

www.g-vo.org

German Astrophysical
Virtual Observatory

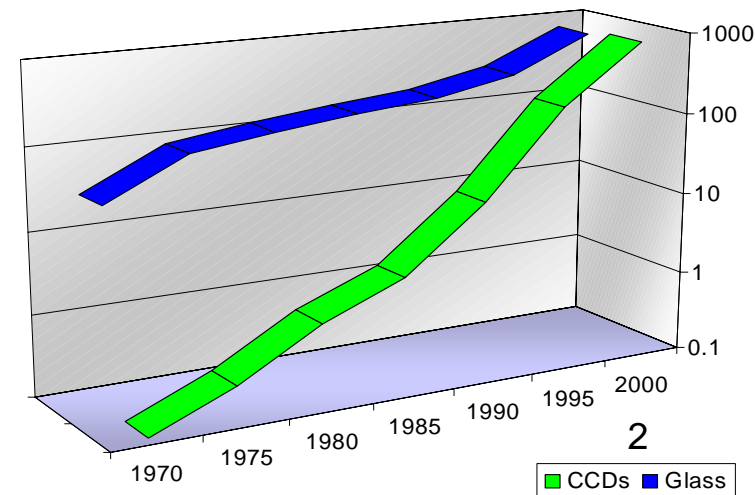
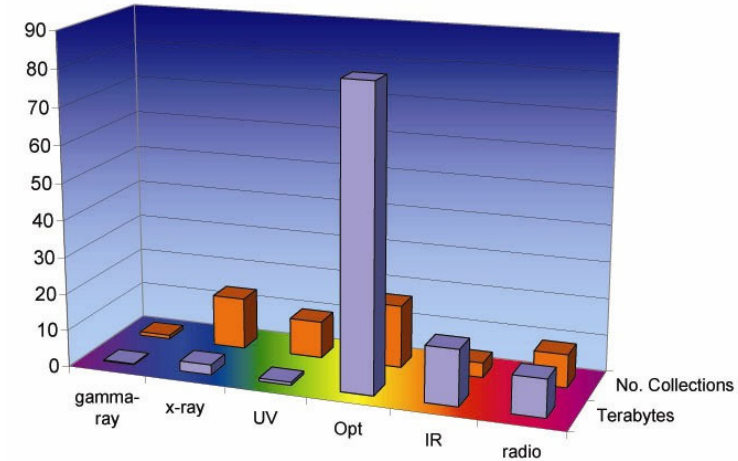
Matthias Steinmetz (AIP) ¹



Leben in einer exponentiellen Welt

- Heute einige 100 TB an astronomischen Daten
 - 1 pixel (byte) / sq arc second
~ 4TB
 - Multi-spectral, zeitaufgelöst,...
⇒ 1PB
- Fundgrube für neue und neuartige Objekte
- Datenvolumen verdoppelt sich jedes Jahr
- Daten werden typischerweise nach 1 Jahr öffentlich

⇒ 50% aller Daten sind öffentlich





Datenzugriff am Limit

FTP und GREP nicht mehr adäquat

- You can GREP 1 MB in a second
- You can GREP 1 GB in a minute
- You can GREP 1 TB in 2 days
- You can GREP 1 PB in 3 years
- You can FTP 1 MB in 1 sec
- You can FTP 1 GB / min (= 1 \$/GB)
- ... 2 days and 1K\$
- ... 3 years and 1M\$
- Oh!, 1PB ~4,000 disks
- Folge:

indices um Suche zu begrenzen
parallele Datensuche und Analyse

⇒ **Databanken**





Trends

CMB Surveys

1990 COBE	1,000
2000 Boomerang	10,000
2002 CBI	50,000
2003 WMAP	1 Million
2008 Planck	10 Million

Angular Galaxy Surveys

1970 Lick	1M
1990 APM	2M
2005 SDSS	200M
2008 VISTA	1000M
2012 LSST	3000M

Time Domain

- QUEST
- SDSS Extension survey
- Dark Energy Camera
- PanStarrs
- SNAP...
- LSST...

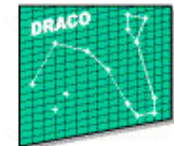
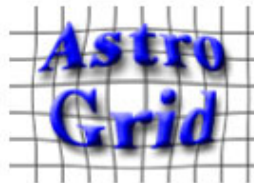
Galaxy Redshift Surveys

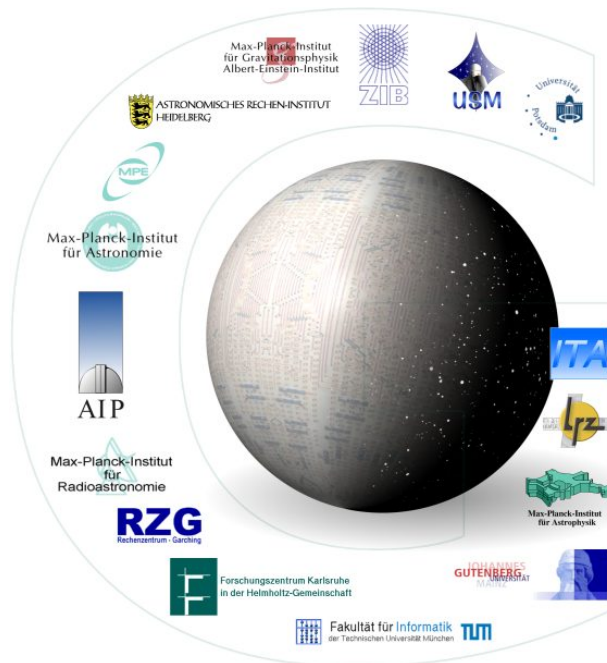
1986 CfA	3500
1996 LCRS	23000
2003 2dF	250000
2005 SDSS	750000

Petabytes/Jahr zum Ende des Jahrzehnts...



International Virtual Observatory Alliance (IVOA)

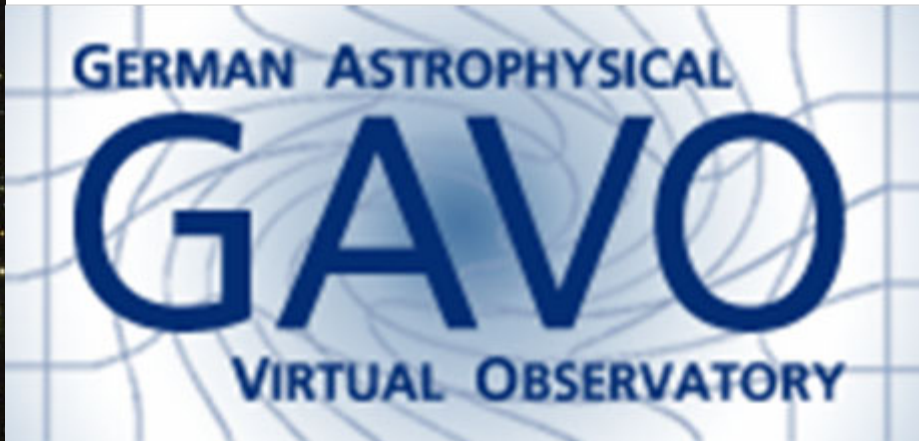




AstroGrid-D

www.gac-grid.org

Informations-
technische
Infrastruktur



GAVO

www.g-vo.org

Astrophysikalische
Applikationen



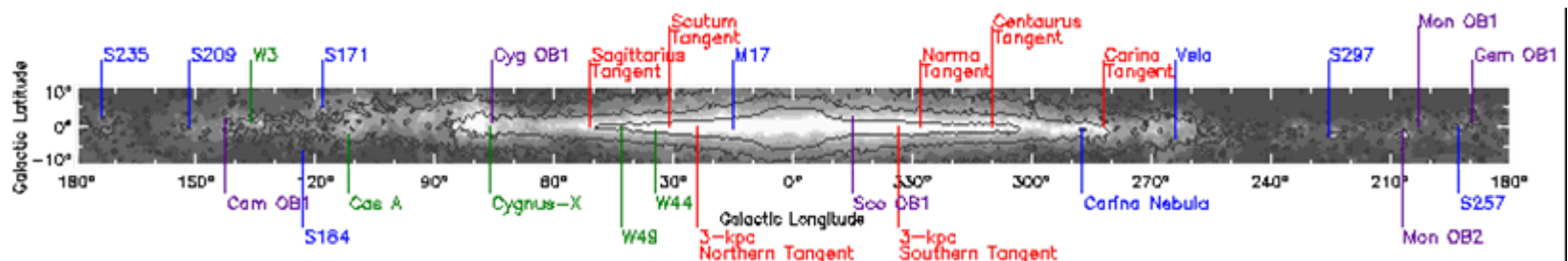
Warum ist Astronomie “special”?

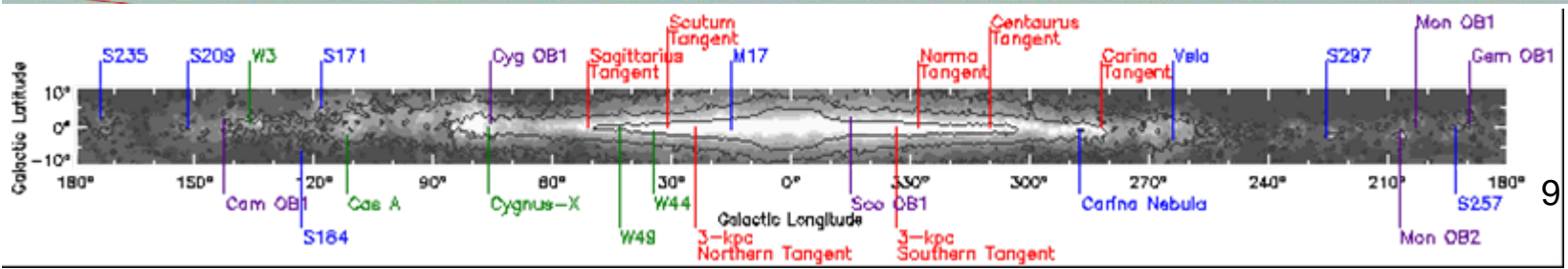
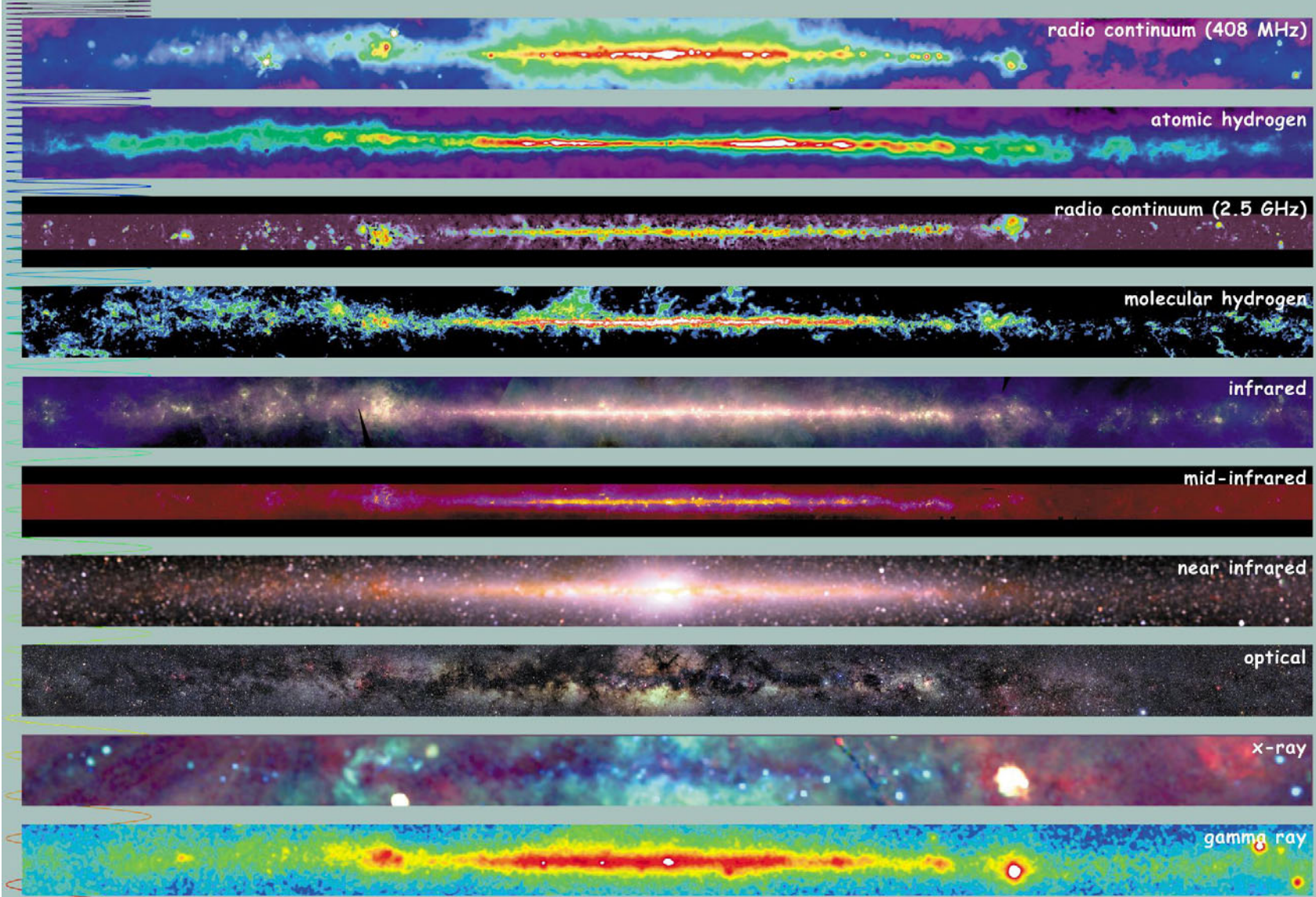
- Hohe Attraktivität für die Öffentlichkeit
- kein kommerzieller Wert
 - keine Rechte, freier Zugang für alle
 - ideal um mit Algorithmen zu experimentieren
- Real und wohl dokumentiert
 - Multidimensional
 - räumlich, zeitlich
- Divers und verteilt
 - Viele verschiedene Instrumente an vielen verschiedenen Orten zu vielen verschiedenen Zeiten
- There is a lot of it (petabytes)



Ein panchromatischer Blick auf die Galaxis

- Eine digitale Galaxie
 - Sternkataloge
 - Interstellares Medium und Staub
 - Vergleich mit theoretischen Modellen (Simulationen)







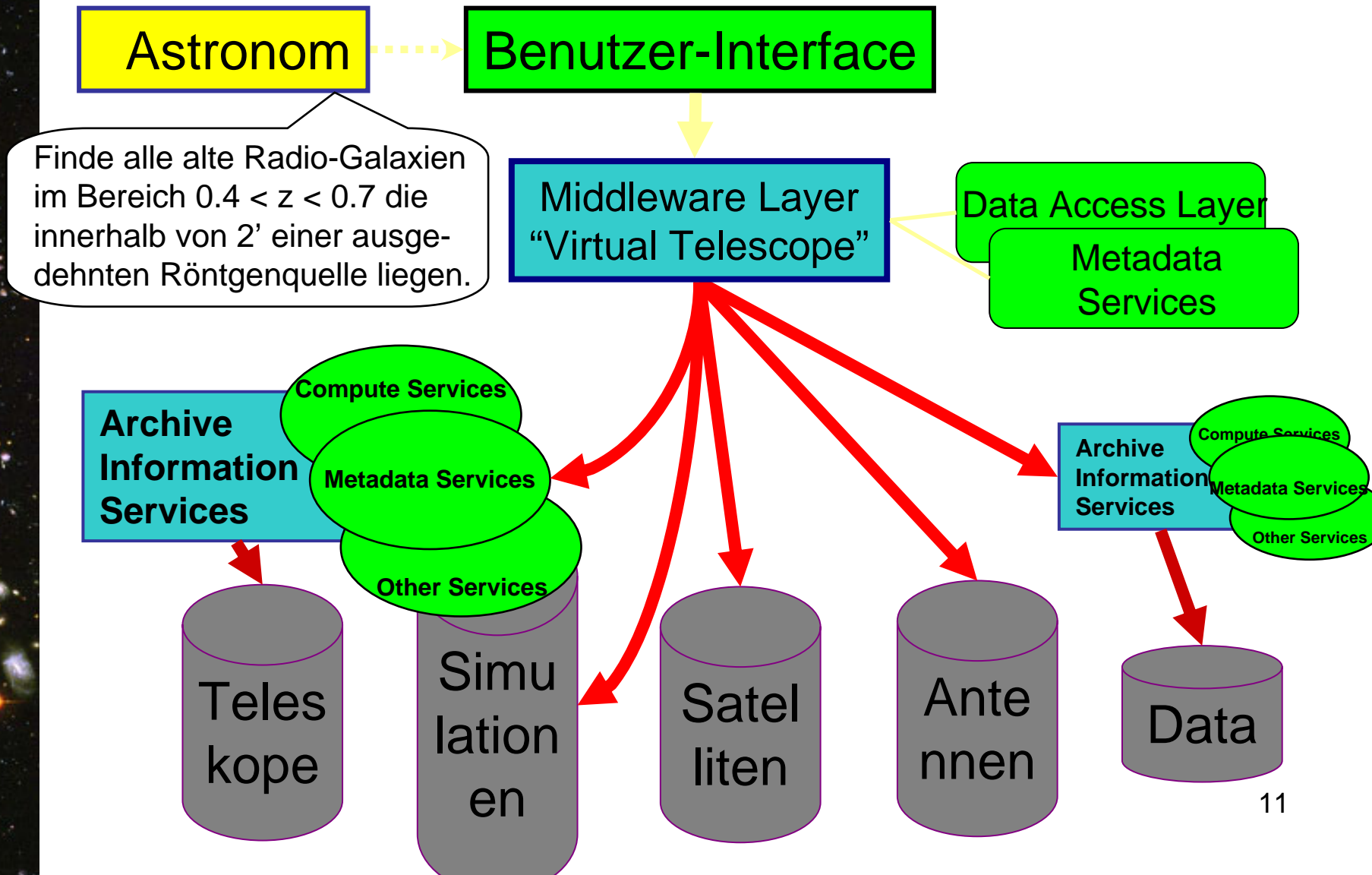
Theorie im VO

- ◆ Publikation von Simulationen
 - Planck-Experiment
 - Millenium-Simulation

- ◆ Theorie-Beobachtung-Interface
 - “Virtual telescope”

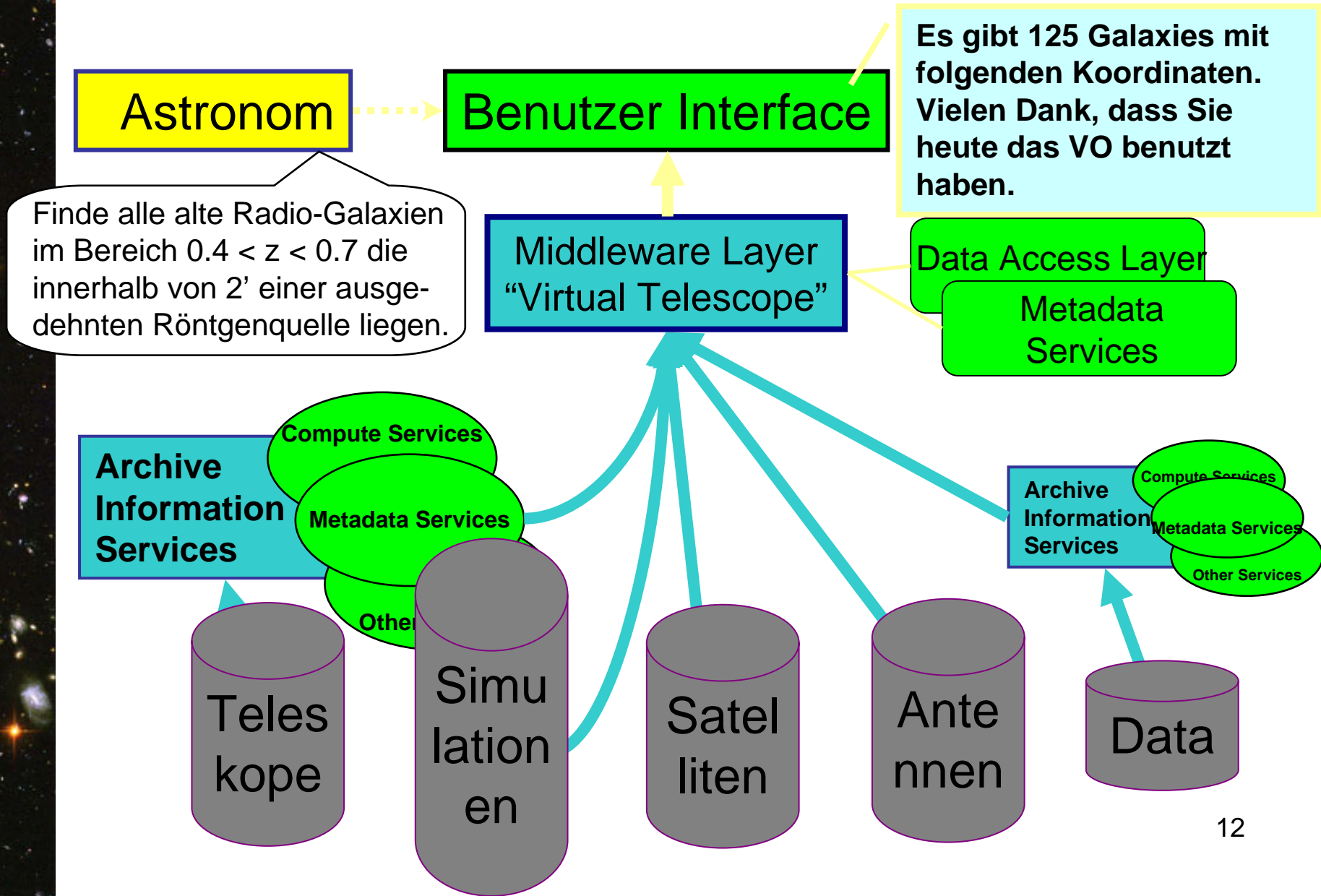


Das Virtual Observatory





Das Virtual Observatory





Download Manager

- Features
 - Tool ...
 - ... greift auf die Registrierdaten an der JHU zu
 - Benutzer ...
 - ... wählt verteilte Kataloge aus
 - ... spezifiziert eine oder mehrere Positionen am Himmel
 - Tool ...
 - ... Abfrage der (remote) Kataloge
 - ... empfängt Datensätze zur Weiterverarbeitung
(to do: slice data cubes)



Download Manager

Multi-Catalog Multi-Cone Search Download Manager - C:\gavo\http\client\html\en\g-serv\software\mcmcs\service-...

File Edit Search Help

Service...	Title	Type	Instru...	Descrip...	Spatia...	Spectr...	Tem...	MaxRe...	MaxSR	Creator	Bas...	sel...
EUVE	Extreme Ultraviolet Explorer	Archive	Scann...	Launch...	circle ...	UV	1992...	100	180.0	Bowyer	http...	<input type="checkbox"/>
EXOSAT	EXOSAT Observatory Data Arc...	Archive	CMA, ...	The Eu...	circle ...	X-ray	1983...	1000	20.0	Europ...	http...	<input type="checkbox"/>
Einstein	Einstein (HEAO-2) Observator...	Archive	IPC, H...	The se...	circle ...	X-ray	1978...	1000	20.0	Smith...	http...	<input type="checkbox"/>
FOC	Faint Object Camera	Archive	FOC	The Fai...	circle ...	UV, O...	1990...	100	180.0	Macch...	http...	<input type="checkbox"/>
FOS	Faint Object Spectrograph	Archive	FOS	The Fai...	circle ...	Optic...	1990...	100	180.0	Harms	http...	<input type="checkbox"/>
FUSE	Far Ultraviolet Spectroscopic E...	Archive	Spectr...	The Fa...	circle ...	UV	1999...	100	180.0	Moos	http...	<input type="checkbox"/>
GCVS	Combined General Catalog of ...	Catalog		The G...	circle ...	Optical	NOT ...	1000	10.0	Kholo...	http...	<input type="checkbox"/>
GHR	Goddard High Resolution Spe...	Archive	GHR	GHR	circle ...	UV	1990...	100	180.0	Wood...	http...	<input type="checkbox"/>
GSC1	Guide Star Catalog 1	Catalog	Schmi...	The Gu...	circle ...	Optical	1950...	10000	1.0	Space...	http...	<input type="checkbox"/>
GSC2.2	Guide Star Catalog 2.2	Survey	Schmi...	GSC2...	circle ...	Optical	1975...	1000	0.5	Space...	http...	<input type="checkbox"/>
GSC2.2...	Guide Star Catalog 2.2	Catalog	Schmi...	GSC2...	circle ...	Optical	1975...	999	1.0	Space...	http...	<input checked="" type="checkbox"/>
HD	Henry Draper Catalogue and ...	Catalog		An upd...	circle ...	Optical	1918...	1000	3.0	Cann...	http...	<input type="checkbox"/>
HIP	Hipparcos Space Astrometry ...	Catalog	Schmi...	Hippar...	circle ...	Optical	1989...	1000	2.0	Hippa...	http...	<input type="checkbox"/>
HSP	High Speed Photometer	Archive	HSP	The Hi...	circle ...	UV, O...	1990...	100	180.0	Code	http...	<input type="checkbox"/>
HST	Hubble Space Telescope	Archive	ACS, ...	Hubble...	circle ...	Optic...	1990...	100	180.0	Space...	http...	<input type="checkbox"/>
HSTG	Hubble Space Telescope	Archive	ACS, ...	Hubble...	circle ...	Optic...	1990...	1000	5.0	Space...	http...	<input type="checkbox"/>
HUT	Hopkins Ultraviolet Telescope	Archive	Spectr...	The Ho...	circle ...	UV	1990...	100	180.0	David...	http...	<input type="checkbox"/>
Hewitt	Revised and Updated Catalog...	Catalog		This is ...	circle ...	Optical	to 19...	1000	10.0	Hewitt...	http...	<input type="checkbox"/>
IUE	International Ultraviolet Explorer	Archive	Spectr...	The Int...	circle ...	UV	1978...	100	180.0	Bogg...	http...	<input type="checkbox"/>

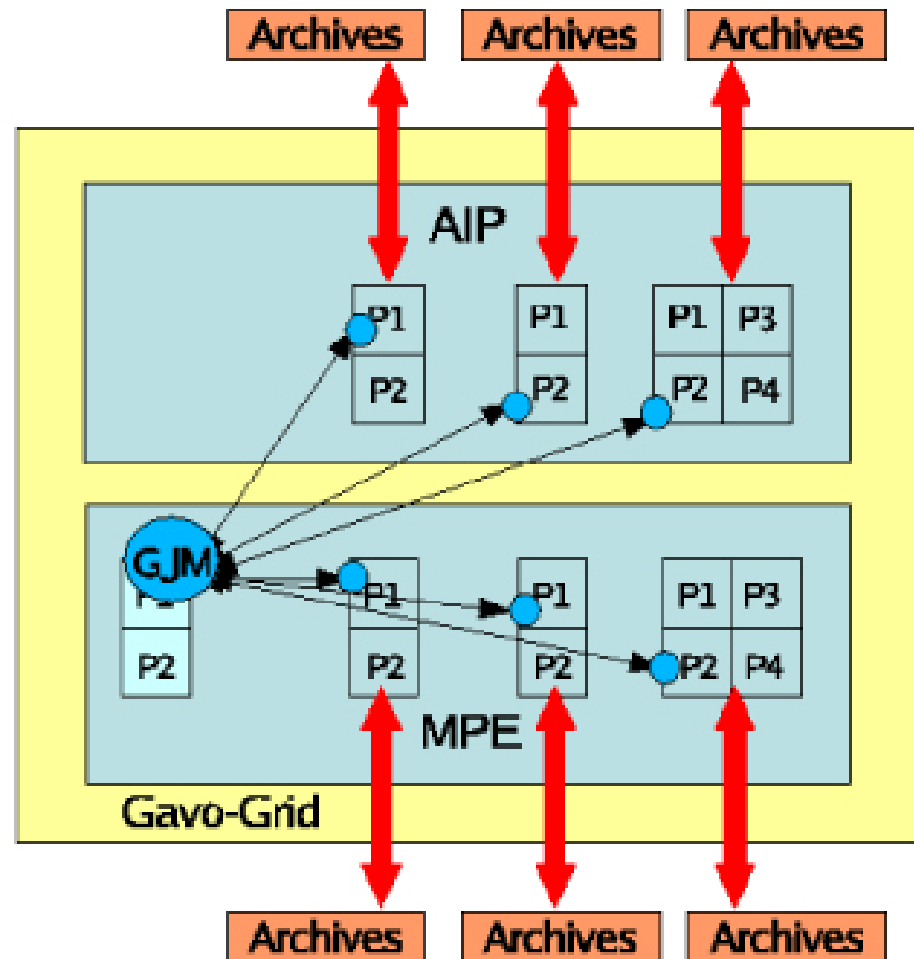
ra	dec	sr	verb	selected
5	0	1	3	<input checked="" type="checkbox"/>
10	0	1	3	<input checked="" type="checkbox"/>

MCMC Search



Matcher

- Wesentlich zum VO data mining
- Prototyp-features
 - “fuzzy” matching zwischen Paaren in der Quellenliste verschiedener Kataloge
 - Berechnung der Zuverlässigkeit einer gegenseitiger Identifikation





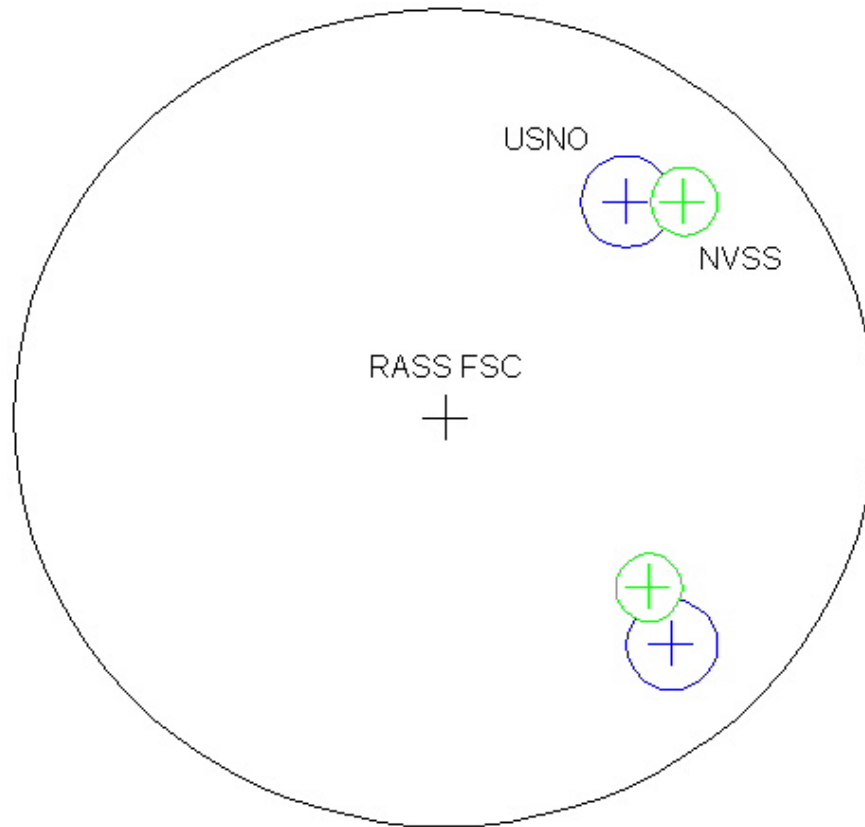
Matcher

The screenshot shows the GAVO Manager application interface. On the left, a file tree under 'rosat' includes a 'CVS' subdirectory with files like '10_RASS.txt', '2511_Tycho2.txt', '323_RASS.txt', '3558_FIRST.txt', 'Match.txt', 'catalog-test.txt', 'catalog.txt', 'control.txt', and 'parameter.txt'. The main window displays a text editor with a list of astronomical records. Each record is a line of text containing various identifiers and numerical values.

```
"ID_1" "RA_1" "DEC_1" "ID_2" "RA_2" "DEC_2" "ID" "ARC_DISTANCE/arcsec" "POSITION_ANGLE" "CHI_SQUARED" "PROBABILITY" "RA" "DEC"
"RBSC_33017001_00045" 0.0525 1.7725 "TYC_2383074" 0.0504 1.7715 "323_RASS::RBSC_33017001_00045&&2511_Tycho2::TYC_2383074"
"RBSC_33017001_00045" 0.0525 1.7725 "TYC_2383074" 0.0509 1.7715 "323_RASS::RBSC_33017001_00045&&2511_Tycho2::TYC_2383074"
"RFSC_33017001_00027" 0.6925 2.1307 "TYC_2395592" 0.6964 2.1304 "323_RASS::RFSC_33017001_00027&&2511_Tycho2::TYC_2395592"
"RFSC_33017001_00027" 0.6925 2.1307 "TYC_2395592" 0.6965 2.1300 "323_RASS::RFSC_33017001_00027&&2511_Tycho2::TYC_2395592"
"RFSC_33017001_00156" 1.5608 -1.8857 "TYC_2299049" 1.5548 -1.8871 "323_RASS::RFSC_33017001_00156&&2511_Tycho2::TYC_2299049"
"RBSC_33017001_00079" 2.3383 0.6360 "TYC_2359761" 2.3405 0.6352 "323_RASS::RBSC_33017001_00079&&2511_Tycho2::TYC_2359761"
"RFSC_33017001_00035" 2.6704 2.0508 "TYC_2395639" 2.6697 2.0555 "323_RASS::RFSC_33017001_00035&&2511_Tycho2::TYC_2395639"
"RFSC_33018001_00019" 3.3879 -3.1519 "TYC_2261556" 3.3863 -3.1503 "323_RASS::RFSC_33018001_00019&&2511_Tycho2::TYC_2261556"
"RFSC_33017001_00059" 3.6521 1.3014 "TYC_2371467" 3.6523 1.2969 "323_RASS::RFSC_33017001_00059&&2511_Tycho2::TYC_2371467"
"RFSC_33018001_00023" 3.7808 -3.3363 "TYC_2261577" 3.7815 -3.3334 "323_RASS::RFSC_33018001_00023&&2511_Tycho2::TYC_2261577"
"RBSC_33017001_00134" 4.4387 -1.0486 "TYC_2311646" 4.4377 -1.0477 "323_RASS::RBSC_33017001_00134&&2511_Tycho2::TYC_2311646"
"RFSC_33018001_00011" 4.8033 -3.0450 "TYC_2261583" 4.8017 -3.0531 "323_RASS::RFSC_33018001_00011&&2511_Tycho2::TYC_2261583"
"RFSC_33018001_00034" 4.8342 -3.6203 "TYC_2249679" 4.8333 -3.6154 "323_RASS::RFSC_33018001_00034&&2511_Tycho2::TYC_2249679"
"RFSC_33017001_00154" 5.1508 -1.8188 "TYC_2299109" 5.1564 -1.8178 "323_RASS::RFSC_33017001_00154&&2511_Tycho2::TYC_2299109"
"RFSC_33018002_00025" 5.7821 -3.6012 "TYC_2249698" 5.7738 -3.6033 "323_RASS::RFSC_33018002_00025&&2511_Tycho2::TYC_2249698"
"RBSC_33018002_00044" 6.2863 -4.5824 "TYC_2225900" 6.2853 -4.5802 "323_RASS::RBSC_33018002_00044&&2511_Tycho2::TYC_2225900"
"RBSC_33016002_00084" 6.3858 4.8976 "TYC_2472538" 6.3872 4.8963 "323_RASS::RBSC_33016002_00084&&2511_Tycho2::TYC_2472538"
"RFSC_33018002_00040" 6.7517 -4.3986 "TYC_2238167" 6.7552 -4.4001 "323_RASS::RFSC_33018002_00040&&2511_Tycho2::TYC_2238167"
"RBSC_33017002_00053" 8.5154 1.0994 "TYC_2371596" 8.5180 1.0965 "323_RASS::RBSC_33017002_00053&&2511_Tycho2::TYC_2371596"
"RBSC_33018002_00043" 8.6225 -4.5553 "TYC_2225949" 8.6240 -4.5465 "323_RASS::RBSC_33018002_00043&&2511_Tycho2::TYC_2225949"
"RBSC_33018002_00023" 8.8112 -3.5931 "TYC_2249754" 8.8111 -3.5928 "323_RASS::RBSC_33018002_00023&&2511_Tycho2::TYC_2249754"
```




Das Identifikationsproblem





OpenSkyQuery

Cross-match your data with numerous catalogs

OpenSkyQuery allows you to cross-match astronomical catalogs and select subsets of catalogs with a general and powerful query language. You can also import a personal catalog of objects and cross-match it against selected databases.



Spectrum Services

Search, plot, and retrieve SDSS, 2dF, and other spectra

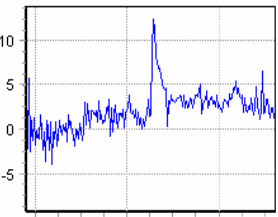
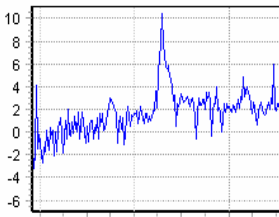
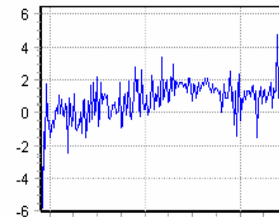
NVO National Virtual Observatory

Spectrum Services | not logged in | [login](#) | [register](#)

home | docs | search | webservice | user

Spectrum Advanced Search Results

Found 447 objects. Displaying from 1 to 12 switch to list mode 1 [Next] [Last]

<input type="checkbox"/> 1. SDSS J101549.00+002020.00 0271-51878-01 class: Qso, Z = 4.4013 ra = 153.954180, dec = 0.338888	<input type="checkbox"/> 2. SDSS J101549.00+002020.00 0271-51878-01 class: Qso, Z = 4.4027 ra = 153.954180, dec = 0.338888	<input type="checkbox"/> 3. SDSS J102043.82+000105.77 0271-51878-01 class: Qso, Z = 4.2073 ra = 155.182580, dec = 0.018269
 details	 details	 details
<input type="checkbox"/> 4. SDSS J102043.82+000105.77 0271-51878-01 class: Qso, Z = 4.2073 ra = 155.182580, dec = 0.018269	<input type="checkbox"/> 5. SDSS J103432.72-002702.57 0273-51957-01 class: Hiz_Qso, Z = 4.3771 ra = 158.636330, dec = -0.450713	<input type="checkbox"/> 6. SDSS J103432.72-002702.57 0273-51957-01 class: Hiz_Qso, Z = 4.3771 ra = 158.636330, dec = -0.450713

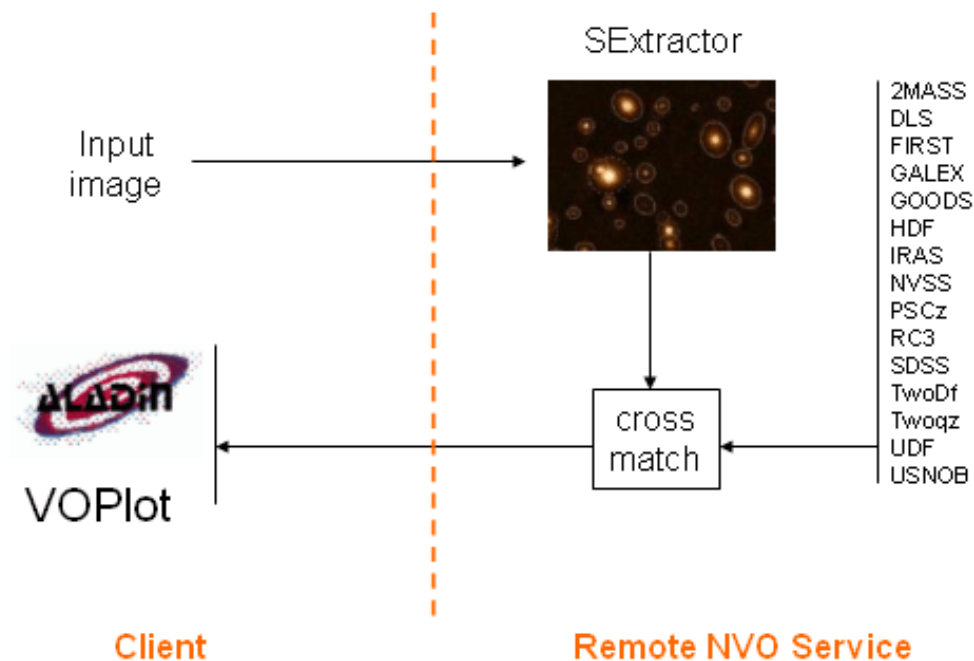
The Spectrum Services web site is dedicated to spectrum related VO services. On this site you will find tools and tutorials on how to access close to 500,000 spectra from the Sloan Digital Sky Survey (SDSS DR1) and the 2 degree Field redshift survey (2dFGRS). The services are open to everyone to publish their own spectra in the same framework. Reading the tutorials on XML Web Services, you can learn how to integrate the 45 GB spectrum and passband database with your programs with few lines of code.



Web Enabled Source Identification with Cross-Matching (WESIX)

Upload images to SExtractor and cross-correlate the objects found with selected survey catalogs.

This NVO service does source extraction and cross-matching for any astrometric FITS image. The user uploads a FITS image, and the remote service runs the SExtractor software for source extraction. The resulting catalog can be cross-matched with any of several major surveys, and the results returned as a VOTable. The web page also allows use of Aladin or VOPlot to visualize results.

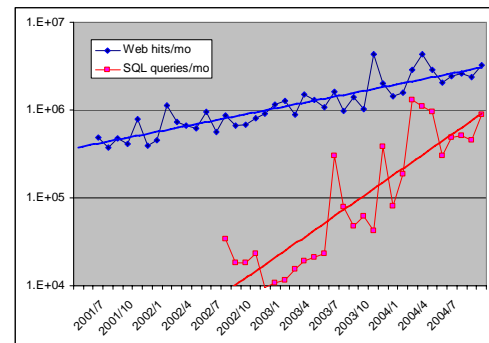
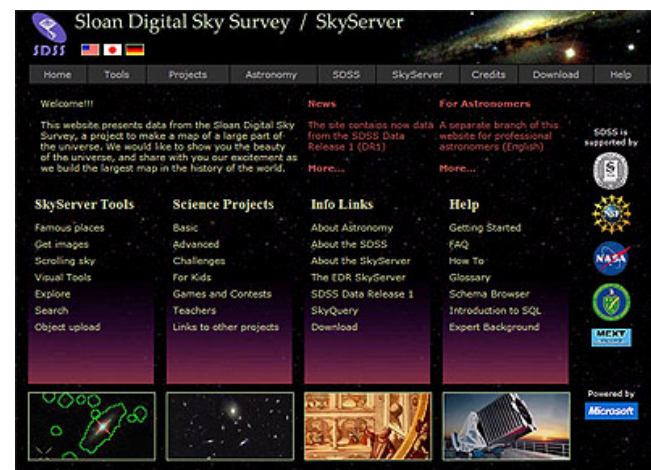




SkyServer

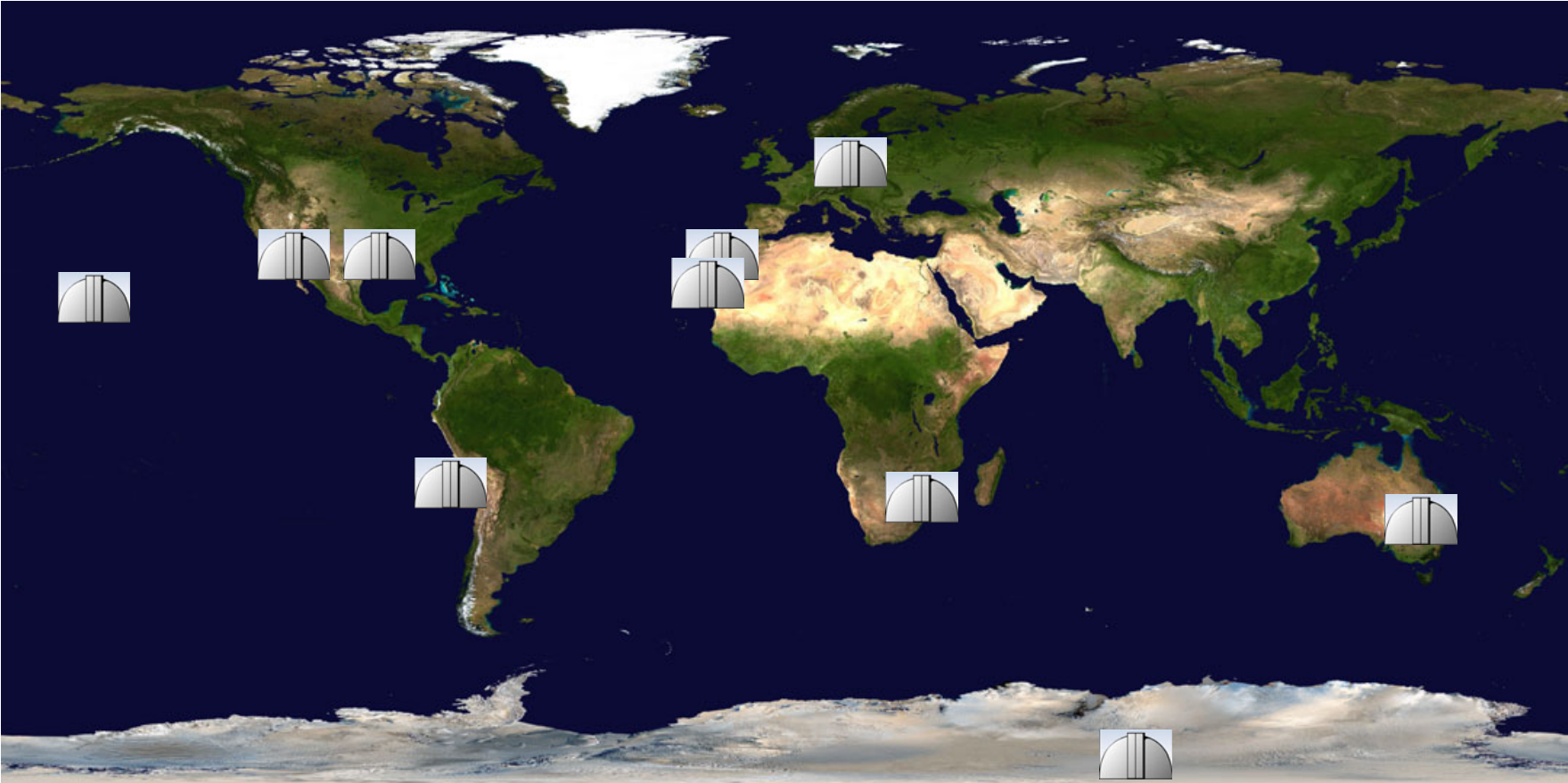
- Sloan Digital Sky Survey: Pixels + Objects
- About 500 attributes per “object”, 300M objects
- Spectra for 1M objects
- Currently 2TB fully public
- Prototype eScience lab
 - Moving analysis to the data
 - Fast searches: color, spatial
- Visual tools
 - Join pixels with objects
- Prototype in data publishing
 - 70 million web hits in 3.5 years

<http://skyserver.sdss.org/>



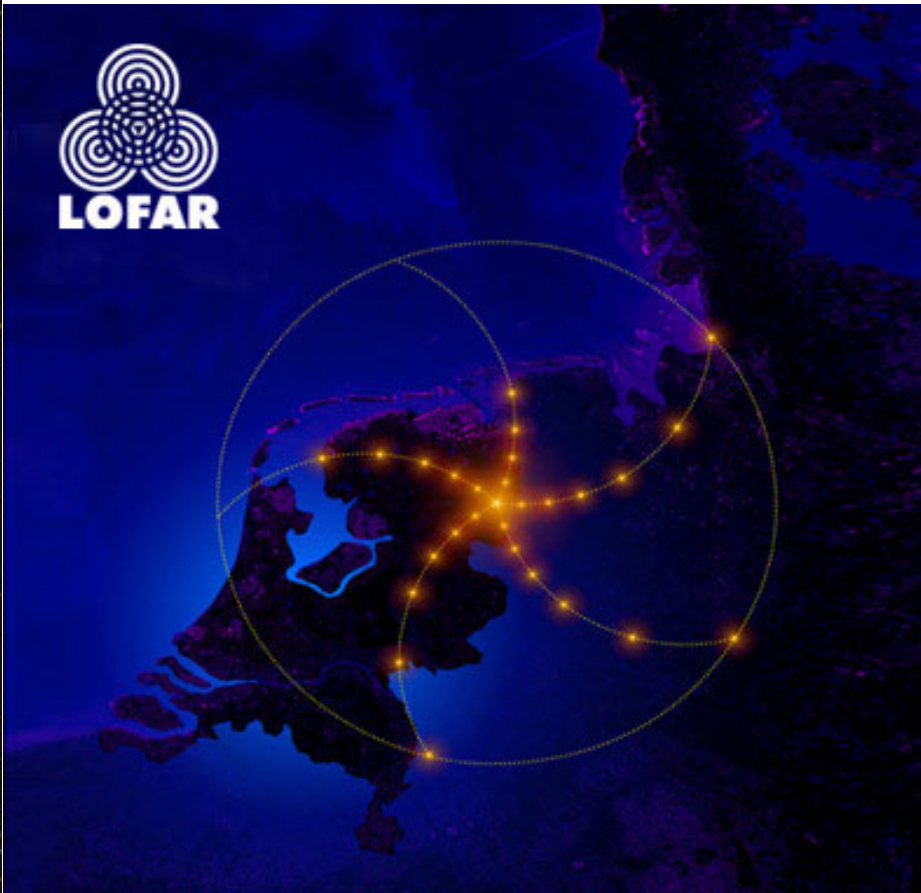


Weltweites Netz robotischer Teleskope





New Science with the VO/GRID



- Radioastronomie bei $\lambda=30\text{m}$
- Teleskopgröße: 1000km
- Lösung: 10000 Antennen verteilt über $\varnothing 1000\text{ km}$
- Laufzeitkorrelation
- All sky

➤ **Verfügbare Computer, Netzwerke & Software definieren Leistungsfähigkeit des Teleskops**



Summary

- Data growing exponentially
- Publishing so much data requires a new model
- Multiple challenges for different communities
 - publishing, visualization, statistics, algorithms, educational
- Information at your fingertips
 - Students see the same data as professional astronomers
- More data coming: Petabytes/year by 2010
 - Need scalable solutions
 - Move analysis to the data!
- Same thing happening in all sciences
 - High energy physics, genomics, cancer research, medical imaging, oceanography, remote sensing, ...
- **eScience**: an emerging new branch of science