Exploiting Virtual Observatory and Information Technology: Techniques for Astronomy

Lecture #6 Goal:
VO Workflows
Science Usage

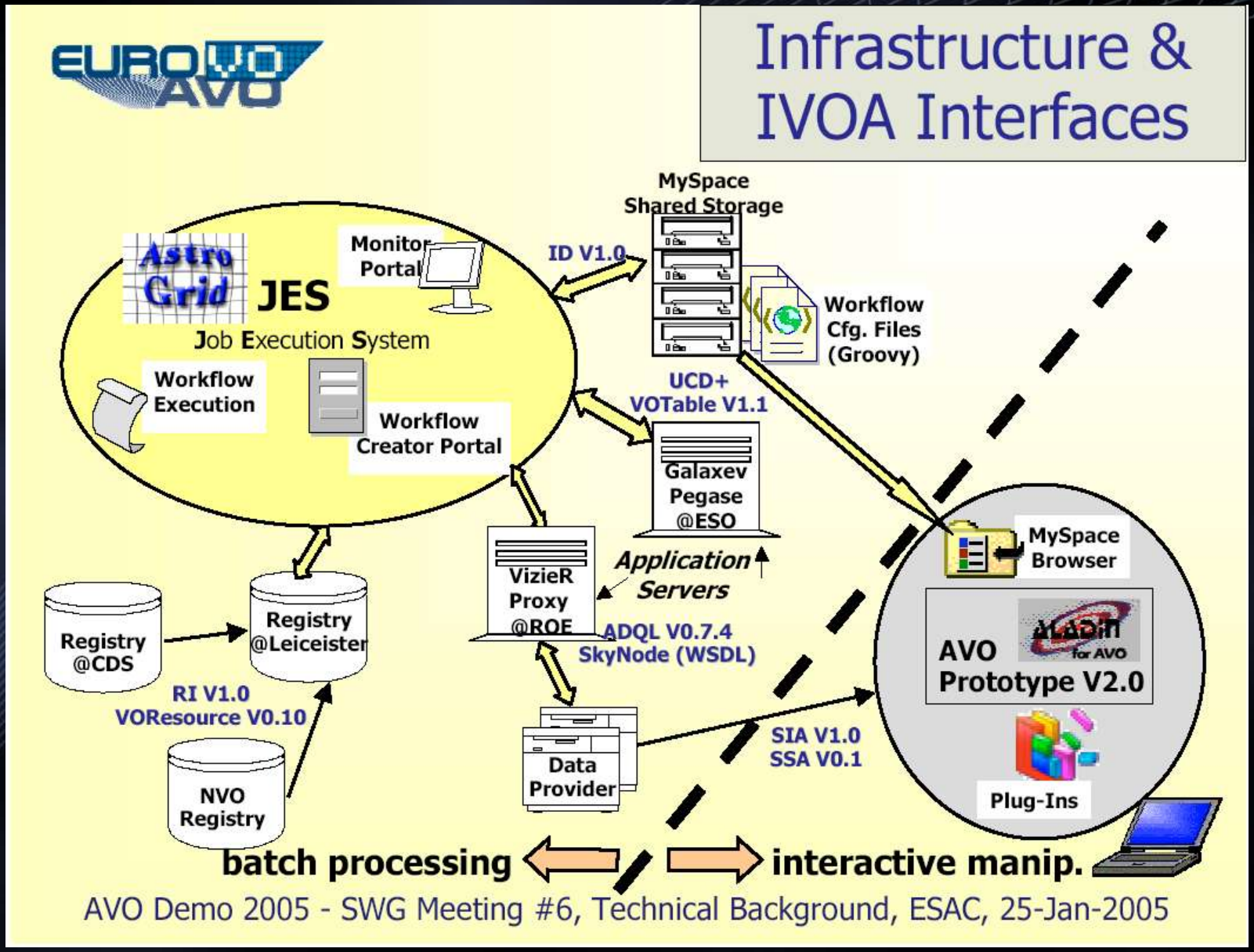
Nicholas Walton
AstroGrid Project Scientist
Institute of Astronomy,
The University of Cambridge

Summary: Lecture #6

- Review of VO Server Side Capabilities
  - Source catalogues created with the NVO
  - Scripting with Aladin
  - GAVO and Theory
  - The AstroGrid system
- Examples of Science Usage
  - Redshift determinations
  - Mining for brown dwarfs
  - Solar flares: movie creation

All lectures plus supporting material at:
http://www.ast.cam.ac.uk/~naw/VO-Course

Recall: server side/ client side ...

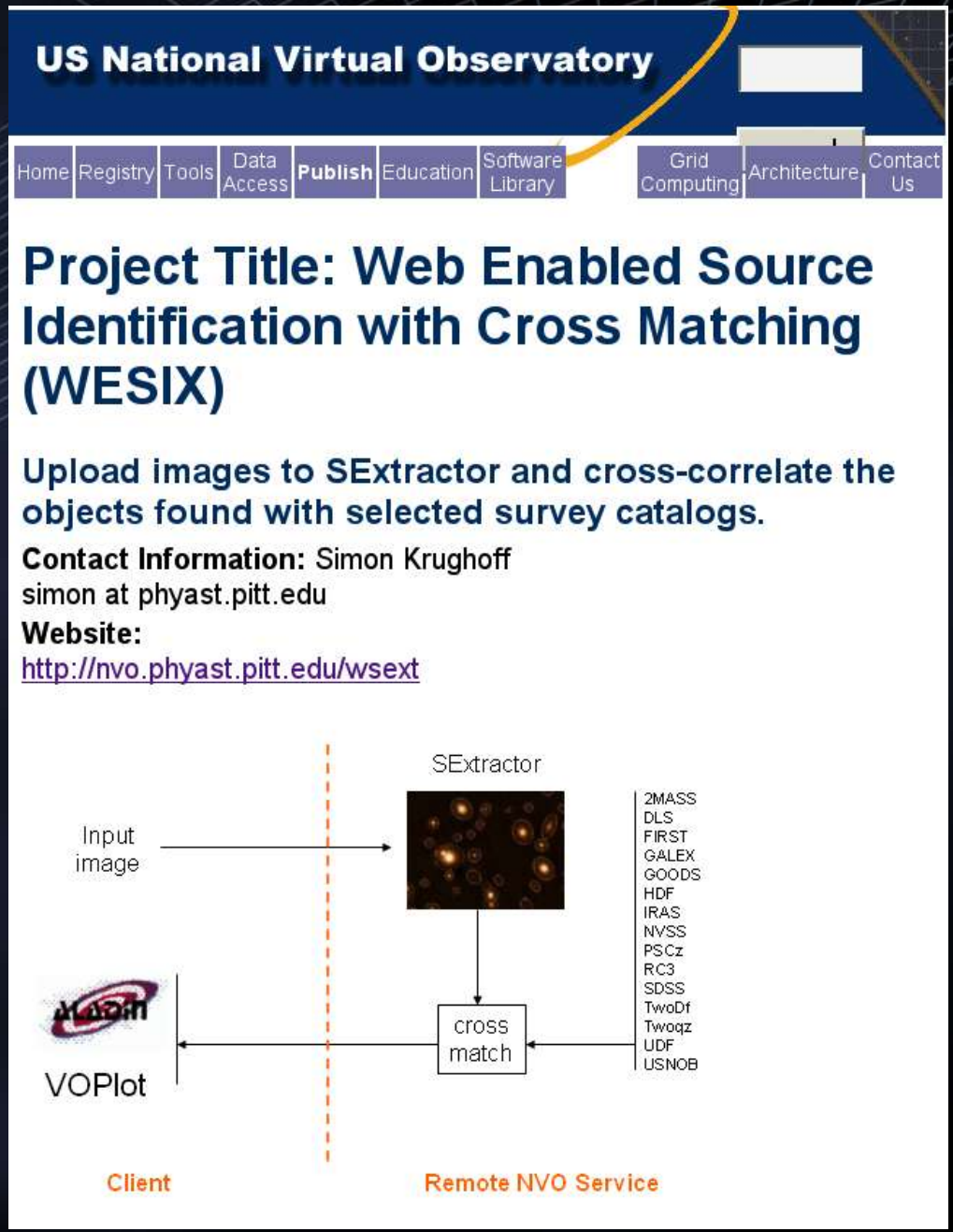


VO Server Side Applications

- A number of server based applications are now available which conform to VO standards
- These applications are aimed at large processing jobs
  - Some require user to remain 'logged in'
  - Asynchronous 'workflow' systems (e.g. AstroGrid) allow complex flows of jobs to be run, capture of 'state' allows user to return to collect results

NVO: WESIX

- Not currently a 'true' workflow example ... but hey ...
- For a single image
  - Create the object catalogue
  - Cross correlate with a number of major catalogue sources
  - Return results and plot



Try this at: http://frank.phyast.pitt.edu:8080/wsext/

WESIX

Home | Help | Plotter | Aladin

University of Pittsburgh

National Virtual Observatory: Web Enabled Source Identification with XMatching

Welcome to the homepage of WESIX

There are just a few steps to getting your source catalog back.
If you are interested in testing out this service,
here is a test file that works.

The following file can be used with tutorial #1

Here is an K band IR image taken with the FLAMINGOS instrument on the Kitt Peak 4m (contributed by Sam Schmidt, University of Pittsburgh). The image is within the footprint of SDSSDR2 and TWOMASS.

The following file can be used with tutorial #2

Here is a deep R band image taken on the 4m telescope at Kitt Peak (contributed by Andrew Hopkins, University of Pittsburgh). The image is within the footprint of SDSSDR2.

Step 1: Specify the file you want to upload

Browse...

Step 2: Select the catalog you would like to crossmatch with.

DLS
FIRST
GALEX
GOODS
HDFN
HDFS

Step 3: Choose whether you want to use the advance

Check this box if you wish to use the advanced pages.

Step 4: Submit your file for processing

Submit

Sponsored all or in part by the National Science Foundation via ITR grants AST0312498 and ACI0121671 and CAREER grant AST9984924. Additional funding provided by NASA through AISR NAG5-11996.
Developed in collaboration with the International Virtual Observatory Alliance.
Contact the NVO Help Desk to report problems and suggestions.

SExtractor Output Fields

Step 5: Select the output fields you would like in your catalog.

You need not select ra (ALPHA) or dec (DELTA). They are already included by default.

NUMBER
FLUX_ISO
FLUXERR_ISO
MAG_ISO
MAGERR_ISO
FLUX_ISOCOR
FLUXERR_ISOCOR
MAG_ISOCOR

Output Fields from SDSSDR2

Step 6: Choose the columns you would like included in your CrossMatched catalog

You need not choose ra and dec as outputs. They are included by default.

column name -------> UCD from SkyNode
(Note: some SkyNodes do not make UCD available)

camcol------->INST_ID
colc------->POS_CCD_X

WESIX

Home | Help | Plotter | Aladin

University of Pittsburgh

National Virtual Observatory : Web Enabled Source Identification with XMatching

Here is some information about this session.

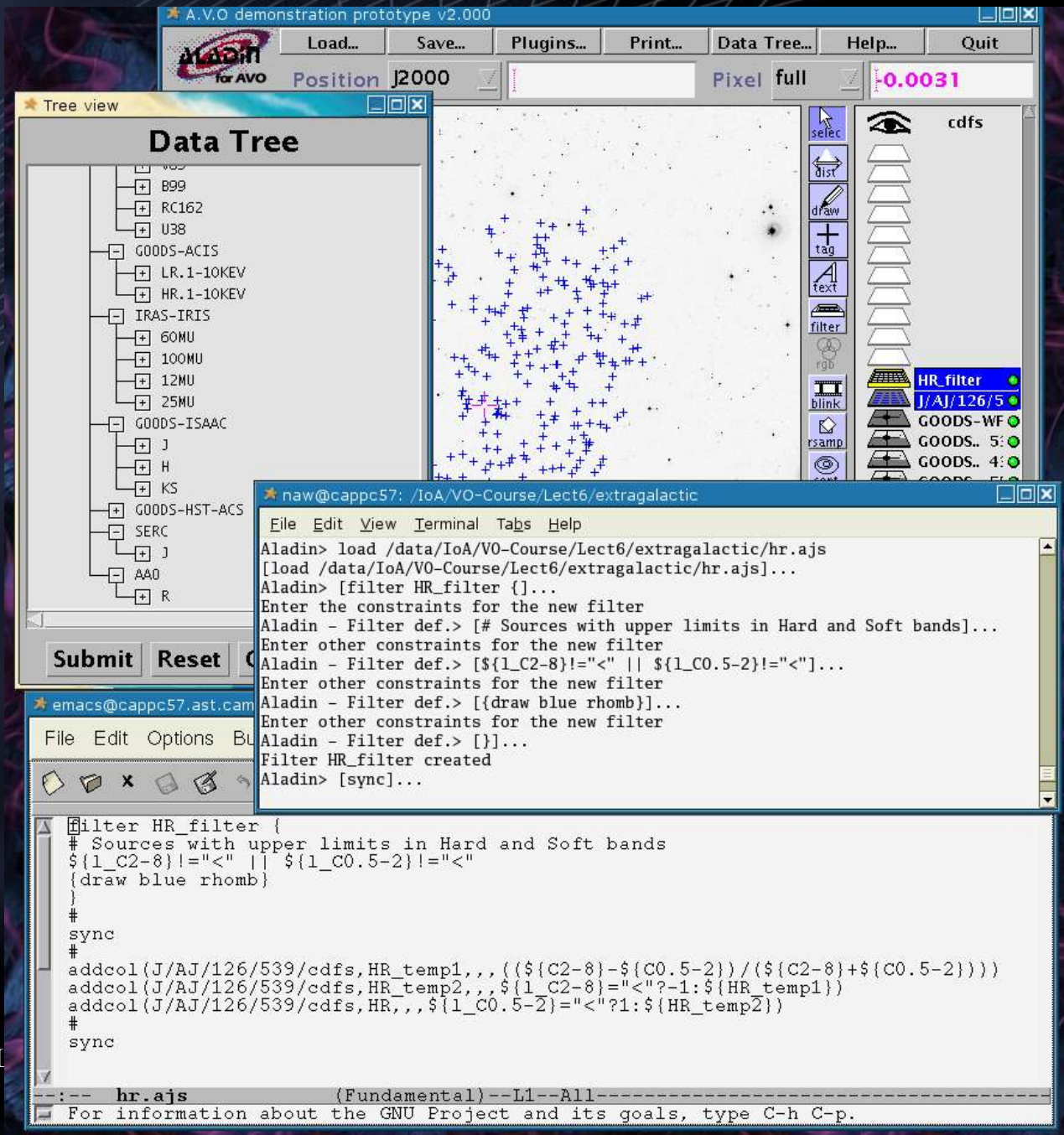
Output Fields:
NUMBER,FLUX_ISO,FLUXERR_ISO,MAG_ISO,MAGERR_ISO,FLUX_ISOCOR,FLUXERR_ISOCOR
Match Fields: colc_g,colc_i,colc_r,colc_u,colc_z
Selected Catalog: SDSSDR2

To begin visualizing your data, click **here**, or click the button on the menu bar that says Aladin.
To plot your data, click **here**, or click the button on the menu bar that says Plotter.

You may also download your catalogs in VOTable format from the following links.
Here is the matched file.
Here is the source catalog file.
Here are the sources from OpenSkyQuery.net in the region of your image.

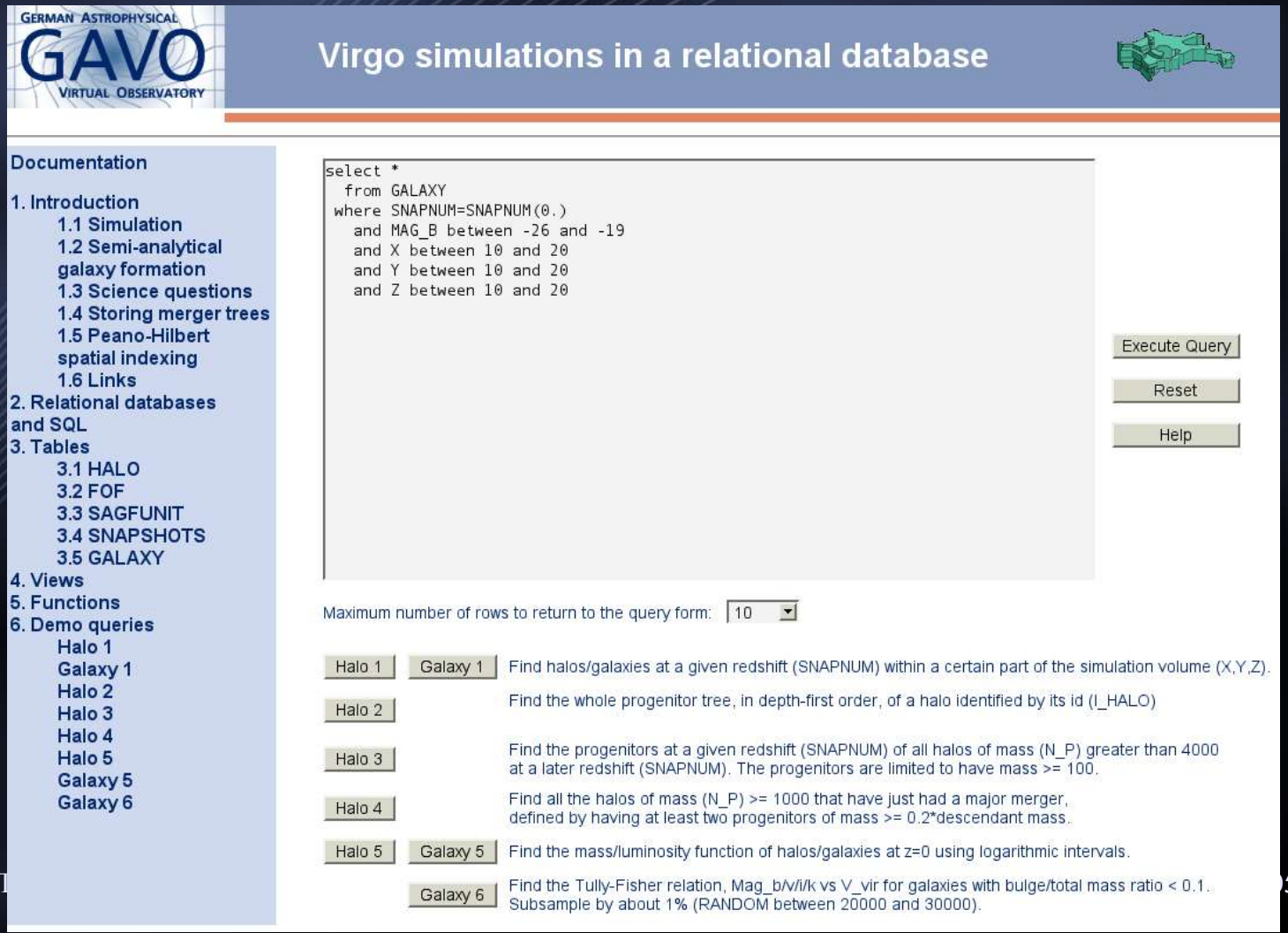
Aladin: Scripting: http://aladin.u-strasbg.fr/java/FAQ.htx#ToC21

- Aladin has a scripting capability
  - Wide range of commands available
- Possible to create PERL scripts controlling Aladin



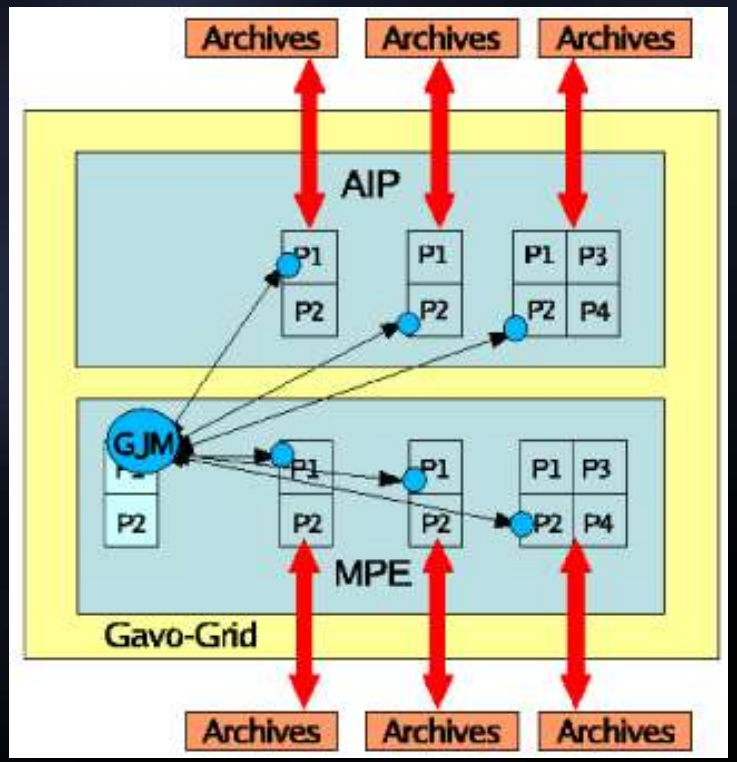
GAVO: theory services

- Access to simulations
  - Main limitation – limited data and non-standard meta-data
- Try it at http://www.g-vo.org/mpasims



GAVO: Cluster Finder

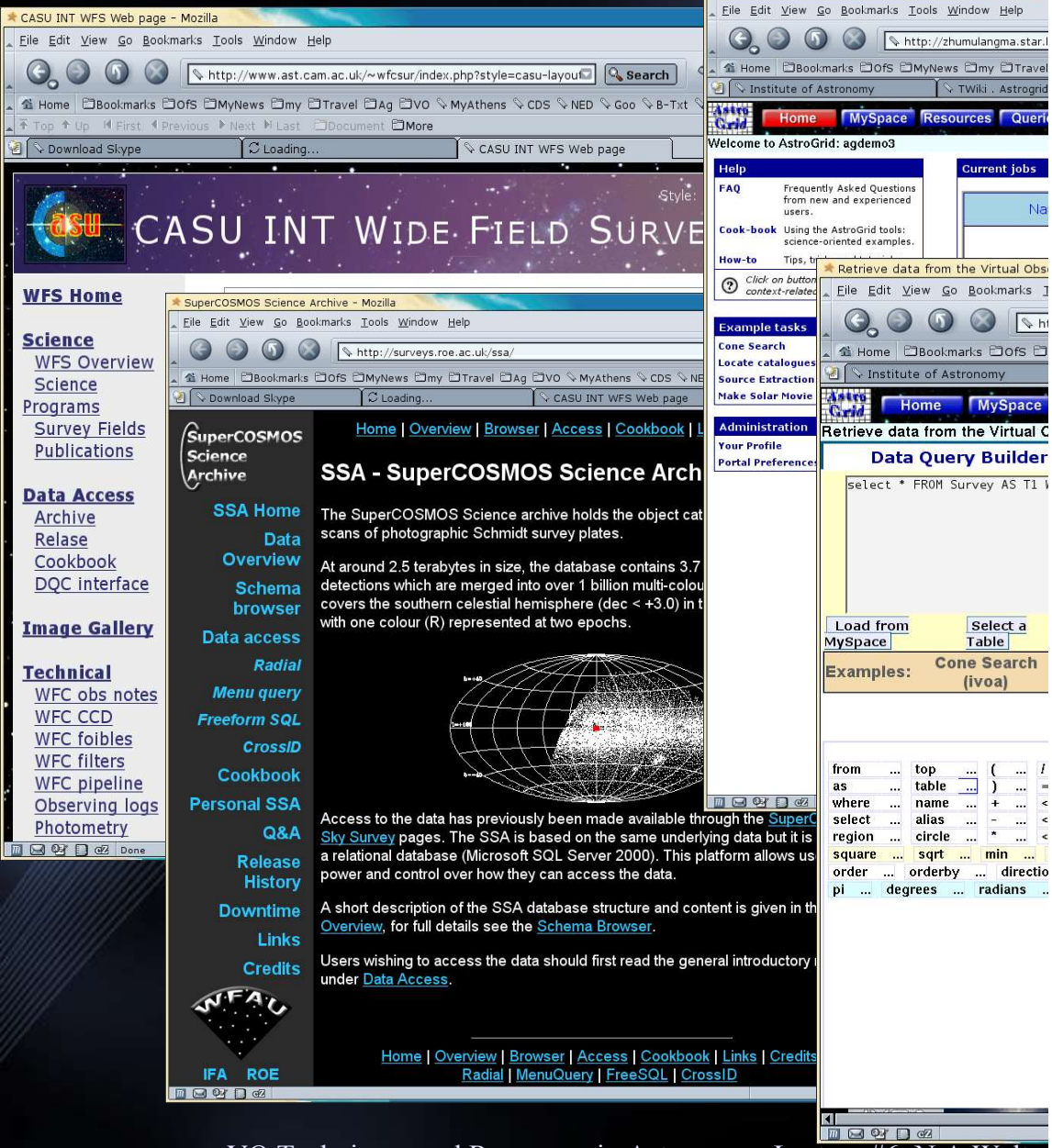
- Cluster finding algorithm: from Rosat RASS and SDSS catalogues
  - Compute intensive process
  - Runs on the GAVO grid: http://gavo3.aip.de:8080/GAVOGRID/
  - http://www.g-vo.org/clusterfinder/
  - Note: under development

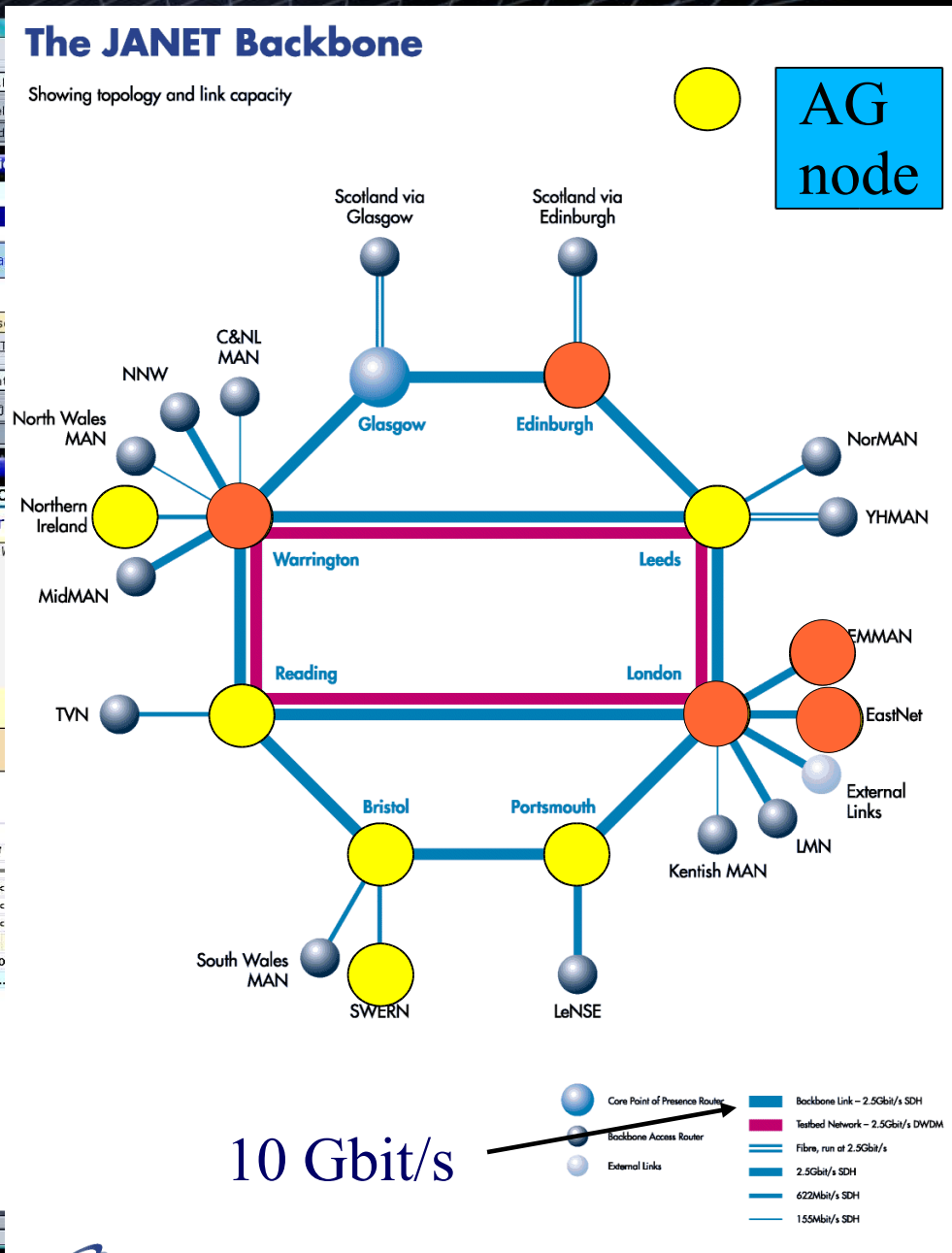


AstroGrid

- Links distributed applications and data
- System based on the idea of creating workflows
  - string sets of processes together
  - carry out the processing steps remotely
  - view the results locally

The AstroGrid Testbed





AstroGrid: Helper Applications

- Range of externally provided applications: tools to analyse and visualise end and intermediate data products
  - VOPlot: handles VOTable data
  - VOPlot3D: handles VOTable data – 3D visualisation
  - Topcat: tabular data and manipulation
  - Treeview: ability to browse MySpace
  - VOSpec: spectral plotting and analysis package
  - Specview: spectral plotting and analysis package
  - Aladin: data visualisation and catalogue access
    - MySpace browser capability (read-only)

But first, a quick word about DataScope, and why we need workflows ...

National Virtual Observatory DataScope

DataScope Help - Feedback

National Virtual Observatory — Hosted at NASA/HEASARC

What do we know about a region of the sky?
Use the Virtual Observatory DataScope to gather and organize information from astronomy archives and data centers around the world.

Enter a position(or name) and the maximum size of the region of sky you are interested in.

Object Name or J2000 Position:		(3c273 *or* 12 29 06, +2 3 8.6 *or* 187.27, 2.05)
Region size (degrees):	0.25	

Send Request | Reset Form

☐ **Bypass cache.** DataScope will issue a fresh request even if an identical request is in the cache.

Recent transient events and requests: (Click on *View* to see cached results.)

View **05h34m31.97s +22°00'52.1"** [$05^h34^m31.97^s$ +22°00'52.1"] (0.250°)	View **crab nebula** [$05^h34^m31.97^s$ +22°00'52.1"] (0.250°)
View **m3** [$13^h42^m11.23^s$ +28°22'31.6"] (0.250°)	View **tadpole galaxy** [$16^h06^m03.93^s$ +55°25'31.5"] (0.250°)
View **NGC 1193** [$03^h05^m55.67^s$ +44°22'59.2"] (0.250°)	View $17^h30^m28.00^s$ -34°41'30.0" (0.100°)

Display:

☑ **Basic Services**	☑ *ADS*	☑ *NED*	☑ *Simbad*	☑ *SkyView*			
☑ **Images**	☑ *Multi*	☑ *Optical*	☑ *Radio*	☑ *IR*	☑ *UV*	☑ *X-ray*	☑ *Other images*
☑ **Tables**	☑ *Observations* ☑ Multi ☑ Optical ☑ Radio ☑ IR ☑ UV ☑ X-ray ☑ Other	☑ *Objects* ☑ Survey ☑ Galaxies ☑ Stars ☑ Other objects	☑ *Other tables*				

You can control the categories of results diplayed by using these checkboxes. Some checkboxes, e.g., *Objects* serve as toggles that control a whole set of categories.

NVO: DataScope

IoA UNIVERSITY OF CAMBRIDGE

DataScope Internals:
http://heasarc.gsfc.nasa.gov/cgi-bin/vo/datascope/init.pl

- User entry of target position/name and search radius
- DataScope queries the registry to determine availability of resources
  - Datascope queries them
  - Returns results
- Results ordered
  - Image archives (SIAP services)
  - Catalogues (cone searches)
    - Observation catalogues
    - Object catalogues
  - List of resources that had no information or where unavailable
- Click through links to many of the returned results
- Possible to download results as a tar file

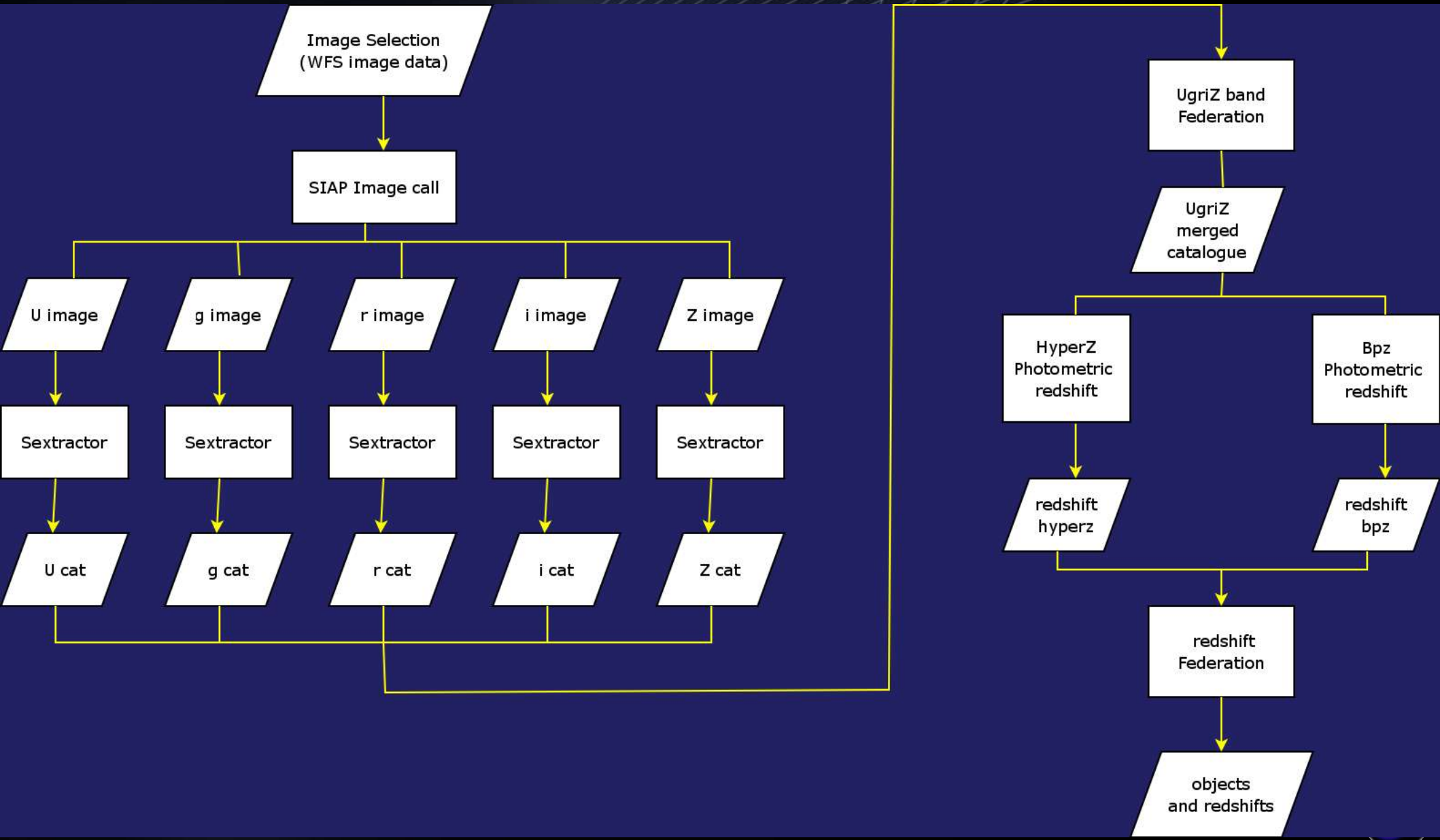
AstroGrid: AstroScope

- AstroGrid science service
  - Release date 30 June 2005 – check it out then
- Key additions to the DataScope
  - Returns info from UK data providers, e.g. CASU at the IoA
  - Importantly – allows download of results to AstroGrid's Myspace – hence results will be available for incorporation into AstroGrid workflows
- Full details (currently under development) at http://wiki.astrogrid.org/bin/view/Astrogrid/AstroScope

Photometric Redshifts – an Extragalactic Case

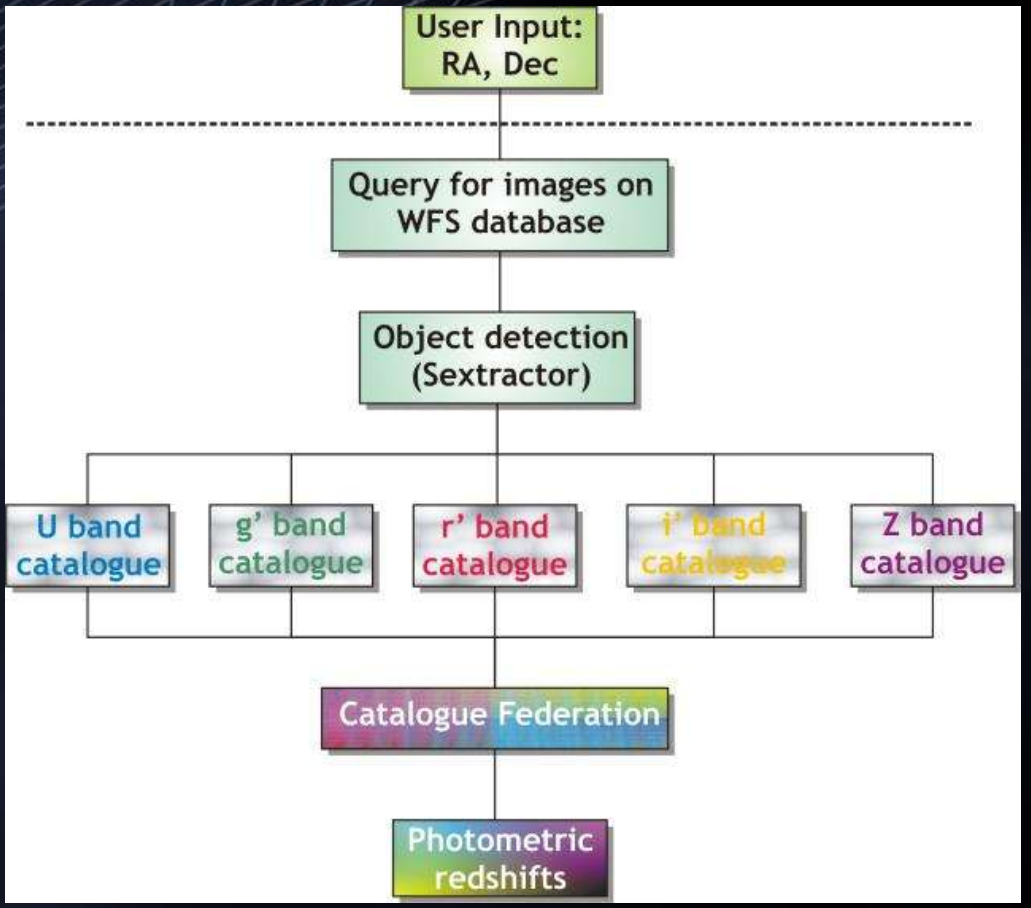
Extragalactic Case Workflow

See: http://wiki.astrogrid.org/bin/view/Astrogrid/AgDemoDec2004Extragalactic



AstroGrid Redshift Science Service

- Packaged workflow
  - Enter RA, Dec, radius
  - System returns objects and redshifts
- User Options
  - Input data (INT-WFS, SDSS)
  - Redshift apps (hyperz, bpz)
- Defaults
  - Source extraction double image mode (r' image ref)
  - Plus lots of other sensible default configs for the cross match, the redshift apps etc.



See - http://wiki.astrogrid.org/bin/view/Astrogrid/RedshiftMaker

Mining for Brown Dwarfs – a Galactic Case

Accessing & processing distributed data

- Select INT-WFS Observing Log entries for Pleiades in *i* and *z* bands with small photometric and pointing errors
  - Cross-match to get *i* and *z* observations of same fields
- Extract Zero-point, Seeing, Exposure time from Logs
- Construct image URLs for Simple Image Access server at Cambridge
- Feed images to SExtractor hosted at JBO
- Cone search 2MASS catalogue at ROE for Pleiades region
- Cross-match lists of extracted i and z sources and 2MASS sources, all held in MySpace at Leicester
- Use TopCat tool to access MySpace files and make colours

Colour-Cutter Tool

- An AstroGrid packaged science service
  - Release date 30 June 2005
  - http://wiki.astrogrid.org/bin/view/Astrogrid/ColourCutter
- User selects a position, radius on the sky and inputs a colour cut selection (e.g. r' – i' > 1.0)
- Service returns list of objects meeting the colour cut critera
- Workflow searchs image data sets, generates object catalogues for the relevant image files, and performs the source selections based on the object colours

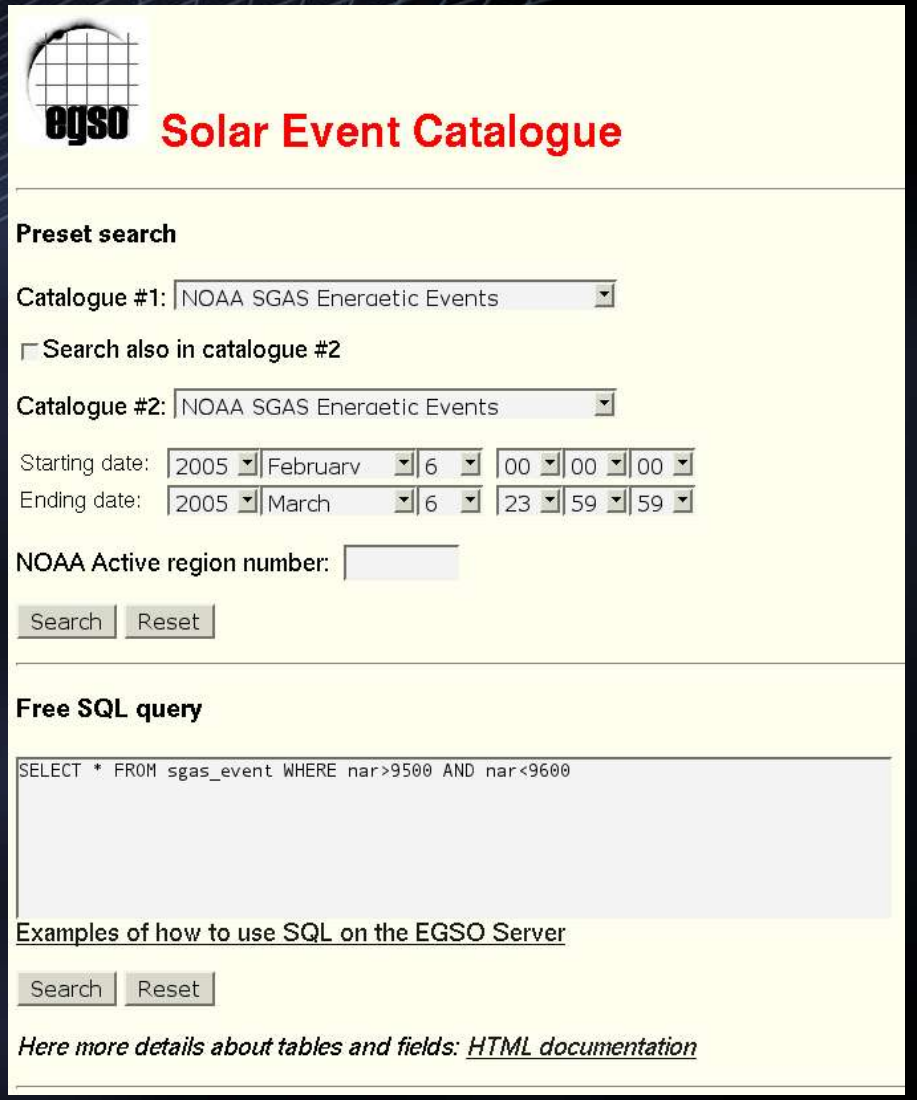
Solar Moviemaker – a Solar Physics Case

Science Case: Movie of solar energetic events

- Obtain a series of movies of solar energetic events (e.g. solar flares).
  - Solar energetic events, for example flares, coronal mass ejections, are routinely identified and event catalogues are compiled.
  - Event info in catalogues: start and end time of event, magnitude, location on the solar disk, Active Region number
- Currently to study a set of energetic events:
  - for each event in list, manually place a request for data to archive
  - Download data, calibrate images (SolarSoft libraries), generate movie

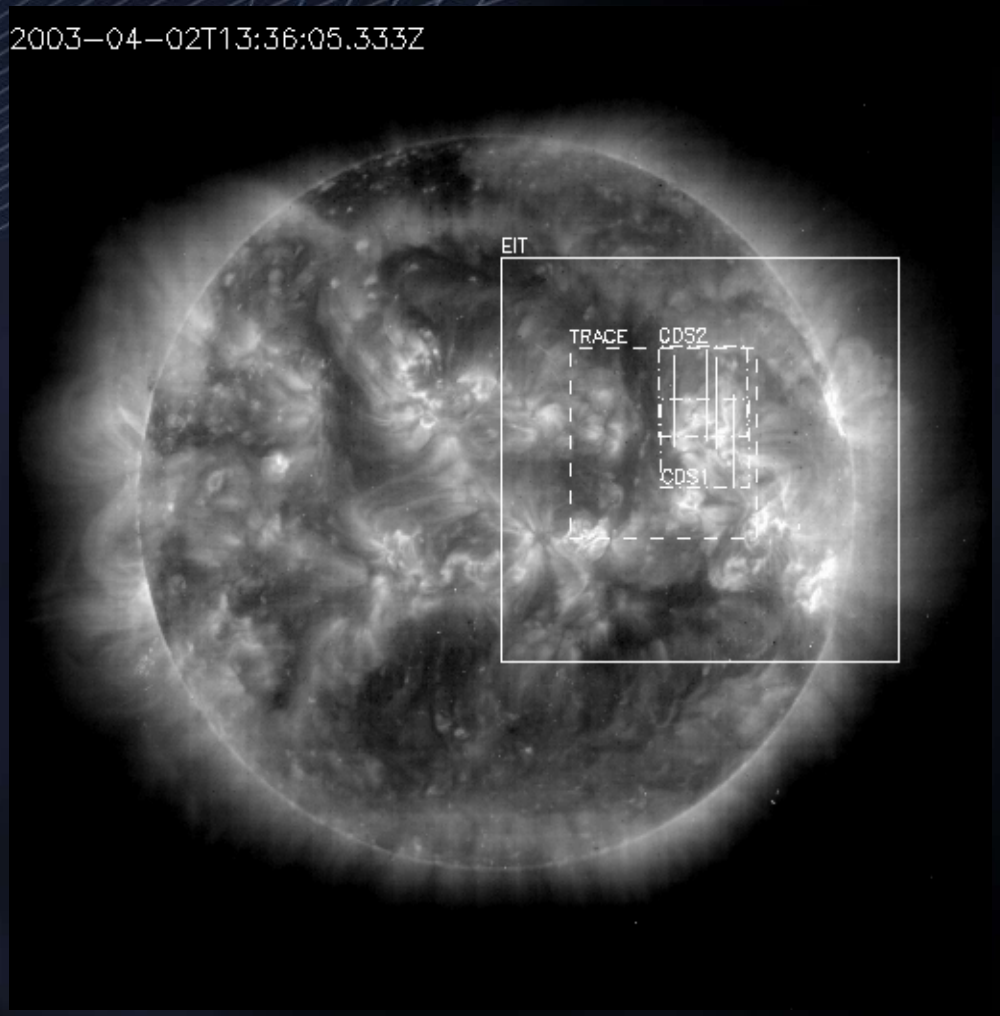
Solar Event Catalogue

- This catalogue provides links to data resources associated with 'solar events'
- The catalogue is available for interactive use at: http://sec.ts.astro.it/sec_ui.php
- The catalogue is queried remotely in this example solar workflow.



Dataset: TRACE

- TRACE is a high resolution (0.5”), high cadence (20s) EUV solar imager.
- SURF TRACE archive at MSSL – hourly files with >100 images

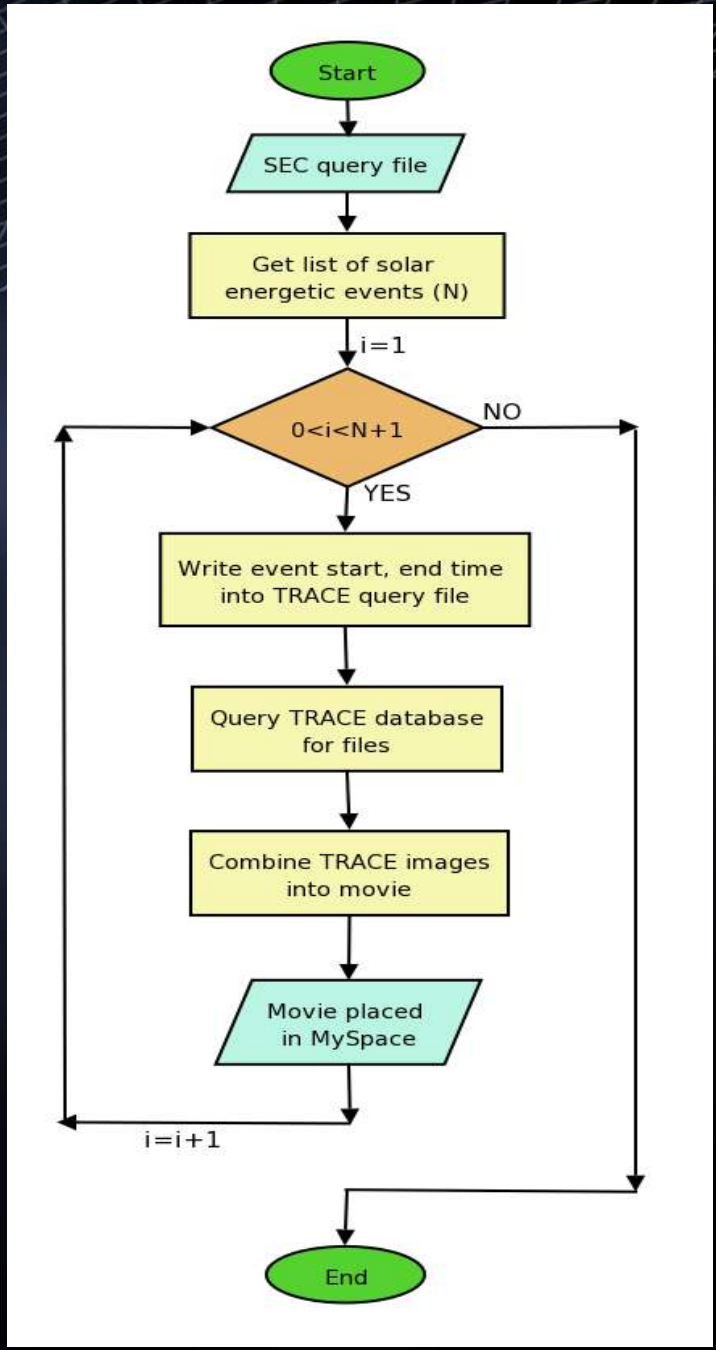


Movie Maker Workflow

- Step A: Query event catalogue
  - obtain VOTable of solar energetic events via query to EGSO/SEC
  - Query can be over active region number, magnitude of event
  - Example query:

```
select * from sgas_event AS s
where s.nar>9850 and
s.nar<9880 and
s.xray_class>'M5
```

- Start loop over events



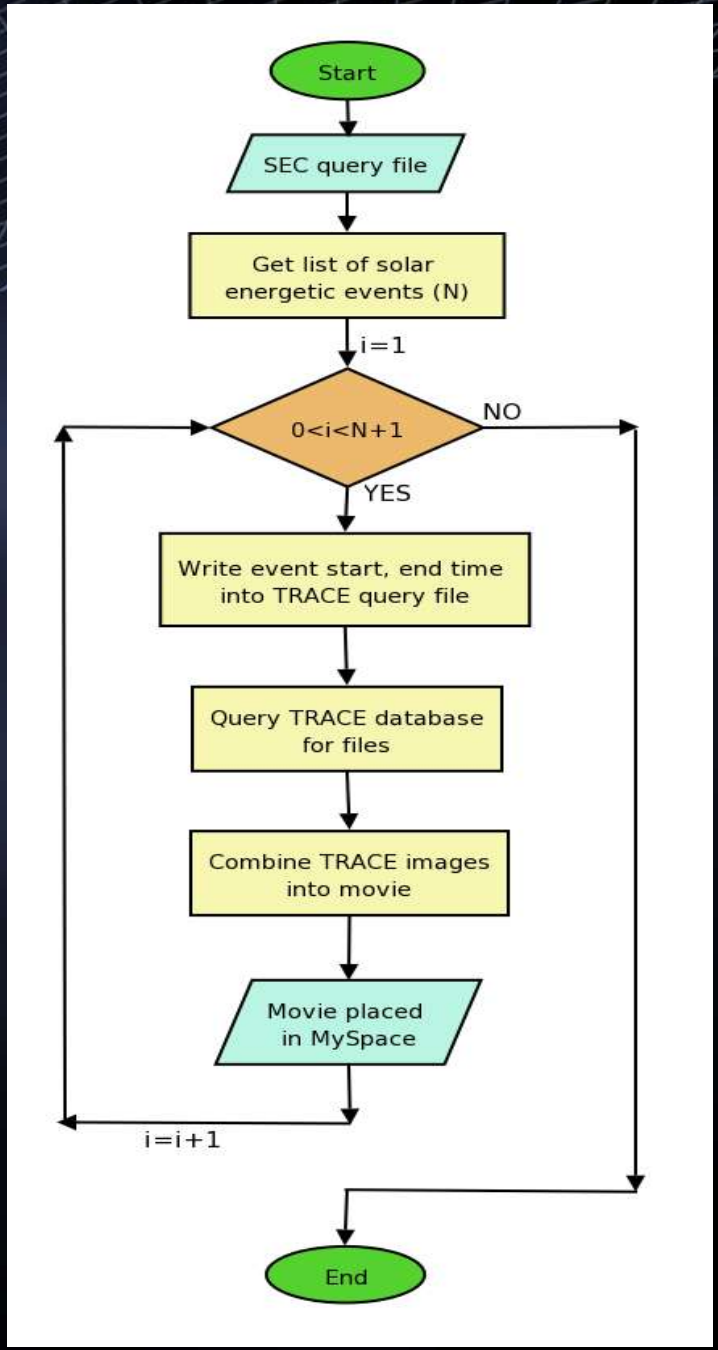
Movie Maker Workflow-contd

- For each event:
- Query TRACE database
  - start time and event time of solar event obtained from SEC query written into query file for TRACE dataset:

```
SELECT * FROM fits_trace AS T1
where T1.Keywords/img_time >=
\'${refstarttime}\'

and T1.Keywords/img_time <
\'${refendtime}\'
```

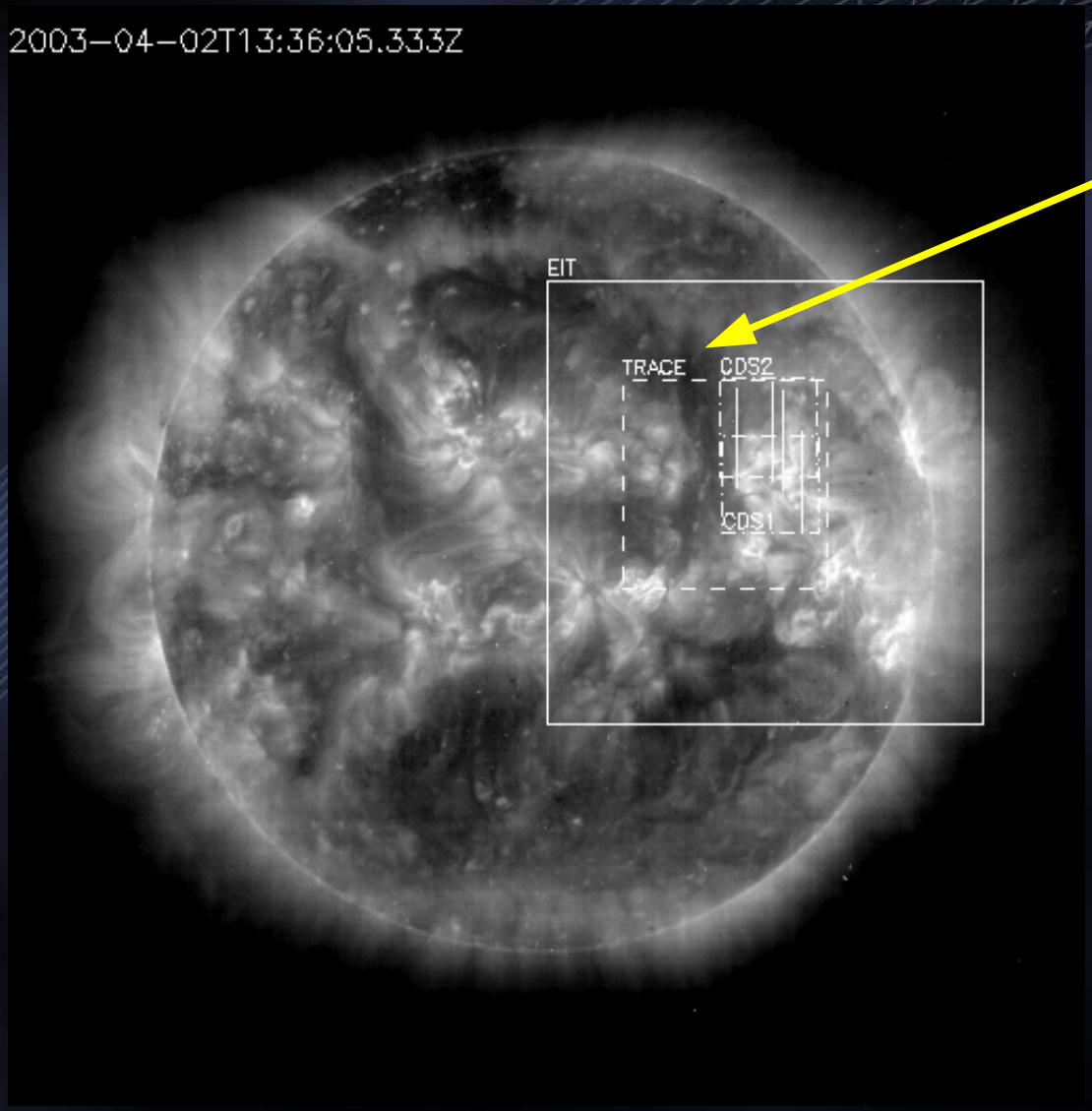
- Start and end time are now workflow variables



Movie Maker Tool

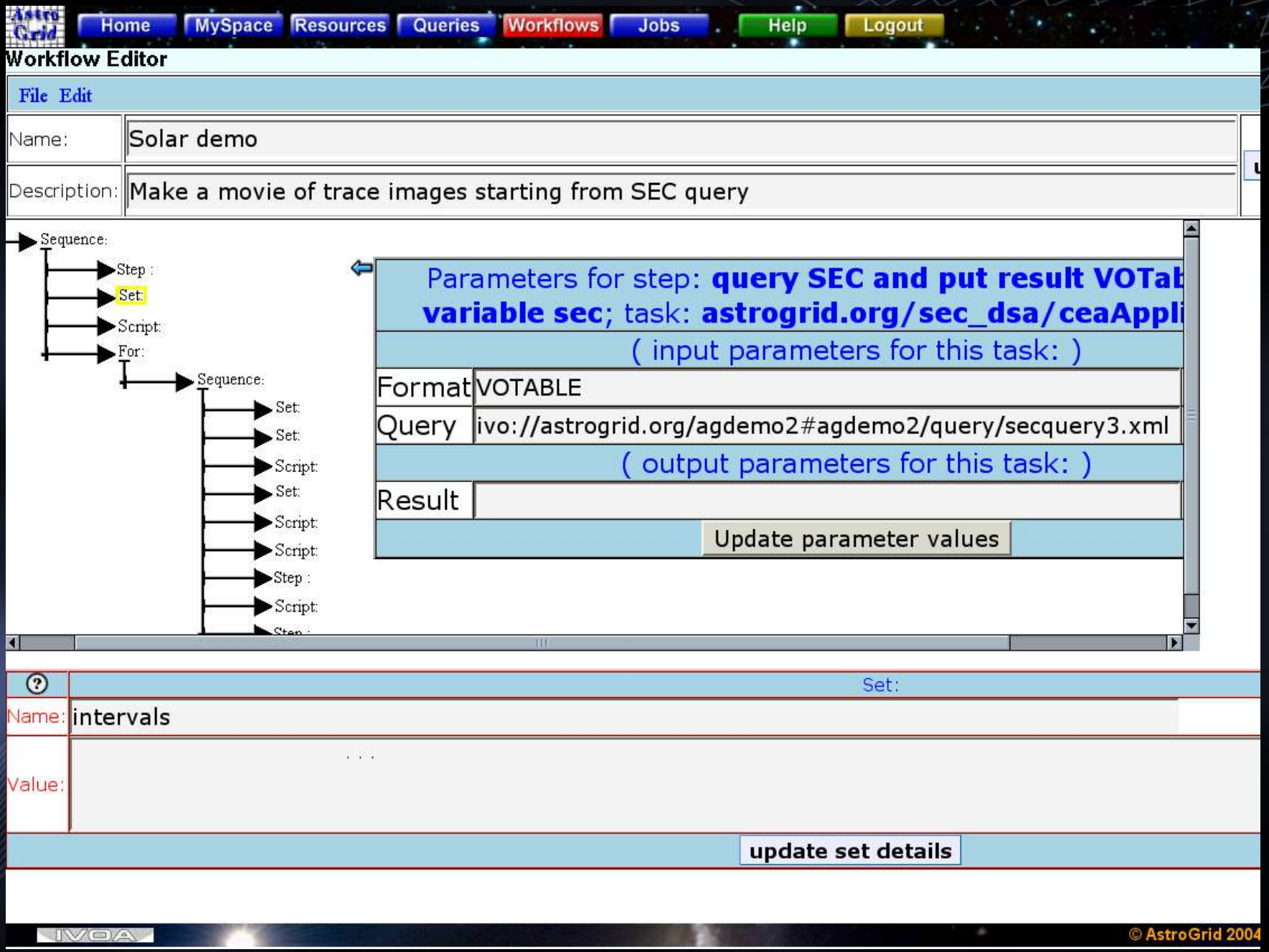
- TRACE files are presented to the movie-maker tool as a series of URLs. CEA downloads URLs as temporary files local to the movie maker tool
- IDL Solarsoft – a multi-instrument IDL library for analysis of solar images
- Shell script generates IDL Solarsoft session
- Solarsoft: read images, select wavelength and apply calibration routines
- One MPEG per hourly file generated – concatenate hourly MPEGs into 1 movie

TRACE Field of View

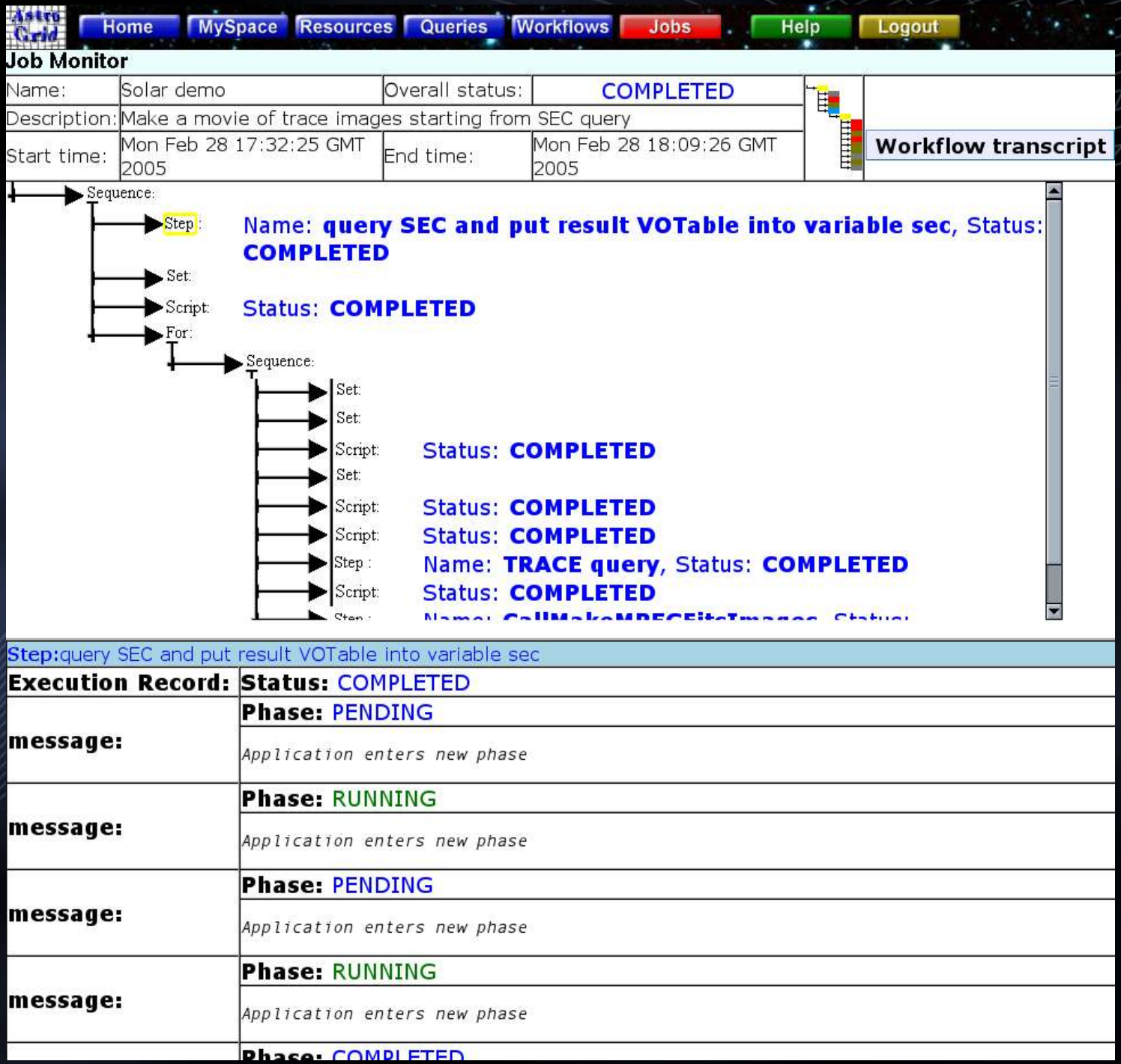


Trace FoV

Solar Movie: WorkFlow



Solar WorkFlow: Log File



Solar Movie Maker

- AstroGrid Science Service
  - Solar movie maker
- Prepackaged workflow
  - User enters NOAA event number or time interval of interest
  - Service returns movie of that event/ time
- Options
  - Data sets (SOHO EIT/ TRACE/ CDS)
  - Time range or event
- Details at http://wiki.astrogrid.org/bin/view/Astrogrid/MovieMaker

AstroGrid Beta Testing

- 1st functional release: 31 March 2005
  - At this stage usernames/ passwords will be available – try out the science services and workflows discussed in this lecture
- Beta testers invited
  - Details linked from http://wiki.astrogrid.org/bin/view/Astrogrid/AgScience
- For further information on AstroGrid or any other issues raised in these lectures, please contact me for further details at naw@ast.cam.ac.uk or Room H37 or tel 37503

Lecture 6: Acknowledgements + Refs

- Refs as noted on the slides

All lecture slides plus supporting material at:
http://www.ast.cam.ac.uk/~naw/VO-Course

Next Lecture:
The Radio/ sub-mm domain – data challenges (John Richer)