Astronomy and Computing 11 (2015) 190–209

Contents lists available at ScienceDirect

Astronomy and Computing

journal homepage: www.elsevier.com/locate/ascom

Full length article

The Virtual Astronomical Observatory: Re-engineering access to astronomical data

R.J. Hanisch [a,b,1], G.B. Berriman [c], T.J.W. Lazio [d,*], S. Emery Bunn [e], J. Evans [f], T.A. McGlynn [g], R. Plante [h]

[a] Virtual Astronomical Observatory, 1600 14th Street NW, Suite 730, Washington, DC 20036, USA
[b] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[c] Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA
[d] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
[e] Center for Advanced Computing Research, California Institute of Technology, Pasadena, CA 91125, USA
[f] Smithsonian Astrophysical Observatory, 60 Garden Street, Cambridge, MA 02138, USA
[g] High Energy Astrophysics Science Archive Research Center, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[h] National Center for Supercomputing Applications, University of Illinois, Urbana, IL 61801, USA

ARTICLE INFO

Article history:
Received 31 October 2014
Received in revised form
11 March 2015
Accepted 16 March 2015
Available online 23 March 2015

Keywords:
Catalogs
Surveys
Virtual observatory tools
Data discovery
Data access
Applications

ABSTRACT

The US Virtual Astronomical Observatory was a software infrastructure and development project designed both to begin the establishment of an operational Virtual Observatory (VO) and to provide the US coordination with the international VO effort. The concept of the VO is to provide the means by which an astronomer is able to discover, access, and process data seamlessly, regardless of its physical location. This paper describes the origins of the VAO, including the predecessor efforts within the US National Virtual Observatory, and summarizes its main accomplishments. These accomplishments include the development of both scripting toolkits that allow scientists to incorporate VO data directly into their reduction and analysis environments and high-level science applications for data discovery, integration, analysis, and catalog cross-comparison. Working with the international community, and based on the experience from the software development, the VAO was a major contributor to international standards within the International Virtual Observatory Alliance. The VAO also demonstrated how an operational virtual observatory could be deployed, providing a robust operational environment in which VO services worldwide were routinely checked for aliveness and compliance with international standards. Finally, the VAO engaged in community outreach, developing a comprehensive web site with on-line tutorials, announcements, links to both US and internationally developed tools and services, and exhibits and hands-on training at annual meetings of the American Astronomical Society and through summer schools and community days. All digital products of the VAO Project, including software, documentation, and tutorials, are stored in a repository for community access. The enduring legacy of the VAO is an increasing expectation that new telescopes and facilities incorporate VO capabilities during the design of their data management systems.

1. Introduction

1.1. Beginnings

The formal Virtual Observatory (VO) program in the United States began with the 2000 Decadal Survey of the National Academy of Science, in which a National Virtual Observatory (NVO) was identified as the top priority small initiative (McKee et al., 2001).

* Corresponding author.
E-mail addresses: robert.hanisch@nist.gov (R.J. Hanisch), gbb@ipac.caltech.edu (G.B. Berriman), joseph.lazio@jpl.nasa.gov (T.J.W. Lazio), janet@cfa.harvard.edu (J. Evans), thomas.a.mcglynn@nasa.gov (T.A. McGlynn), rplante@illinois.edu (R. Plante).
[1] Current address: Office of Data & Informatics, Material Measurement Laboratory, National Institute of Standards & Technology, Gaithersburg, MD 20899, USA.

http://dx.doi.org/10.1016/j.ascom.2015.03.007

The NVO is the committee's top-priority small initiative. NVO involves the integration of all major astronomical data archives into a digital database stored on a network of computers, the provision of advanced data exploration services for the astronomical community, and the development of data standards and tools for data mining. The committee recommends coordinated support from both NASA and the NSF, since NVO will serve both the space- and ground-based science communities.

The NVO project and parallel projects in Europe and the UK were formulated through a series of meetings, beginning with "Virtual Observatories of the Future" (Brunner et al., 2001), held at the California Institute of Technology in 2000 June.

At the 2002 conference, "Toward an International Virtual Observatory" (Quinn and Górski, 2004), held in Garching, Germany, the International Virtual Observatory Alliance[2] (IVOA) was formed with the NVO, the Astrophysical Virtual Observatory (AVO, ESO), and AstroGrid (UK) as founding partners. R. Hanisch, the then-NVO Project Manager, was the first chair of the IVOA Executive Committee. In the subsequent decade, the IVOA has grown to have 21 member national projects.

The IVOA patterned itself on the World-Wide Web Consortium[3] (W3C) and adopted its process for the development of standards (Working Drafts → Proposed Recommendations → Recommendations) with the actual standards documents developed by a set of working groups. (See Section 3.1 for more details.) A Virtual Observatory Working Group was established under Commission 5 of the International Astronomical Union (IAU) in order to give IVOA Recommendations official status within the IAU, but this process has not been used in practice since there was already global acceptance of IVOA standards.

The NVO project focused on standards and infrastructure development, working closely in the context of the IVOA, and implemented a number of prototype science applications to demonstrate the utility of the underlying VO standards. NVO also ran an active program of engagement with the astronomical community through annual summer schools of one-week duration, exhibits at American Astronomical Society meetings, and the production of a major reference book, *The National Virtual Observatory: Tools and Techniques for Astronomical Research* (Graham et al., 2007). In a demonstration of this book's value, it was translated into Mandarin by members of the VO-China project.

The NVO project was funded by the National Science Foundation's Information Technology Research program, starting in 2001, and included organizations in astronomy and computer science. Its funding came to a planned close in 2008, after demonstrating the technology framework for supporting a VO.

1.2. Program

In 2010, the successor to the NVO, the Virtual Astronomical Observatory (VAO), was begun to sustain and evolve those technologies successfully demonstrated by the NVO as part of an operating virtual observatory. While there were numerous management and logistical barriers to the establishment of the VAO, the National Science Foundation (NSF) and National Aeronautics and Space Administration (NASA) agreed to fund the project jointly, with NSF support directed through the VAO, Limited Liability Company, and NASA support provided directly to the participating NASA data centers.

The VAO, LLC, was created as a 50–50 collaboration between the Association of Universities for Research in Astronomy (AURA) and the Associated Universities, Inc. (AUI), with an independent Board of Directors. This management structure was chosen deliberately so that the VAO would be perceived as belonging to the research community and have dedicated oversight. Executive authority within the VAO was provided by the Director, who worked with a Program Manager, Project Scientist, and Project Technologist. In order to provide advice on priorities for research tools, a Science Council was established. Within the VAO, a Program Council consisting of senior management representatives from each VAO member organization was also established. The Program Council worked with the VAO management to map Science Council priorities onto available resources and expertise, and thus to develop the annual program plan. Work packages for all organizations, whether funded by NSF or NASA, were agreed with the Director and Program Manager. The program plan covered all work at all organizations regardless of the source of funding.

Table 1 shows the VAO program history and funding. As a result of two major reviews, NSF and NASA redefined program priorities and reduced the overall budget from an original plan of $27.5M ($20M NSF + $7.5M NASA) to $16.5M ($11M NSF + $5.5M NASA). In addition to simple reductions in funding, these reviews were often accompanied by recommended changes in the direction of the project, and, ultimately, the project duration was reduced by seven months. Consequently, some activities that were started or intended to be started were reduced in scope or stopped early to respond to the combination of lower funding and recommended changes in direction. A specific example of this change in direction and cessation of activities was the Time Series Search Tool (Section 2.4), which was unable to be brought to the desired level of maturity.

1.3. Major accomplishments

The accomplishments of the NVO and VAO are extensive and will be described in further detail in the following sections of this paper. At a summary level, however, we note the following accomplishments:

- Major contributor to IVOA standards. Appendix B contains a list of IVOA standards to which NVO/VAO staff contributed. The list includes standards recommended by the IVOA Executive Committee and those submitted to the Executive Committee for recommendation.
- Leadership within the IVOA, within the executive, Working Groups, and Interest Groups.
- High-level science applications for data discovery, integration, analysis, and catalog cross-comparison.
- Scripting toolkits that allow scientists to incorporate VO data directly into their reduction and analysis environments.
- A robust operational environment in which VO services worldwide are routinely checked for aliveness and compliance with IVOA standards.
- Community engagement through AAS meetings, summer schools (NVO), and community days (VAO).
- Comprehensive web site with on-line tutorials, announcements, links to both US and internationally developed tools and services.
- Take up of VO standards and infrastructure within essentially every major data center and survey project in the United States, with approximately 1M VO-based data requests per month and some 2000 unique users.
- Prudent fiscal management, with overall management expenses kept below 15% and the project completed with an unspent balance of funds of less than 1% (for an $11M [lifetime] budget over 4 years).

[2] http://www.ivoa.net/.
[3] http://www.w3.org/.