Fig. 1

Fig. 2

1. The VO – What? Why?

(vgl. Fig. 1)

Markus Demleitner
msdemlei@ari.uni-heidelberg.de

(vgl. Fig. 2)

The Virtual Observatory (VO) is (or will be), a
comprehensive set of
data and **services**
relevant to **astronomy**
accessible from **clients** of **your choice**
regardless of where you are and
preserving products of digital astronomy.

2. "comprehensive"

The VO intends to allow access to basically all astronomical data, present and past.

Right now: About 15000 resources like

- VizieR catalogs
- Lots of space missons
- Many observatory collections
- Theory data like synthetic spectra
- Much more

But **much** is still missing (e.g., much of ESO's data).

Fig. 3

3. Data and Services

While the VO is about data, much of it is concerned with services.

A service is

- a piece of software accessible via a network
- with a well-defined interface
- allowing access to some data collection.

Important: Service users ("clients") must be able to figure out how to operate the service and find out as much as possible about the data contained.

4. Astronomy

Well, of course. . .

(vgl. Fig. 3)

. . . but we also pave the way for similar endeavours in other fields; cf. the current ASTROTROP project for VO tech in tropical rainforest research.

5. Clients and Choice

"Web pages" aren't really what the VO is about. It is about standard interfaces to data.

This means: A single program (possibly web-based) can operate all kinds of archives and services. Many such programs are listed at http://ivoa.net.

It also means: A given service can be operated by any client speaking the VO languages – you get to choose or use libraries like pyVO in your own programs.

Fig. 4

6. The big equalizer

It used to be that you had to go to the big observatories to get top-notch data.

Converseley, chances someone would see and use your data if you weren't there weren't terribly good.

The VO already delivers excellent data to anywhere and anyone in the world.

And with All-VO searches and increased adoption of Registry use, everyone gets a more uniform view of the data taken.

7. Preservation

We're currently losing historical observations at an unprecedented rate: All the tapes from the 80ies and 90ies are deteriorating.

Linus Torvalds:

> Only wimps use tape backup: real men just upload their important stuff on ftp, and let the rest of the world mirror it.

If data is to survive, it must be in living services not far from spinning disks.

(Yes, there's more to it, but the living part is vital)

8. Who's doing this?

The IVOA is an association of numerous national or institutional projects.

Twice a year, we meet to work on depeloping the standards.

(vgl. Fig. 4)

Astrogrid UK, Armenian VO, Australian VO, Brazilian VO, China-VO, Virtual Observatory Italy, US-VAO, VO India, Ukranian VO, Spanish VO, South African Astroinformatics Allicance, Russian VO, Nuevo Observatorio Virtual Argentino, Japanese VO, Hungarian VO, German Astrophysical VO, VO France, ESA VO, Canadian VO


Fig. 5

9. Another Reason

That's all nice, but the VO really is

not a project of choice

but

a project of necessity:

- There's too many data collections out there to know them all
- With many data sources are available relevant to a particular research project, writing extra code for each is not feasable any more
- Most importantly: We cannot move data (any more)

10. Fifteen Years Ago: HIPPARCOS

This stack of books is the HIPPARCOS catalog of 118218 Stars. 30 cm of shelf space, say.

That's 15 Megabyte – easily moved, and small by today's standards.

(vgl. Fig. 5)

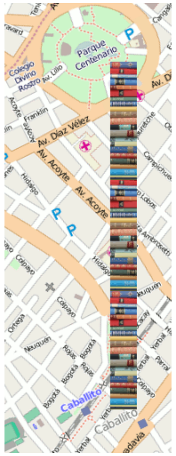
Fig. 6

11. Now: PPMXL

A modern star catalog like PPMXL has about 10^9 objects; that's 8000 HIPPARCOS catalogs or a couple of 10 Gigabyte. Transport of this is marginally possible, but you'd rather avoid that. This is a shelf from Parque Centenario to Caballito.

But to make that data of that size "workable", you need indices and clever organization, preferable a database. This typically is not something Scientists enjoy doing, in particular if they want to bring together several such resources and a couple of smaller ones.

(vgl. Fig. 6)

12. Soon: LSST

LSST is going to produce 60 PB (that's Petabyte) of Images and 30 PB of catalogs. That's a stack of books 3 million kilometers high.

On the image, there's Earth, Moon, and a third of the stack of books required.

There's no way we can move this kind of data any time soon. We'll have to move the smarts (analysis).

(vgl. Fig. 7)

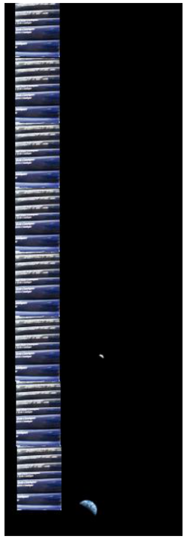
Fig. 7

13. Standards help

Standardized protocols let scientists

- find and
- use

remote resources, just moving data to their desktops when it's destilled down to what's relevant to the research at hand.

Learn more at http://ivoa.net

14. Later

In this afternoon's talk:
VO technologies – what to use, how to use it.

See you then!

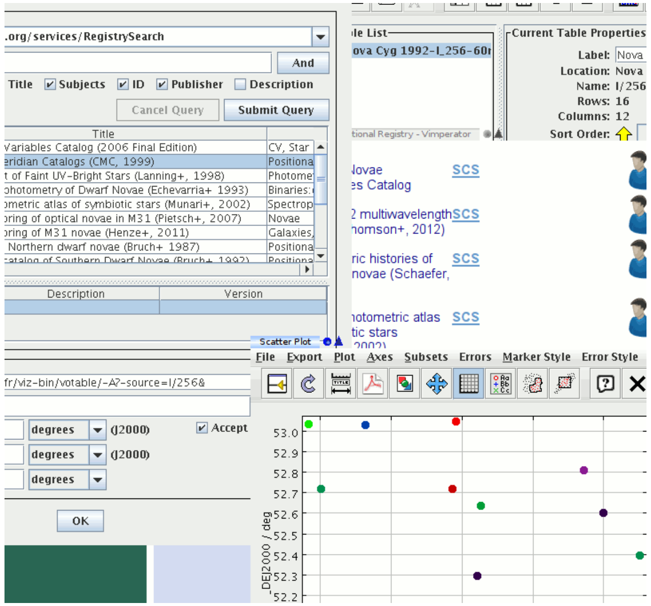

Fig. 8

15. Action

A little illustration involving WIRR and TOPCAT, using

- Registry
- Cone Search
- SAMP

Data discovery in the VO starts at the registry. There are various interface to it, but let's start at http://dc.g-vo.org/WIRR.

Look for resources mentioning "dwarf nova" and having a Service Capability of Cone Search.

Start TOPCAT open its VO/Cone Search window. In WIRR, select "Register with SAMP Hub", ok the warning and then send "Tabular Services" (this uses a protocol calles SAMP).

At "Object Name", say "Nova Cygni 1992", hit "Resolve", enter something (say, 1 degree), then select any of the services and hit Ok. Re-open the dialog if necessary and try another service. Some may give empty results, but TOPCAT can talk to them all. That's because they all speak a standard protocol (it's called SCS).

This may not be exciting in and of itself – but being able to use one interface for lots of different services is one of the key functions of the VO.

(vgl. Fig. 8)