FIRST SCIENCE FOR THE VIRTUAL OBSERVATORY

THE VIRTUAL OBSERVATORY WILL REVOLUTIONISE THE WAY WE DO ASTRONOMY, BY ALLOWING EASY ACCESS TO ALL ASTRONOMICAL DATA AND BY MAKING HANDLING AND ANALYSING DATASETS AT VARIOUS LOCATIONS ACROSS THE GLOBE MUCH SIMPLER AND FASTER. WE REPORT HERE ON THE STATUS OF THE VIRTUAL OBSERVATORY IN EUROPE AND ON THE FIRST SCIENCE RESULT COMING OUT OF IT, THE DISCOVERY OF ~30 SUPERMASSIVE BLACK HOLES THAT HAD PREVIOUSLY ESCAPED DETECTION BEHIND MASKING DUST CLOUDS.

PAOLO PADOVANI[1], MARK G. ALLEN[2], PIERO ROSATI[3], NICHOLAS A. WALTON[4]

[1]ST-ECF, EUROPEAN SOUTHERN OBSERVATORY
[2]CENTRE DE DONNÉS ASTRONOMIQUES DE STRASBOURG, FRANCE
[3]EUROPEAN SOUTHERN OBSERVATORY
[4]INSTITUTE OF ASTRONOMY, CAMBRIDGE, UK

ASTRONOMY in the 21st century is facing the need for radical changes. When dealing with surveys of up to ~1,000 sources, one could apply for telescope time and obtain an optical spectrum for each one of them to identify the whole sample. Nowadays, we have to deal with huge surveys (e.g., the Sloan Digital Sky Survey [SDSS], the Two Micron All Sky Survey [2MASS], the Massive Compact Halo Object [MACHO] survey), reaching (and surpassing) 100 million objects. Even at, say, 3,000 spectra at night, which is only feasible with the most efficient multi-object spectrographs and for relatively bright sources, such surveys would require more than 100 years to be completely identified, a time which is clearly much longer than the life span of the average astronomer! But even taking a spectrum might not be enough to classify an object. We are in fact reaching fainter and fainter sources, routinely beyond the typical identification limits of the largest telescopes available (approximately 25 mag for 2–4 hour exposures), which makes "classical" identification problematic. These very large surveys are producing a huge amount of data: it would take about two months to download at 1 Mbytes/s (a very good rate for most astronomical institutions) the Data Release 2 SDSS images, about two weeks for the catalogues. The images would fill up about 1,000 DVDs. And the final SDSS will be about three times as large. These data, once downloaded, need also to be analysed, which requires tools which may not be available locally and, given the complexity of astronomical data, are different for different energy ranges. Moreover, the breathtaking capabilities and ultra-high efficiency of new ground- and space-based observatories have led to a "data explosion", with astronomers world-wide accumulating of the order of a Terabyte of data per night. Finally, one would like to be able to use all of these data, including multi-million-object catalogues, by putting this huge amount of information together in a coherent and relatively simple way, something which is impossible at present.

All these hard, unescapable facts call for innovative solutions. For example, the observing efficiency can be increased by a clever pre-selection of the targets, which will require some "data-mining" to characterise the sources' properties before hand, so that less time is "wasted" on sources which are not of the type under investigation. One can expand this concept even further and provide a "statistical" identification of astronomical sources by using all the available, multi-wavelength information without the need for a spectrum. The data-download problem can be solved by doing the analysis where the data reside. And finally, easy and clever access to *all* astronomical data worldwide would certainly help in dealing with the data explosion and would allow astronomers to take advantage of it in the best of ways.

THE VIRTUAL OBSERVATORY

The name of the solution is the *Virtual Observatory* (VO). The VO is an innovative, evolving system, which will allow users to interrogate multiple data centres in a seamless and transparent way, to best utilise astronomical data. Within the VO, data analysis tools and models, appropriate to deal also with large data volumes, will be made more accessible. New science will be enabled, by moving Astronomy beyond "classical" identification with the characterisation of the properties of very faint sources by using all the available information. All this will require good communication, that is the adoption of a common language between data providers, tool users and developers. This is being defined now using new international standards for data access and mining protocols under the auspices of the recently formed *International Virtual Observatory Alliance* (IVOA: http://ivoa.net), a global collaboration of the world's astronomical communities.

One could think that the VO will only be useful to astronomers who deal with colossal surveys and huge amounts of data. That is not the case. The World Wide Web is equivalent to having all the documents of the world inside one's computer, as they are all reachable with a click of a mouse. Similarly, the VO will be like having all the astronomical data of the world inside one's desktop. That will clearly benefit not only professional astronomers but also anybody interested in having a closer look at astronomical data. Consider the following example: imagine one wants to find *all* the observations of a given source available in *all* astronomical archives in a given wavelength range. One also needs to know which ones are in raw or processed format, one wants to retrieve them and, if raw, one wants also to have access to the tools to reduce them on-the-fly. At present, this is extremely time consuming, if at all possible, and would require, even to simply find out what is available, the use a variety of search interfaces, all different from one another and located at different sites. The VO will make all this possible very easily.

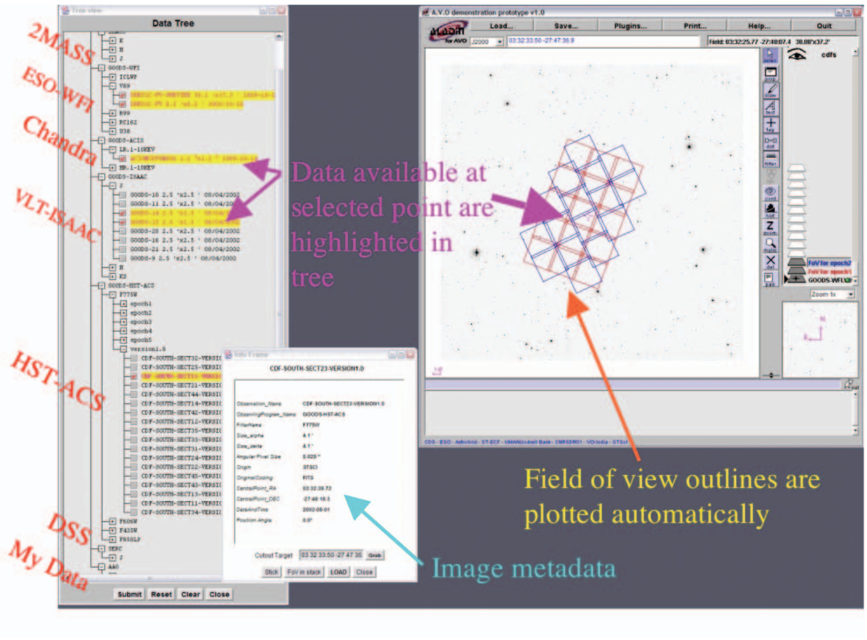


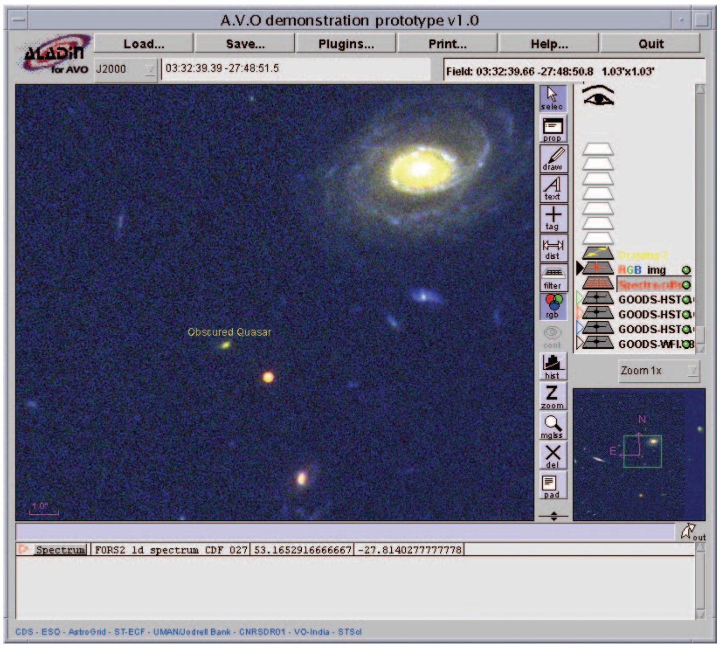
Figure 1: The AVO prototype in action. An ESO/WFI image of the GOODS southern field, overlaid with the HST/ACS data field of view outlines. The "data-tree" on the left shows the images available in the *Aladin* image server. Data available at selected coordinates get highlighted in the tree. Metadata information is also accessible. The user's own data can also be loaded into the prototype. This is based on the use of IVOA agreed standards, namely the Data Model, descriptive Metadata, and data interchange standards.

THE VO IN EUROPE: THE ASTROPHYSICAL VIRTUAL OBSERVATORY

The status of the VO in Europe is very good. In addition to seven current national VO projects, the European funded collaborative *Astrophysical Virtual Observatory* initiative (AVO: http://www.euro-vo.org) is creating the foundations of a regional scale infrastructure by conducting a research and demonstration programme on the VO scientific requirements and necessary technologies. The AVO has been jointly funded by the European Commission (under the Fifth Framework Programme [FP5]) with six European organisations participating in a three year Phase-A work programme. The partner organisations are ESO in Munich, the European Space Agency, AstroGrid (funded by PPARC as part of the United Kingdom's E-Science programme), the CNRS-supported Centre de Données Astronomiques de Strasbourg (CDS) and TERAPIX astronomical data centre at the Institut d'Astrophysique in Paris, the University Louis Pasteur in Strasbourg, and the Jodrell Bank Observatory of the Victoria University of Manchester. The AVO is the definition and study phase leading towards the Euro-VO – the development and deployment of a fully fledged operational VO for the European astronomical research community.

The AVO project is driven by its strategy of regular scientific demonstrations of VO technology, held on an annual basis in coordination with the IVOA. For this purpose progressively more complex AVO demonstrators are being constructed. The current one is an evolution of *Aladin*, developed at CDS, and has become a set of various software components, provided by AVO and international partners, which allows relatively easy access to remote data sets, manipulation of image and catalogue data, and remote calculations in a fashion similar to remote computing.

The AVO held its second demonstration, "AVO 1st Science", on January 27–28, 2004 at ESO. The demonstration was truly multiwavelength, using heterogeneous and complex data covering the whole electromagnetic spectrum. These included: MERLIN, VLA (radio), ISO [spectra and images] and 2MASS (infrared), USNO, ESO 2.2m/WFI

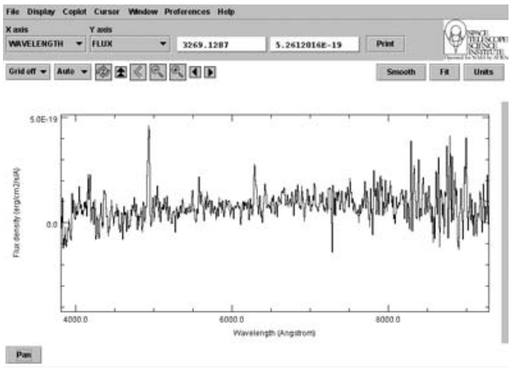


J2000 03:32:39.39 -27:48:51.5

Field: 03:32:39.66 -27:48:50.8 1.03'x1.03'

Spectrum | FORS2 1d spectrum CDF 027 | 53.1652916666667 | -27.8140277777778

Figure 2: An example of the direct links between imaging and spectral data: an obscured source selected via its X-ray (Chandra) properties, imaged by the HST/ACS, and identified through an ESO/VLT FORS2 spectrum (below) as a type 2 quasar at redshift 3.06 (Szokoly et al. 2004). The great majority of our new QSO 2 candidates are too faint to be classified even by the VLT or Keck. The colour-composite image has been generated on-the-fly from HST images in three different bands and the spectrum is displayed using SpecView, an application plugged-in into the AVO prototype.



X axis: WAVELENGTH

Y axis: FLUX

3269.1287

5.2612016E-19

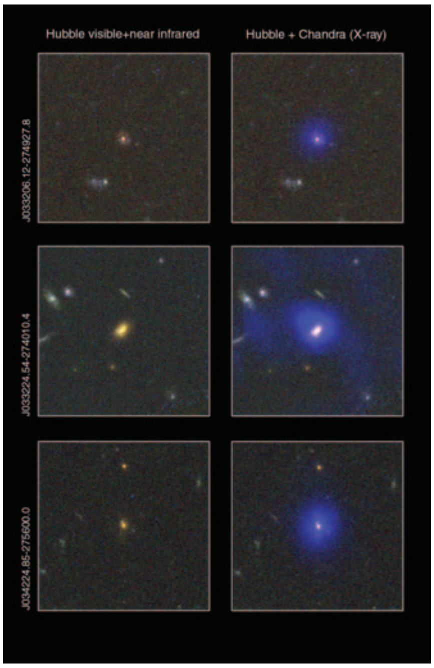


Figure 3: Colour-composite images of the host galaxies of three of the newly found dust-enshrouded supermassive black holes. To the left, images taken with the Advanced Camera for Surveys on-board the NASA/ESA Hubble Space Telescope. To the right these images are overlaid with images (in blue) from NASA's Chandra X-ray Observatory.

and VLT/FORS [spectra], and HST/ACS (optical), XMM and Chandra (X-ray) data and catalogues. Two cases were dealt with: an extragalactic case on obscured quasars, centred around the Great Observatories Origin Deep Survey (GOODS) public data, and a Galactic scenario on the classification of young stellar objects.

The extragalactic case was so successful that it turned into the first published science result fully enabled via end-to-end use of VO tools and systems.

DISCOVERING OPTICALLY FAINT, OBSCURED QUASARS WITH VO TOOLS

How did we get a scientific paper out of a science demonstration? The extragalactic science case revolved around the two GOODS fields (Giavalisco et al. 2004), namely the Hubble Deep Field-North (HDF-N) and the Chandra Deep Field-South (CDF-S), the most data-rich, deep survey areas on the sky. Our idea was to use the AVO prototype to look for high-power, supermassive black holes in the centres of apparently normal looking galaxies.

Black holes lurk at the centres of active galaxies (AGN) surrounded by dust which is thought to be, on theoretical and observational grounds (see, e.g., Urry & Padovani 1995; Jaffe et al. 2004), distributed in a flattened configuration, torus-like. When we can look down the axis of the dust torus and have a clear view of the black hole and its surroundings these objects are called "type 1" AGN, and display the broad lines and strong UV emission typical of quasars. "Type 2" AGN, on the other hand, lie with the dust torus edge-on as viewed from Earth so our view of the black hole is totally blocked by the dust over a range of wavelengths from the near-infrared to soft X-rays.

While many dust-obscured low-power black holes, the Seyfert 2s, have been identified, until recently few of their high-power counterparts were known. This was due to a simple selection effect: when the source is a low-power one and therefore, on average, closer to the observer, one can very often detect some features related to narrow emission lines on top of the emission from the host galaxy, which qualify it as a type 2 AGN. But when the source is a high-power one, a so-called QSO 2, and therefore, on average, further away from us, the source looks like a normal galaxy.

Our approach was to look for sources where nuclear emission was coming out in the hard X-ray band, with evidence of absorption in the soft band, a signature of an obscured AGN, and the optical flux was very faint, a sign of absorption. One key feature was the use of a correlation discovered by Fiore et al. (2003) between the X-ray-to-optical ratio and the X-ray power, which allowed us to select QSO 2s even when the objects were so faint that no spectrum, and therefore no redshift, was available.

We used a large amount of data: deep X-ray (Chandra) and optical (HST/ACS) catalogues, and identifications, redshifts, and spectra for previously identified sources in the CDF-S and HDF-N based on VLT and Keck data. Using the AVO prototype made it much easier to classify the sources we were interested in and to identify the previously known ones, as we could easily integrate all available information from images, spectra, and catalogues at once (see Fig. 1 and 2). One interesting feature is the prototype catalogue cross-matching service, which can access all entries in Vizier, at CDS, and allowed us to cross-correlate a variety of catalogues very efficiently.

Out of the 546 X-ray sources in the GOODS fields we selected 68 type 2 AGN candidates, 31 of which qualify as QSO 2 (estimated X-ray power > 10^{44} erg/s; see Fig. 3). Our work brings to 40 the number of QSO 2 in the GOODS fields, an improvement of a factor ~4 when compared to the only nine such sources previously known. These sources are very faint ($\langle R \sim 27 \rangle$) and therefore spectroscopic identification is not possible for the large majority of objects, even with the largest telescopes currently available. By using VO methods we are sampling a region of redshift - power space so far unreachable with classical methods. For the first time, we can also assess how many QSO 2 there are down to relatively faint X-ray fluxes. We find a surface density greater than 330 deg^{-2} for $f(0.5–8\ \text{keV}) \geq 10^{-15}\ \text{erg cm}^{-2}\ \text{s}^{-1}$, higher than previously estimated.

The identification of a population of high-power obscured black holes and the active galaxies in which they live has been a key goal for astronomers and will lead to greater understanding and a refinement of the cosmological models describing our Universe. The paper reporting these results has just been published (Padovani et al. 2004).

USING THE AVO PROTOTYPE

The AVO prototype can be downloaded from the AVO Web site at http://www.-euro-vo.-org/twiki/bin/view/Avo/SwgDownload. We encourage astronomers to download the prototype, test it, and also use it for their own research. For any problems with the installation and any requests, questions, feedback, and comments you might have please contact the AVO team at twiki@euro-vo.org. (Please note that this is still a prototype: although some components are pretty robust, some others are not.)

FUTURE DEVELOPMENTS

The next AVO demonstration event is to be held in January 2005 and this will see the rollout of the first version of the Euro-VO portal, through which the European astronomer will gain secure access to a wide range of data access and manipulation capabilities. The science drivers for this next demonstration release are being developed by the AVO science team with input from the AVO Science Working Group. The AVO Phase-A study is now being completed. The Science Reference Mission, on which the AVO team is also working, will form the scientific basis for the wider scale development of the Euro-VO.

ACKNOWLEDGEMENTS

The AVO was selected for funding by the Fifth Framework Programme of the European Community for research, technological development and demonstration activities, contract HPRI-CT-2001-50030. We thank the AVO team for their superb work, without which these results could have not been obtained, and the many people who have produced the data on which our paper was based, particularly the GOODS, CDF-S, and Penn State Teams.

REFERENCES

Fiore, F., Brusa, M., Cocchia, F., et al. 2003, A&A, 409, 79
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJ, 600, L93
Jaffe, W., Meisenheimer, K., Röttgering, H. J. A., et al. 2004, Nature, 429, 47
Padovani, P., Allen, M. G., Rosati, P., & Walton, N. A. 2004, A&A, 424, 545
Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, ApJS, in press [astro-ph/0312324]
Urry, C. M., & Padovani, P. 1995, PASP, 107, 803