

The Virtual Observatory – enabling data access – key science drivers

Nicholas Walton

AstroGrid and Euro-VO Technology Centre
Project Scientist
(Institute of Astronomy,
University of Cambridge)



The Virtual Observatory

- The Challenge: Facilitate Better Science
 - Science: range of astronomy, solar to cosmology
 - Technical: global system
 - Political: global community
- Building a solution: The Virtual Observatory
 - standards
 - IVOA
 - GGF
 - implementations
- Context for this workshop
 - AstroGrid: an example VO implementation

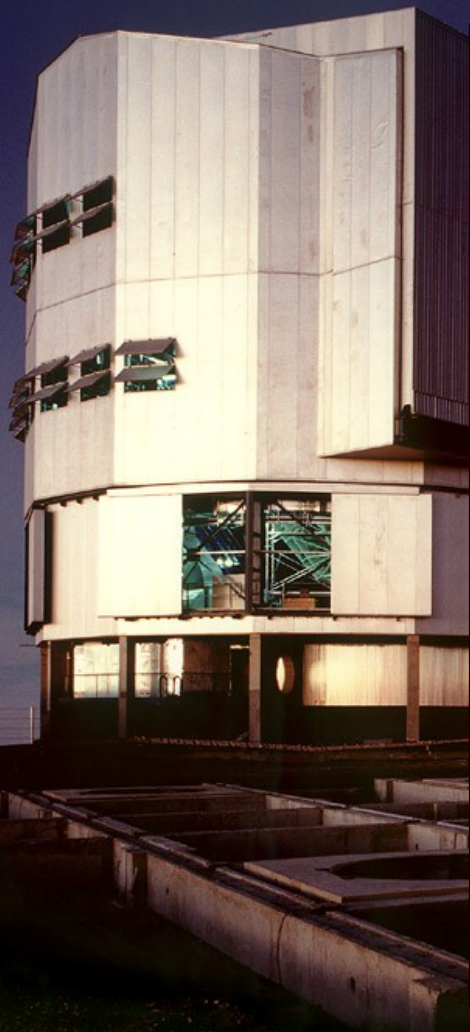
The Challenge of Data

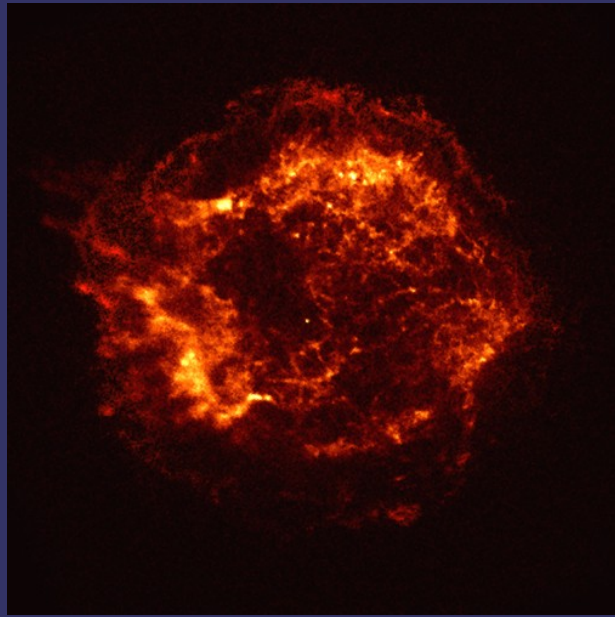
- Astronomy is an observational science
- Progress is made via understanding gained from the study of the cosmos
- Powerful observatories exist producing observational data across the wavelength domain
 - data comes in many formats
 - levels of complexity
- Data & applications held globally: USA, Europe, Asia
 - heterogeneous data archives
- Research partnerships are also global
- Connecting researchers with data and applications is the challenge for the Virtual Observatory

The Need for Virtual Observatories: Managing Technological Change



- The massive **Growth of Data**
- Number + size of telescopes
 - Optical: ESO's 4x8m VLT, 2x8m Gemini
 - X-ray: XMM-Newton
 - sub-mm: ALMA
- Increase in **size and multiplex** capabilities of instrumentation:
 - Infra-Red: VISTA > 100 GB/nights
 - Radio: e-Merlin > data rates ~320 Gbps
 - **All sky at 0.1 arcsec – 100 TB**

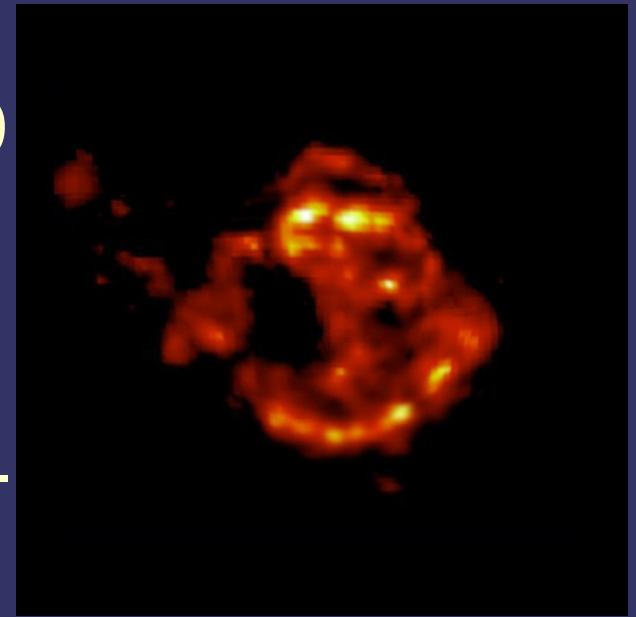




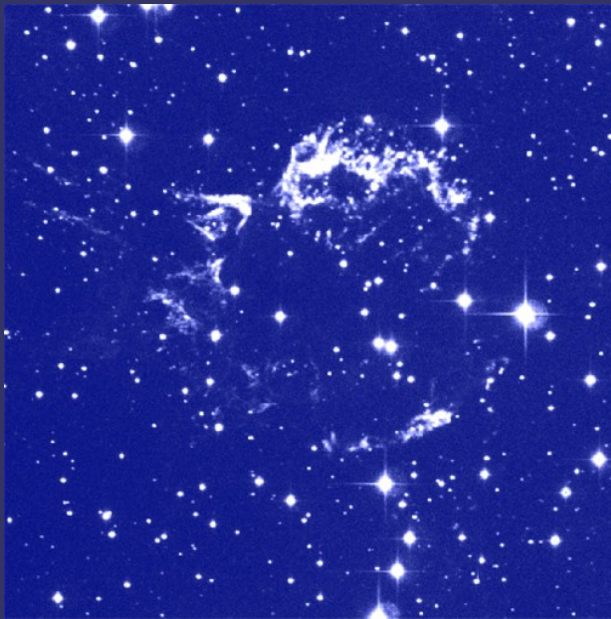
Shocks seen in the X-ray
Chandra image

Supernova Remnant Cassiopeia-A – a 300 year old Supernova

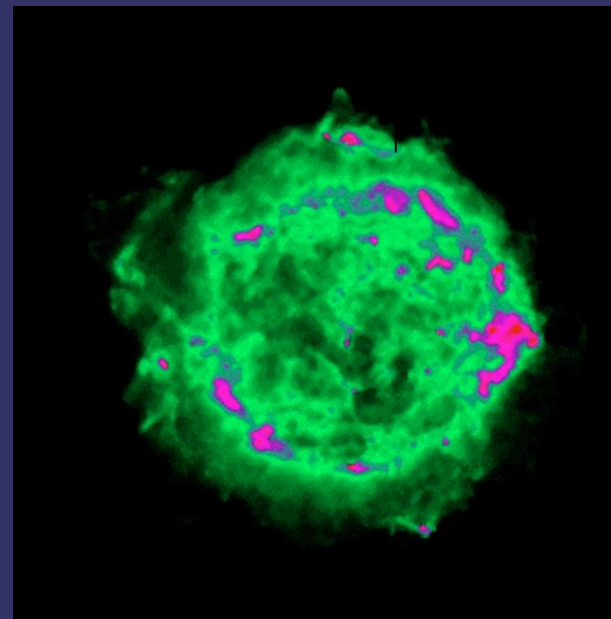
The Challenge and Opportunity of multi- Wavelength data:



Dust shows in the IR



Heavy elements
seen in the
optical



Mapping e^- s in
the magnetic
field as revealed
by Radio data

Images from Chandra Science Centre

New & Improved Science from VO's: Space Weather

What happens to the Earth's magnetosphere during a coronal mass ejection ?

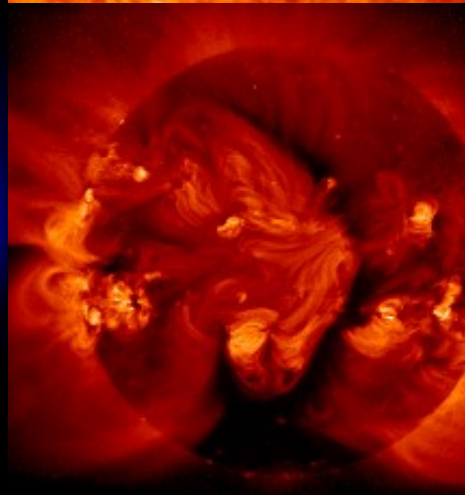
Event imaged by space based solar observatory

Effect detected later by satellites and ground radar

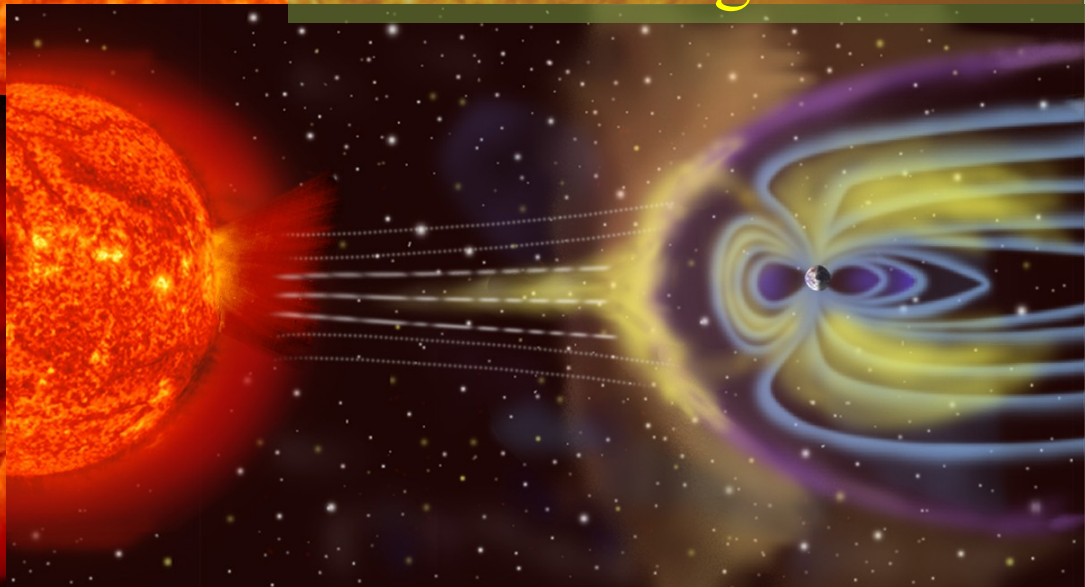


2002/03/12 07:00

SOHO/EIT – EUV

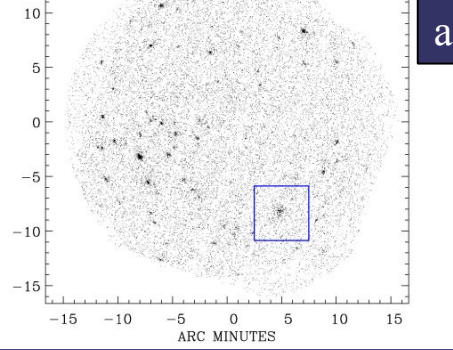
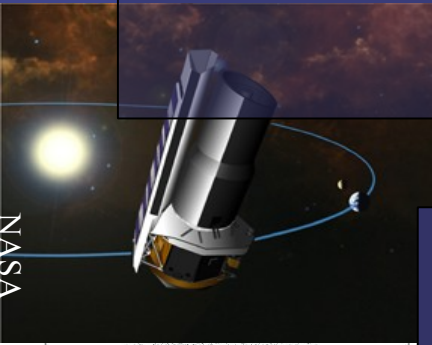


Yohkoh – Xray



NASA: Living With a Star – <http://lws.gsfc.nasa.gov>

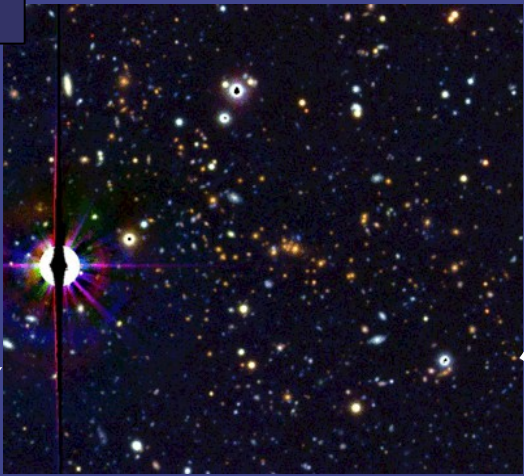
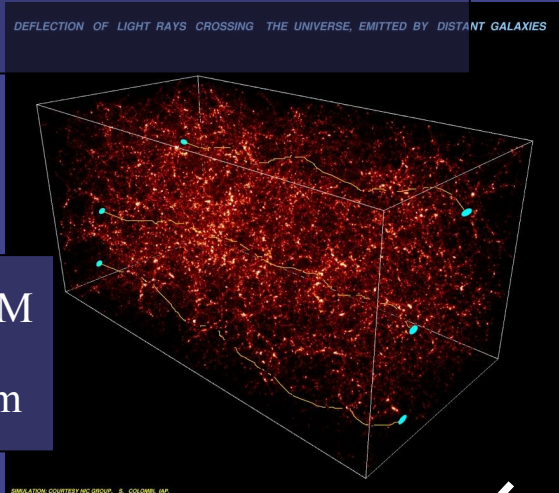
New & Improved Science: Cosmology



Multiple large image sources: registration & association

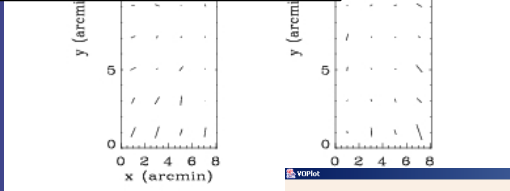
Automatic cluster finding techniques

Multi-TB Λ CDM models, e.g. Millennium Sim

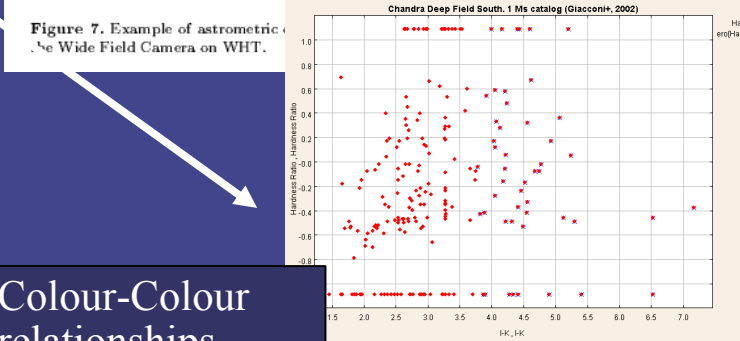
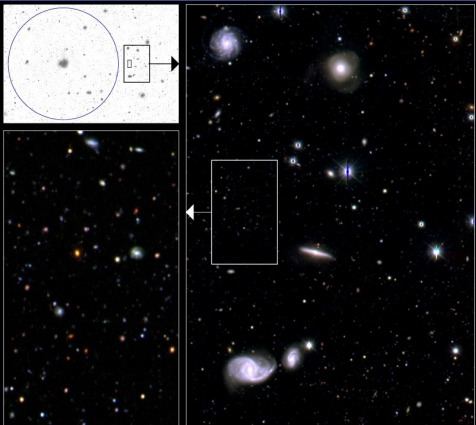


X-ray cluster: Chandra X-ray (Mullis) overlaid on a deep BRI image (Clowe & Luppino).

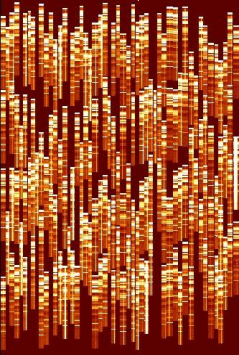
Generate Shear Maps c.f. CDM models > DM distribution with redshift



Remove star correlate gal with z



Source ID from multiplexed spectral data

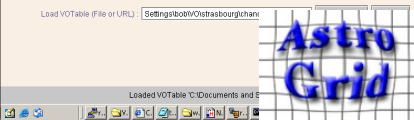


Colour-Colour relationships classification in multi-phase space



Drivers - May 10

7/01/0



Science Challenge

- Multi-Wavelength data required
 - key science: e.g. planets, large scale structure, galaxy formation: all needs a combination of data
- More data required
 - survey telescopes create ~TB/night
- More use of database organised resources used

System needs to be fast and easy for the astronomer to use and affordable for the data providers to operate

System use drives increased (sustainable) funding for operations and development of data access system

Solutions: a Flexible Framework

- Create a system that recognises:
 - no one data provider or repository: thus data interoperability
 - application provision
- Requires
 - a system built upon agreed interoperability standards
- Exploits
 - wider IT developments: Grid and WS technologies
 - power of XML/ SOAP etc
 - access to high speed networks
 - but note: backbones ~10Tb/s, desktops ~100Mb/s
 - reduced costs of h/w: all data now on spinning disks

Building the Virtual Observatory:

Global scope

International partnerships

Agreeing interoperability standards

Building regional implementations

Based on new computational technologies

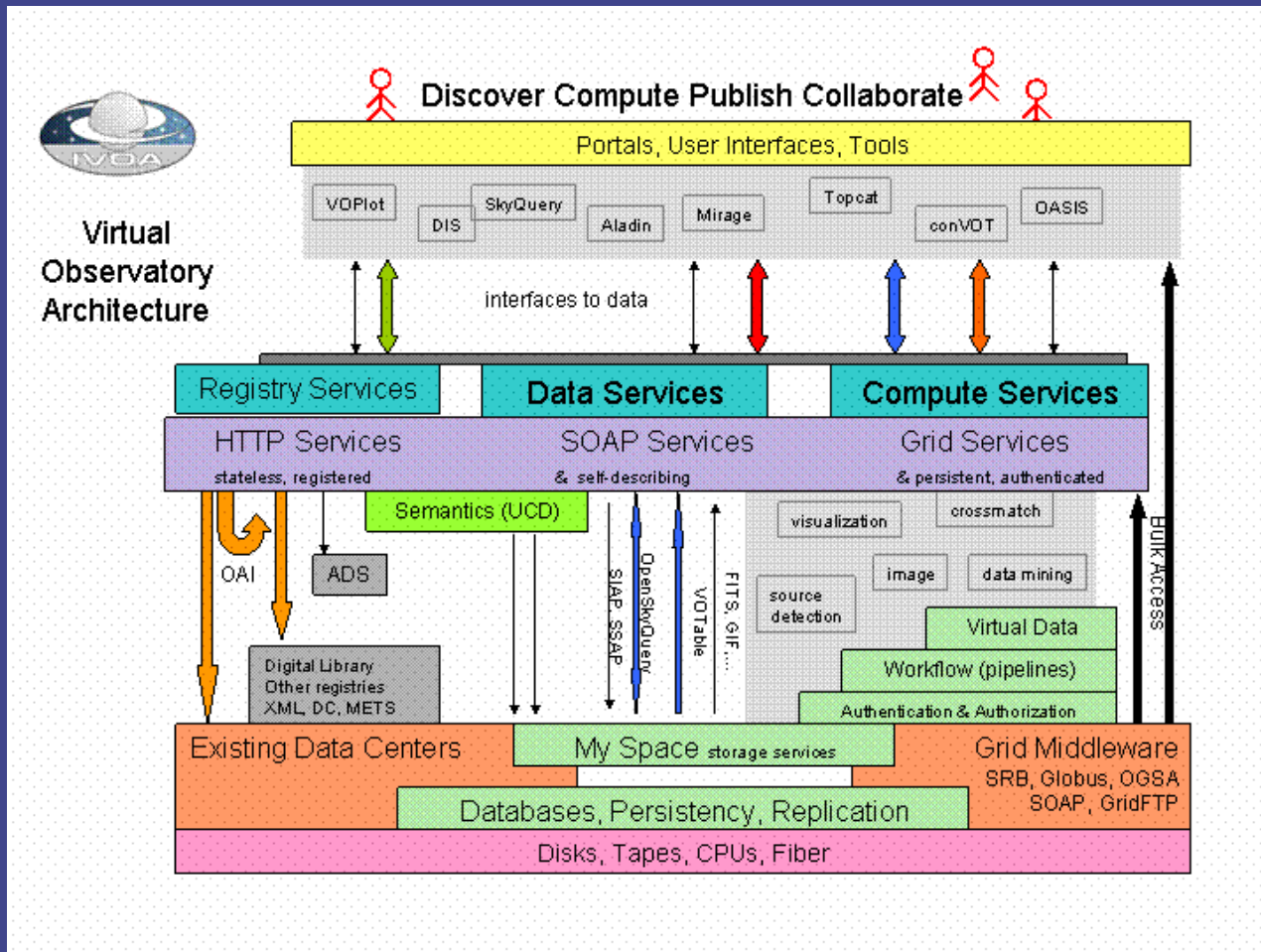
Deployed on the fastest networks

IVOA: Stds Enabling Interoperability

- The International Virtual Observatory Alliance
<http://www.ivoa.net>
- A global partnership
- Projects represent global astronomy data providers
- IVOA a forum for interoperability standards
- VO projects build on these standards
- Global reach



IVOA Architecture Analysis



Analysis of a VO:

- Multi-layer
- Complex
- User interfaces thru a portal
- Astro-apps interface to a VO abstraction layer
- Lower level middleware provided by the 'grid' world
 - e.g. SRB
- Hardware at bottom layer

IVOA has working groups to address 'astro' specific 'boxes'

Ref: IVOA Architecture Overview: Williams et al, 2004

IVOA Note 2004-06-14:

<http://www.ivoa.net/Documents/Notes/IVOArch/IVOArch-20040615.html>

VO Standard Areas

- Standard vocabulary (semantics)
- Standard ontology describing how terms are related
- Standard data model (encoding format) for each type of measurement
- Standard query language for issuing spatial, temporal, and semantic queries across the catalogs.
- Standard access services for retrieving catalog records or image cutouts.
- Standard mechanisms for interacting with storage systems (VOspace)
- Standard authentication/ authorisation mechanisms
- Standard event notification services.

IVOA Working Groups: <http://www.ivoa.net/forum>

- Registry:
 - how to 'register' resources: concept of VOResources
- Data Access Layer:
 - Standards for remote data access: e.g. SIAP, SSA
- Data Model:
 - Standards for the actual data: e.g. XML'ing of FITS
- VO Query Language:
 - Standards for 'astro' database access: e.g. Openskyquery, 'circle'
- Unified Content Descriptors: std Vocabulary
 - Standards for common ways of describing data: metadata
- VOTable:
 - XML representation of tabular data
- Grid & Web Services:
 - Interfaces to Grid and Web Service stds: e.g. 'Heartbeat'

GGF and Astronomy

- Current IVOA implementations largely utilise web services, however GGF standards are becoming relevant in a number of areas:
 - Authorisation
 - Transport, e.g. GridFTP
 - Persistent Storage
 - Job Submission
- Use of grid service standards actively under investigation by a number of VO projects
 - AstroGrid
 - USA's NVO

Science Driven Virtual Observatories



US-VO: NVO <http://www.us-vo.org>

- Partnership of major data/ compute centres in the USA:
 - IPAC, NASA-HESARC, NASA-JPL, NRAO, NOAO, SDSS, SAO, STScI
 - SDSC, NCSA, Pittsburgh
 - Globus, MS
- Webservices based
- Initially tools
 - Now moving to an architecture

The screenshot shows the NVO website with the following content:

- Navigation:** Home, Registry, Tools, Data Access, Publish, Education, Software Library, Grid Computing, Architecture, Contact Us.
- News:** NVO at the AAS, January 2005 VO Science Session at San Diego AAS Meeting, NVO News Archive.
- About:** What is the NVO? Who is Involved? Science Objectives, NVO in Use.
- Community:** NVO Meetings, International VO Alliance, NVO Summer School.
- Documents:** Recent NVO Documents: How to Publish to the NVO, Quarterly Report, October-December 2004, NVO-TeraGrid, First Year Results.
- Main Content:**

NVO - Facilitating Scientific Discovery

NVO's objective is to enable new science by greatly enhancing access to data and computing resources. In conjunction with the January 2005 meeting of the American Astronomical Society, NVO is releasing a first set of software tools and applications that make it easy to locate, retrieve, and analyze data from archives and catalogs worldwide.

 - [The NVO Registry Portal at STScI](#) finds source catalogs, observation logs, image archives, and other astronomical resources registered with the NVO.
 - [NVO DataScope](#) helps you discover and explore astronomical data from repositories around the world.
 - [OpenSkyQuery](#) lets you cross-match your data with numerous catalogs.
 - [NVO Spectrum Services](#) allow you to search, plot, and retrieve Sloan, 2dF, and other spectra.
 - [The Web-Enabled Source Identification with Cross-Matching \(WESIX\)](#) package lets you upload images to SExtractor and cross-correlate the objects found with selected survey catalogs.
 - [How to publish your data collections to the NVO](#) provides software libraries and sample code of VO Services for people who want to write their
- Team:** NVO Team Meeting, 25-26 April 2005, LSST Observatory, Tucson, AZ.
- IVOA:** IVOA Interop Workshop, Kyoto (16-20 May). Detailed information will soon be available at the [JVO](#) web site.
- DataScope:** Find images and catalog objects.

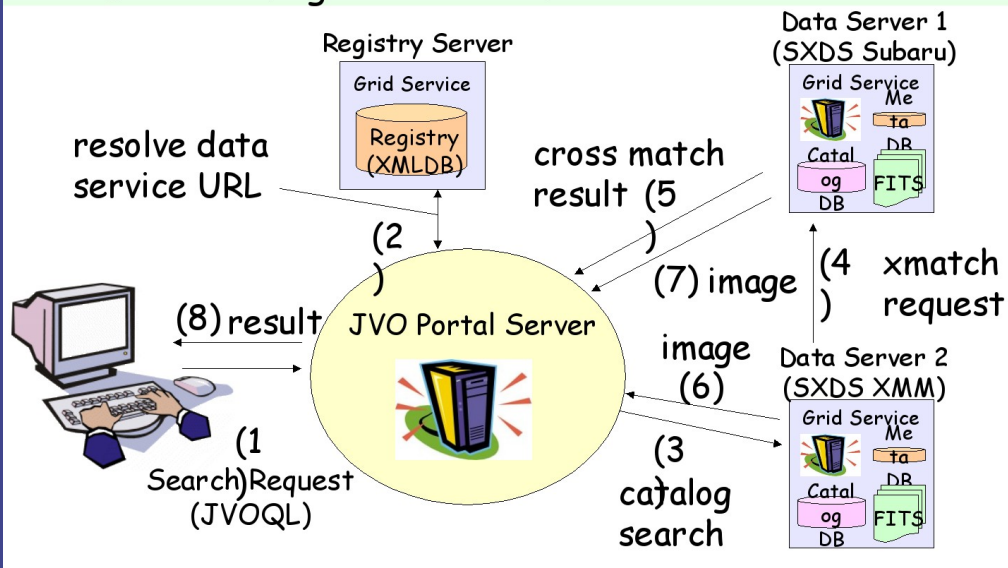
Japan-VO: JVO <http://jvo.nao.ac.jp>

- Partners

- NAOJ, JAXA, ICRR, Ochanomizu U, Osaka U, Titech, Fujitsu
- Data from Suburu, Astro-F, Nobeyama, Alma
- 1st JVO prototype based on Globus Toolkit 2 – 'grid'
- 2nd JVO prototype based on GT3 – grid services, improved performance

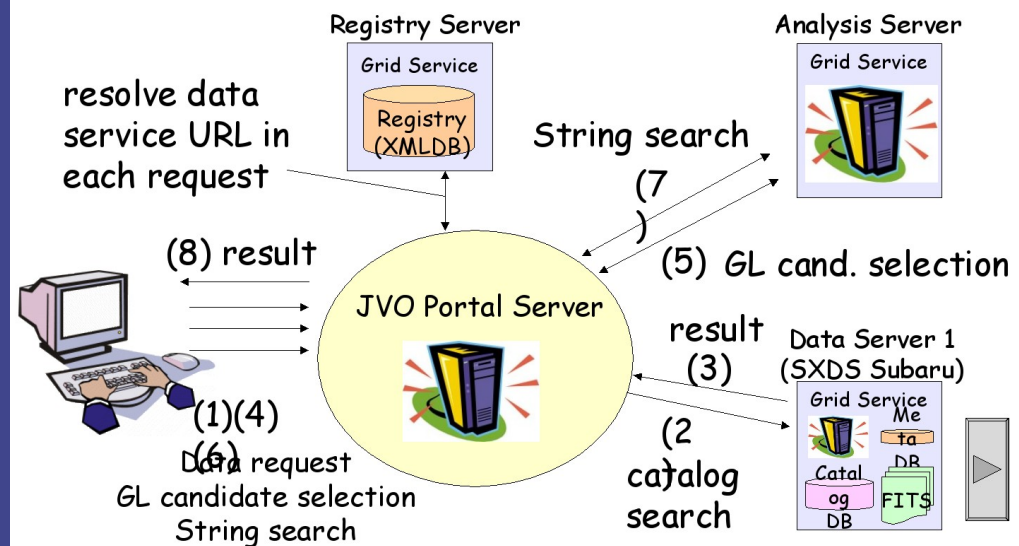
Demo 1: Cross match & Image request

Cross match of the optical and X-ray catalogs of SXDS and image retrievals.



Demo 2: Cosmic String Search

Data request to the SXDS optical catalog, GL candidate selection, String search by pattern recognition.



AstroGrid: UK's Virtual Observatory

Empowerment of scientists

- Improve the quality, ease, speed and cost effectiveness of on-line astronomy
- Make comparison and integration of data seamless
- Removing barriers to multi-wavelength astronomy
- Enable access to very large data sets

Project: 2001-2007: key data and resource providers in consortium





Euro-VO: VOTC Technology Centre



Infrastructure, new tools, resource
discovery, data mining+visualisation



AstroGrid: An example implementation



AstroGrid 2006.2 Release: Apr 2006

The image shows a screenshot of a Mozilla Firefox browser window displaying the AstroGrid website. The browser's address bar shows the URL <http://www.astrogrid.org/maven/docs/tmp/>. The website features a large banner for "ASTROGRID 2006.2" and a navigation menu on the left with sections like "About AstroGrid", "User's Guide", "Installer's Guide", "Developer's Guide", "Reference", "Further Information", "Other Formats", "Previous Releases", "About this documentation", and "Grid Status". The main content area is titled "AstroGrid Release 2006.2" and contains the text: "Astrogrid is pleased to announce software release 2006.2. From this site you will find useful information as follows:" followed by three icons: "Use AstroGrid", "Download and Install AstroGrid Services", and "Develop against AstroGrid".

Overlaid on the bottom right of the browser window is a screenshot of the "AstroGrid Workbench" application. The application window has a menu bar with "File", "Edit", and "Help". Below the menu bar is a toolbar with "Applications", "PLASTIC", and "ACR". The main area displays a row of application icons: "AstroScope", "HelioScope", "Launch", "Canned Workflows", "Workflow", "Lookout", "Myspace", and "Resources". At the bottom of the application window, it says "Not Logged In" and "© -2006,".

<http://www.astrogrid.org/launch>

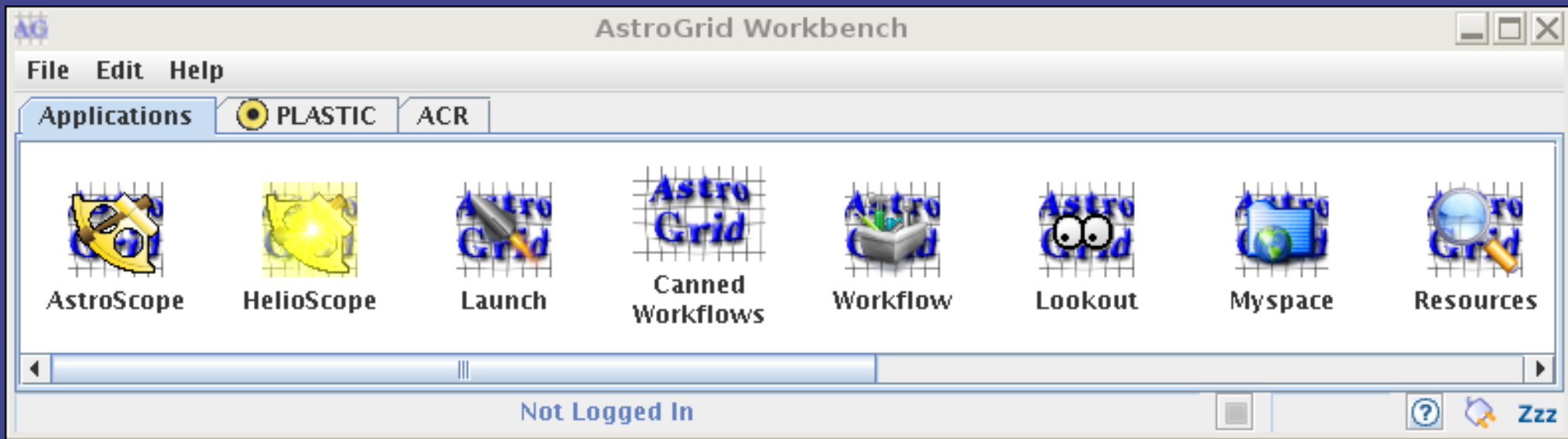
AG System

The screenshot displays the AG System interface, which is a collection of astronomical data analysis tools. The main window is titled "AstroGrid Workbench" and contains several panes:

- Applications:** A toolbar with icons for AstroScope, HelioScope, Launch, Canned Workflows, Workflow, Lookout, Myspace, and Resources.
- TOPCAT:** A table management tool showing a table named "RC3_sub1.xml" with 23,022 rows. It includes a "Spherical Plot" showing a 3D visualization of the data points on a sphere.
- VOSpec:** A spectral analysis tool showing a "VOSpec Spectra Viewer" with a plot of flux versus wavelength. The plot shows a dense cluster of red points on the right and a sparse cluster of blue points on the left.
- AstroScope:** A search tool showing a "Search Results" tree diagram. The tree is centered on "Catalogs" and branches out to various astronomical databases and surveys, including ADIL, ROSAT, Einstein, Chandra, and XMM-Newton.

The interface also includes a file explorer on the left, a status bar at the bottom, and a taskbar with various system icons and application windows.

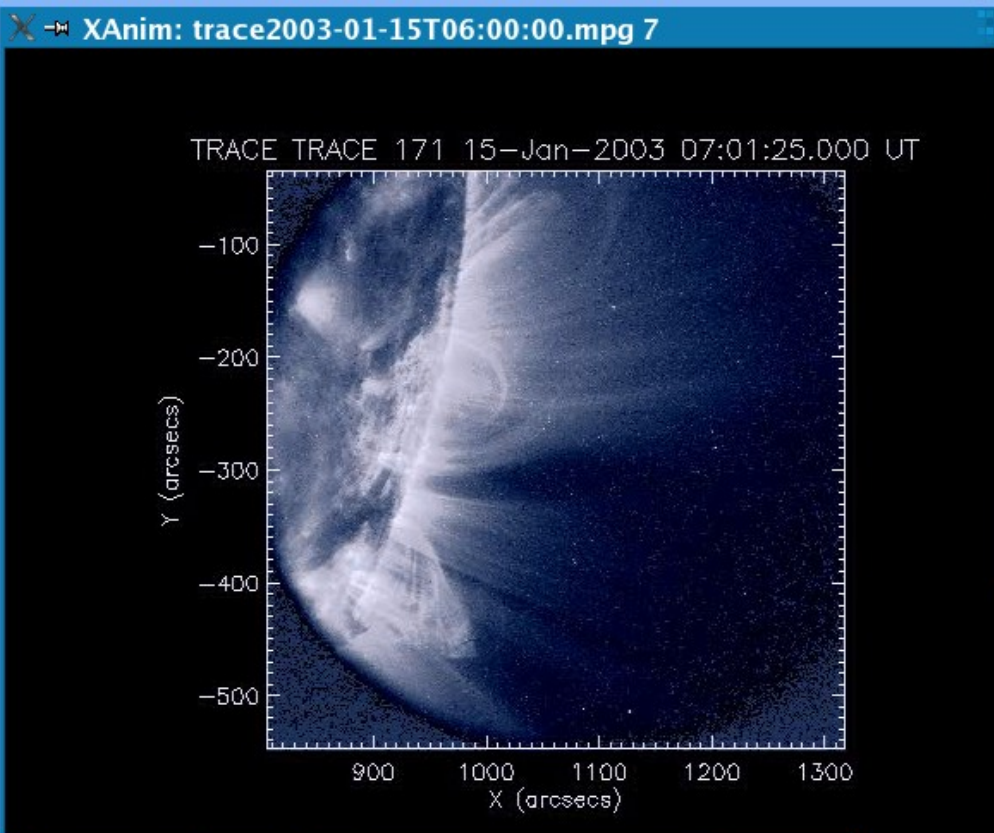




Workbench

- User Interface to VO services
- Delivery via Java Webstart technology
- Components
 - Registry
 - Find Data
 - Work with Apps
 - Workflows
 - Client Visualisation
- Enables Science

Solar Movie Maker



Jobs Monitor

Job

Refresh Rate 0 1 2 3 4 5 6 7 8 9 10

Name	Start	Finish	Status
eit query	Thu Apr 28 10:55:12 BS...	Thu Apr 28 10:55:46 BS...	COMPLETED
CDS query	Sun Apr 03 10:37:35 BS...	Sun Apr 03 10:41:38 BS...	COMPLETED
TimeMovieMaker	Fri Apr 29 17:03:48 BST ...		RUNNING
CDS query	Sun Apr 03 10:46:01 BS...	Sun Apr 03 10:49:53 BS...	COMPLETED
CDS query	Sun Apr 03 10:55:53 BS...	Sun Apr 03 10:57:10 BS...	COMPLETED
eit query	Tue Apr 26 08:44:49 BS...	Tue Apr 26 08:45:51 BS...	COMPLETED
SEC query	Tue Apr 26 09:15:24 BS...	Tue Apr 26 09:15:46 BS...	COMPLETED
CDS query	Fri Apr 01 13:42:28 BST ...	Fri Apr 01 13:45:02 BST ...	COMPLETED
solar demo	Mon Apr 04 18:40:54 BS...	Mon Apr 04 19:27:12 BS...	COMPLETED

Astrogrid Dashboard

Modules Help

- System
- AstroGrid
- UserInterface
- Scripting
- ExternalApps

User Interface

JobMonitor

VospaceBrowser

RegistryBrowser

ParameterizedWorkflow

Run...

Parameter Editor

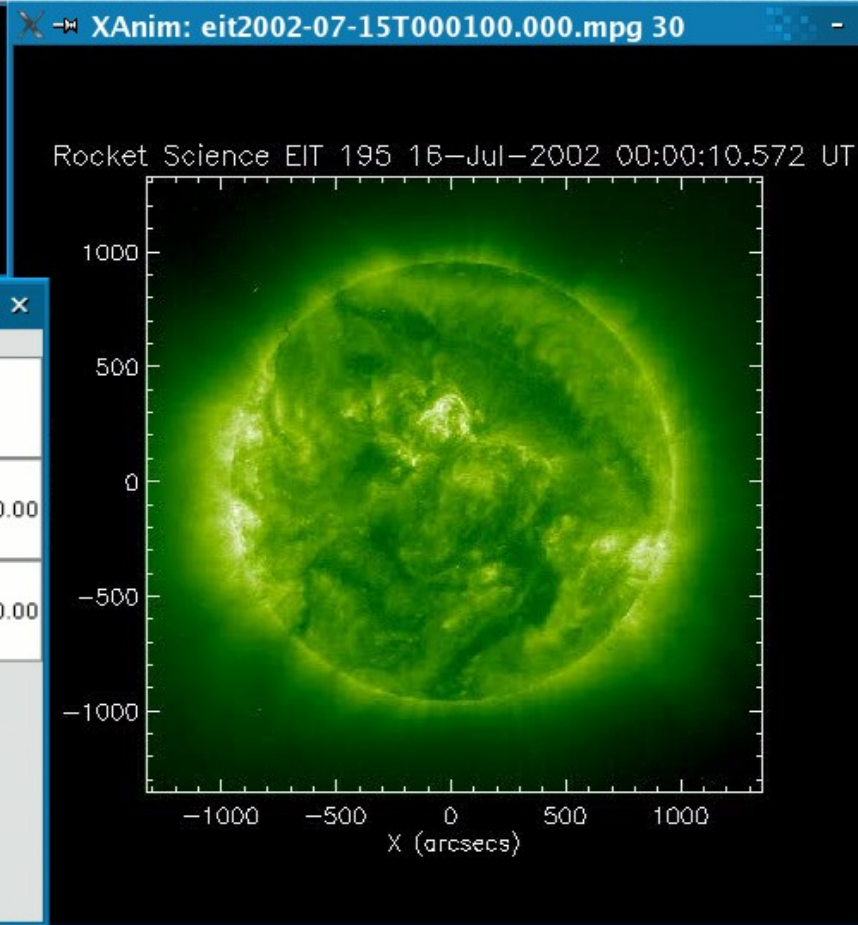
Solar Movie Maker (Time)

Instrument Name: eit

Start time for movie: 2002-07-28T01:00:00.00

End time for movie: 2002-07-28T06:00:00.00

Execute Cancel



Wider Relevance

- Use of IVOA stds vs use of more generic GGF stds
 - implementations in use – supporting large data flows, complex computational environments
- Comparison with other domains, e.g. Medical
 - workflows and application environment
 - server based system with pluggable clients for data visualisation
 - potential use from medicine to literature
- Supporting People Systems and Technical Systems
- Following talks and presentations look at specific areas where GGF standards may have relevance in supporting the construction of the Virtual Observatory

Closing

- The Virtual Observatory is developing a science driven infrastructure to allow access to distributed data and applications required to perform analysis and interpretation
 - 2006 is the year of rollout of significant systems in the USA, Europe, Asia and the UK
- Key Links
 - Astro-RG @ GGF: <https://forge.gridforum.org/projects/astro-rg/>
 - IVOA: <http://www.ivoa.net>
 - AstroGrid: <http://www.astrogrid.org>
 - NVO: <http://www.us-vo.org>
 - Euro-VO: <http://www.euro-vo.org>