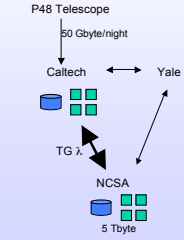
Building and Using Services

Roy Williams
California Institute of Technology

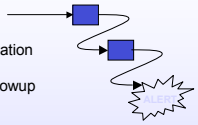
- Palomar-Quest
- Making a simple portal
- VOTable and UCD
- Large scale processing
- VIM as a VO mashup

Palomar-Quest Survey

Caltech, NCSA, Yale



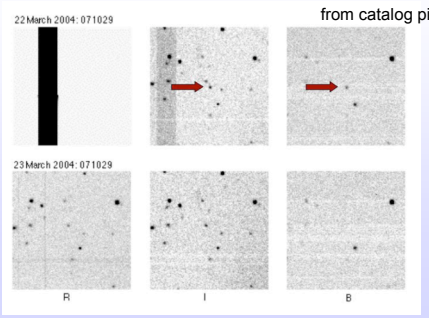
Transient pipeline
computing reservation
at ***sunrise***
for immediate followup
of transients



Synoptic survey
image coaddition
for ultrafaint detection

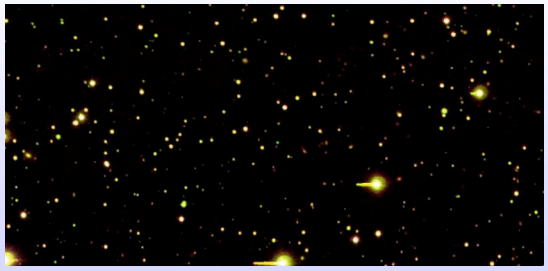
Transient from PQ

from catalog pipeline

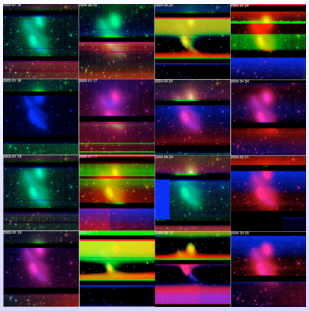


PQ stacked images

from image pipeline



Synoptic Image Stack



Wide-area Mosaicking (Hyperatlas)

An NVO-Teragrid project

Caltech

DPOSS 15º

Griffith Observatory "Big Picture"

Palomar-Quest as a VO

- Images available by SIAP
- Catalog by Skynode
- Casjobs installed for batch database queries
- VOEvents used for transients (click)
- NESSSI used for secure, async services

Making a portal for a command line application

The command-line application

```
$ mycode -apple 56 -banana 5346 -orange SDSS
```

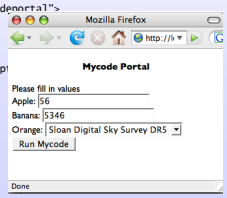
(1) Make HTML form

```
<center> <h4>Mycode Portal</h4> </center>

Please fill in values<br/>
<form method=GET action="http://localhost/cgi-bin/mycodeportal">

Apple: <input name="apple"><br/>
Banana: <input name="banana"><br/>
Orange: <select name="orange"><br/>
  <option value="SDSS">Sloan Digital Sky Survey DR5</opt
  <option value="2MASS">2MASS All-Sky Catalog</option>
</select>
<input type=submit value="Run Mycode">

</form>
```



Making a portal for a command line application

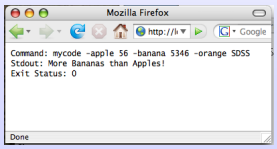
(2) Make CGI wrapper

```
import cgi
form = cgi.FieldStorage()

cmd = "mycode -apple %s -banana %s -orange %s" ]
     % (form["apple"], form["banana", form["orange"])

print "Content-type: text/plain\n"
print "Command %s" % cmd

pipe = os.popen(cmd)
print "Stdout %s", pipe.read()
print "Exit status %s", pipe.close()
```



More VOTable

```
<VOTABLE version="v1.0">
 <RESOURCE type="results">
   <DESCRIPTION>Results from query to  NASA/IPAC Extragalactic Database (NED) .... </DESCRIPTION>
   <TABLE ID="NED_MainTable" name="Searching NED within 0.3 arcmin of 178.542980, 10.796330">
     <DESCRIPTION>Main information about object (Cone Search results)</DESCRIPTION>
     <PARAM ucd="time.equinox" datatype="char" value="J2000.0" name="Equinox"/>
     <PARAM ucd="pos.system.coord" datatype="char" value="Equatorial" name="CoordSystem"/>
     <FIELD ucd="meta.id" datatype="int" ID="main_col1" name="No.">
       <DESCRIPTION>A sequential object number applicable to this list only.</DESCRIPTION>
     </FIELD>
     <FIELD ucd="meta.id;meta.main" datatype="char" arraysize="30" ID="main_col2" name="Object Name">
       <DESCRIPTION>NED preferred name for the object</DESCRIPTION>
     </FIELD>
     <FIELD ucd="pos.eq.ra;meta.main" datatype="double" ID="main_col3" unit="degrees" name="pos_ra_equ">
       <DESCRIPTION>Right Ascension in degrees (Equatorial J2000.0)</DESCRIPTION>
     </FIELD>
     <FIELD ucd="pos.eq.dec;meta.main" datatype="double" ID="main_col4" unit="degrees" name="pos_dec_equ">
       <DESCRIPTION>Declination in degrees (Equatorial J2000.0)</DESCRIPTION>
     </FIELD>
.....
```

History

- Created at CDS Strasbourg
  - (the first VO prototype)
- Harvested
  - From 5000 tables, 20000 table columns
  - To create ~450 UCD words
- Example
  - pos.eq.ra means right ascension
  - think of a UCD as a “semantic type”:

```
int n;
float x;
pos.eq.ra alpha;
```

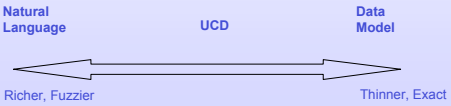
UCD is a semantic type (not a name)

example: table of double stars

name	alpha1	delta1	alpha2	delta2
ucd	**pos.eq.ra**	**pos.eq.dec**	**pos.eq.ra**	**pos.eq.dec**

Summary

- UCD is inherently fuzzy
  - does not contain all metadata
- UCD is a description, not a unique name
- UCD already works
  - easy to shoot at!
- UCD is mined from large amount of metadata
  - phenomenological, not opinion



General Syntax of IVOA UCD

- [namespace:] w.o.r.d; [namespace:] w.o.r.d; ....
- Character set is
  - A-Z, a-z, 0-9, hyphen, + 3 special
  - No white space
  - Case insensitive
- Three special characters
  - . period for hierarchy
  - : colon for namespaces
  - ; semicolon for separating words

UCD Tree

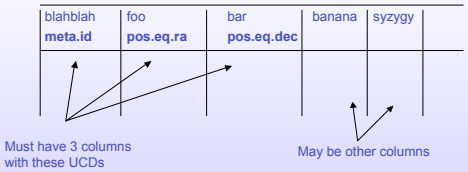
arith	Quantities related to arithmetic and mathematics, including count, difference, ratio
em	The electromagnetic spectrum
meta	Quantities related to metadata, such as identifiers, flags, notes, URL, and
instr	Quantities related to an instrument; typical sub-levels are telescope, observatory, etc.
obs	Observation methods such as detector, filter, plate, spectrograph, exposure time, etc.
phot	All photometric measurements, organized according to the wavelength; includes polarization.
phys	Generic physical quantities, such as length, velocity, mass, and including atomic & molecular concepts and properties, temperature, pressure, gravity, etc...
pos	Position in the sky, reference frames; including equatorial, galactic etc coordinates; geocentric, heliocentric etc; and precession and nutation.
spect	Quantities related to spectroscopic measurements
src	Properties of the observed source of radiation: source classifications and morphology, extension in the sky, variability,
stat	Statistical quantities and quantities related to model fitting, including concepts such as error, maximum, residuals.
time	Quantities related to time.

Multiple words

Quantity	UCD
Magnitude	**phot.mag**
Flag (weather, reliability, etc)	**meta.code**; phot.mag
Error on magnitude	**stat.error**; phot.mag
Maximum temperature of an instrument	**phys.temperature**; instr; stat.max
Error of a V-band magnitude	**stat.error**; phot.mag; em.opt.V

Cone Search

- service request is a cone
- service response is a table:



Must have 3 columns with these UCDs

May be other columns

VOTable/VOEvent

- Uses UCD to indicate what has been observed

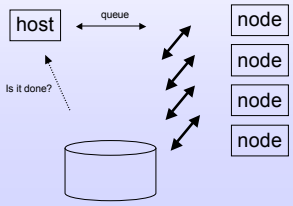
```
<Group name="SQUARE_GALAXY_FLUX">
    <Param name="counts" value="73288" unit="ct" ucd="phot.count"/>
    <Param name="peak" value="1310" unit="ct/s" ucd="arith.rate;phot.count"/>
  </Group>
  <Param name="seeing" value="2" unit="arcsec" ucd="instr.obsty.site.seeing"/>
```

Unified Content Descriptor (UCD)

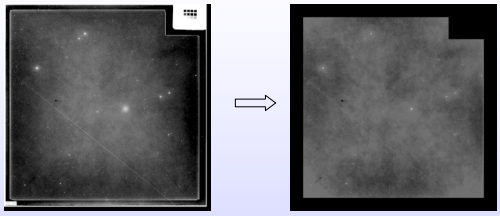
- Show UCD browser
  - http://vizier.u-strasbg.fr/UCD/

Large Scale Processing

- Pushing terabytes through the Grid
  - Keeping the nodes busy
  - Recovering from error
- Example: DPOSS flattening
  - 2700 images, each 1 Gbyte



DPOSS flattening



Source Target

2650 x 1.1 Gbyte files
Cropping borders
Quadratic fit and subtract

Driving the Queues

```
for f in os.listdir(inputDirectory):
# if the file exists, with the right size and age, then we keep it
        ofile = outputDirectory +"/"+ f
        if os.path.exists(ofile):
            osize = os.path.getsize(ofile)
            if osize != 1109404800: print " -- wrong target size, remaking", osize
            else:
                time_tgt = filetime(ofile)
                time_src = filetime(file)
                if time_tgt < time_src:
                    print(" -- target too old or nonexistant, making")
                else:
                    print " -- already have target file "
                    continue
        cmd = "qsub flat.sh -v \"FILE=" + f +"\""
        print " -- submitting batch job: ", cmd
        os.system(cmd)
```

VIM

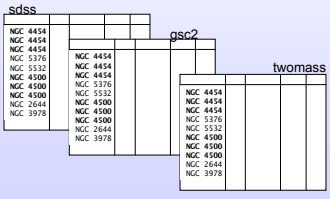
Resources

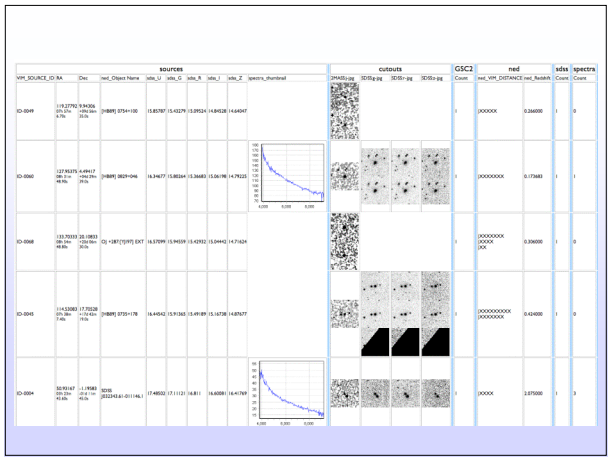
catalog or other position-based dataset
exposed by cone or skynode service
example: SDSS
example: Abell galaxy cluster catalog

Customer provides **Sources** table

```
187.209, -1.938, NGC 4454
208.826, 59.506, NGC 5376
214.218, 10.807, NGC 5532
187.844, 57.964, NGC 4500
130.384,  4.971, NGC 2644
179.042, 60.522, NGC 3978
…
…
```

Multicone **resources** provide data tables



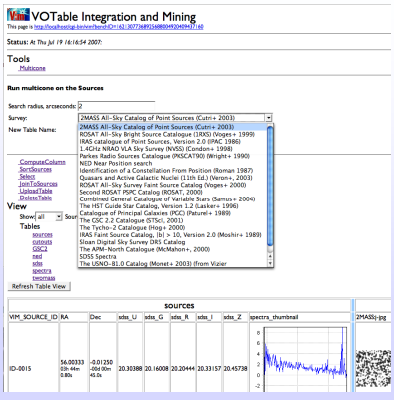


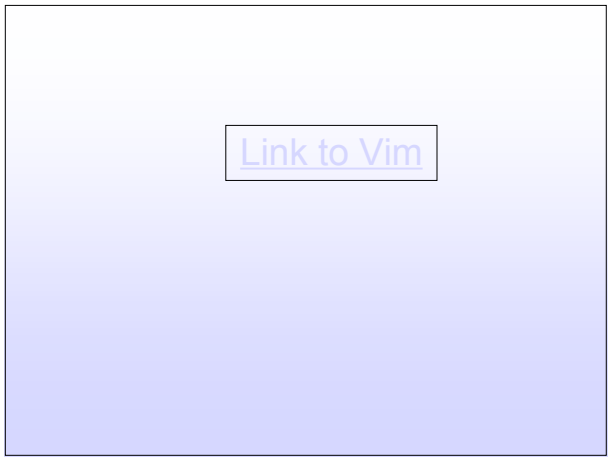
VIM

- View
  - Select attributes to view for each table
- Tools
  - Multicone and upload
    - Bring in new data
  - Compute new column
  - Sort and select
  - Join
    - Best or All

Table metadata

	i	Name	ID	ucd	unit	datatype
☐	0	GSC2_VIM_SOURCE_ID				char
☐	1	GSC2_VIM_DISTANCE				char
☐	2	GSC2__RAJ2000	Right ascension (FK5) Equinox=J2000.0 Epoch=J2000.000, proper motions tak	POS_EQ_RA_MAIN	deg	double
☐	3	GSC2__DEJ2000	Declination (FK5) Equinox=J2000.0 Epoch=J2000.000, proper motions taken	POS_EQ_DEC_MAIN	deg	double
☐	4	GSC2__r	Distance from center (RA(J2000=049.300833, DE(J2000=-07.980278)	POS_ANG_DIST_GENERAL	deg	double
☐	5	GSC2_USNO-B1.0	Designation of the object (1)	ID_MAIN		char
☐	6	GSC2_RAJ2000	Right Ascension at Eq=[2000, Ep=]2000 (2)	POS_EQ_RA_MAIN	deg	double
☐	7	GSC2_DEJ2000	Declination at Eq=[2000, Ep=]2000 (2)	POS_EQ_DEC_MAIN	deg	double
☐	8	GSC2_e_RAdeg	Mean error on RAdeg*cos(DEdeg) at Epoch	ERROR	mas	short
☐	9	GSC2_e_DEdeg	Mean error on DEdeg at Epoch	ERROR	mas	short
☐	10	GSC2_Epoch	Mean epoch of observation (2)	TIME_EPOCH	yr	float
☐	11	GSC2_pmRA	Proper motion in RA (relative to Y54.0)	POS_EQ_PMRA	mas/yr	int
☐	12	GSC2_pmDE	Proper motion in DE (relative to Y54.0)	POS_EQ_PMDEC	mas/yr	int
☐	13	GSC2_Ndet	[0,5] Number of detections (7)	NUMBER		unsignedByte
☐	14	GSC2_B1mag	? First blue magnitude	PHOT_PHG_B	mag	float
☐	15	GSC2_R1mag	? First red magnitude	PHOT_PHG_R	mag	float
☐	16	GSC2_B2mag	? Second blue magnitude	PHOT_PHG_B	mag	float
☐	17	GSC2_R2mag	? Second red magnitude	PHOT_PHG_R	mag	float
☐	18	GSC2_Imag	? Infrared (N) magnitude	PHOT_PHG_I	mag	float




Link to Vim