VAO

Robert Hanisch
Space Telescope Science Institute
Director, VAO

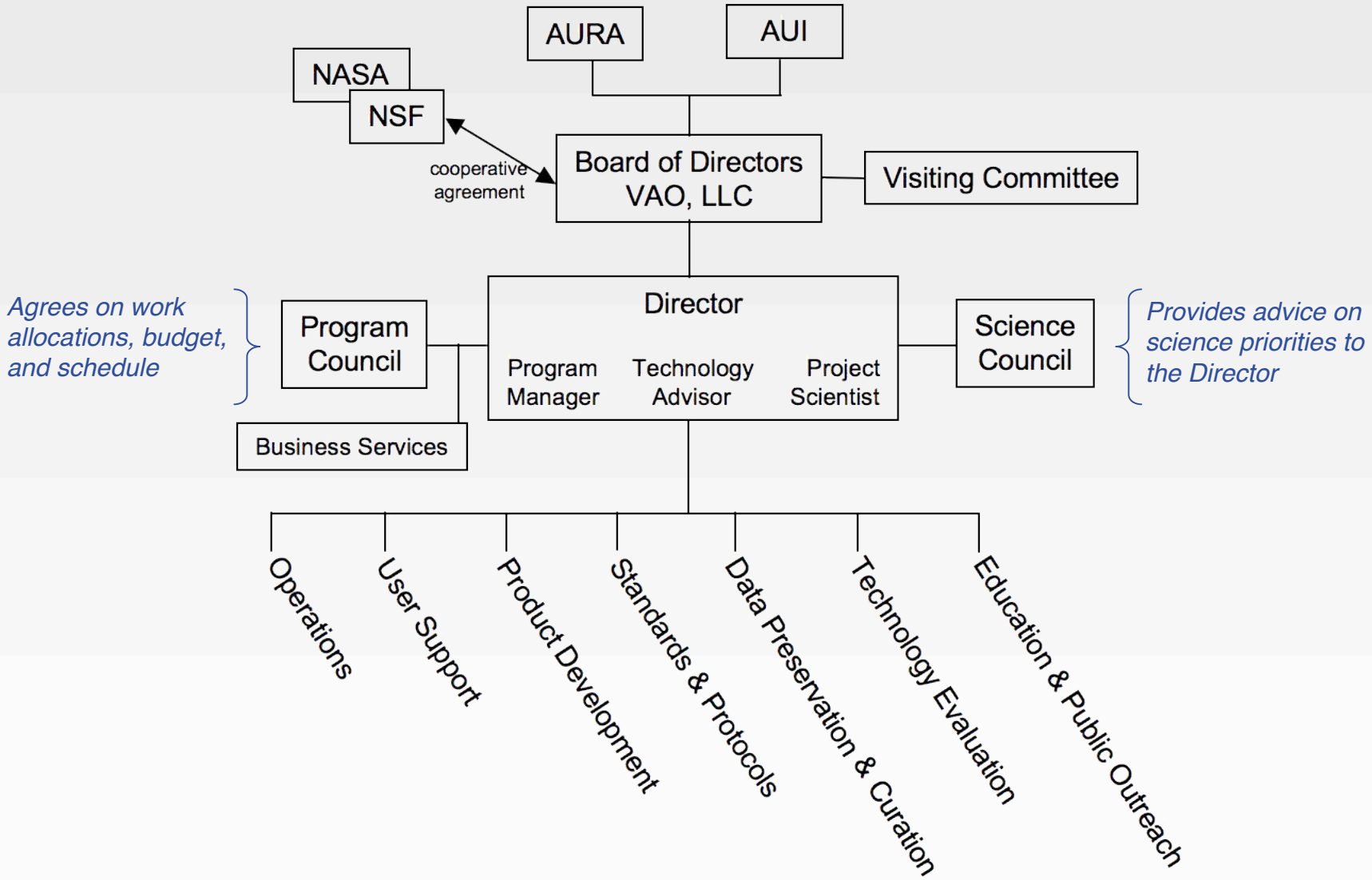
History

- VAO proposal submitted to NSF/NASA in April 2008
- NSF/NASA informed us of intent to fund in spring 2009
- NASA funding began in summer of 2009
- NSF issued award on 15 May 2010
- Program Plan developed over the summer, submitted to NSF/NASA on 13 August 2010 (no feedback yet)
- VAO team has regrouped in the past several months

Details

- Funding is $5.5M/year for five years, subject to annual performance review
  - $4M/year from NSF: funds SAO, STScI (75%), JHU, NCSA, Caltech, NOAO, NRAO, AUI (financial management)
  - $1.5M/year from NASA: funds HEASARC, IPAC, STScI (25%)
  - Covers ~27 FTE over the ten organizations
- VAO is managed by the VAO,LLC (limited liability company) co-owned by AUI (operates NRAO and ALMA) and AURA (operates NOAO and STScI)
  - VAO has its own Board of Directors

Organization

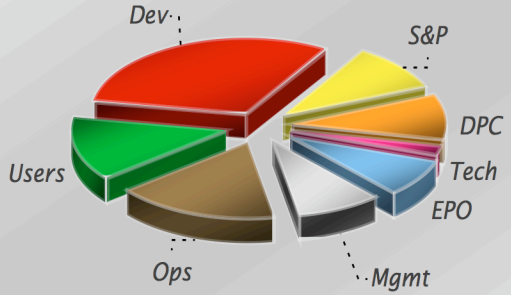


Management

- Governance:
  - Board of Directors (Jay Gallagher, U. Wisconsin, chair)
  - Director – annual operating plan, facility conduct, appointing Senior Personnel (Robert Hanisch, STScI)
  - Program Manager – day-to-day running of VAO. Convenes Program Council to negotiate work plans. (Bruce Berriman, IPAC)
  - Project Scientist – scientific oversight. (Dave De Young, NOAO)
  - Technology Advisor – identifying new technologies. (Alex Szalay, JHU)
  - Science Council Head – convenes Science Council for advice on scientific priority. (Pepi Fabbiano, SAO)
  - Business Manager – managing business and financial matters. (Marie Huffman, AUI)

Scope and functions

- Seven major areas of activity
  - Operations: Tom McGlynn, HEASARC, Ani Thakar, JHU
  - User Support: Betty Stobie, NOAO, Maria Nieto-Santisteban, JHU
  - Product Development: Ray Plante, NCSA, Gretchen Greene, STScI
  - Standards and Protocols: Matthew Graham, Caltech, Doug Tody, NRAO
  - Data Preservation and Curation: Arnold Rots, SAO, Joe Mazzarella, NED (Alberto Accomazzi, SAO/ADS)
  - Technology Evaluation: Ashish Mahabal, Caltech
  - Education and Public Outreach: TBD, STScI



Operations

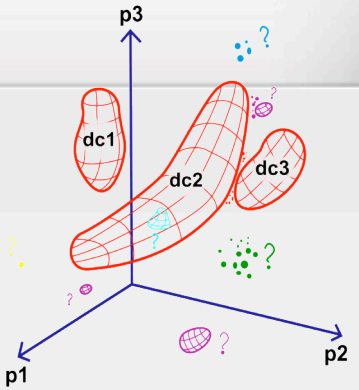
- ~17% of overall effort
- Maintenance of distributed services
  - Monitoring
  - Adherence to standards
  - Work with service providers to repair, revise
  - Provide status to users and technical team
- Quality assurance
  - Unit testing
  - Applications and service templates
  - Metadata quality evaluation, correction
  - Software revision control
  - Web browser compatibility, multi-platform support
  - Usage logging
- Facility support
  - Problem report ticket system
  - Hardware support (preservation facility)

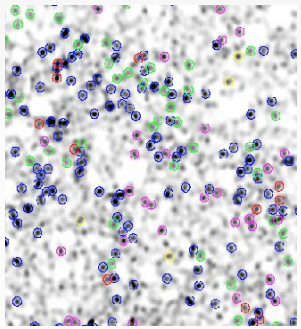
User Support

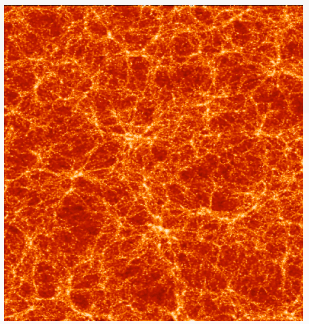
- ~14% of overall effort
- End-user testing
  - Readiness reviews, assessment of meeting science requirements
  - System integration, regression testing
  - Usability, user interface design
  - Quality of documentation
- Training and advocacy
  - Summer schools, professional outreach events
  - Newsletters, tutorials, use-cases, recipes
  - Collection of feedback
- Help desk, web site
  - Documentation
  - User forum

Product Development

- ~31% of overall effort
- Science applications and products
  - Support and enhance existing applications
  - Integrate VO access into existing desktop applications
  - Cross-correlation
  - Data-mining
  - Visualization
  - Theoretical models and simulations
- Service toolkits/templates
- Infrastructure
- Registry/Directory
- Robust software development process







Standards and Protocols

- ~10% of overall effort
- International standards process with IVOA
  - Protocols, service definitions
  - Data models
  - Registry
  - Semantics
  - Application frameworks, work flows, grid computing
  - Prototyping

SCS, SIAP, SSAP, TAP, SLAP, SAMP,
STC, UCD, UType, Identifiers,
SSO, UWS, ADQL,SDM, SLDM,
VOTable, VOResource, VODataService
VOEvent, Vocabularies

Data Preservation and Curation

- ~10% of overall effort
- Repository for community-produced high-level data products
  - Images, spectra, time series, etc., published in journals
  - Data collections that are currently privately hosted
  - Collaboration with NSF OCI DataNet program, JHU-led Data Conservancy team
- Cross-repository linking
  - Papers and bibliographic records, archives, repository
- Curation standards (in collaboration with Standards and Protocols activity)

Technology Evaluation

- ~2% of overall effort
- Monitor new technologies relevant to VO infrastructure
- Select promising technologies for evaluation; install and test
- Make recommendations to Product Development for adoption
- Assist User Support team and end-users in scaling-up applications

Education & Public Outreach

- ~8% of overall effort
- Train education and outreach specialists in use of, and capabilities of, VAO
  - EPO developer workshops
  - Establish partnerships with leaders in the EPO community
- Continue and strengthen relationship with Microsoft WWT and GoogleSky
- Build EPO-focused website
- Assist in creation of EPO-friendly data products
- Inform public of how IT / network enables data access and research

Management

- ~9% of overall effort
- Director, program manager, project scientist, business manager, chair of Science Council
- Costs for Board, Science Council, and Program Council meetings
- Costs associated with the VAO, LLC (banking, legal, audit, tax filing, insurance, office space, etc.)

Science initiatives

- The VAO has selected seven science initiatives that were endorsed by the Science Council as providing maximal scientific impact in the astronomy community:
  1. Development of a dedicated VAO Portal
  2. Scalable cross-matching between catalogs of sources
  3. Building and Analyzing Spectral Energy Distributions
  4. Time Domain Astronomy: (a) Periodograms and light curve analyses; (b) Transient event services
  5. Data Linking and Semantic Astronomy
  6. Desktop Tool Integration
  7. Data Mining and Statistical Analysis

Science deliverables

- Four areas selected for science deliverables in Year 1 (assume start date = Oct 1 2010).
  - Selected because they can take advantage of work already done

Science Deliverable	Delivery Date	Lead
Portal that supports search, visualization, filtering and data access across all data sets accessible to the VAO	Jun 30, 2011	Tom Donaldson, STScI
SED service that collects and plots multi-wavelength data and supports interactive visualization attributes of data	July 30, 2011	Janet Evans, SAO
Deliver **cross-matching engine** that supports cross-matches across at least two large catalogs	August 30, 2011	John Good, IPAC & Tamas Budavari, JHU
Time Series Astronomy: Deliver periodogram service and light curve classification service for data sets at NStED, TSC (Harvard)	September 30, 2011	John Good, IPAC

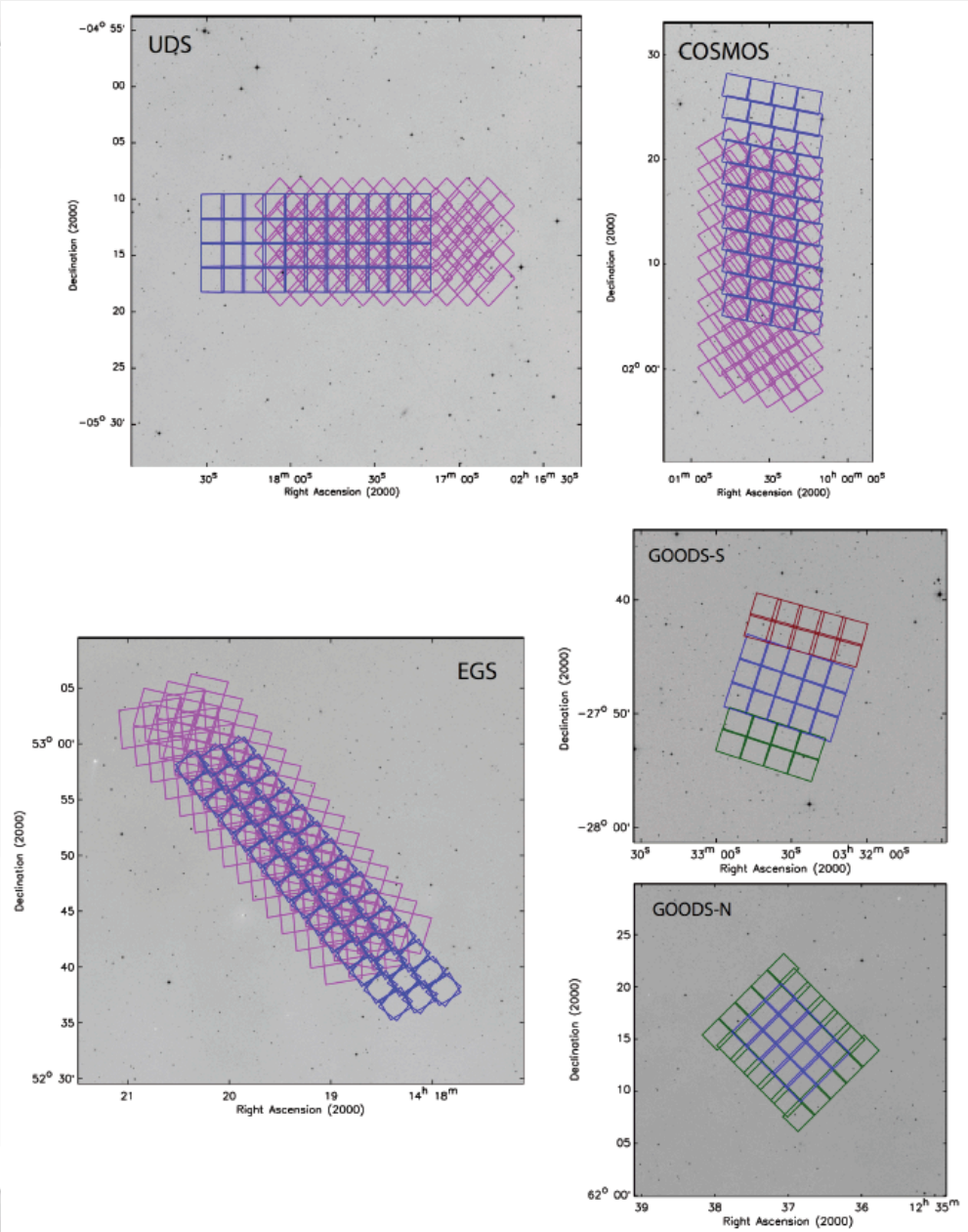
Science deliverables

- Four science initiatives will undergo a study period during Year 1:
  - Time Domain Astronomy (Transients)
  - Data Linking and Semantic Astronomy
  - Desktop Tool Integration
  - Data Mining and Statistical Analysis
- The goals of these studies are to make recommendations on science deliverables for Year 2+ that will be evaluated by the Science Council.

Science collaborations

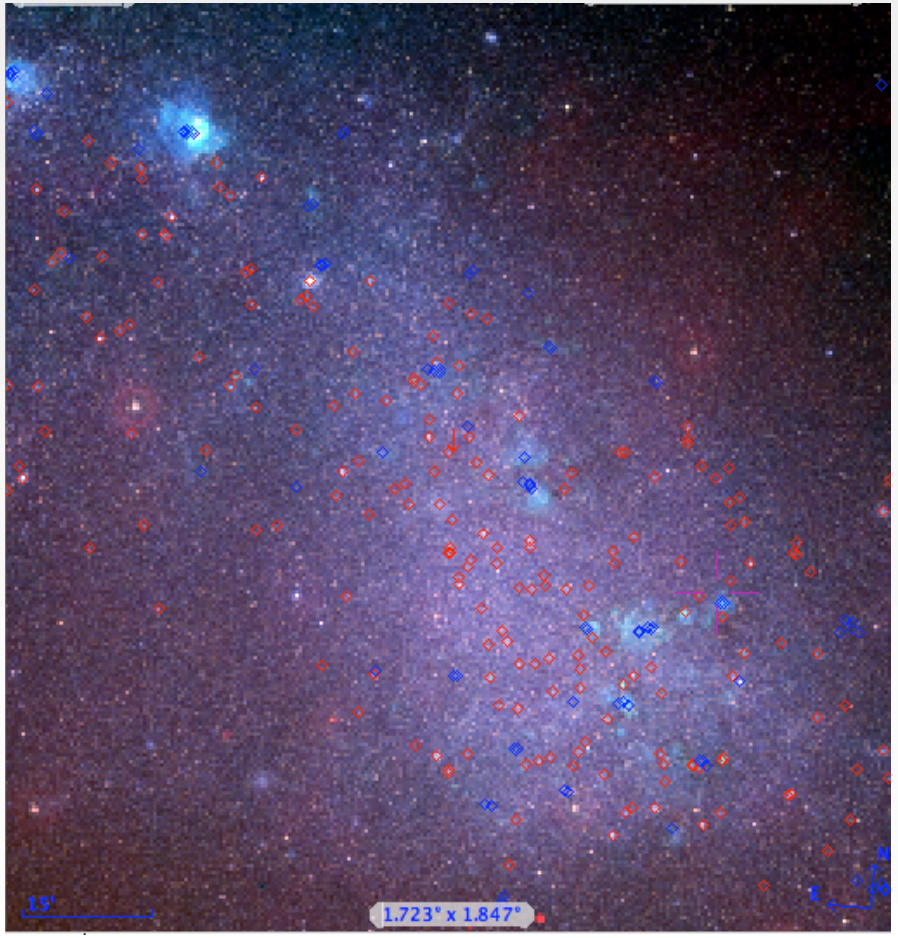
- CANDELS: Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey
  - HST multi-cycle (3-year) treasury program, S. Faber and H. Ferguson, CoPIs, >100 members of science team
  - Multi-wavelength (radio to x-ray) study of >250k galaxies with $1.5 < z < 8$
  - Understand initial epoch of star formation, disk formation, first generation of interactions and mergers, role of AGN formation in galaxy evolution
- SED-informed cross-matching
- VOEvent notices (supernovae)
- Image cut-out services

CANDELS fields



Small Magellanic Cloud

- Construct 3-dim model of SMC based on period-luminosity data on 3,000+ Cepheid variables
  - Construct SEDs for ~100M objects in 10x10 deg FOV
  - Stellar population study of a dwarf galaxy
  - Effects of galaxy interactions in dwarf systems
  - B. Madore (Carnegie) PI
- Test of scalable cross-matching and large-scale SED construction



Decadal Survey and VO

- Programs begun or in progress from previous DS not reconsidered
- Importance of archives shines through, including incorporating data management into overall project costs and planning
- No explicit mention of data mining, but is to some extent assumed
- Time Domain is identified as a discovery space

Shameless plug

- IAU Symp. 285, *New Horizons in Time Domain Astronomy*
- 19-23 Sept 2011
- Oxford, UK
- R. Hanisch and E. Griffin, co-chairs of SOC
  - also Masatoshi Ohishi, Rob Seaman, Tara Murphy, et al.

www.physics.ox.ac.uk/iaus285