## Data Mining, ADQL, & National Vir

# The National Virtual Observatory's OpenSkyQuery Utility

by Richard "Doc" Kinne, KQR 2008 AAVSO Fall Conference Nantucket, MA







# Need for Standard Database Operations

- For Business
- For Astronomy

#### SQL - The Win!

- Standard Query Language
- Used With...
  - SQL Anywhere Commercial
  - PostgreSQL
  - mySQL used by Amazon.com...and the AAVSO!

#### SQL - What Is It?

- A Means of Interacting with, and creating, a Standard Relational Database.
- Relational Database? A linked series of 2dimensional tables – linked, stacked Excel spreadsheets.
- SQL a language that allows you to create, modify, read, and query those tables of data.

### **SQL Example Syntax**

 create database varstar; use varstar; create table observation name varchar(30), mag float, dec varchar(30), ra varchar(30), jd float not null } type=InnoDB;

## **SQL Example Syntax**

 use varstar; insert into observation values ("Z Uma", 7.4, "57:52:18.0", "11:56:30.2", 2454728.4);

# **SQL Example Syntax**

use varstar;
 delete from observation where name = "Z UMa";

#### **ADQL**

- ADQL Astronomical Data Query Language.
- A Subset of SQL 92
- Differences?
  - Intended for read-only queries only. Actually very simple.
  - Astronomical Extensions have been added.
    - Spatial areas: Regions & Catalog crossmatching

### **ADLQ – Example Syntax**

- Select o.objld, o.ra, o.r, o.type, t.objld from SDSS:PhotoPrimary o, TWOMASS:Photoprimary, t where xmatch(o,t) < 2.5 and Region ('Circle J2000 181.3 -0.76 6.5') and o.type=3;
- What the heck does this mean?

#### **ADLQ – Example Explanation**

- Select o.objld, o.ra, o.r, o.type, t.objld from SDSS:PhotoPrimary o, TWOMASS:Photoprimary, t where xmatch(o,t) < 2.5 and Region ('Circle J2000 181.3 -0.76 6.5') and o.type=3;
- SDSS:Photoprimary is a table now called "o"
- TWOMASS:Photoprimary is a table now called "t"
- Select the objld, ra, r, and type columns out of "o"
- Select the objld column out of "t"
- These are the columns in the two tables we're going to look at.

#### **ADLQ – Astronomical Extension**

- Select o.objld, o.ra, o.r, o.type, t.objld from SDSS:PhotoPrimary o, TWOMASS:Photoprimary, t where xmatch(o,t) < 2.5 and Region ('Circle J2000 181.3 -0.76 6.5') and o.type=3;
- XMATCH(o,t)<2.5</li>
  - Function will calculate the chi-squared probability that the objects in these two catalogs are actually the same object.
  - <2.5 means that the detection will be rejected if the standard deviation between the two is more than 2.5.

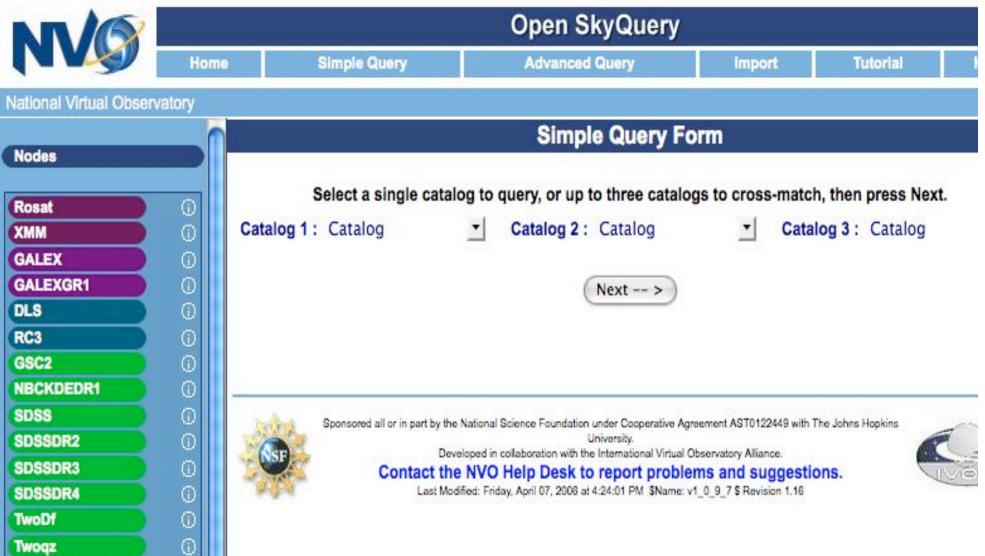
#### ADLQ – Astronomical Extensions

- Select o.objld, o.ra, o.r, o.type, t.objld from SDSS:PhotoPrimary o, TWOMASS:Photoprimary, t where xmatch(o,t) < 2.5 and Region ('Circle J2000 181.3 -0.76 6.5') and o.type=3;
- Region ('Circle J2000 181.3 -0.76 6.5') does the position for the current record fall in this region of the sky?
- Circle J2000, RA, DEC, r of circle in dec. degrees.

### **OpenSkyQuery**

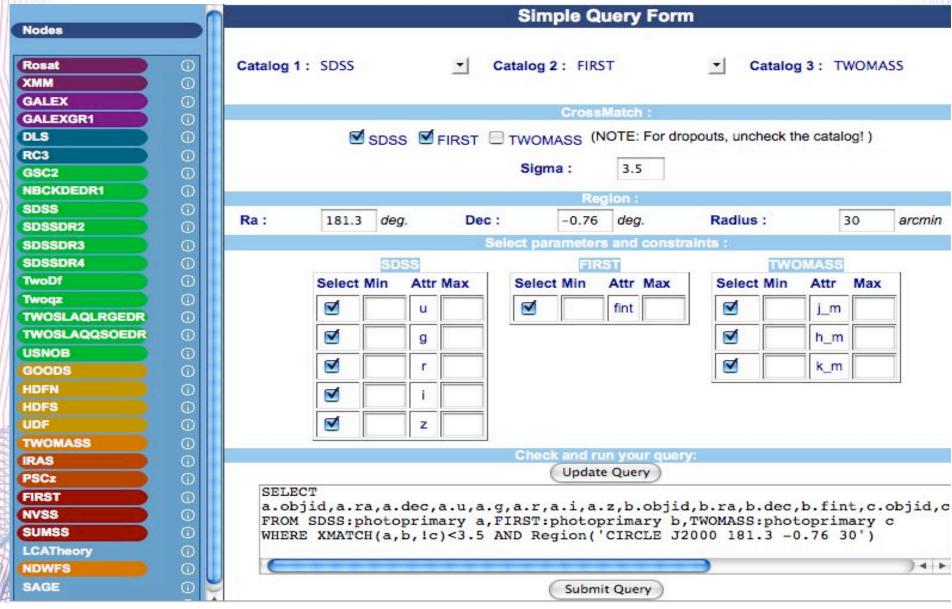
- OpenSkyQuery is a Data Mining Utility of the Virtual Observatory.
- OpenSkyNode supports distributed queries! DBs are hosted & maintained by their institutions but readable by everyone.
- In order for a DB to be part of it, it needs to have a SkyNode interface and be published to the VO Registry.
- Once this is done, it appears on a list of databases that can be searched via OpenSkyQuery.
- OpenSkyQuery is at <a href="http://openskyguery.net/">http://openskyguery.net/</a>

# OpenSkyQuery – A Simple Example



TWOSLAQLRGEDR

# OpenSkyQuery – A Simple Example



# OpenSkyQuery – A Simple Example

HTML \$ Save

44 row(s)

sdss_objid	sdss_ra	sdss_dec	sdss_u	sdss_g	sdss_r	sdss_i	sdss_z	fir
587722982277513779	181.280964329843	-0.775161775579332	26.59848	23.48722	22.30811	21.48366	21.06459	)
588848898839478741	181.241753429695	-0.895253982374432	22.75468	22.23607	21.08645	20.3742	20.21942	2
588848899376415277	181.395560517706	-0.618968107715356	22.60993	22.11696	20.3791	19.39062	18.85863	3
588848899376415271	181.393855203689	-0.61742038695532	22.54627	22.63787	21.0656	20.0659	19.78067	7
588848899376415277	181.395560517706	-0.618968107715356	22.60993	22.11696	20.3791	19.39062	18.85863	3
588848898839413066	181.158631722352	-0.873301373473849	21.84024	22.21916	21.38232	20.8104	21.38801	
588848898839544419	181.449753252085	-0.860114981262776	23.96865	21.61563	20.03169	19.22029	18.84507	7
588848898839543917	181.452000649341	-0.860159402334619	23.65605	22.46628	21.36011	20.83372	20.81125	5
588848899376349612	181.222654063796	-0.574898585088707	25.42542	22.21821	21.58511	21.27709	21.3112	2
588848899376349206	181.257280892827	-0.537363871624612	19.77114	19.62425	19.36931	19.3394	19.29524	1
588848899376284517	181.100151529517	-0.615921152282603	26.13067	24.42188	22.47922	21.92833	21.48851	L
587722982277644787	181.546871355706	-0.734921765861926	22.5851	22.94553	21.21888	20.39603	19.54589	)
588848898839544345	181.408718502499	-0.993146105058267	22.74373	21.7	20.01053	19.32947	18.89978	3
588848898839544346	181.408058073042	-0.992703320749044	24.96557	21.72807	20.22829	19.74157	19.07533	3
588848898839609792	181.534254735305	-0.873295134776541	24.41793	22.79921	23.10035	21.95683	22.10049	)
588848898839610404	181.536192688924	-0.878429757406259	23.38247	22.94795	22.39088	22.20344	21.41267	7
500000000000000000000000000000000000000	101 575500000000	0.0000000000000000000000000000000000000	00 7 1000	00.00005	20 75704	20.02507	10 10000	TH

#### Resources

- US National Virtual Observatory http://www.us-vo.org/
- The National Virtual Observatory: Tools and Techniques for Astronomical Research – Available from the US VO website.
- OpenSkyQuery Tutorial at the website http://www.openskyquery.org/
- Latest ADQL Specification http://www.ivoa.net/Documents/latest/ADQL.html

