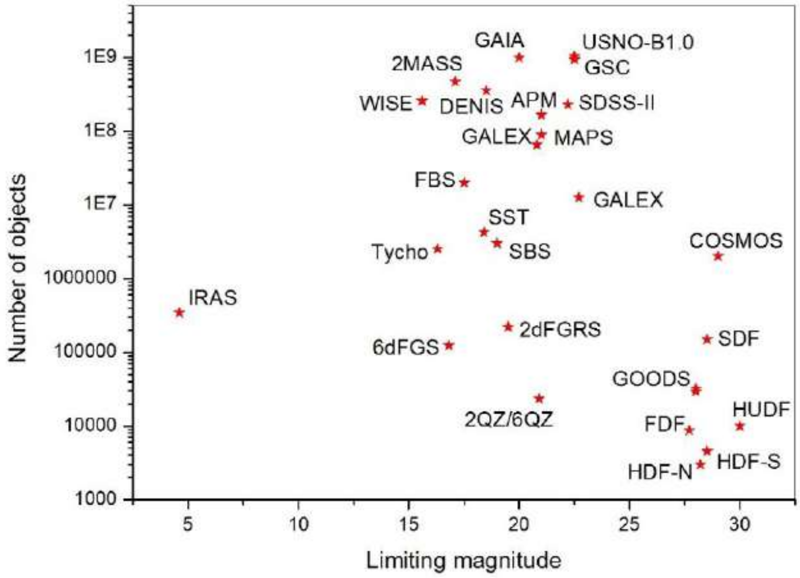


Figure 3. Number of objects in different astronomical surveys given by the limiting magnitude of surveys. Deep surveys are located in the right bottom, as relatively smaller number of objects have been discovered in them.

Table 2. Numbers of catalogues astronomical objects in different wavelengths and the main surveys and catalogs providing these numbers.

| Wavelength range | Wavelengths | Main surveys and catalogs | Number of catalogued objects |
|---|---|---|---|
| **Gamma-rays** | <0.1 Å | Fermi, INTEGRAL | 10 000 |
| **X-rays** | 0.1 – 100 Å | ROSAT, Chandra, XMM | 1 500 000 |
| **UV** | 100 – 3000 Å | GALEX, Hubble | 100 000 000 |
| **Optical** | 3000–10000 Å | DSS1/2, SDSS, Gaia | 2 400 000 000 |
| **NIR** | 1 – 10 $\mu$m | 2MASS, DENIS | 600 000 000 |
| **MIR** | 10 – 100 $\mu$m | WISE, Spitzer | 600 000 000 |
| **FIR** | 100 – 300 $\mu$m | IRAS, AKARI, Spitzer | 4 000 000 |
| **Submm/mm** | 0.3 – 10 mm | Herschel, ALMA | 1 000 000 |
| **Radio** | 1 cm – kms | NVSS, FIRST | 2 000 000 |

optical, UV and near-IR parts of the spectrum. Available data from wide variety of space missions (including HST) and integrated ground-based surveys (DSS, FIRST, etc.).

Another database in Strasbourg (France), VizieR, maintains all published astronomical catalogs and lists and tables from the published papers. Some 30,000 catalogs and tables are available and cross-correlation tools (X-Match service) are provided as well. Most of the data are VO compliant and are given in VO format.

## Astrophysical Virtual Observatories

Various useful tasks and tarnsformations are being done by Vos, such as combining data from different wavelengths and building multiwavelength Spectral Energy Distributions (SED, Fig. 7, left panel), overlapping images and comparing objects and sources from different databases (Aladin, Fig. 7, right panel), overlapping spectra for the same object to compare changes (Fig. 8, left panel), combining spectral data for the same object from observations of different telescopes and building combined spectra (VOSpec, Fig. 8, right panel), etc.

There is the **International Virtual Observatory Alliance (IVOA)** created in 2002 and unifying 23 VO projects (21 national and 2 European ones). The **Armenian Virtual Observatory (ArVO)** created in 2005 is part of this consortium.


doi:https://doi.org/10.52526/25792776-22.69.2-179

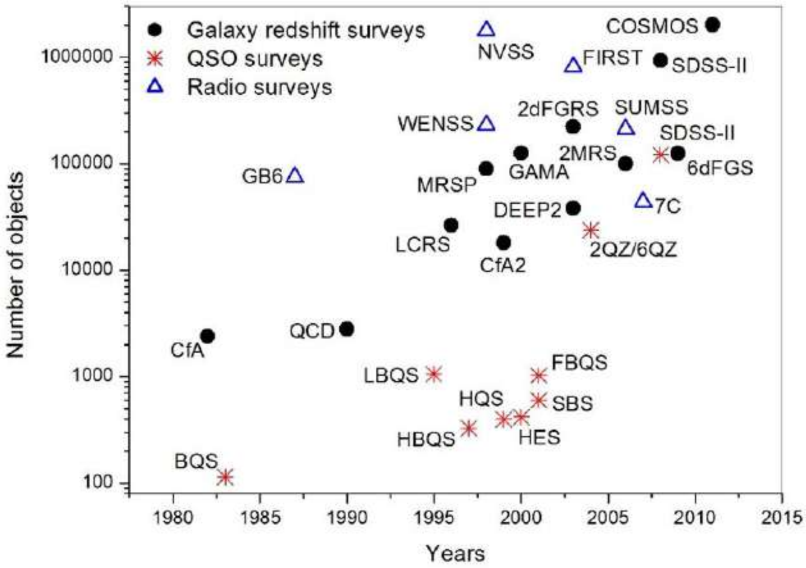


Figure 4. Number of objects in different astronomical surveys given by the years of surveys. Galaxy redshift surveys, QSO and radio surveys are given in different icons.

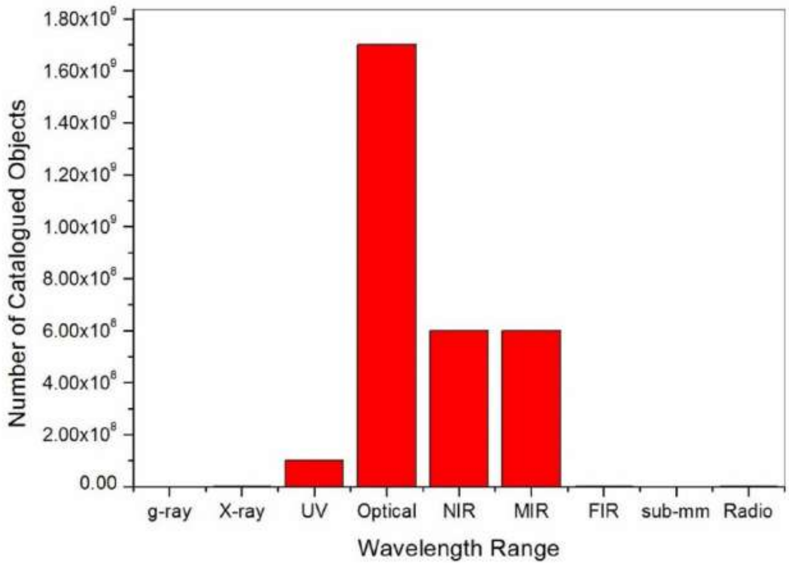


Figure 5. Numbers of astronomical objects so far discovered in different wavelength ranges, from gamma-rays to radio.

## Summary

Astronomy is one of the science disciplines related to Big Data and Data Science. **Big Data** are often characterized as 4 Vs, namely:

1) **Volume.** Quantity of generated and stored data. The size of the data determines the value and potential insight, whether it can be considered big data or not. In Astronomy, this condition is well maintained, as the Universe provides the Biggest Data. At present some 3 billion objects have been catalogued and each has numerous data related to its spatial and physical characteristics; astrometry, photometry, variability, spectroscopy, polarimetry, etc. E. g., some spectra show hundreds of spectral lines with line parameters and overlapped profiles totaling thousands of data.

2) **Variety.** The type and nature of the data. This helps people who analyze it to effectively use the resulting insight. Big Data draws from text, images, audio, video; plus it completes missing pieces through data fusion. In Astronomy, most of the data is in the form of images and spectra, as well as photometric, polarimetric data, etc. And these are quite different as well.

3) **Velocity.** In this context, the speed at which the data is generated and processed to meet the demands


doi:https://doi.org/10.52526/25792776-22.69.2-179

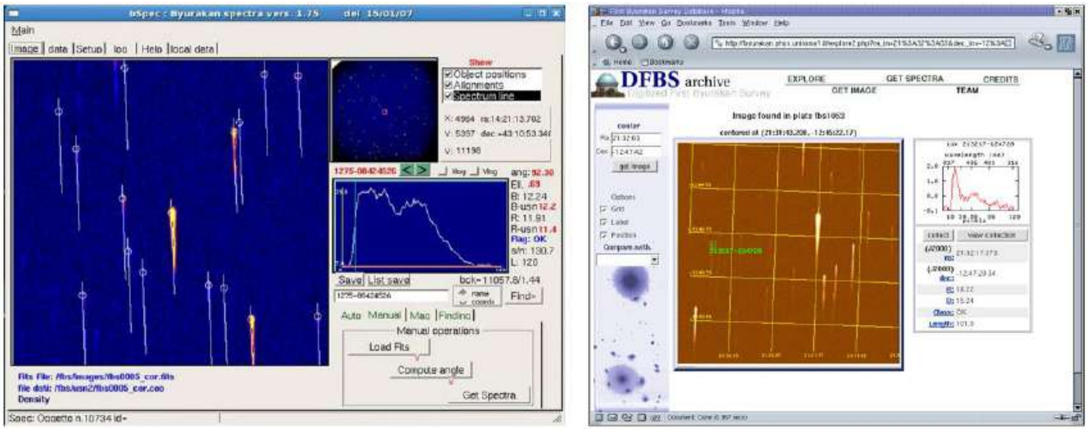


Figure 6. Dedicated software bSpec for extraction and analysis of the DFBS spectra (left panel) and DFBS webpage interface (right panel)

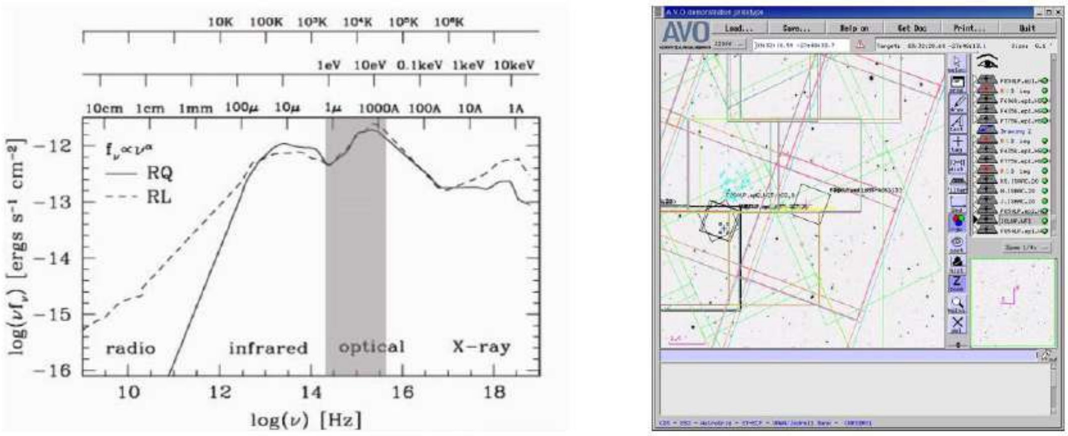


Figure 7. VO methods in astronomy: SEDs and image overlapping.

and challenges that lie in the path of growth and development. Big Data is often available in real-time. There are astronomical telescopes and receivers giving terabytes of data per night, accumulating peta and at present also exa bytes per year.

4) **Veracity.** The data quality of captured data can vary greatly, affecting the accurate analysis. Especially important is that many astronomers permanently verify data and each discovery is being checked by many others. Moreover, in VOs, various data are being combined and used together. This way, data become more homogenous and confident.

In Table 3, we give data volumes in different astronomical survey projects, showing the correspondence of astronomical data to the definition of Big Data.

Table 3. Data volumes in different astronomical projects.

| **Surveys, Projects** | **Short** | **Range** | **Information Volume** |
|---|---|---|---|
| **Digitized First Byurakan Survey** | DFBS | opt | 400 GB |
| **Digital Palomar Observatory Sky Survey** | DPOSS | opt | 3 TB |
| **Two Micron All-Sky Survey** | 2MASS | NIR | 10 TB |
| **Green Bank Telescope** | GBT | radio | 20 TB |
| **Galaxy Evolution Explorer** | GALEX | UV | 30 TB |
| **Sloan Digital Sky Survey** | SDSS | opt | 140 TB |
| **SkyMapper Southern Sky Survey** | SkyMapper | opt | 500 TB |
| **Panoramic Survey Telescope and Rapid Response System, expected** | PanSTARRS | opt | ~40 PB |
| **Large Synoptic Survey Telescope, expected** | LSST | opt | ~200 PB |
| **Square Kilometer Array, expected** | SKA | radio | ~4.6 EB |


doi:https://doi.org/10.52526/25792776-22.69.2-179

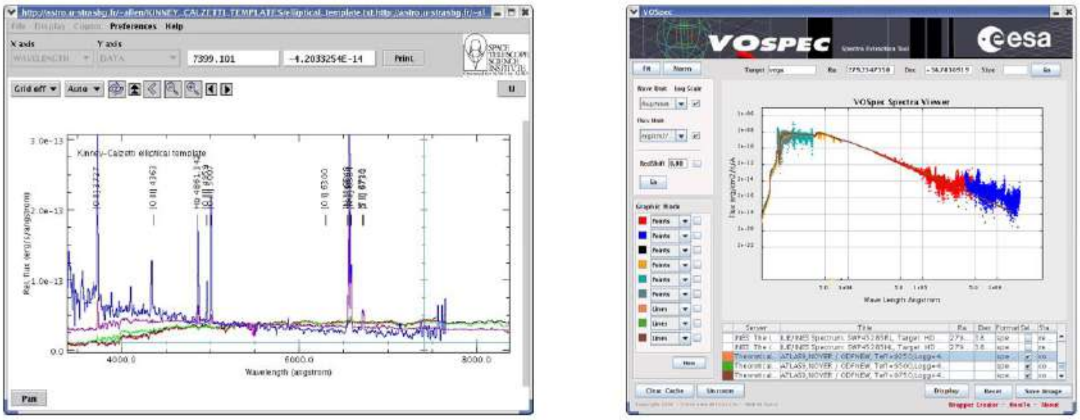


Figure 8. VO methods in astronomy: spectra overlapping and spectra combining and analysis.

### Acknowledgements

This work was partially supported by the Republic of Armenia Ministry of Education, Science, Culture, and Sports (RA MESCS) Committee of Science, in the frames of the research projects No. 15T-1C257 and 21AG-1C053. This work was made possible in part by research grants from the Armenian National Science and Education Fund (ANSEF) based in New York, USA (grants astroex-4193, astroex-4195, astroex-2347 and astroex-2597).

## References

1988, Infrared Astronomical Satellite (IRAS) Catalogs and Atlases.Volume 1: Explanatory Supplement. Vol. 1

Abrahamyan H. V., Mickaelian A. M., Knyazyan A. V., 2015, Astronomy and Computing, 10, 99

Acero F., et al., 2015, Astrophys. J. Suppl. Ser. , 218, 23

Ahumada R., et al., 2020, Astrophys. J. Suppl. Ser. , 249, 3

Bianchi L., Herald J., Efremova B., Girardi L., Zabot A., Marigo P., Conti A., Shiao B., 2011, Astrophys. Space. Sci. , 335, 161

Bird A. J., et al., 2010, Astrophys. J. Suppl. Ser. , 186, 1

Cabanela J. E., Humphreys R. M., Aldering G., Larsen J. A., Odewahn S. C., Thurmes P. M., Cornuelle C. S., 2003, Publ. Astron. Soc. Pac. , 115, 837

Calchi Novati S., et al., 2017, VizieR Online Data Catalog, p. J/ApJ/814/92

Colless M., et al., 2001, Mon. Not. R. Astron. Soc. , 328, 1039

Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, Astron. J. , 115, 1693

Cutri R. M., et al., 2003, 2MASS All Sky Catalog of point sources.

Cutri R. M., et al., 2013, Explanatory Supplement to the AllWISE Data Release Products, Explanatory Supplement to the AllWISE Data Release Products, by R. M. Cutri et al.

DENIS Consortium 2005, VizieR Online Data Catalog, p. B/denis

D'Elia V., et al., 2013, Astron. Astrophys. , 551, A142

Evans I. N., et al., 2010, Astrophys. J. Suppl. Ser. , 189, 37

Gaia Collaboration et al., 2021, Astron. Astrophys. , 649, A1

Garn T., Green D. A., Hales S. E. G., Riley J. M., Alexander P., 2007, Mon. Not. R. Astron. Soc. , 376, 1251

Gold B., et al., 2011, Astrophys. J. Suppl. Ser. , 192, 15

Gregory P. C., Scott W. K., Douglas K., Condon J. J., 1996, Astrophys. J. Suppl. Ser. , 103, 427

Hagen H. J., Engels D., Reimers D., 1999, Astron. and Astrophys. Suppl. Ser. , 134, 483

Hartman R. C., et al., 1999, Astrophys. J. Suppl. Ser. , 123, 79

Helfand D. J., White R. L., Becker R. H., 2015, Astrophys. J. , 801, 26

Helou G., Walker D. W., 1985, IRAS small scale structure catalog

Høg E., et al., 2000, Astron. Astrophys. , 355, L27


doi:https://doi.org/10.52526/25792776-22.69.2-179

Huchra J. P., Vogeley M. S., Geller M. J., 1999, Astrophys. J. Suppl. Ser. , 121, 287

Ishihara D., et al., 2010, Astron. Astrophys. , 514, A1

Lane W. M., Cotton W. D., van Velzen S., Clarke T. E., Kassim N. E., Helmboldt J. F., Lazio T. J. W., Cohen A. S., 2014, Mon. Not. R. Astron. Soc. , 440, 327

Lasker B. M., Doggett J., McLean B., Sturch C., Djorgovski S., de Carvalho R. R., Reid I. N., 1996, in Jacoby G. H., Barnes J., eds, Astronomical Society of the Pacific Conference Series Vol. 101, Astronomical Data Analysis Software and Systems V. p. 88

Lasker B. M., et al., 2008, Astron. J. , 136, 735

Markarian B. E., Lipovetsky V. A., Stepanian J. A., Erastova L. K., Shapovalova A. I., 1989, Soobshcheniya Spetsial'noj Astrofizicheskoj Observatorii, 62, 5

Massaro E., Mickaelian A. M., Nesci R., Weedman D., 2008, The Digitized First Byurakan Survey

Mauch T., Murphy T., Buttery H. J., Curran J., Hunstead R. W., Piestrzynski B., Robertson J. G., Sadler E. M., 2003, Mon. Not. R. Astron. Soc. , 342, 1117

McGlynn T. A., White N. E., Scollick K., 1994, in American Astronomical Society Meeting Abstracts #184. p. 27.08

McMahon R. G., Irwin M. J., Maddox S. J., 2000, VizieR Online Data Catalog, p. I/267

Mickaelian A. M., 2016a, Baltic Astronomy, 25, 75

Mickaelian A. M., 2016b, Astronomy Reports, 60, 857

Mickaelian A. M., et al., 2007, Astron. Astrophys. , 464, 1177

Monet D., 1998, USNO-A2.0

Monet D., Canzian B., Harris H., Reid N., Rhodes A., Sell S., 1998, VizieR Online Data Catalog, p. I/243

Monet D. G., et al., 2003, Astron. J. , 125, 984

Moshir M., 1989a, IRAS Faint Source Survey, Explanatory supplement version 1 and tape

Moshir M., 1989b, in Bulletin of the American Astronomical Society. p. 768

Oliver S. J., et al., 2012, Mon. Not. R. Astron. Soc. , 424, 1614

Pojmanski G., 1998, ACTAA, 48, 35

Simpson J. M., et al., 2015, Astrophys. J. , 807, 128

Skrutskie M. F., et al., 2006, Astron. J. , 131, 1163

Stepanian J. A., 2005, Rev. Mex. Astron. Astrofis. , 41, 155

Tsvetkov M. K., Stavrev K. Y., Tsvetkova K. P., Mutafov A. S., 1994, IAU Commission on Instruments, 5, 47

Vlemmings W. H. T., Ramstedt S., O'Gorman E., Humphreys E. M. L., Wittkowski M., Baudry A., Karovska M., 2015, Astron. Astrophys. , 577, L4

Voges W., et al., 1999, Astron. Astrophys. , 349, 389

Wisotzki L., Christlieb N., Bade N., Beckmann V., Köhler T., Vanelle C., Reimers D., 2000, Astron. Astrophys. , 358, 77

Xmm-Newton Survey Science Centre C., 2013, VizieR Online Data Catalog, p. IX/44

Yamamura I., Makiuti S., Ikeda N., Fukuda Y., Oyabu S., Koga T., White G. J., 2010, VizieR Online Data Catalog, p. II/298


doi:https://doi.org/10.52526/25792776-22.69.2-179