



# Virtual Observations 2016 is about

- BigData
- Data bases
- Information Systems in Astronomy
- Data mining
- Data processing
- Data federations

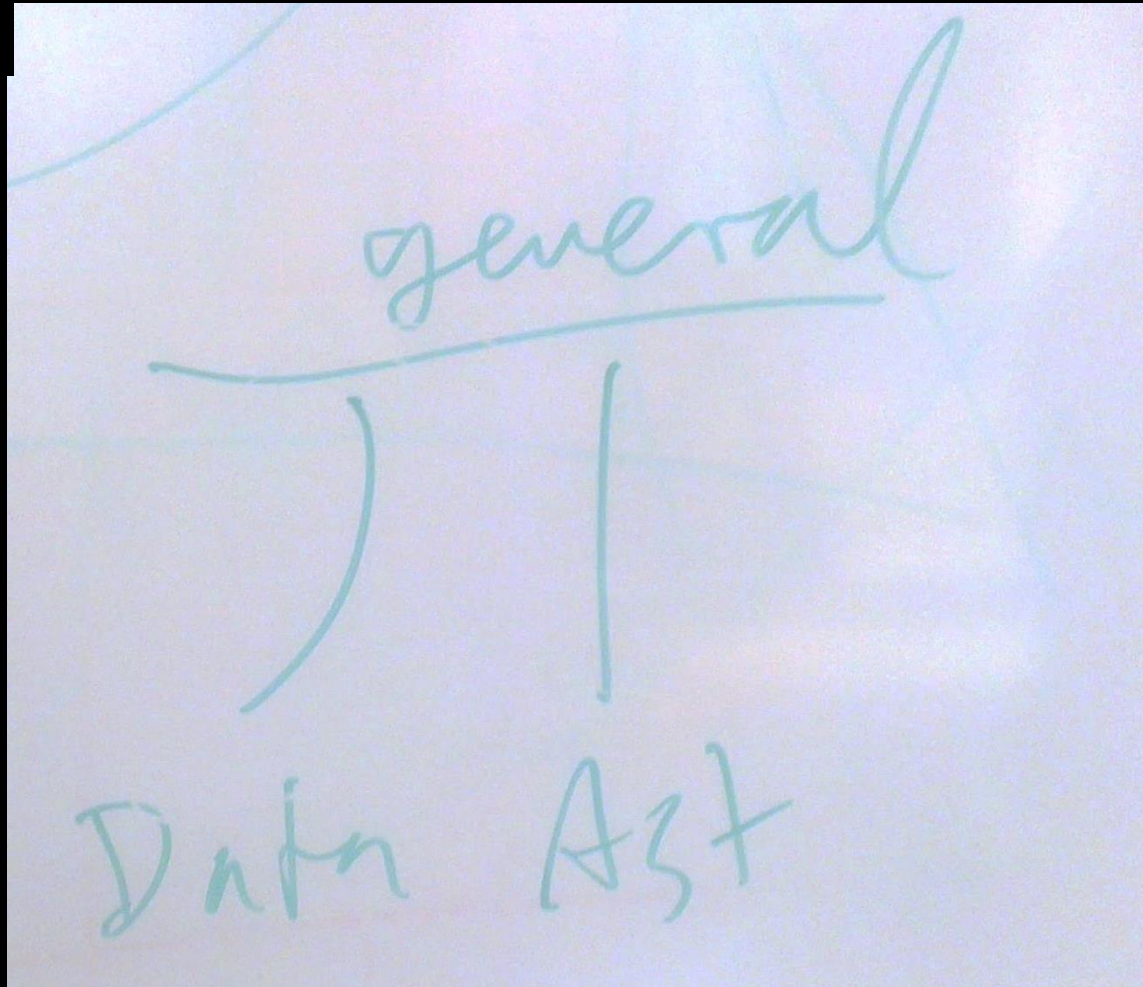


# Big Data Machines

- Lofar 2010 [www.lofar.org](http://www.lofar.org)
- VST-OMEGACAM 2011 [www.astro.rug.nl/~omegacam](http://www.astro.rug.nl/~omegacam)
- ALMA 2013 [www.eso.org/sci/facilities/alma](http://www.eso.org/sci/facilities/alma)
- GAIA 2014 [www.sci.esa.int/gaia/](http://www.sci.esa.int/gaia/)
- EUCLID 2020 [www.euclid-ec.org](http://www.euclid-ec.org)
- LSST 30Tb/night [www.lsst.org](http://www.lsst.org)
- SKA 2022 [www.skatelescope.org](http://www.skatelescope.org)
- The virtual observatory, Euro-VO, IVOA



# DS & CS - data scientist



# e-science

- Beyond “workstation science” of the 80-90’s
- Distributed services
- Distributed communities
- Distributed archives
- p2p networks – KAZAA- NAPSTAR
  - Share cpu
  - Share storage
  - Share info / meta data /knowledge



# Datacentric > 2015's living archives communities - data federations



- 1-100 Pbyte
  - Compute – local– grid -local
  - Data Store - distributed
  - Data manage - database
- } data scientist





**KiDS**

**VLT Survey Telescope (VST)**

2.6m telescope

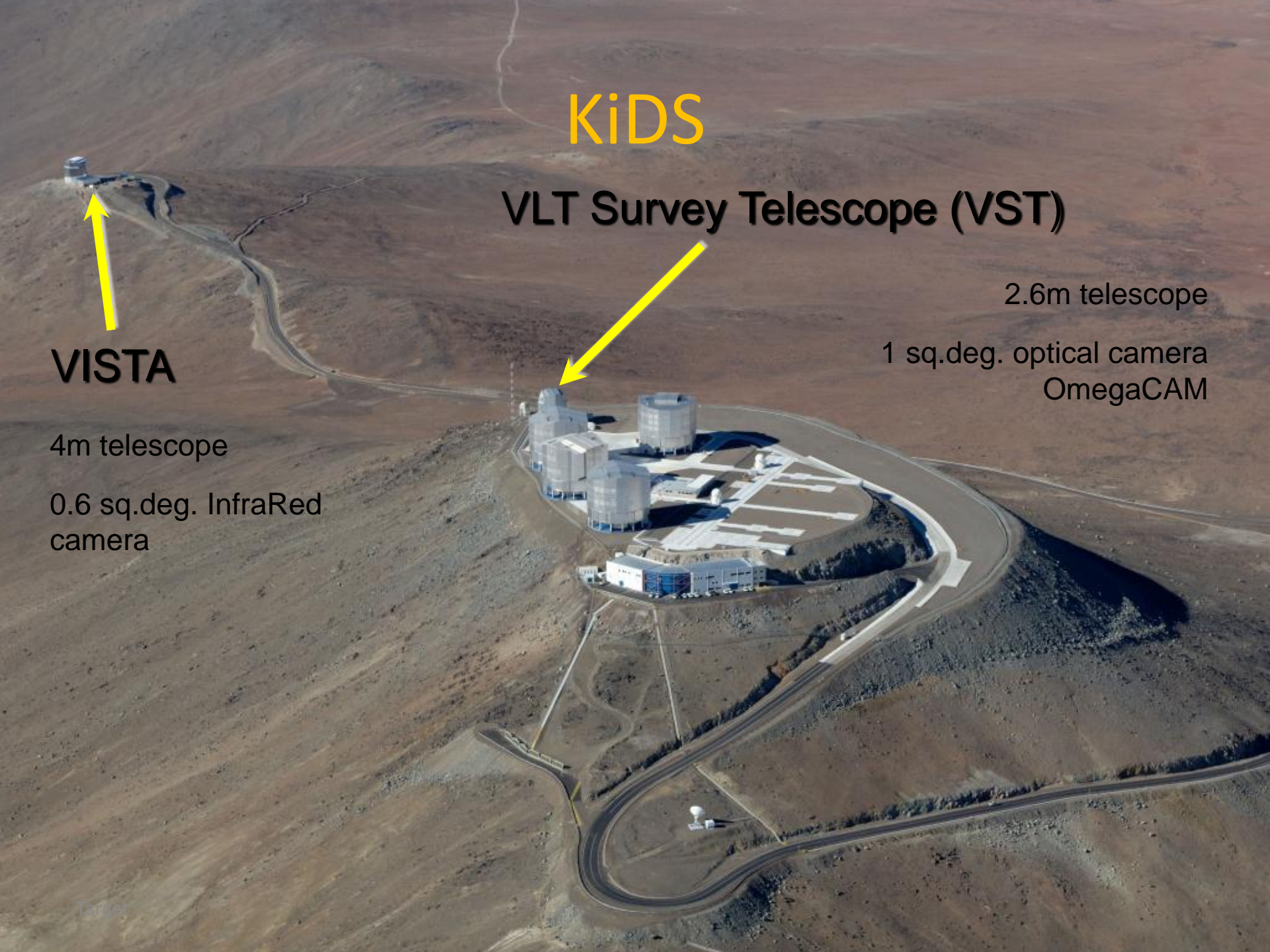
1 sq.deg. optical camera  
OmegaCAM

**VISTA**

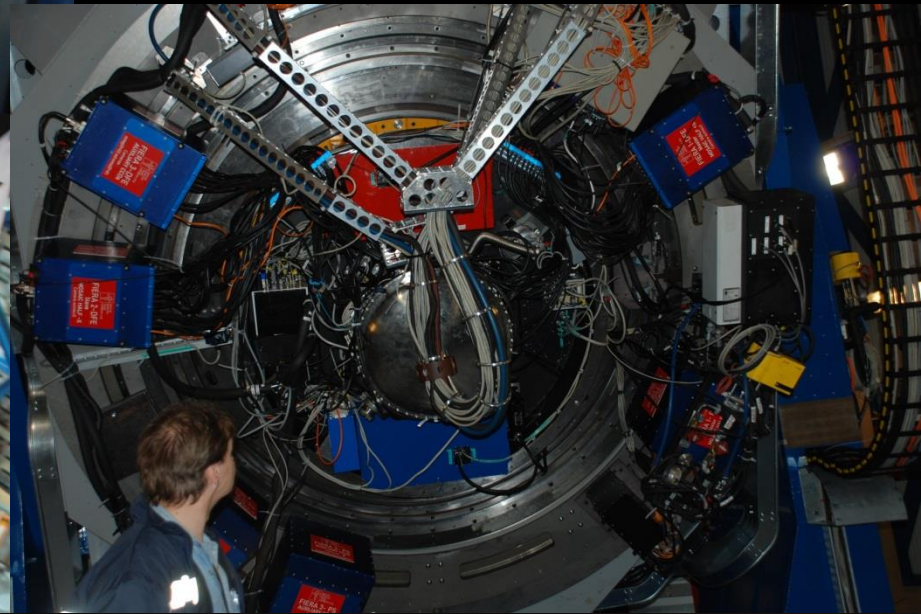
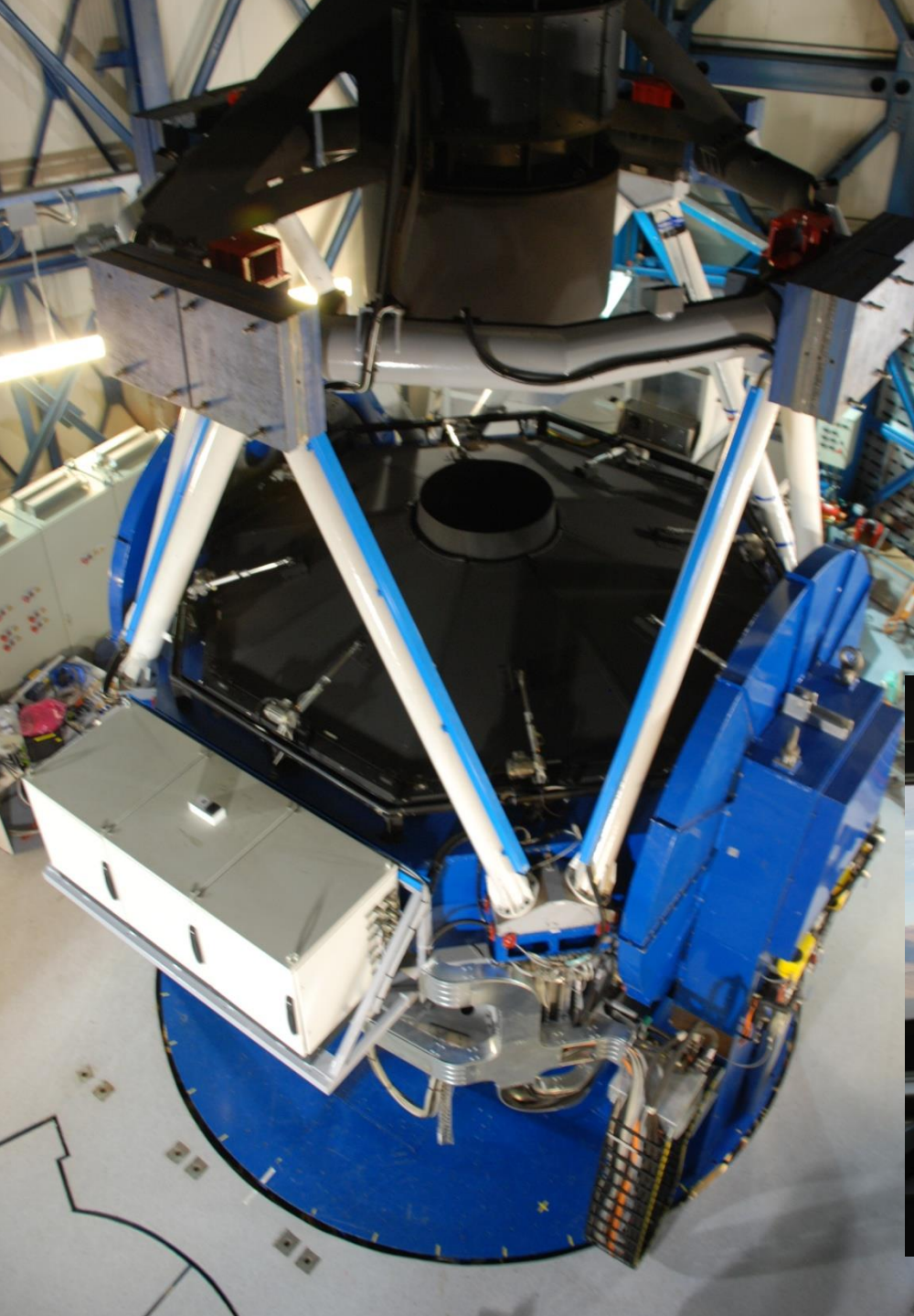
4m telescope

0.6 sq.deg. InfraRed  
camera

Target









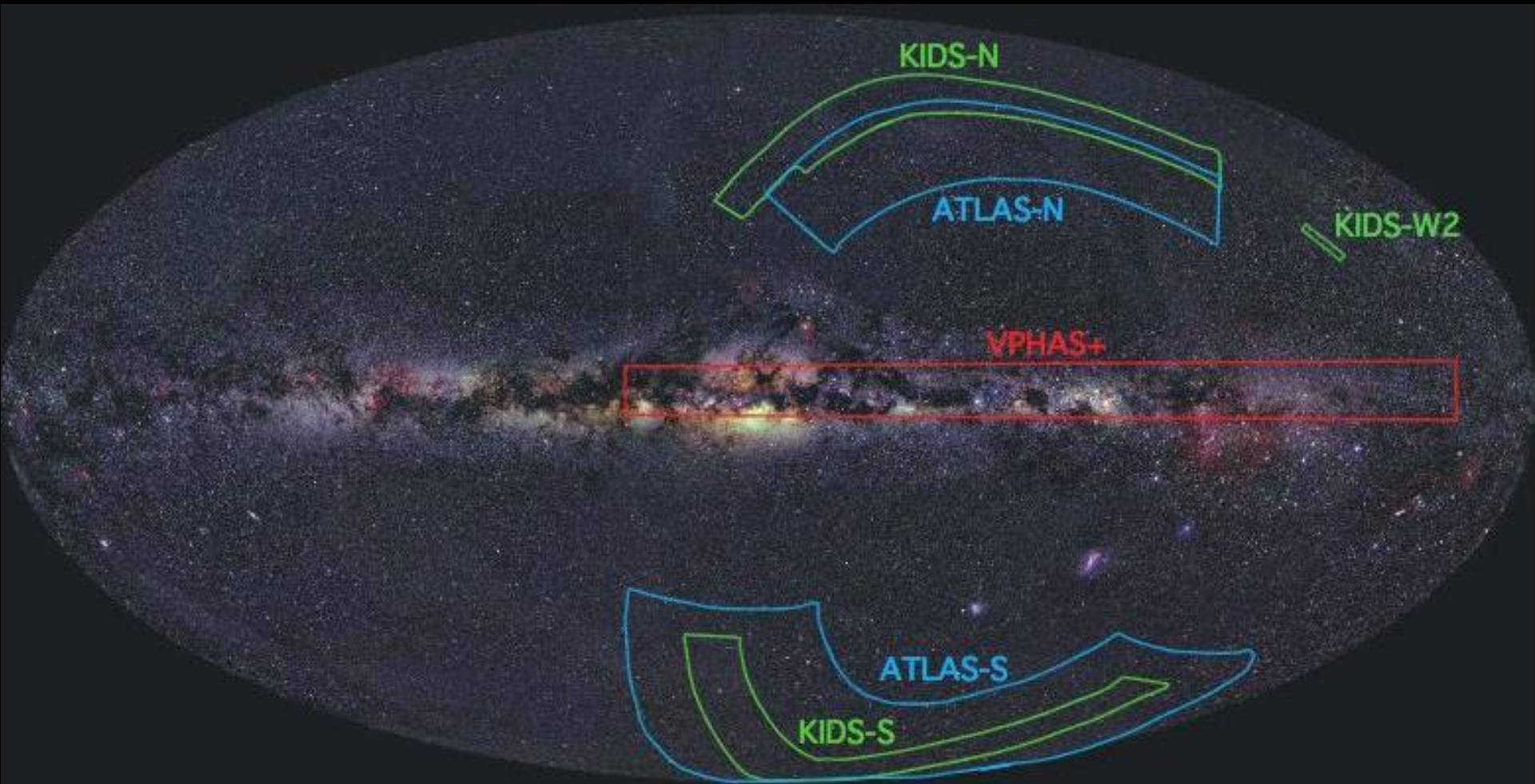




Leo triplet

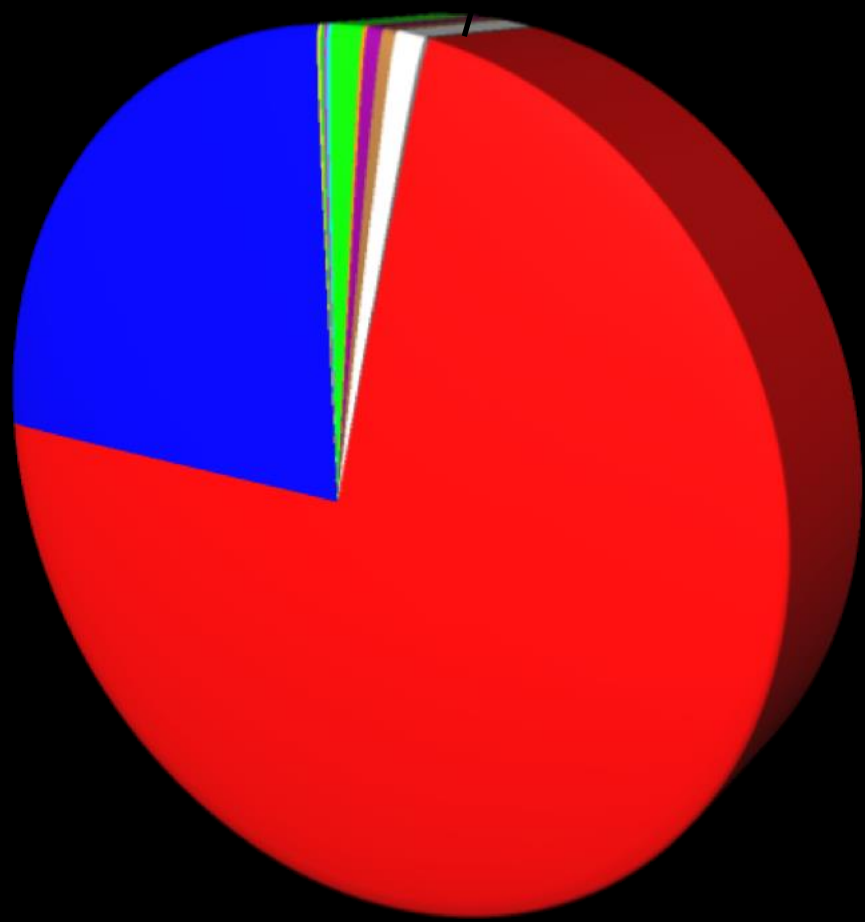


# ESO public surveys

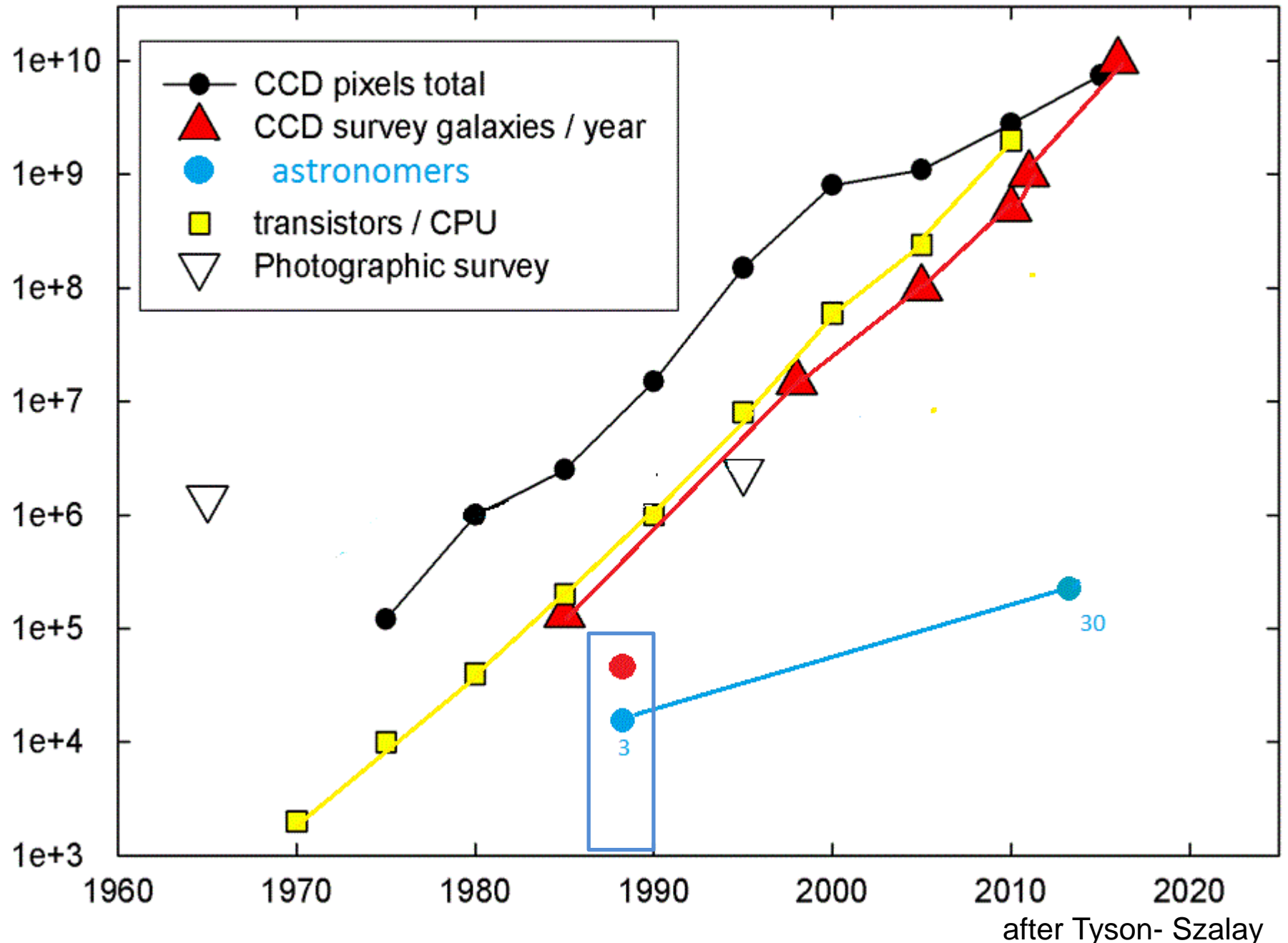




# Paranal Monthly Data Rates 2007 statistics



# Trends in Optical Astronomy Survey Data



# Basics- Surveys

- Defined area on sky
- Homogeneous
  - Survey limit
    - Flux (magnitude)
    - Size
    - Surface brightness
    - distance
- Quality control

# Basics - time

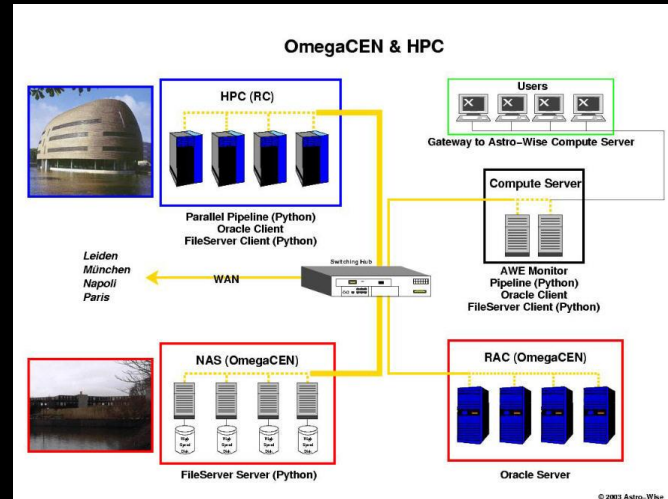
Everything changes in time

- Physical changes
- Our inside in modeling
- Methods, code, bugs

# basics- pipelines

- Workflow
- What triggers a pipeline?
  - Data items
  - Operators
  - users

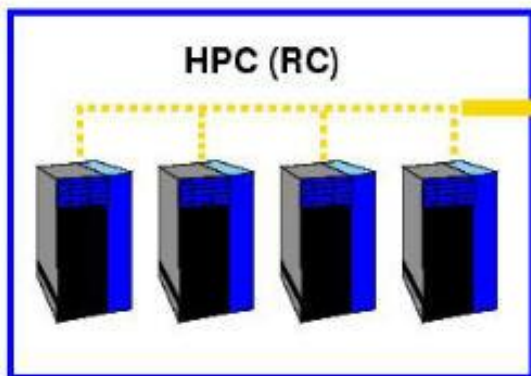
# Datacentric 2003-2015's local networks and internet



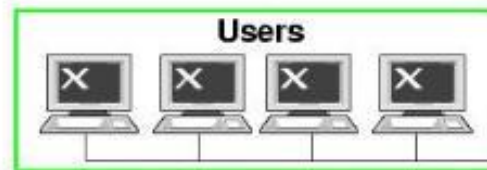
2003 Rug-CIT

- 1-100 Tbyte – Pbytes
  - DPU- Distributed Processing Unit
  - Distributed Data server
  - Data manage - database
- } User

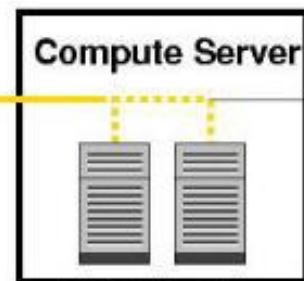
# VST - Virtual Survey Telescope



Parallel Pipeline (Python)  
Oracle Client  
FileServer Client (Python)

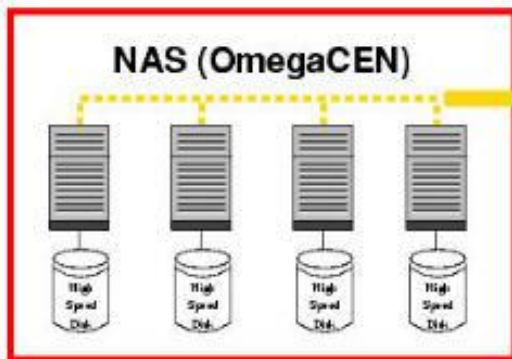


Gateway to Astro-Wise Compute Server

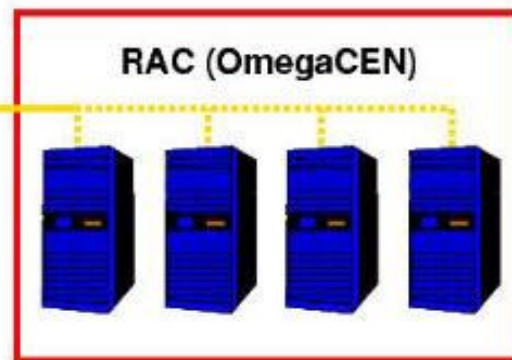


AWE Monitor  
Pipeline (Python)  
Oracle Client  
FileServer Client (Python)

Leiden  
München  
Napoli  
Paris



FileServer Server (Python)

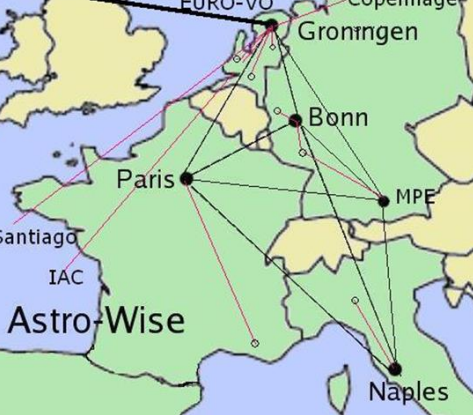


Oracle Server



VO and Datamining





# Astronomical Wide-field Imaging System for Europe

## publications

Science and technical publications

## team

Organizational structure, telephone numbers etc.

## events archive

Archive of past meetings, presentations etc.

## mailing lists

News

[archive](#) [subscribe](#)

Issues

[archive](#) [subscribe](#)

## web services

- [GoWise Search Engine](#)
- [Database Viewer](#)
- [Calibration Validation](#)
- [Target Processor](#)

a partnership of



(co-ordinator) [OmegaCEN - NOVA at Kapteyn Institute, Groningen - NL](#)

[Netherlands Research School for Astronomy, Leiden - NL](#)

[Argelander-Institut für Astronomie, Bonn - D](#)

[Osservatorio Astronomico di Capodimonte, Napoli - I](#)

[Universitäts-Sternwarte München - D](#)

[ESO, Garching bei München - D](#)

[Terapix, IAP, Paris - F](#)

An on-going project which started from a FP5 RTD programme funded by the EC Action "Enhancing Access to Research Infrastructures".

## Astro-WISE Online

### Overall storage and user statistics

Online storage: 1.6 PB (=1600 TB)

Number of files stored: 10717913

Database accounts: 156

Total queries<sup>1</sup>: 514980167

<sup>1</sup>sum for all databases since their last restart

### Status of services at Astro-WISE nodes

[Bonn](#) ● ●

[Groningen](#) ● ●

[Leiden](#) ● ● [Nijmegen](#) ● ●

[München](#) ● ●

[Napoli](#) ● ●

[golf details](#)

Updated: 24 Sep 2011 01:42:02



# Basics - Information systems

- Pipeline design
- Standards: Fits, table format- VO Standards
- Protocols
- Project management- sociology
- Data model
- Data base
- Integrated/ distributed File systems
  - Grid FTP, AstroWise, Hadoop, Cloud, Dropbox
  - Distributed computing

Grids, cloud computing

# Quality control

- Distributed
- Shared over the whole community
- web based
- OmegaCAM calibration plan
- OmegaCAM observing strategies

# Astro-WISE information system – fully datacentric

All data beyond pixel data is Metadata

all pixel data  $\leftrightarrow$  data servers

all Metadata  $\leftrightarrow$  database

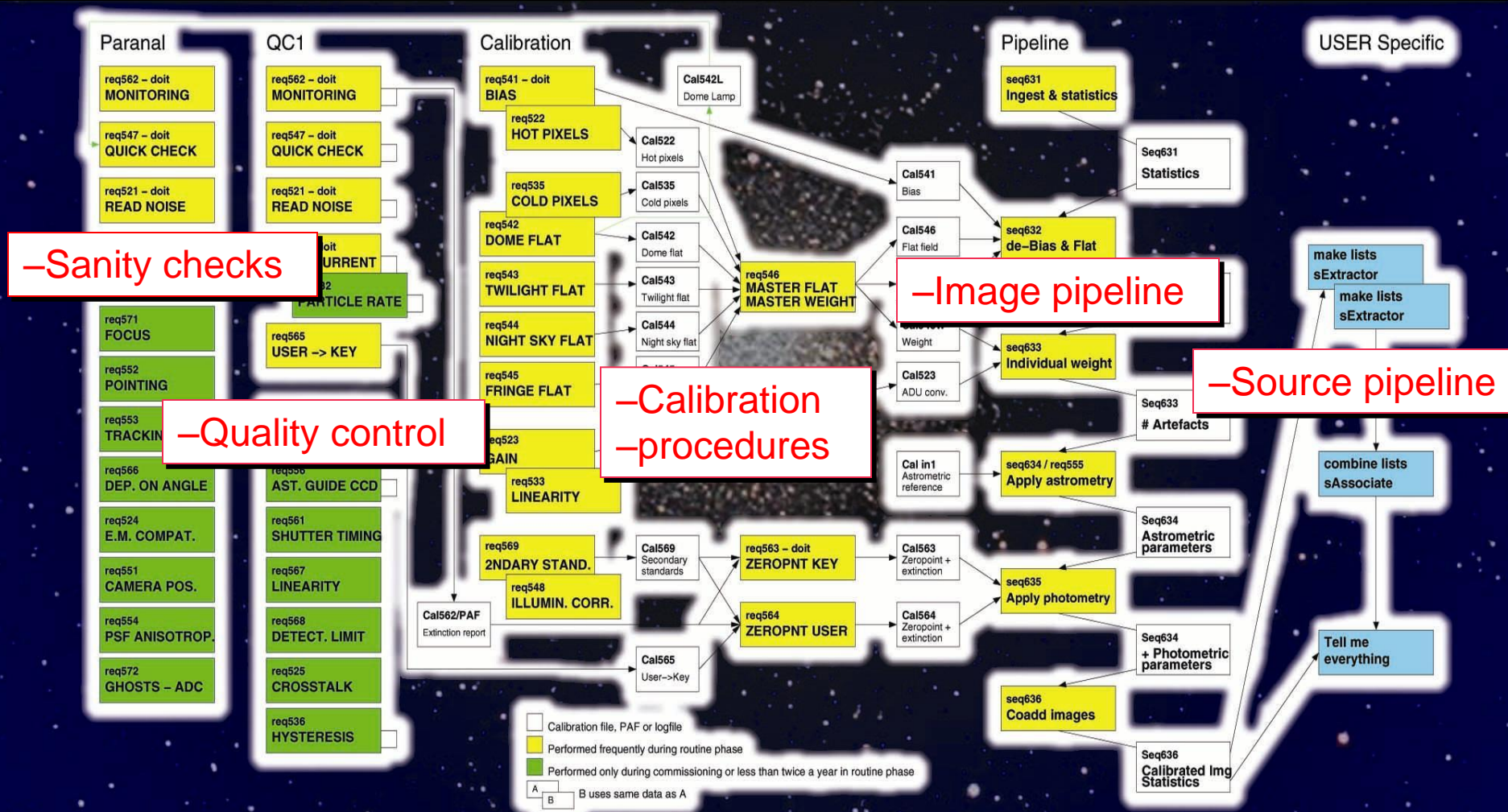
compute clusters / GRIDs all I/O to db

- all components scalable
- all components EU distributed

# Central role db

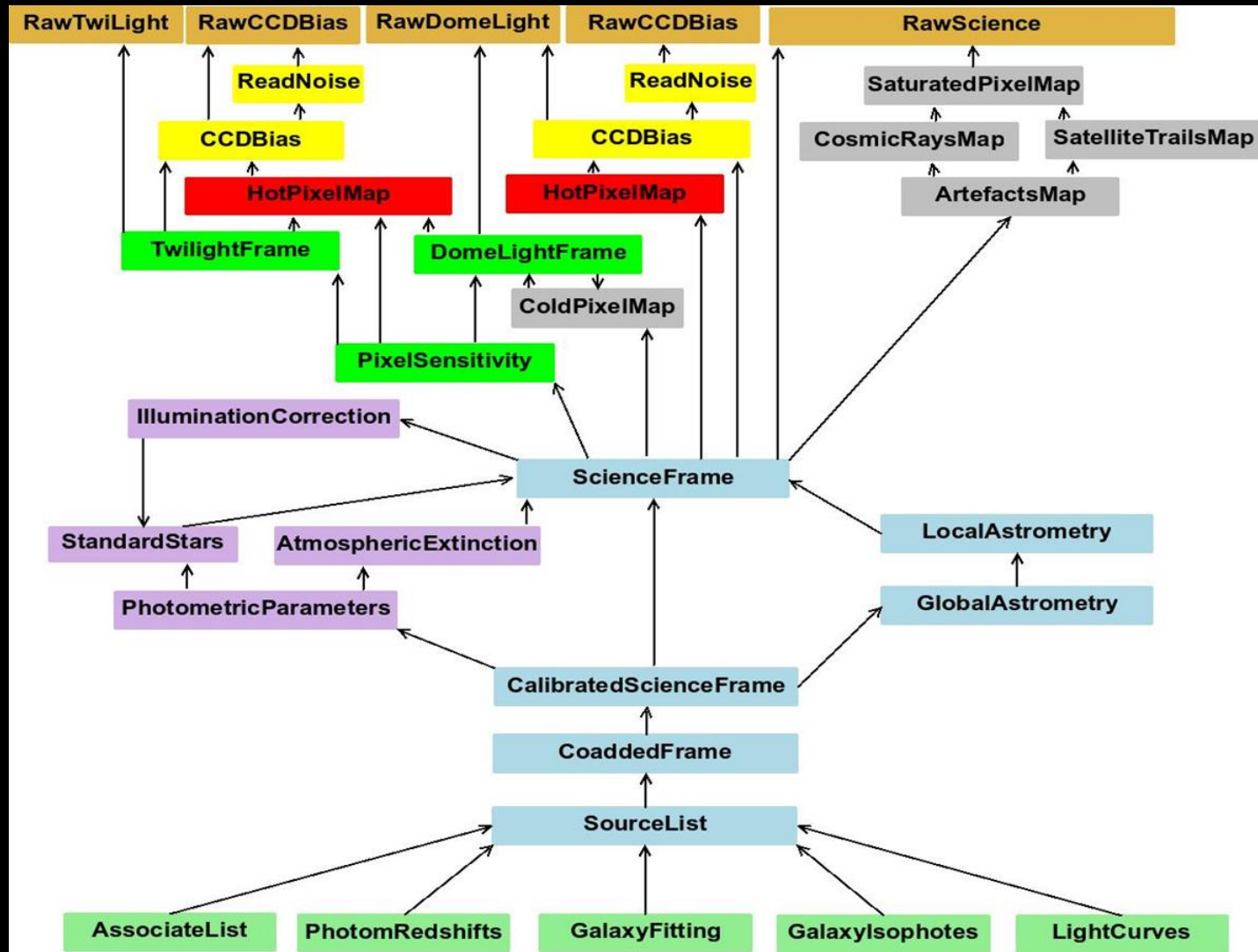
- All I/O via db ; metadata, source catalogues
- Objects persistent in db
- Quality control
- data server access via global filename (key) in db
- Security
- (parallel)processing
- Webservers
- Synchronized real time National Nodes

# Data Model



# Data model -> ++ model -> Target Diagram

## Data-centric /backward chaining



Target



# Target processing ++ the make metaphor

```
awe> targethot=HotPixelMap.get(date='2003-02-14', chip='A5382')
```

The processing chain is:

ReadNoise <-- Bias <-- HotPixels

- > class HotPixelMap(ProcessTarget):
- >> def self.make()
  
- > class ProcessTarget():
- >> def exist() # does the target exist?
- >> def uptodate() # is each dependency up to date?
- >> def get(date, chip) # if not exist/up-to-date then make()

# Persistency dbobject

```
Class DBMeta                                # python<->db
    def __new__ # makes any derived Class persistent
    def __call__ # instantiate persistent object - attributes

Class DBObject:
    __metaclass__ = DBMeta
    object_id = persistent('The object identifier', oidtype) #unique

# make it
    example = DBObject()
    example.commit()

# get it
    oid = example.object_id
    result = DBObject(object_id = oid)
```

# Persistency dataobject

```
from astro.database.DBMain import DBObject, persistent
class DataObject(DBObject):
    filename = persistent('File part of this object', str, '')
```

```
example = DataObject(pathname='example.txt')
example.store()
example.commit()
```

```
g = DataObject.filename.like('example*')
```

# AstroWise paradigm

<b>“Classical” paradigm</b>	<b>Target processing - Awe</b>
<b>Forward chaining</b>	<b>Backward chaining</b>
waterfall model	User hunts upstream
TIER architecture	
driven by input raw data	Driven by query of user
Process in pipeline	Process in bits and pieces on the fly
workflow	Backward chaining
Operators push data	User pulls data
Results in releases	Provide information system
Static archives – publish	Dynamic archives –publish Internet
Raw data - obsolete	Raw data is sacred

# Object view

Show [uplinks](#) of this ReducedScienceFrame

Show [quality view](#) of this ReducedScienceFrame

[Re-process](#) this ReducedScienceFrame

[show hidden attributes](#)

[expand All](#)

▣ ReducedScienceFrame(fa00388133145bfee043c216a9c325a8)

- AIRMEND 1.186
- AIRMSTRT 1.174
- creation\_date 2014-05-22 18:41:01
- DATE 2012-05-17 04:35:09
- DATE\_OBS 2012-05-17 04:29:06
- EXPTIME 360.0
- filename [\(preview\) Sci-JMCFARLAND-OMEGACAM-----OCAM\\_r\\_SDSS-ESO\\_CCD\\_#80-Red---Sci-56799.7783076-ad1e0002aeb303b8612f11ca724f565df2](#)
- globalname None
- is\_valid 1
- NAXIS1 2048
- NAXIS2 4100
- OBJECT KIDS\_212.0\_0.5
- process\_status 1
- psf\_radius 0.525272135701
- quality\_flags 0
- scale\_factor 0.0

▣ astrom

▣ bias [\(up\)](#) [\(link\)](#)

- creation\_date 2012-05-21 09:26:42
- filename [\(preview\) Cal-EHELMICH-OMEGACAM-----ESO\\_CCD\\_#80-Mas---Bias-56068.3934298-2e0fe8bdea1eb9a0d3a82598a652962b31fbd307.f](#)

Target globalname None

- is\_valid 1
- NAXIS1 2048

# Supported data sources



Here you can find quick links to different subsets of all public data in our database.

## WFI@2.2m data

[raw science](#) [reduced science](#) [coadded science](#)  
[sourcelists](#) [world](#)

Find data for the WFI instrument.

## WFC@INT data

[raw science](#) [reduced science](#) [coadded science](#)  
[sourcelists](#) [world](#)

Wide-Field Camera on the Isaac Newton Telescope at La Palma.

## SUP@Subaru data

[raw science](#) [reduced science](#) [coadded science](#)  
[sourcelists](#) [world](#)

Suprime-Cam data from the Subaru telescope on Mauna Kea, Hawaii.

## HST ACS data

[coadded science](#) [sourcelists](#) [world](#)

Data from the Advanced Camera for Surveys instrument aboard

## WENSS data

[radio images](#) [world](#)

Data of the Westerbork Northern Sky Survey.

## ESO-LV data

[ESO-LV Images](#) [ESO-LV Catalogue](#) [world](#)

Browse the images in the LV-Catalogue, query them using RA and DEC, description or ESO name.  
or

Browse the ESOLV-Catalogue, query all objects using numerous properties.

## SDSS DR4 data

[SDSS DR4 subset SourceList](#) [world](#)

Browse an exemplary subset of the SDSS DR4 catalog locally. This subset has been inserted into an Astro-WISE SourceList. Complete ingestion of DR4 is pending.

## SDSS DR5 data

[SDSS DR5 subset SourceList](#) [world](#)

Browse an exemplary subset of the SDSS DR5 catalog locally. This subset has been inserted into an Astro-WISE SourceList.

+ WFCAM, MegaCAM, LBC, ISAAC, LOFAR

# Quality control- flags

- Quality flags

System quality\_flags (method verify)

User is\_valid (method inspect by user)

0 = bad

1 = OK

2 = Qualified - ready for delivery

- Context - privileges

1 Mydb

user\_CalFile

2 Project, eg KIDS

Project favorite flag

project CalFile

3 AstroWise

Awe Calfile

4 World



# QC - calibration scientist monitoring

Calibration Timestamps - Netscape

File Edit View Go Bookmarks Tools Window Help

http://calts.astro-wise.org:8878/

Home Google OCam OCen EV NOS AE AA Ilse PyDoc AweSQL Awe CVS AweNews AweCalts Router Start Lyc AWE SQLform...

New Tab Calibration Timestamps

Astro-Wise  
Calibration Timestamps

width : 1024 1280  
author : wjvriend  
user : awevalentyn

521 Readout Noise  
522 Hot Pixelmap  
523 CCD Gain  
535 Cold Pixelmap  
**541 Master Bias**  
542 Master Domeflat  
543 Master Twilightflat  
544 Nightsky Flat  
545 Fringe flat  
546 Master Flatfield  
548 Illumination  
548F Illumination Coef.  
563+564 Zeropoint  
565 Band pass transformation  
631 RawScienceFrame

Instrument: WFI Chip: ccd51 Filter: <none>

year: 2000 quarter: 3 month: <none> week: <none>

Only good data (no flags set)

Table / Graph

Timestamp start - end 01 Jul 2000 30 Sep 2000 Creation date

ccd51

01 Jan 1990-01 Jan 2030		01 Jan 1990
01 Aug 2000-02 Aug 2000		07 Jan 2005
03 Aug 2000-07 Aug 2000		07 Jan 2005
07 Aug 2000-08 Aug 2000		07 Jan 2005
08 Aug 2000-20 Aug 2000		10 Jan 2005
20 Aug 2000-21 Aug 2000		10 Jan 2005
21 Aug 2000-22 Aug 2000		10 Jan 2005
22 Aug 2000-23 Aug 2000		10 Jan 2005
23 Aug 2000-24 Aug 2000		10 Jan 2005
24 Aug 2000-25 Aug 2000		10 Jan 2005
25 Aug 2000-26 Aug 2000		10 Jan 2005
02 Aug 2000-03 Aug 2000		14 Jan 2005
02 Aug 2000-03 Aug 2000		14 Jan 2005

Total calibration files: 13

Legend:

- used data
- eclipsed data
- quality\_flags <> 0
- super\_flag <> 0

# 4 Lines script

```
#Find ScienceFrames for a ccd named ccd53 and filter
```

```
awe> q = (ReducedScienceFrame.chip.name == 'ccd53') and  
(ReducedScienceFrame.filter == '#841')
```

```
# From the query result, get the rms of the sky in image
```

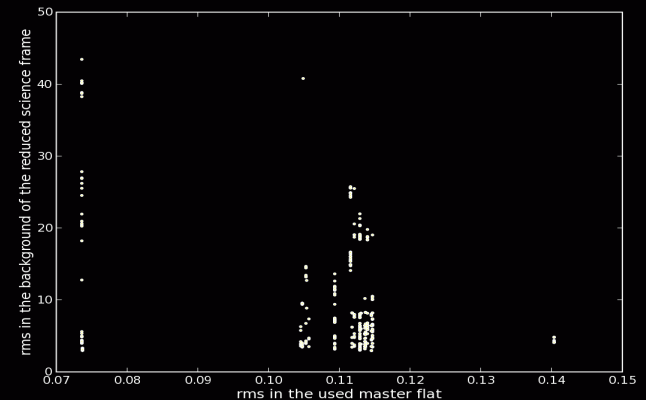
```
awe> x = [k.imstat.stdev for k in q]
```

```
# get the rms of the used Masterflat
```

```
awe> y = [k.flat.imstat.stdev for k in q]
```

```
# Make a plot
```

```
awe> pylab.scatter(x,y)
```



# CLI

```
AWE prompt - Mozilla Firefox
File Edit View History Bookmarks Tools Help
AWE prompt
process.astro-wise.org/AWE
Most Visited openSUSE Getting Started Latest Headlines

Welcome to the online Ast
version 0.5
21 apr 2008
NEW : see the Code S
and use Tab for completi
author Willem-Jan
user AWANONYM

Welcome to the Astro-WISE Envir
|22-Sep-11 OmegaCAM-VST Public Survey observ
mid October 2011
| You are running the AWBASE version

Importing Astro-WISE packages. Please wait...
ERROR: Traceback (most recent call last):
File "<input>", line 4, in <module>
File "/astro-wise/awehome/AWBASE/astro/recipe
dpus = [name[0] for name in dpu_get_info(i
TypeError: 'NoneType' object is not iterable
Initializing Distributed Processing Unit...

Error executing query: SELECT NAME
FROM AWOPER.DPU_SERVER_TABLE

Exception: ORA-00942: table or view does not

Current profile:
- username : AWANONYMOUS
- database : db.astro.rug.astro-wise.org
- project : OMEGACAM@VST

awe>

dpu = Processor('dpu.hpc.rug.astro-wise.org')
fr_list = (CoaddedRegriddedFrame.instrument.name == 'VIRCAM') & \
(CoaddedRegriddedFrame.creation_date > datetime.datetime(2010,8,18)) & \
(CoaddedRegriddedFrame.creation_date < datetime.datetime(2010,8,25)) & \
(CoaddedRegriddedFrame.is_valid == 1) & \
(CoaddedRegriddedFrame.filename.like('*EHELMICH*'))

pars={}
pars['SourceList.sexconf.DETECT_MINAREA']=4
pars['SourceList.sexconf.DETECT_THRESH']=3.0
pars['SourceList.sexconf.DEBLEND_MINCONT']=0.00005
pars['SourceList.sexconf.DEBLEND_NTHRESH']=16

for frame in fr_list:
pars['SourceList.sexconf.SEEING_FWHM']=frame.psf_radius
print pars['SourceList.sexconf.GAIN'], pars['SourceList.sexconf.SEEING_FWHM']
dpu.run('SourceList',f=[frame.filename],p=pars,sp=sp,C=1,dpu_time=10800)

Welcome to the Astro-WISE Environment

|08-Mar-11 OmegaCAM commissioning starting late March 2011. For more
details visit http://www.astro-wise.org/~omegacam

| You are running the AWBASE version

Importing Astro-WISE packages. Please wait...

Initializing Distributed Processing Unit...

Current profile:
- username : awabelikov
- database : db.astro.rug.astro-wise.org
- project : ALL
- current privileges : 1 (MyDB)

WARNING: Program swarp version mismatch. Expected version 2.19.1, got 2.17.6
awe>
```

# Re-processing view



**Astro-WISE Homepage**

**Target Processor**

**Contact**  
Willem-Jan Vriend

**DB User**  
awevalentyn

**Help**  
Getting Started

**Project**  
KIDS

**Instrument**  
OMEGACAM

**State**

1. Preselect Target
2. Specify Target
3. Select Target(s)
4. Process or Query

**Options**

Preferences

Process Parameters

Target

Upload Code

Job overview

## Query results

- A treeview is given of the target(s). This treeview gives an overview of the target's dependencies. Green dependencies mean a newer version exists.
- For each target only one chip is shown, but with the 'view all chips' link all chips can be queried. For the processing to be processed.
- For each target the target number, filter, date and chip are shown. With the process link on the right the target will be processed.
- On the bottom the messages, logs and statistics can be viewed.

Query Depth Calibration  Science

0	OCAM_r_SDSS SloanR	23 Feb 2014 00:06:54	ESO_CCD_#65 shown	<a href="#">view all chips</a>	<a href="#">process</a>	<a href="#">query</a>
---	--------------------	----------------------	-------------------	--------------------------------	-------------------------	-----------------------

- (+) (-)  1.0 ReducedScienceFrame (to be built)
  - 2.1 BiasFrame
  - 2.2 ColdPixelMap (rebuild)
  - 2.3 MasterFlatFrame (rebuild)
    - 3.1 ColdPixelMap (rebuild)
    - 3.2 DomeFlatFrame (rebuild)
    - 3.3 HotPixelMap (new version)
      - 4.1 BiasFrame (new version)
      - 3.4 NightSkyFlatFrame (new version)
      - 3.5 TwilightFlatFrame (new version)
    - 2.4 FringeFrame (null)
    - 2.5 HotPixelMap
      - 3.1 BiasFrame
    - 2.6 IlluminationCorrectionFrame
    - 2.7 RawScienceFrame

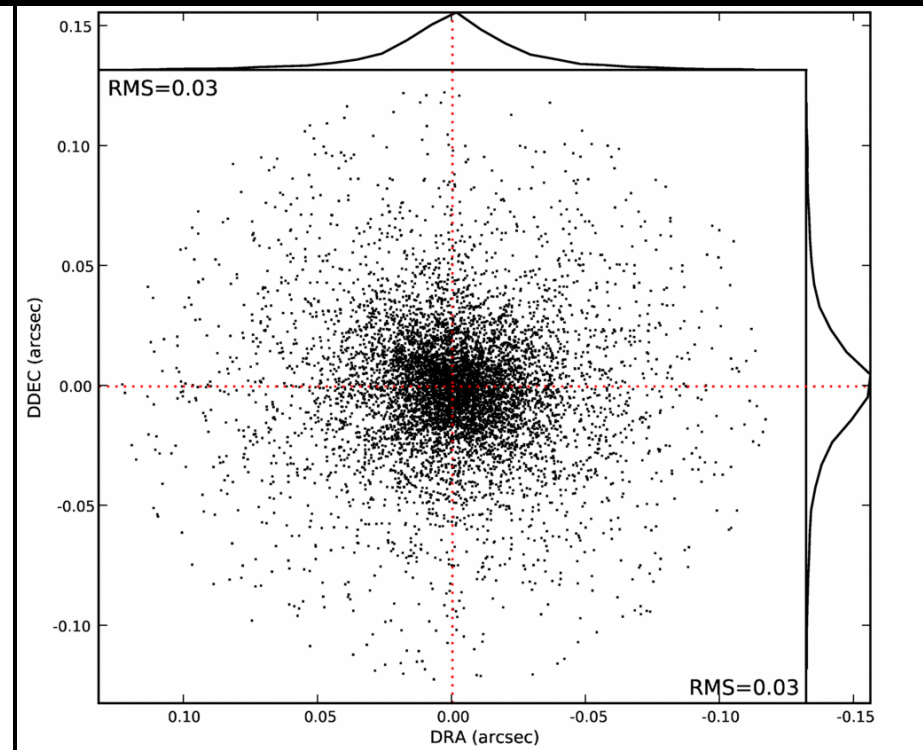
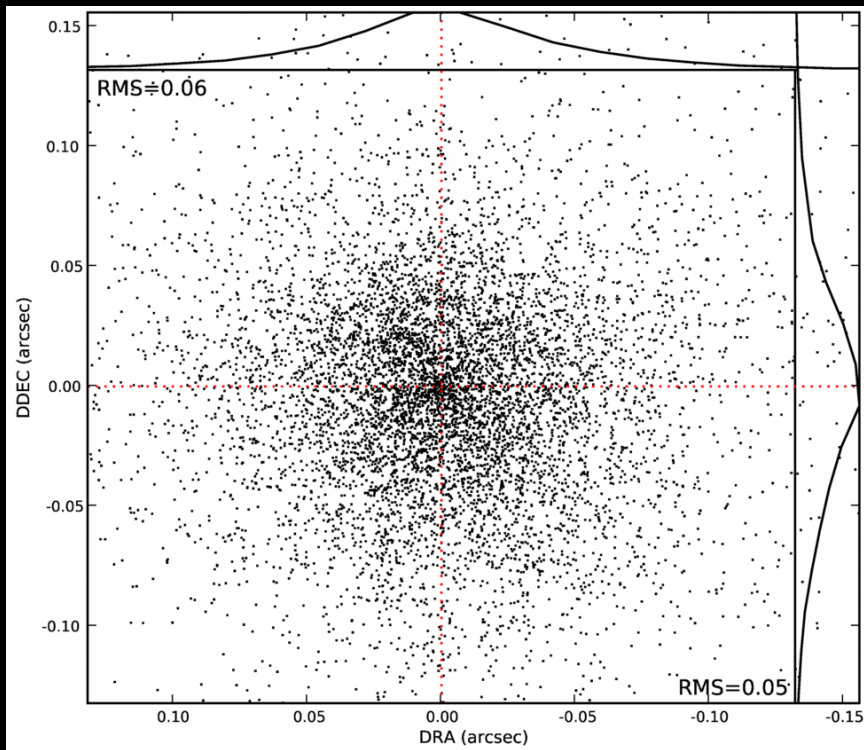
**Object Details** [Close](#)

Name	Value
type	RawScienceFrame
filename	OMEGACAM.2014-02-23T00:06:54.704_1.fits
chip	ESO_CCD_#65
filter	OCAM_r_SDSS
Object	STD,ZEROPOINT
quality_flags	0
is_valid	1
observation	23 Feb 2014 00:06:54
creation	23 Feb 2014 06:38:38
Links	<a href="#">Download</a> <a href="#">Preview</a> <a href="#">DBview</a> <a href="#">Quality</a>

Log File  
+ Messages

# Global Astrometry

reproducibility



Local solution

Global solution

(Internal regrid-to-regrid residuals of one WFI N=4 dither)



### Quality of REGRIDDEDFRAME:

Sci-EVALENTYN-WFI---#842-ccd50-Regr---Sci-54566.3131050-f56144d965b5e765b40bdec3d685fe595215d52b.fits

ASTROWISE REVIEW CALS PROCESS

no previous comments

OBname: zwoVARD055 project: WFI@2.2m

is\_valid = 1: valid

#### Processing Details

creation_date	2008-04-10 07:31:02
is_valid	1
quality_flags	0
Privileges	4

#### Image Statistics Details

mean	+9.327e+01
median	+8.887e+01
stdev	+2.945e+03
min	-6.036e+06
max	+3.727e+06

#### Local Astrometry Details

creation_date	2008-04-10 07:30:42
is_valid	1
quality_flags	0
RMS	0.252
SEEING	0.856
NREF	317
SIG_DRA	0.209
SIG_DDEC	0.178
MEAN_DRA	-0.001
MEAN_DDEC	0.002

#### Photometry Details

creation_date	2008-03-29 20:15:06
is_valid	1
quality_flags	0
zeropoint	24.759
zp_error	0.000
zp_origin	derived
num_sources	173

#### Observational Details

DATE_OBS	2002-03-18 03:35:21	OB SERVER	UNKNOWN
MJD_OBS	52351.1495509	EXPTIME	299.9178
OBJECT	Bnc-W	AIRMBSTR	1.246
R.A.	13 25 30.0000	AIRMEND	1.246
Dec.	-31 35 44.8808	Filter	#842
	mag_id		JohnsonB

Chip ccd60 of instrument WFI



RegriddedFrame

2438 X 4873 pixel

8.13 X 16.24 arcmin

WeightFrame

2438 X 4873 pixel

8.13 X 16.24 arcmin

AstrometricParameters

Processing Details

creation_date	2014-05-22 18:41:01
is_valid	1
quality_flags	0
Privileges	2

Image Statistics Details

mean	+5.932e+02
median	+5.924e+02
stdev	+1.948e+01
min	+4.986e+02
max	+6.901e+02

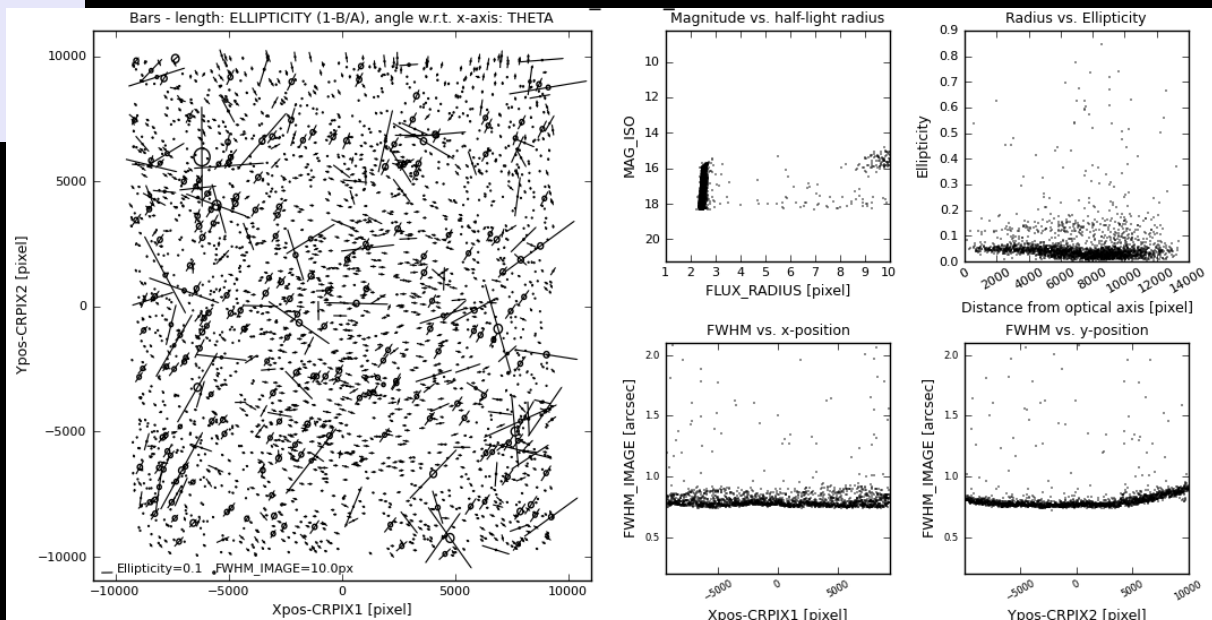
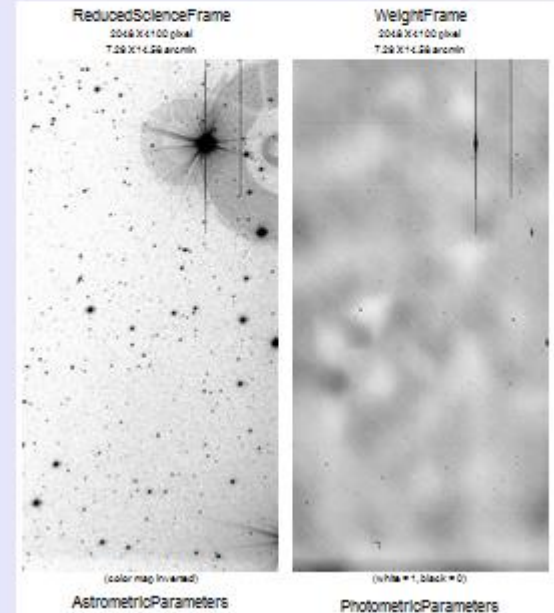
Astrometry Details  
(derived from a global solution)

creation_date	2014-05-23 00:22:00
is_valid	1
quality_flags	0
SEEING	0.527
RMS	0.396
NREF	36
SIG_DRA	0.253
SIG_DDEC	0.277
MEAN_DRA	0.086
MEAN_DDEC	0.024
RMS_OVERLAP	
N_OVERLAP	
SIG_DRA_OVERLAP	
SIG_DDEC_OVERLAP	
MEAN_DRA_OVERLAP	
MEAN_DDEC_OVERLAP	

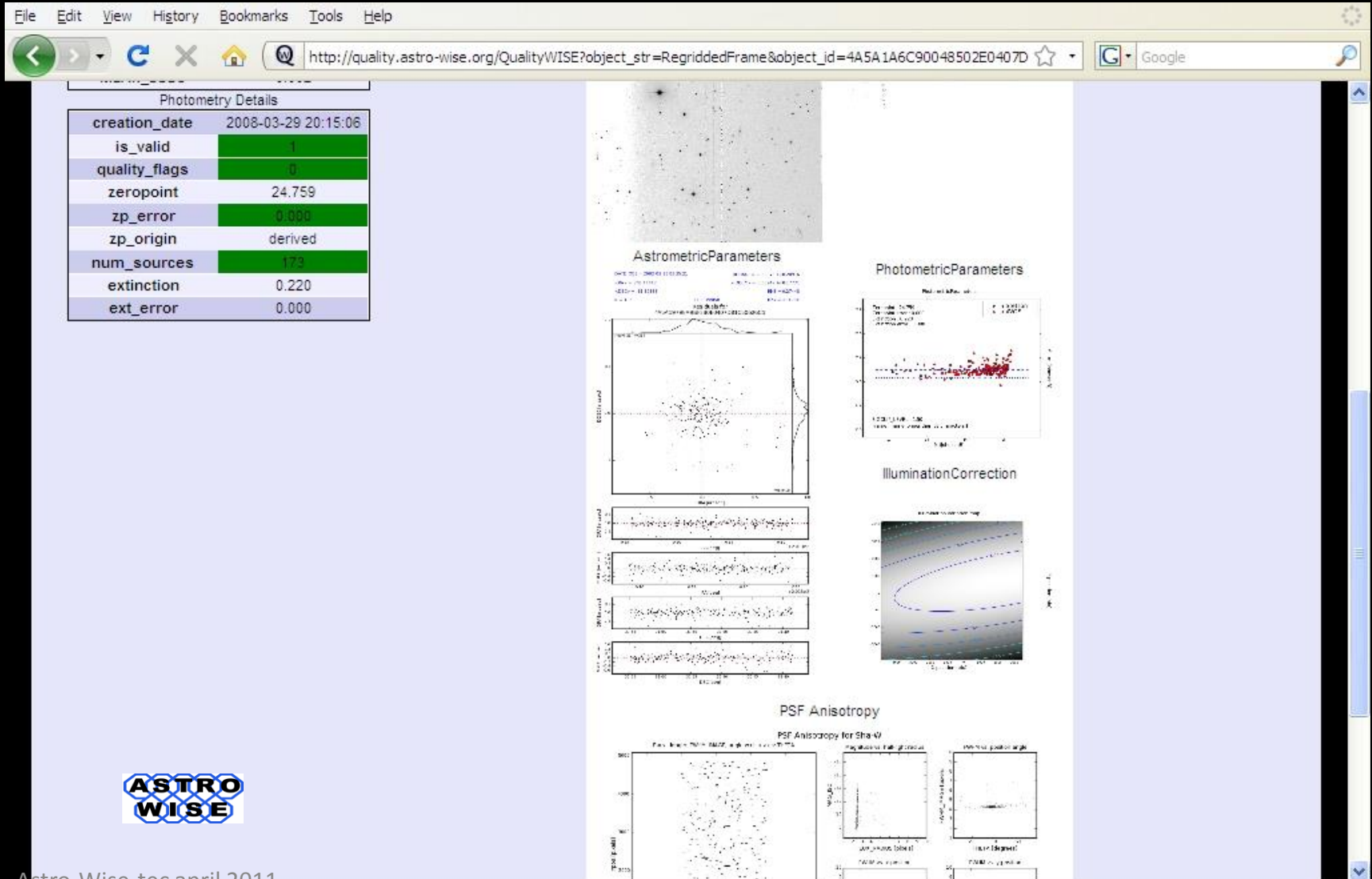
Photometry Details

creation_date	2013-10-06 16:56:48
is_valid	1
quality_flags	0
zeropoint	24.729
zp_error	0.000
zp_origin	derived
num_sources	0
extinction	0.102
ext_error	0.002

Observational Details			ChIP ESO CCD #98 of instrument
DATE_OBS	2013-05-17	OBSERVER	UNKNOWN
TIME_OBS	04:29:04		
INSTRUMENT	24004-1666617	EXPTIME	240.0
OBJECT	NGC_2120_02	AIRMSZT	1.174
obs_id	732411	AIRMSD	1.166
R.A.	212.0	Filter	CCD_M_9099
Dec.	0.5	mag_id	SloanR



# Quality view

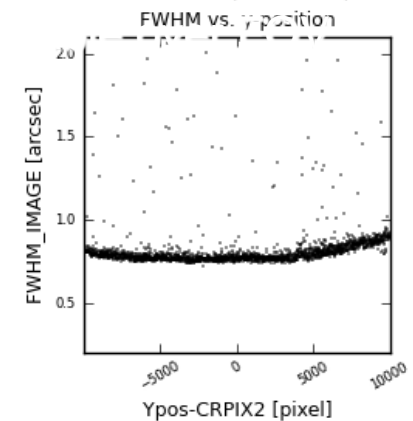
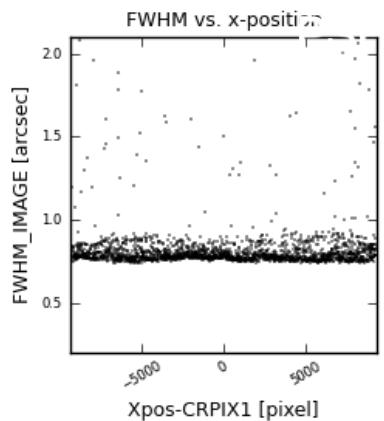
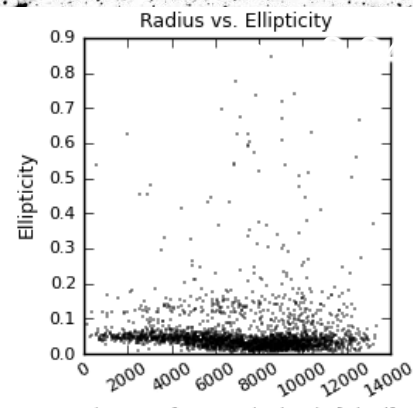
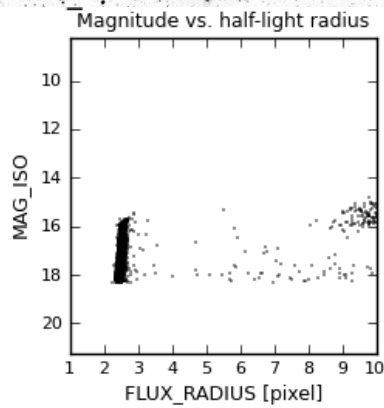
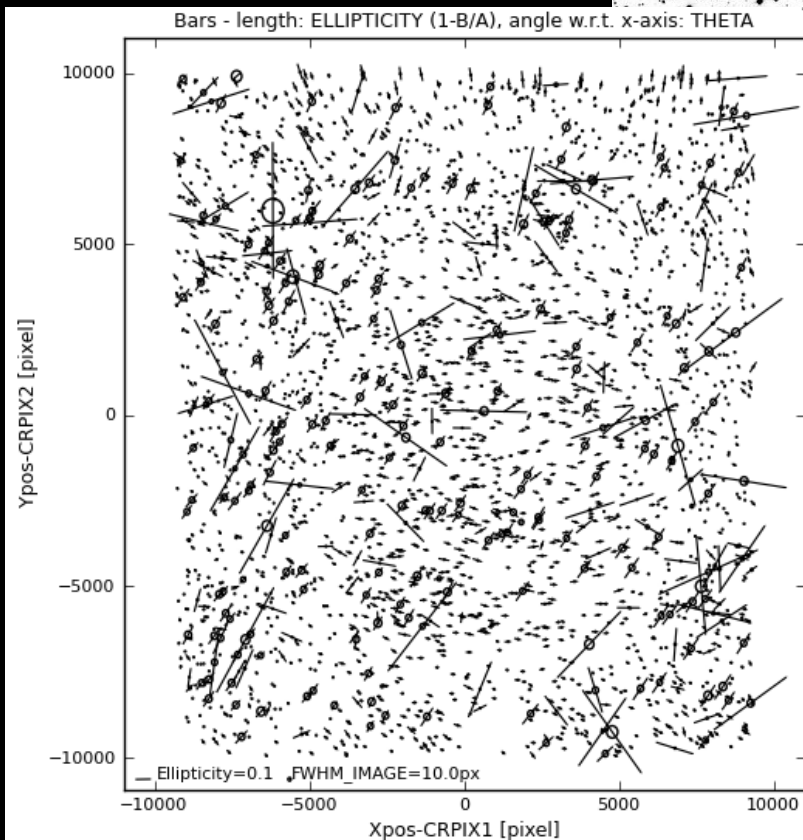
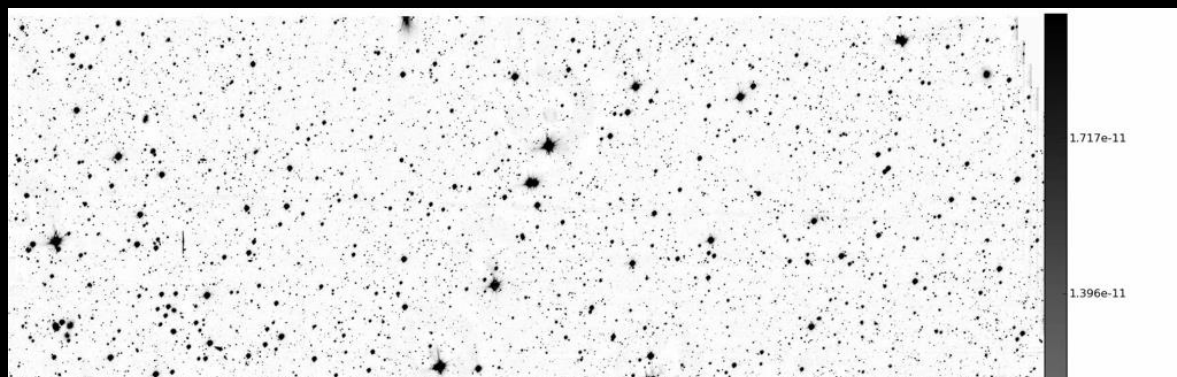




# KiDS: data quality

KIDS\_129.0\_-0.5

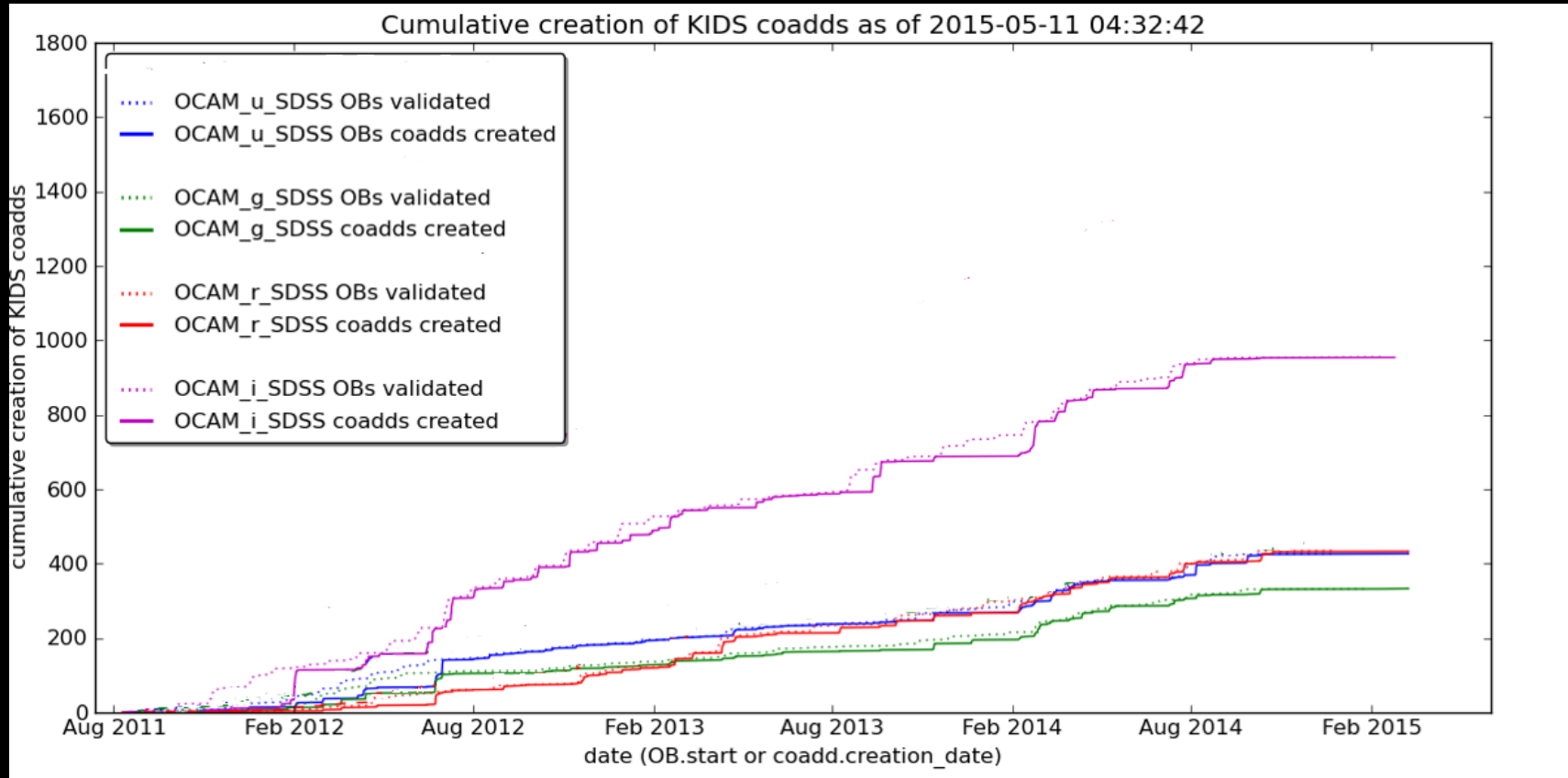
r



# KiDS/VIKING survey status

dd May 2015: ~375sq.deg in ugri(ZYJHK)

DR1 DR2 delivered

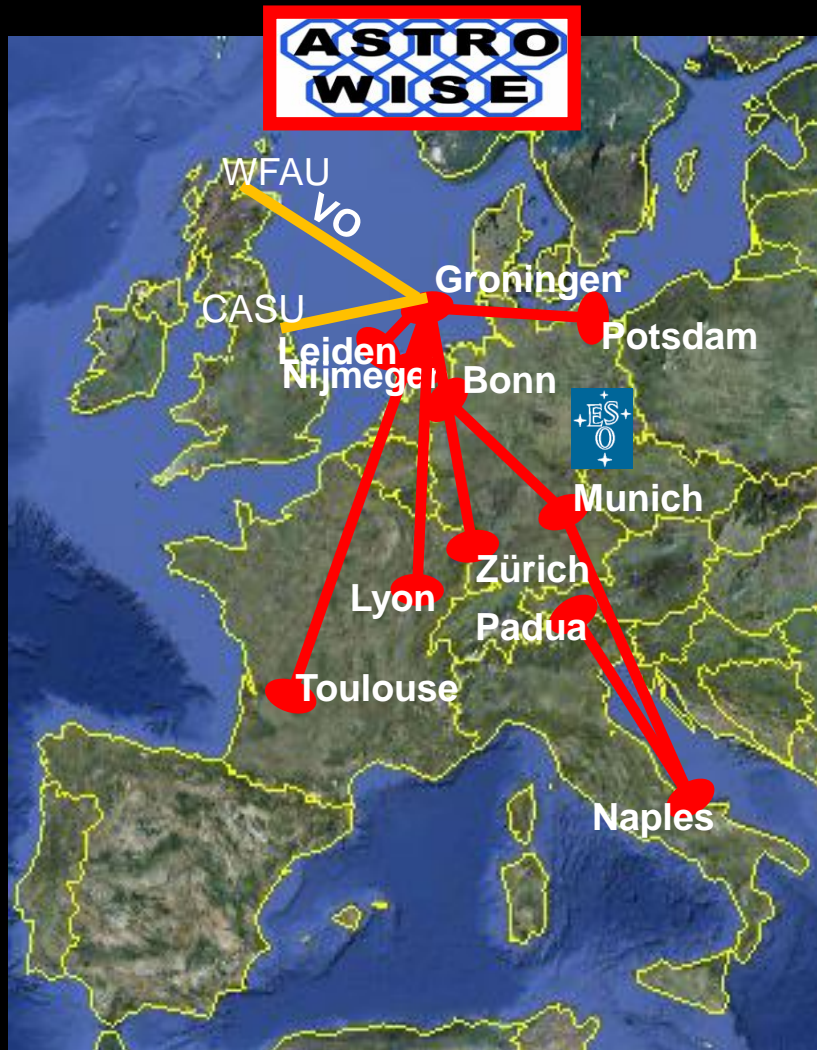


# the avalanche

- on-the fly re-processing for everything
- 5LS: 5 Lines Script Awe> prompt Python
- Trend analysis Awe > prompt
- All dependent bits are traced “tell\_me\_everything\_tool
- Administration for parallel processing -compute GRID
  - Connect to EGEE - Grid
- Global solutions
  - astrometry/photometry
- Build-in workflow
- Fully user tunable – own provided script
- Context: projects/surveys, instruments, mydb
- Publish directly in EURO-VO

# AstroWise

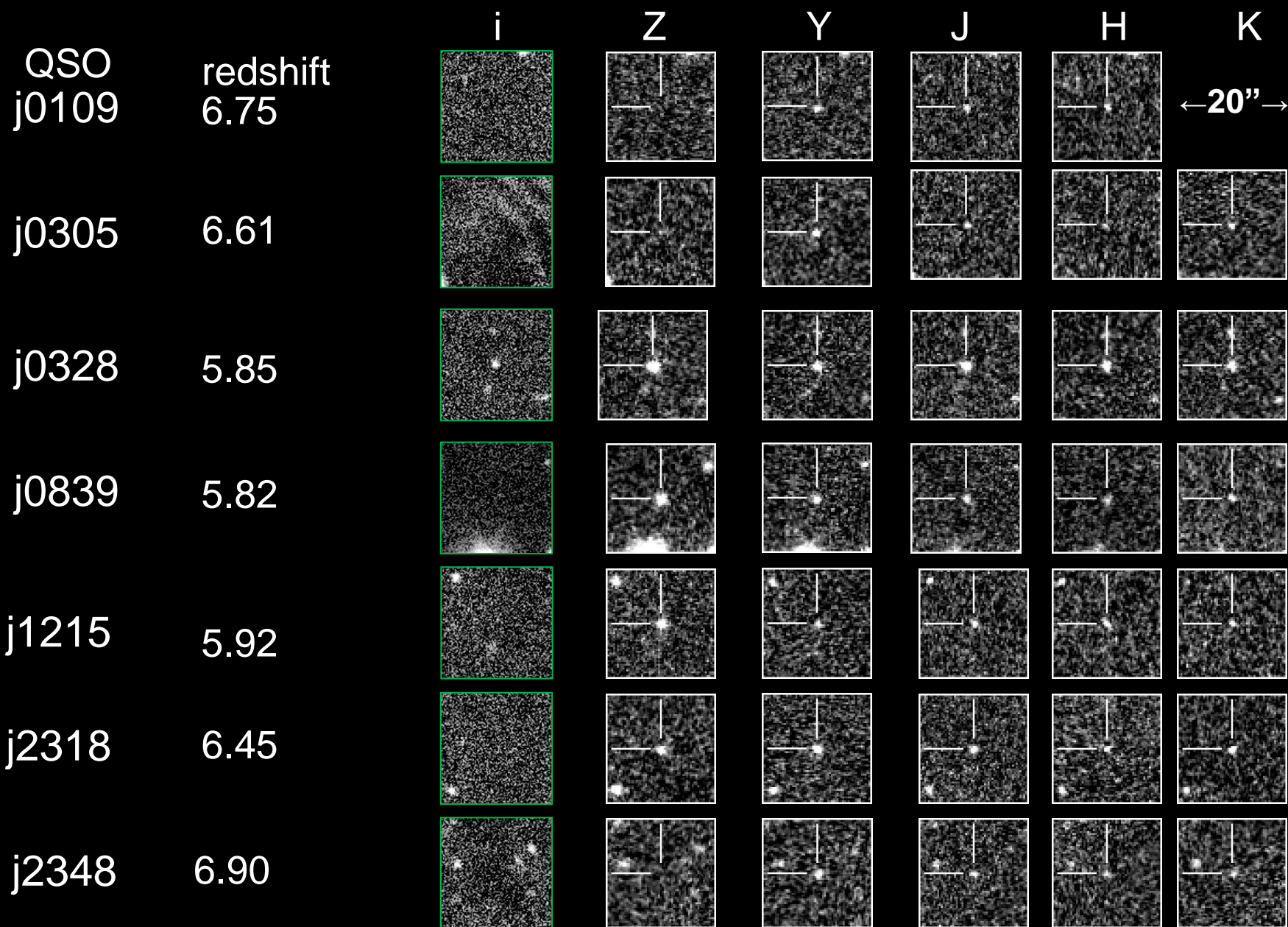
## living archives – data production

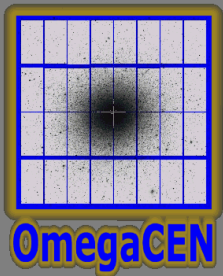


- Modeling dependencies
  - 64 bit identifier
  - Terabytes of pointers
  - Extreme datalineage ++
  - Distributed-access-process-calibrate- QC -analysis
  - Distributed- Kids etc
- 11 nodes - 260 users  
2Pb, 45M file  
10 Giga table entries



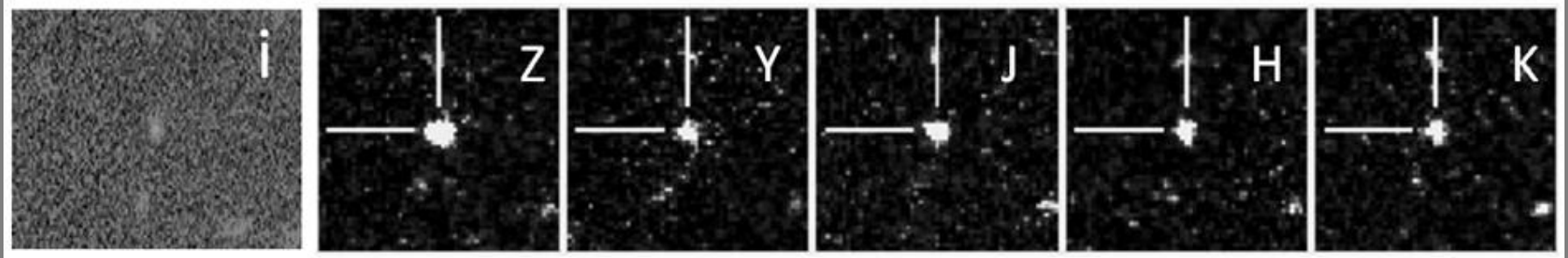
# High redshift QSOs



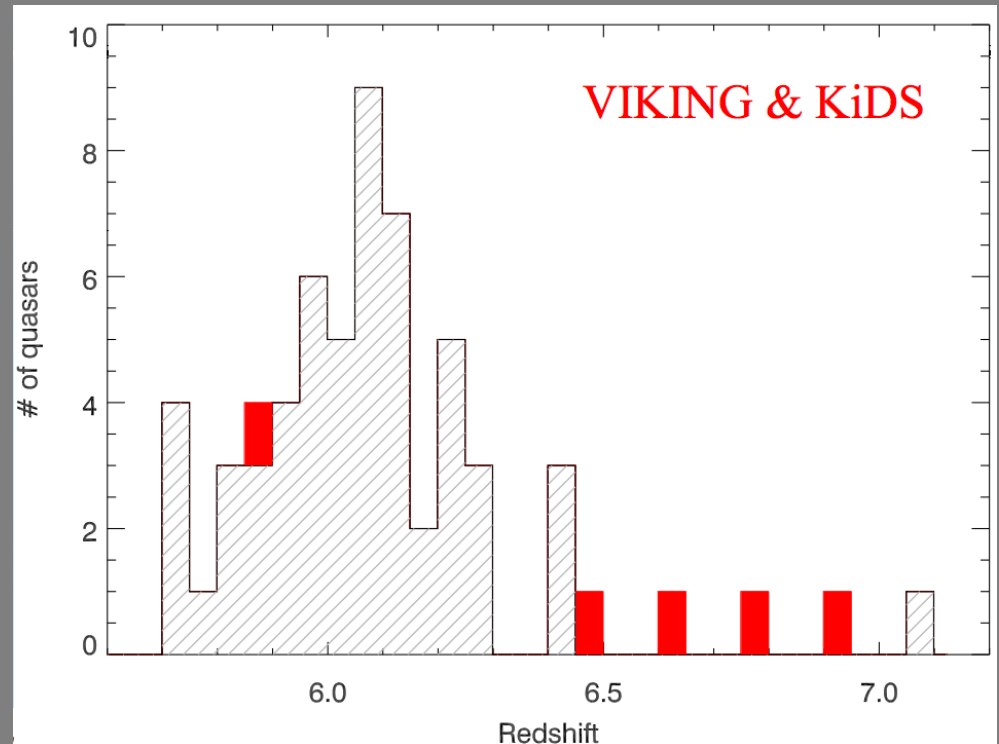
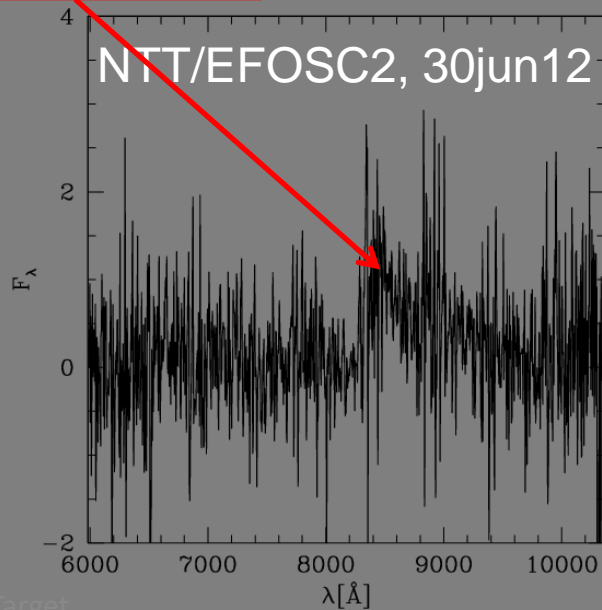


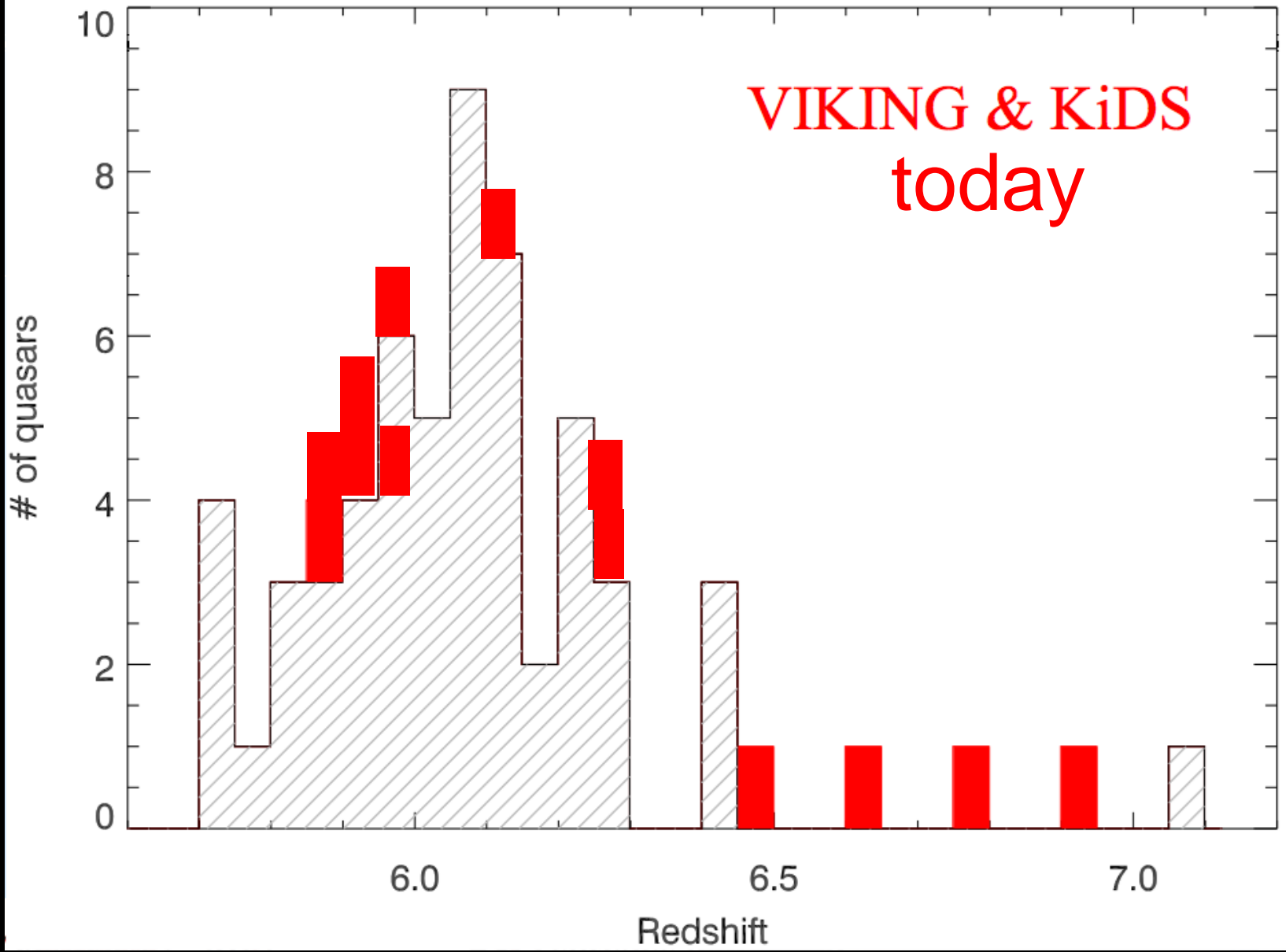
# Most distant quasars in the universe

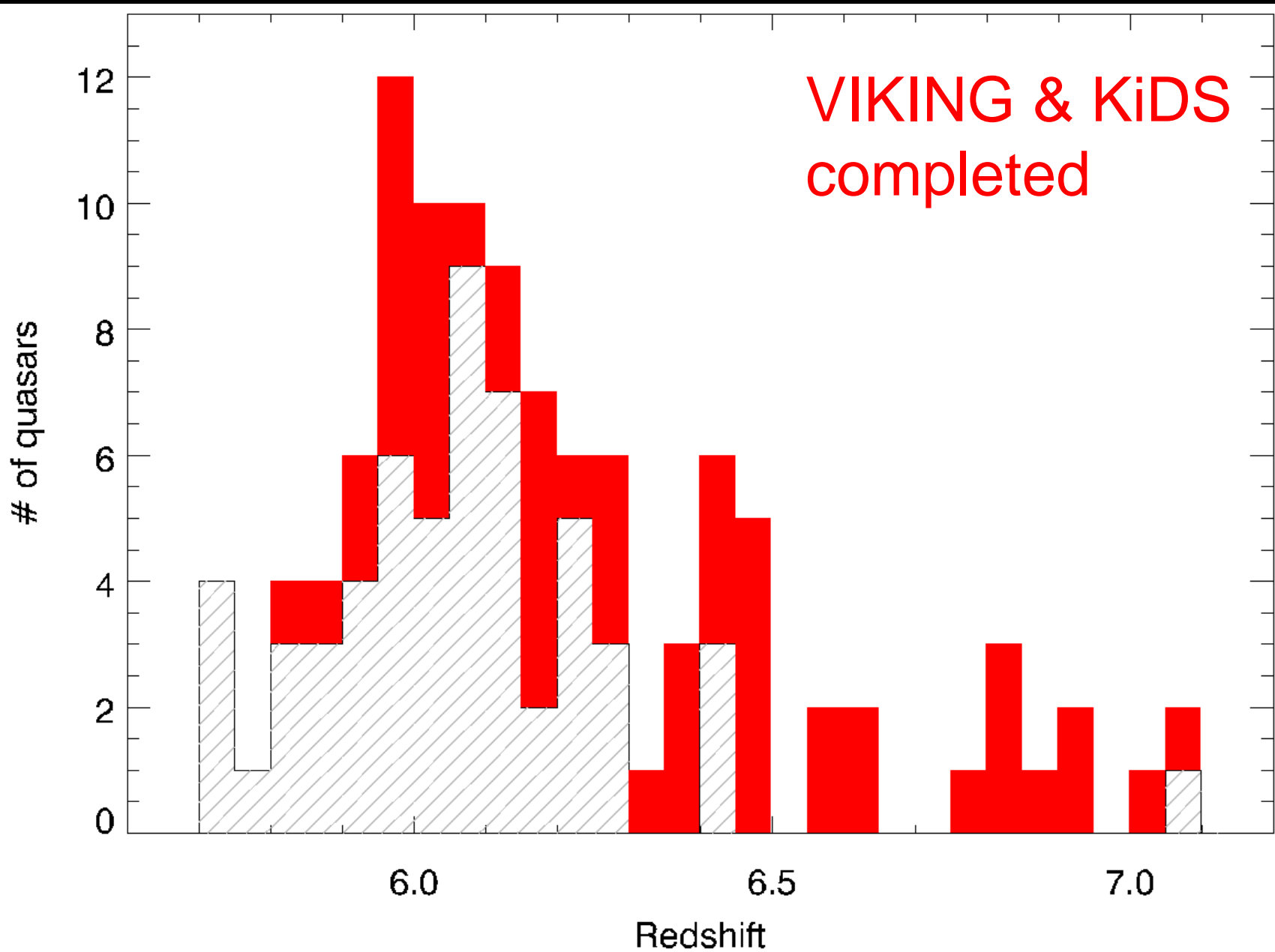
KiDS QSO at  $z \sim 5.8$  in 5 months



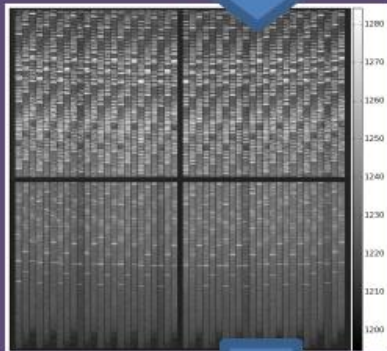
Ly-break+alpha



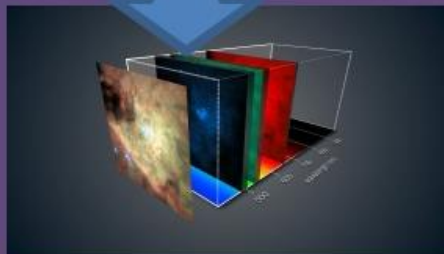








Raw 2D detector pixels



Calibrated 3D pixel-cubes

## MUSE-WISE Information system

- 7 Nodes: Goettingen, Groningen, Leiden, Potsdam, Lyon, Toulouse and Zurich
- Each node has a dataserver and DPU
- Consortium: ~80 people
- 2 Databases: production and backup
- QC and QA integrated

Re-use:



## From DRS to WISE

### XML

```
<xml>  
<frame tag="MASTER_BIAS">  
<parameter name="Observer">
```

Generate Python from XML

### Python

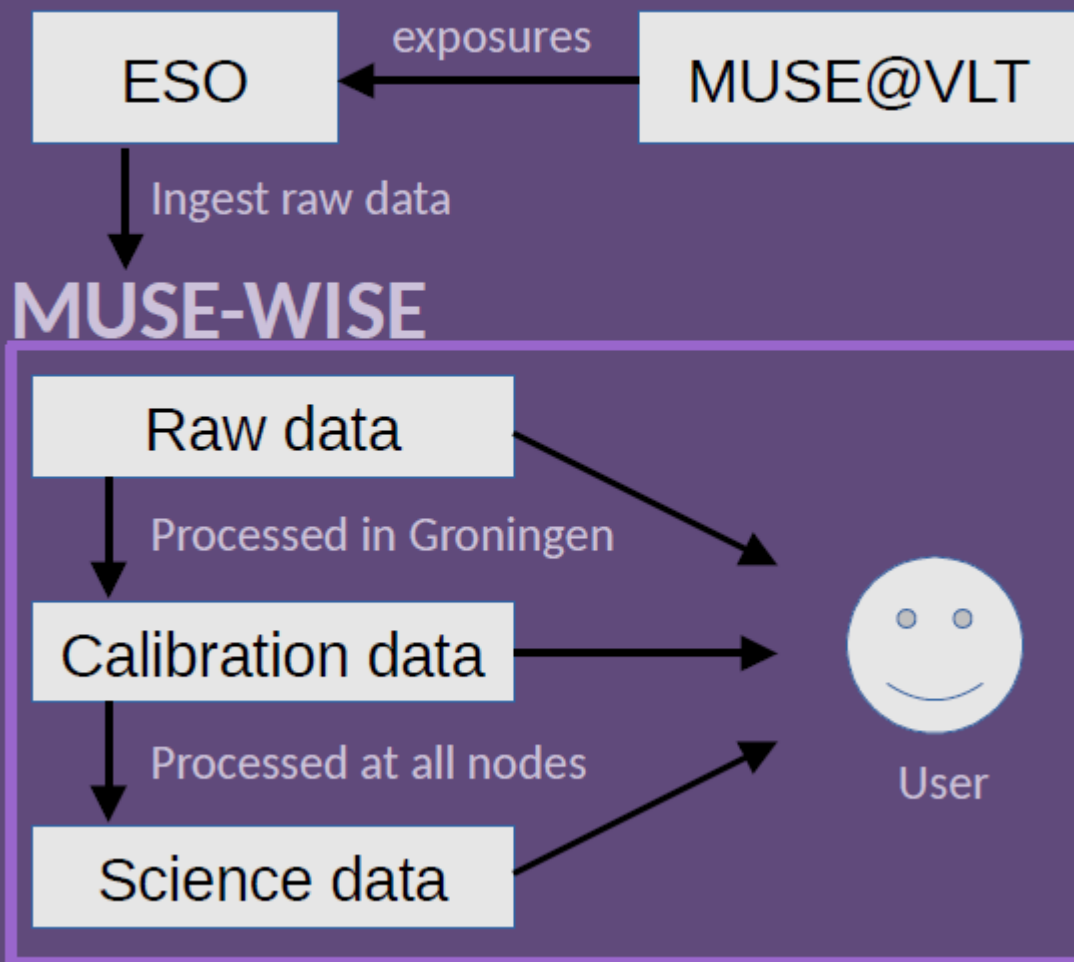
```
Class MASTER_BIAS(DataObject):  
    Observer = persistent(str)
```

Generate Database from Python

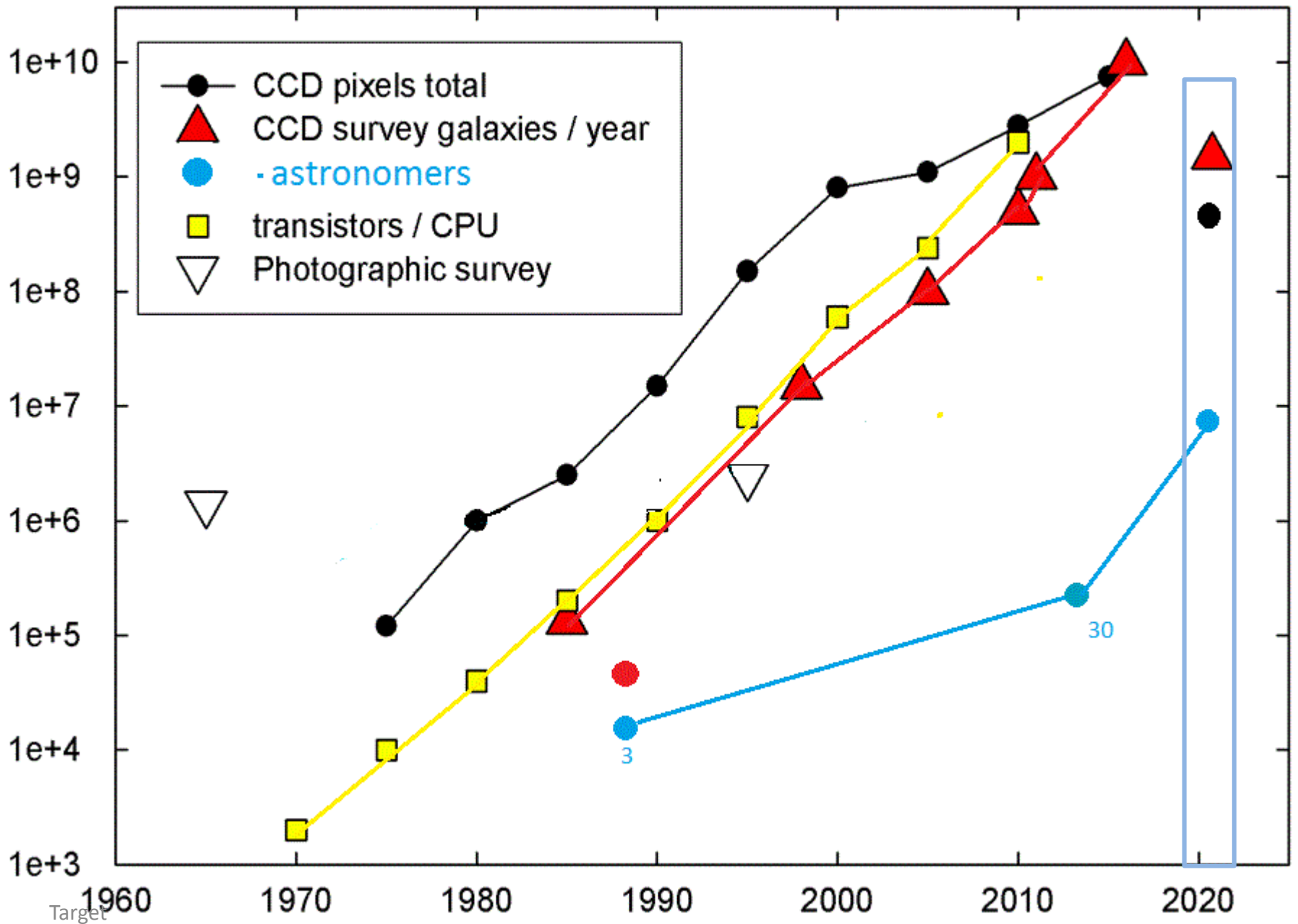
### Database

```
TABLE MASTER_BIAS  
Column ID  
Column Observer  
Column Filename
```

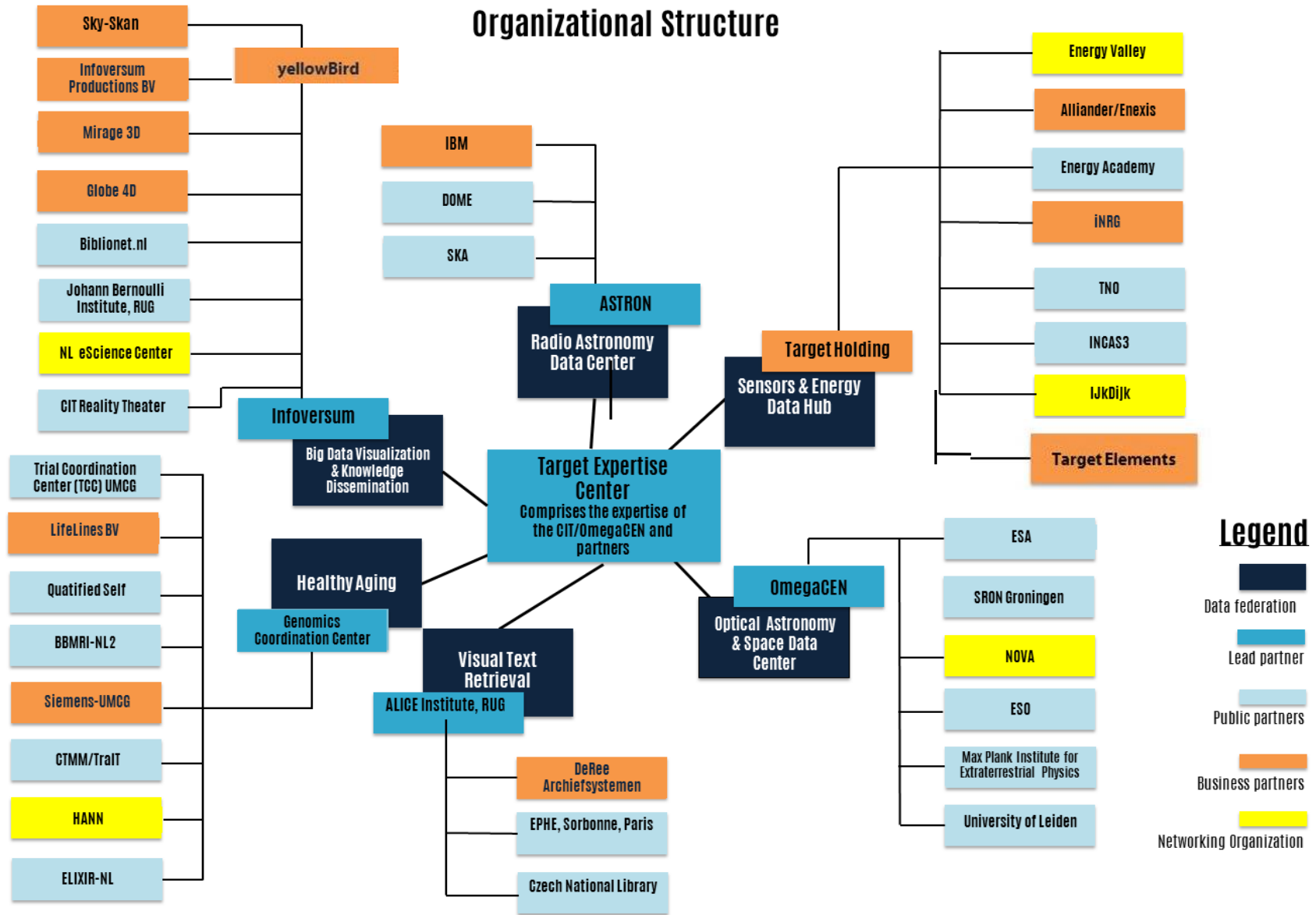
## Data-flow



# Trends in Optical Astronomy Survey Data

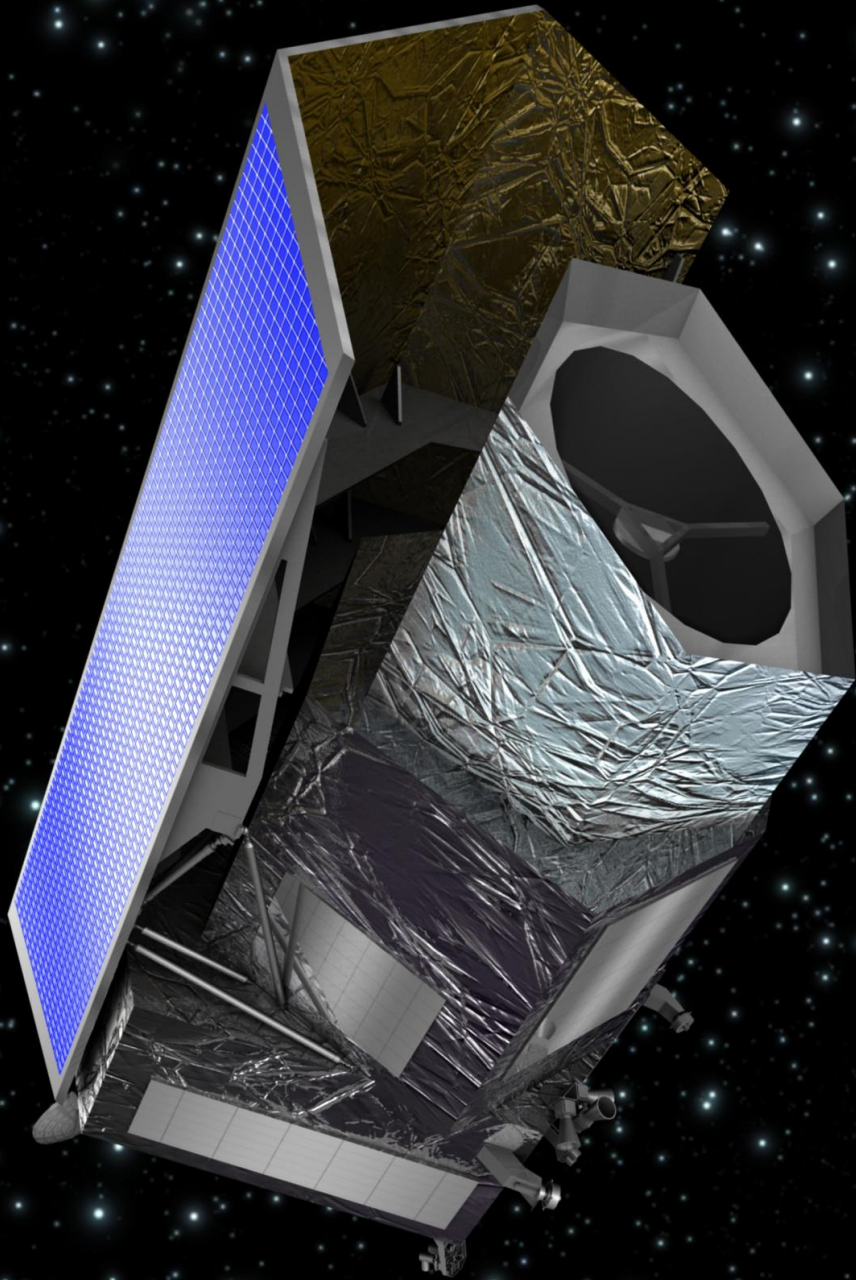


# Data federations



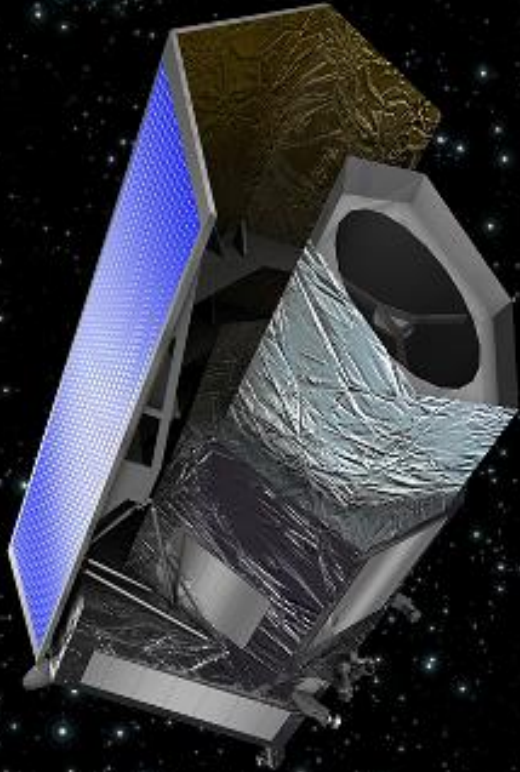


# Euclid



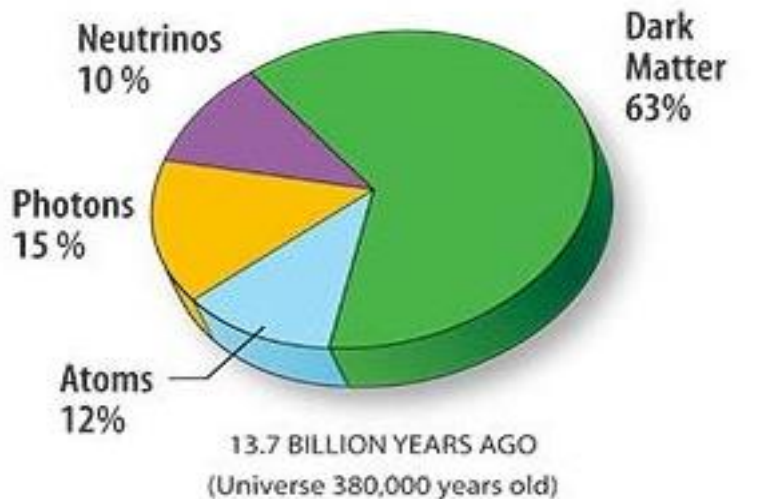
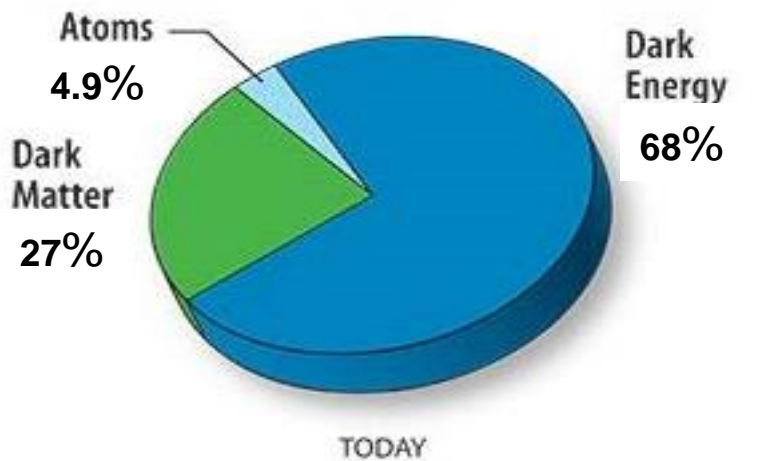


# Euclid ESA

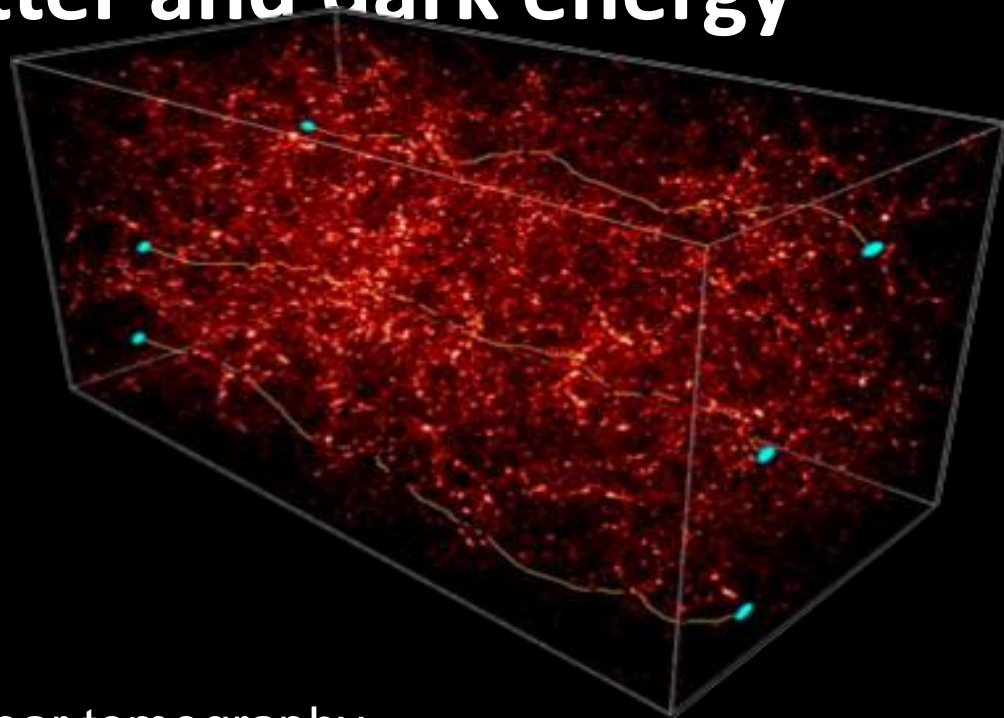


- Due for launch in 2020
- 1300 registered users in  
130 institutes 13 countries
- Euclid Archive System (EAS) is data  
centric information system with  
many of the WISE concepts
- EAS prototype uses Astro-WISE code  
hosted in the Euclid SDC-NL in  
Groningen

# Weak gravitational lensing as probe of dark matter and dark energy



Target



Shear tomography

light rays deflected  $\sim 3'$

by Large Scale Structure

gravitational shear as function of source redshift  
tomographically reconstruct 3D gravity field

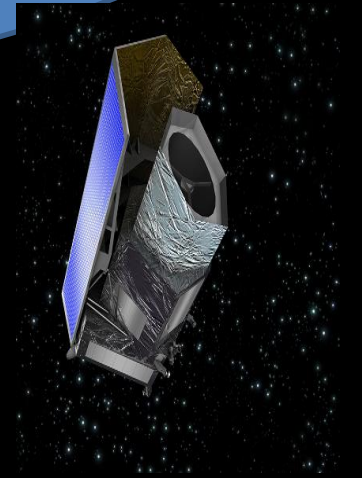
KiDS:  $< 100 \cdot 10^6$  redshifts

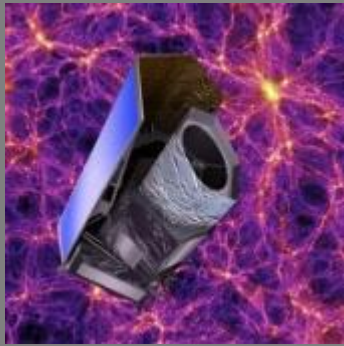
EUCLID:  $1 \cdot 10^9$  redshifts

# Distributed communities access-process-calibrate-analyse publish

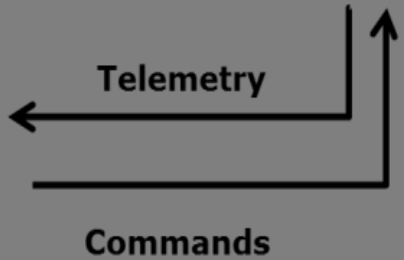
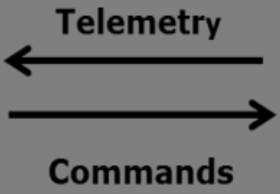
## Euclid:

- 1300 registered members and growing
- 800 researchers
- 130 laboratories/departments
- 13 European countries contributing: Austria, Denmark, France, Finland, Germany, Italy, Netherlands, Norway, Spain, Switzerland, Portugal, Romania, UK + USA IPAC
- NASA/US: provides the IR detectors.





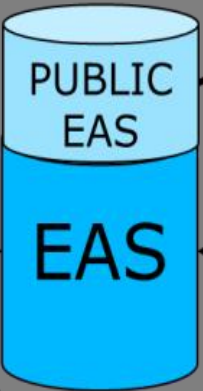
Mission Operations Centre



Observation planning  
Instrument commands

Science Operations Centre

Raw TM  
Level 1 data

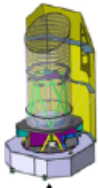
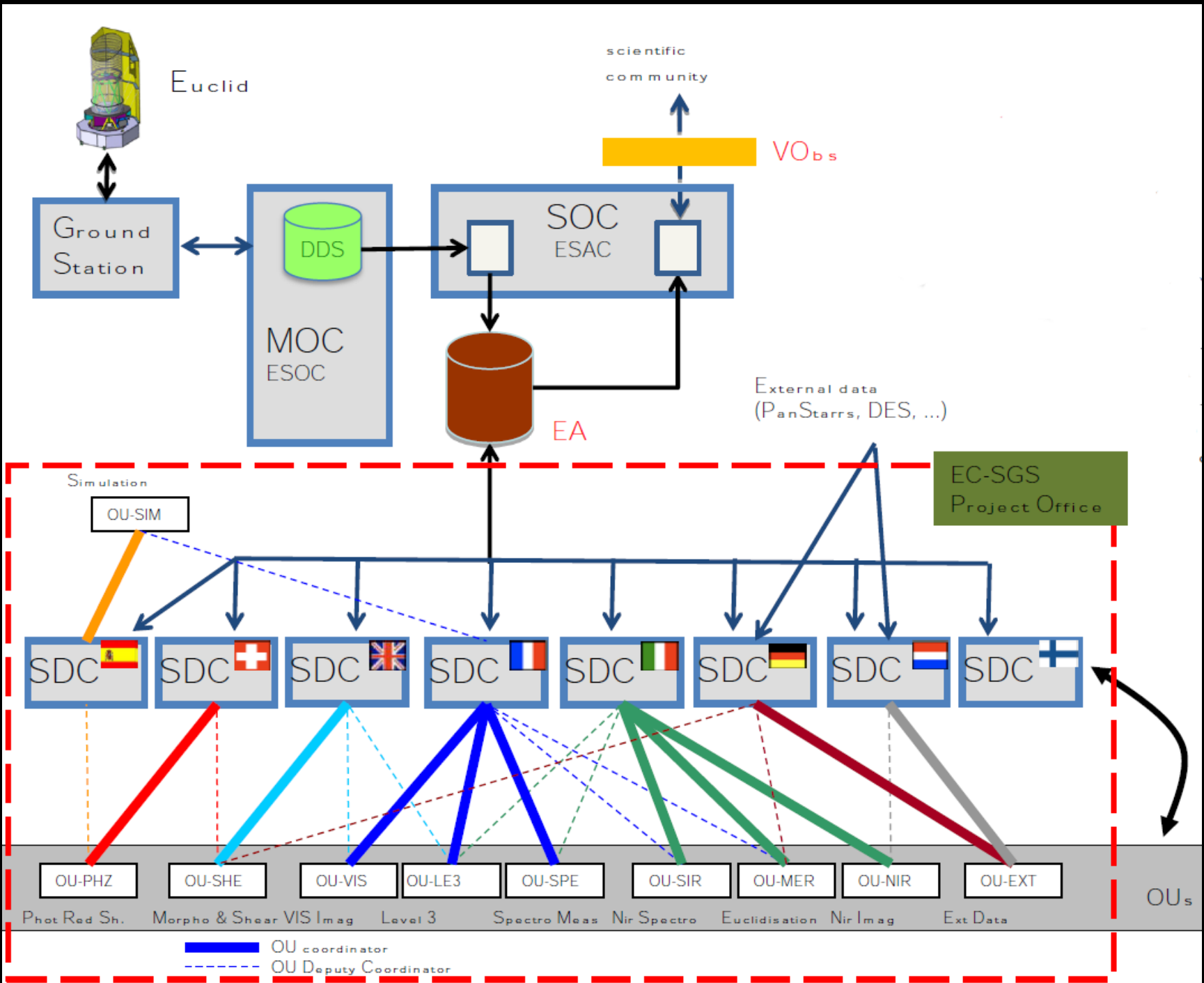


Data products

Euclid Consortium  
Science Ground  
Segment

Data products to external  
community





Euclid

scientific community

VO<sub>bs</sub>

Ground Station

DDS

SOC ESAC

MOC ESOC

EA

External data (PanStarrs, DES, ...)

Simulation

OU-SIM

EC-SGS Project Office

SDC

SDC

SDC

SDC

SDC

SDC

SDC

SDC

OU-PHZ

OU-SHE

OU-VIS

OU-LE3

OU-SPE

OU-SIR

OU-MER

OU-NIR

OU-EXT

OU<sub>s</sub>

Phot Red Sh.

Morpho & Shear VIS Imag

Level 3

Spectro Meas

Nir Spectro

Euclidisation

Nir Imag

Ext Data

— OU coordinator

- - - OU Deputy Coordinator

# Data federations ingredients

Project management – Sociology (– GAIA !)

Standards                  Data model

Compute server      data server                  database

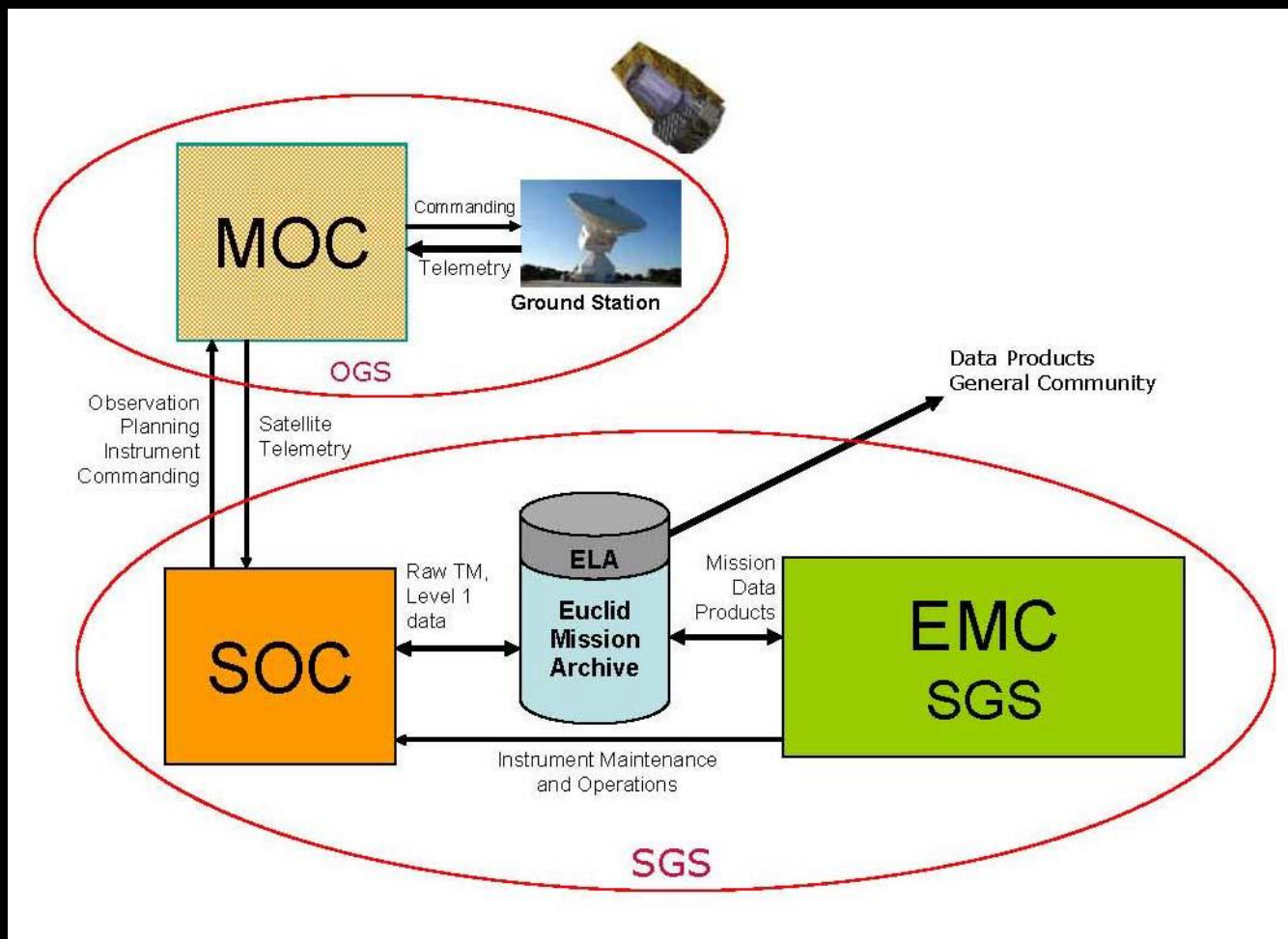
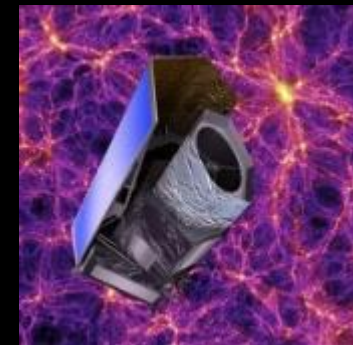
+ Various abstraction layers

- ✓ zoo of hardware / hardware operations
- ✓ datacentric
- ✓ full reproducibility – lineage
- ✓ QC – re-processing

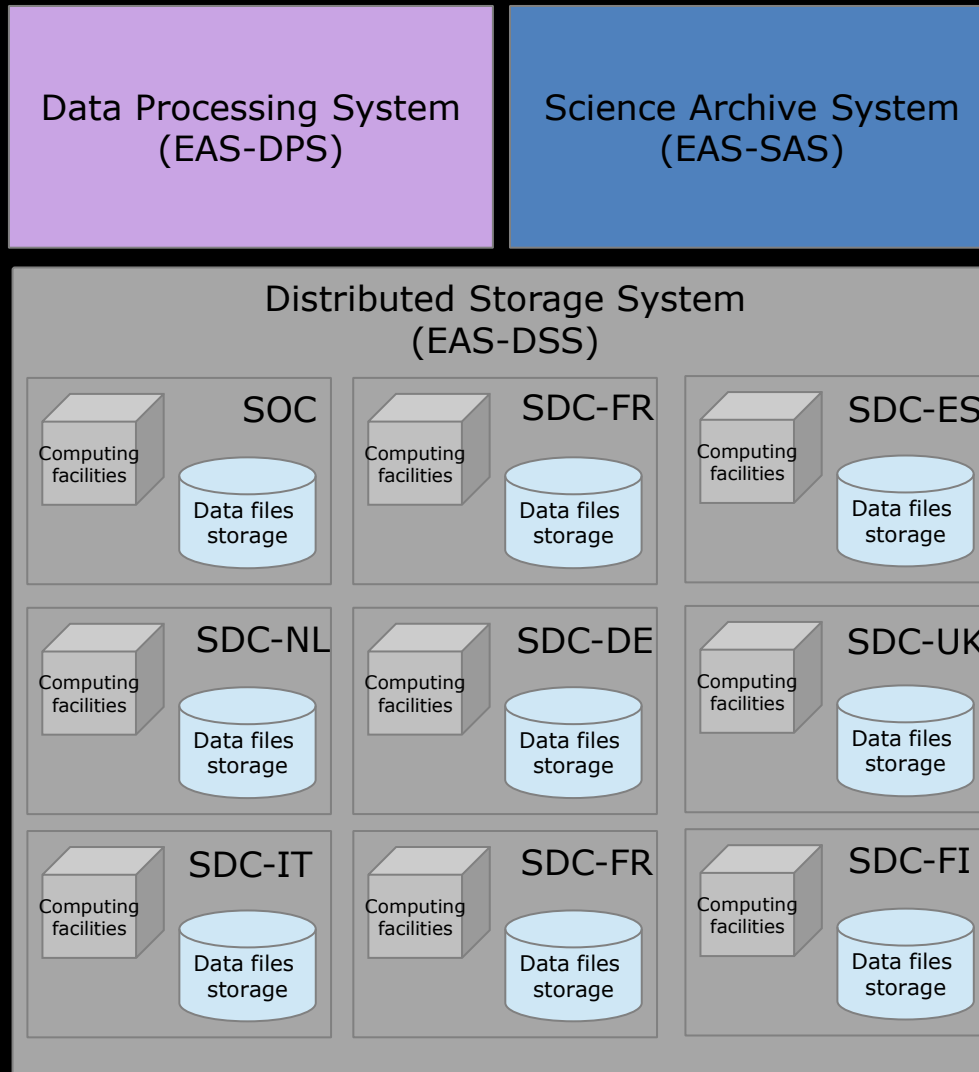
# Euclid Archive System

## Target – ESA

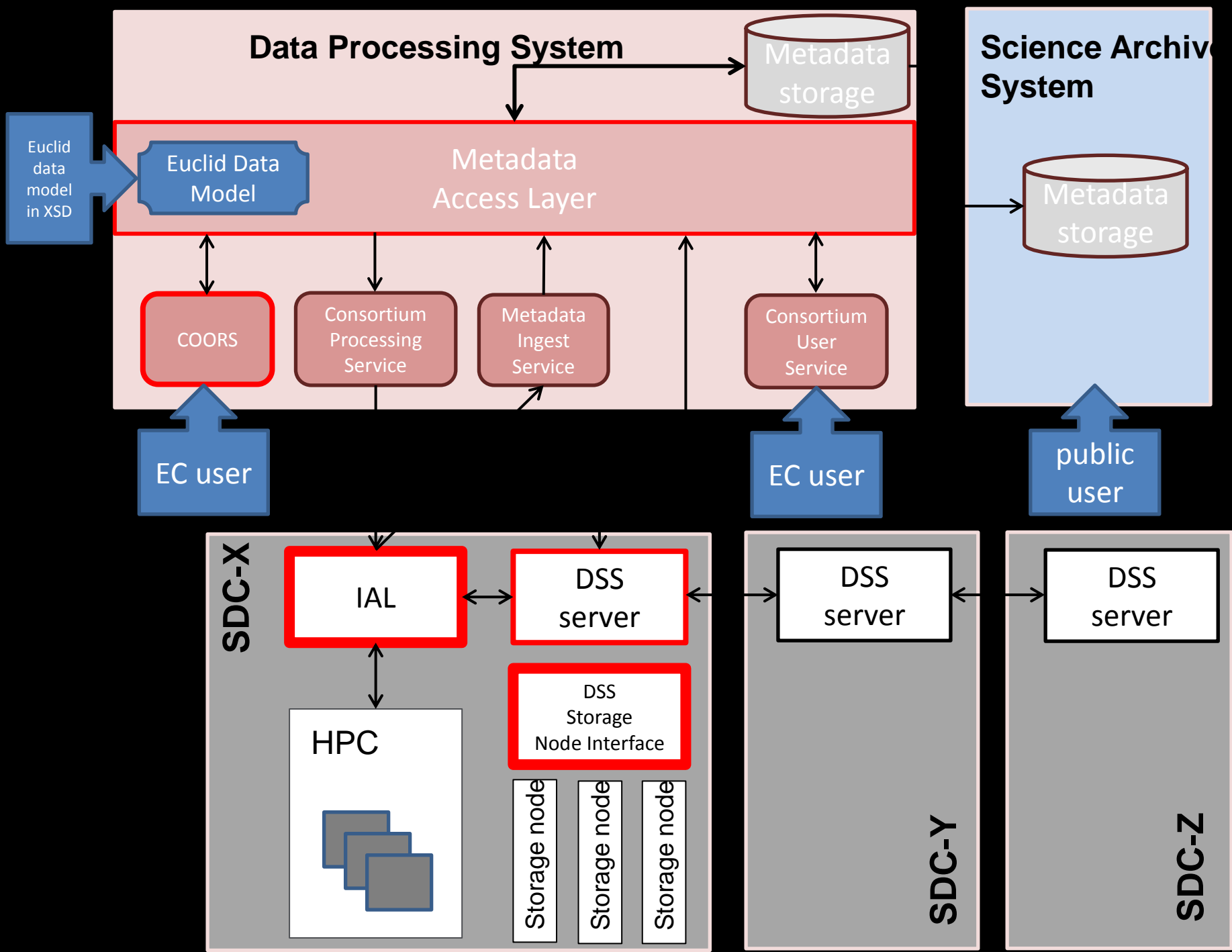
### data centric approach

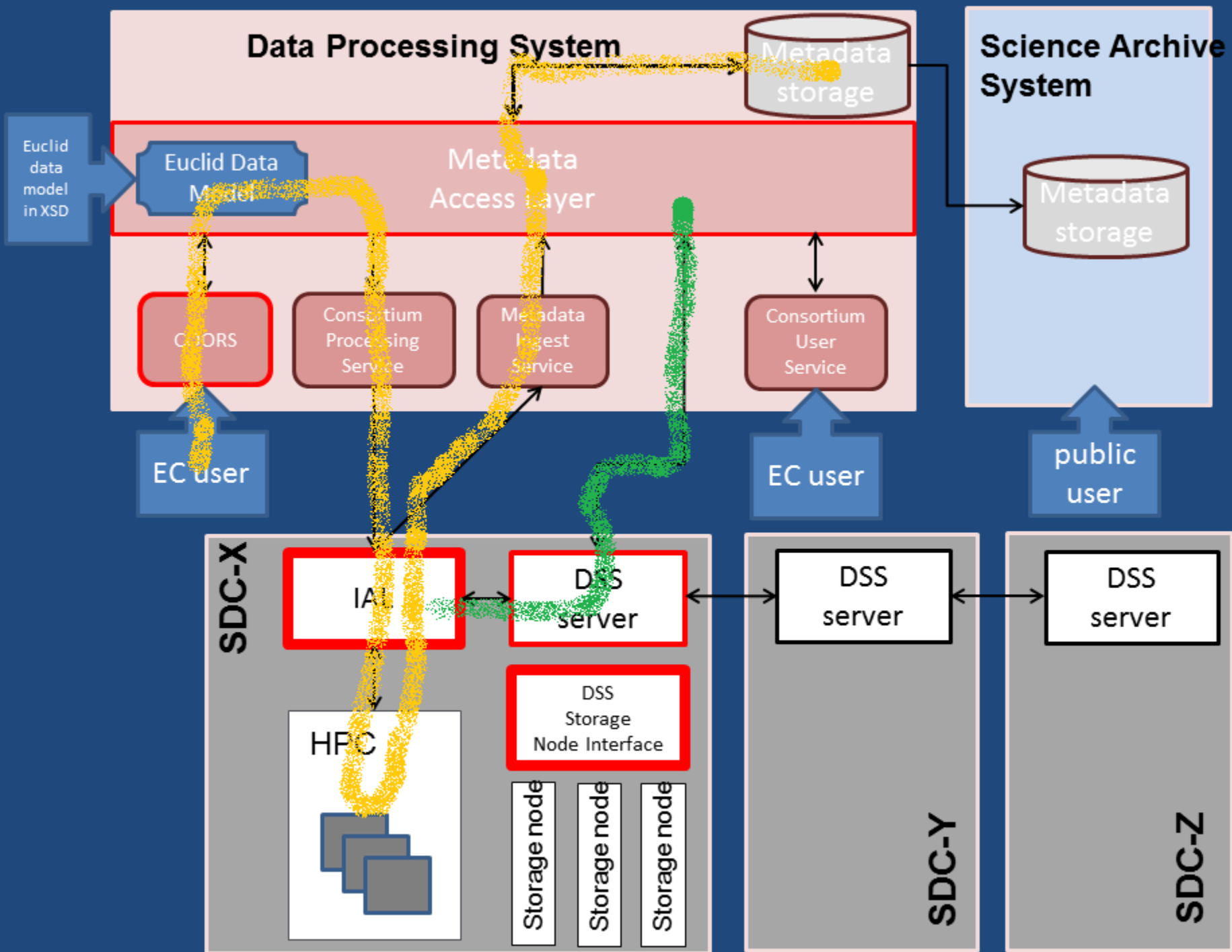


# EAS Design

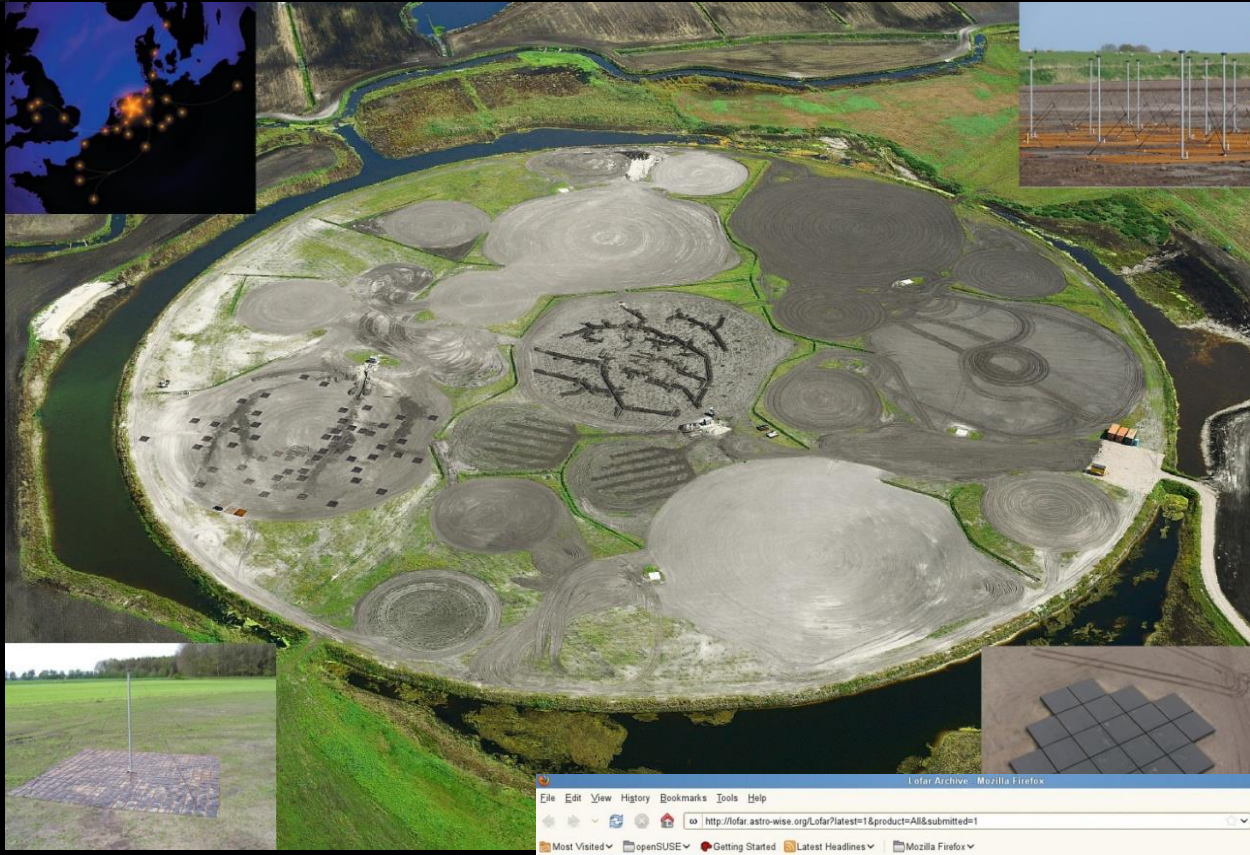








# Astro-WISE – LOFAR LTA



LoFar Archive - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://lofar.astro-wise.org/LoFar/latest=1&product=All&submitted=1

Most Visited openSUSE Getting Started Latest Headlines Mozilla Firefox

LoFar Archive

Home Help login (anonymous) project (test-lofar) Search Show Latest

16 results for UVMeasurement

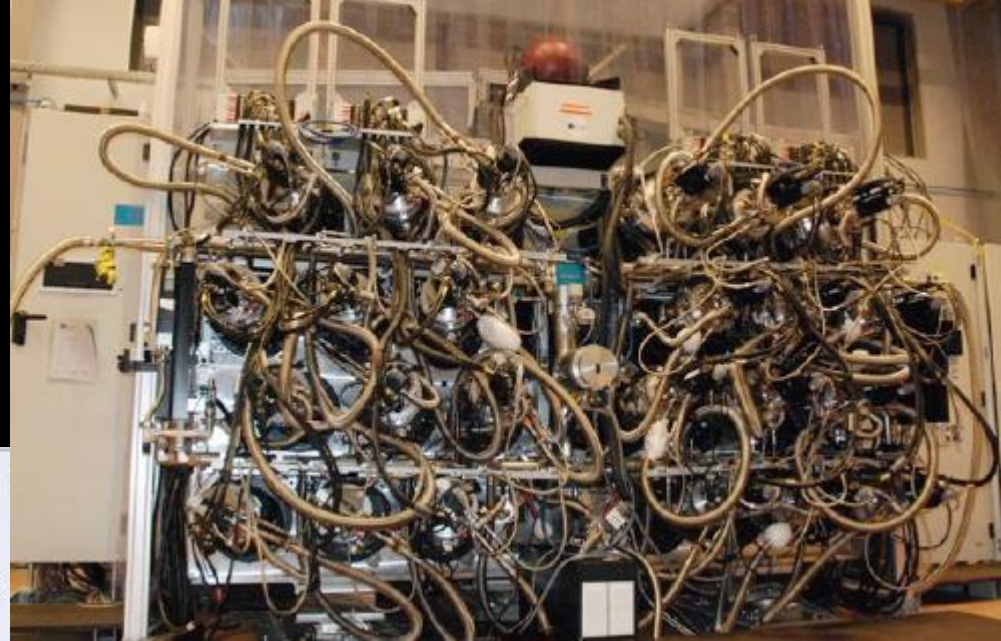
Project	antennaSelection	creationDate	instrumentFilter	measurementIdentifier	measurementType	startT
1	test-lofar	HBA Both	2010-04-06 12:38:53	170-230 MHz	10162	Target
2	test-lofar	HBA Both	2010-04-06 10:26:43	170-230 MHz	10199	Target
3	test-lofar	HBA Both	2010-04-02 22:59:47	170-230 MHz	10015	Target
4	test-lofar	HBA Both	2010-04-02 22:51:32	170-230 MHz	10016	Target
5	test-lofar	HBA Both	2010-04-02 22:26:19	170-230 MHz	10018	Target
6	test-lofar	HBA Both	2010-04-03 04:02:23	170-230 MHz	10017	Target
7	test-lofar	HBA Both	2010-03-31 13:33:00	170-230 MHz	10014	Target
8	test-lofar	LBA Sparse Even	2010-03-15 16:06:41	10-90 MHz	10009	Target
9	test-lofar	HBA Both	2010-02-12 16:57:20	10-90 MHz	71	Calibration
10	test-lofar	HBA Both	2010-02-12 15:50:19	170-230 MHz	79	Target
11	test-lofar	LBA Outer	2009-11-26 17:34:12	30-80 MHz	450	Target
12	test-lofar	LBA Outer	2009-11-26 16:48:39	30-80 MHz	447	Target
13	test-lofar	LBA Outer	2009-11-26 16:39:13	30-80 MHz	442	Target
14	test-lofar	LBA Outer	2009-11-26 15:31:13	30-80 MHz	438	Target
15	test-lofar	LBA Outer	2009-11-26 15:05:07	30-80 MHz	435	Target
16	test-lofar	LBA Outer	2009-11-26 14:56:53	30-80 MHz	432	Target

5 Petabyte/year



# MuseWISE

ESO VLT  
IFS 24 IFU  
100 Gb/night



[Home](#) | [Contact](#) | [Help](#) | [user awjbrinchmann](#) | [project INM](#) | [Preferences](#) | [Tables](#)

## Welcome to the MuseWise DBView Web Service

The following table lists the versions of MuseWise components

Component	Version
Muse-WISE version	0.03.01
muse2wise version	1.58
QC version	v0.0.1
musep version	0.06.00
CPL version	0.5

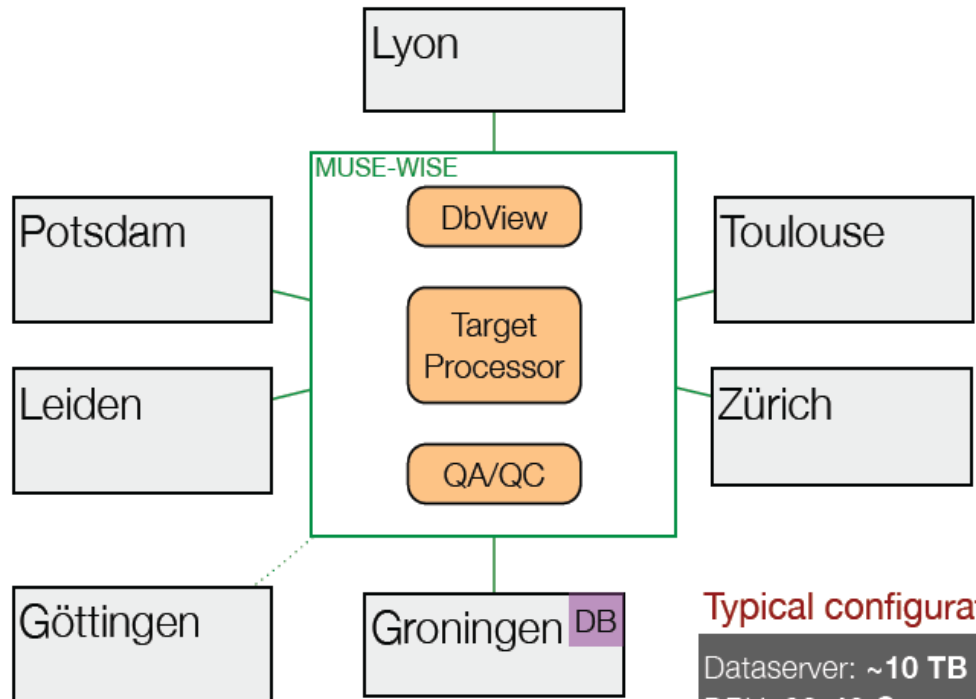
Please choose a table category to start querying

- [All tables](#)
- [Raw Frames](#)
- [External Products](#)
- [Processed Calibration Products](#)
- [Processed Science Products](#)

empowered by



Target



### Typical configuration

Dataserver: ~10 TB  
DPU: 32-48 Cores  
256 GB Mem



# Monk -L. Schomaker 372 books

NI National Archive- Czech, Sorbonne, Harvard, Stanford, Uppsala


The Reading Monk - Mozilla Firefox

Restand Beveiken Beeld Geschiedenis Blgdwijzers Extra Help

http://rugtest5.service.rug.nl:9001/Monk

WebServices < Target < Rijksunivers... Astro-WISE Processing Book flights The Reading Monk

MONK



amsterdam Submit

There are 152 results. Showing results 1 - 10 for amsterdam

[show paragraph](#) [show page](#) [query image](#)

*onderwijzer J. G. Blijm te Amsterdam  
Bestand fia  
Vervolg op bladz 64.*

onderwijzer J.G. Blijm te Amsterdam \_ \_ Bestand fia \_ \_ Vervolg op bladz 64.

[show paragraph](#) [show page](#) [query image](#)

*Sociaal Etablissement te Amsterdam  
Huis No. 12 op de Hoorn*

Sociaal Etablissement te Amsterdam

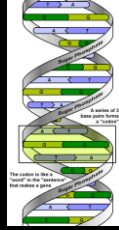
[show paragraph](#) [show page](#) [query image](#)

*Het Sijver te Amsterdam tot het keuren van  
Hout*

# Lifelines – Healthcare 165.000 persons- 30 years LifeWise



Molecular



Physiology



Care



GCC

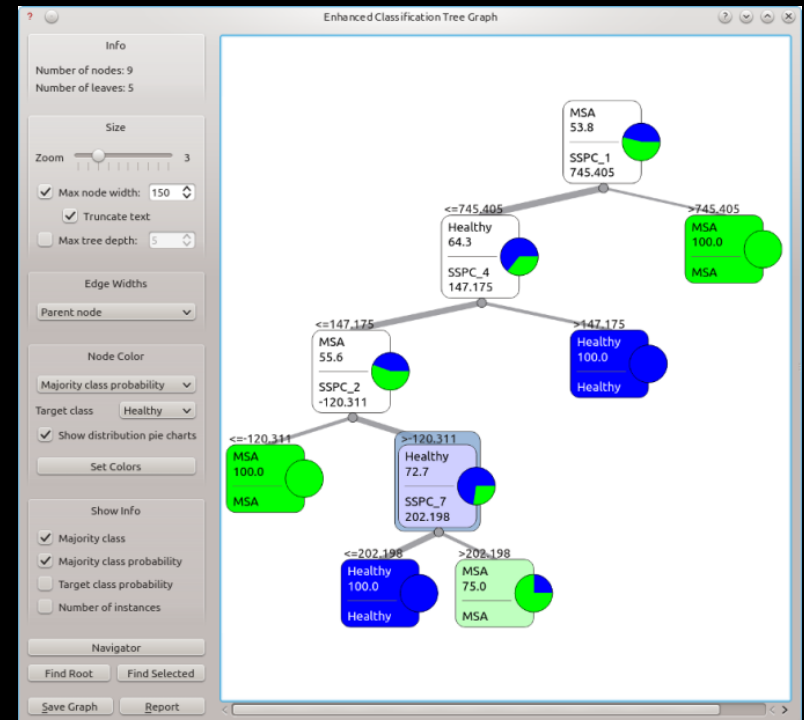
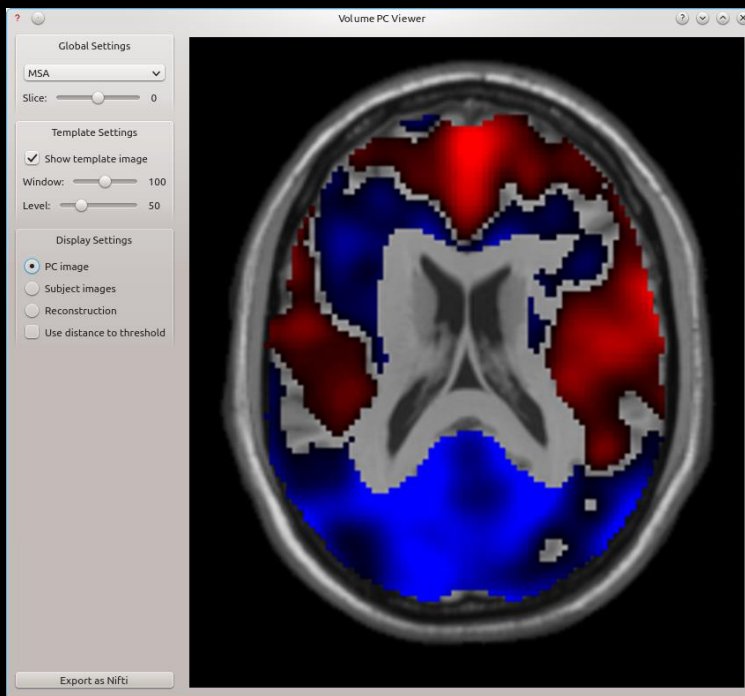
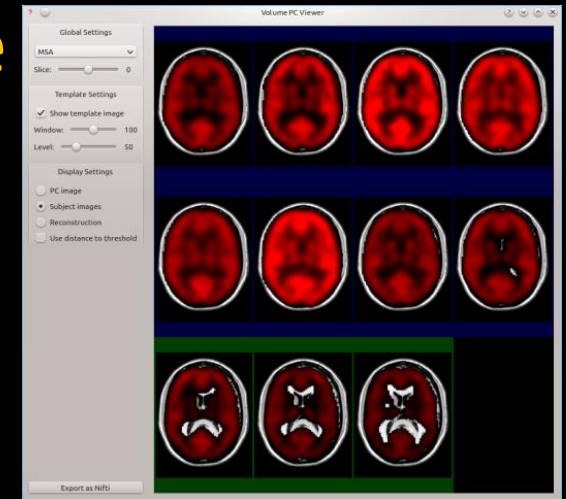
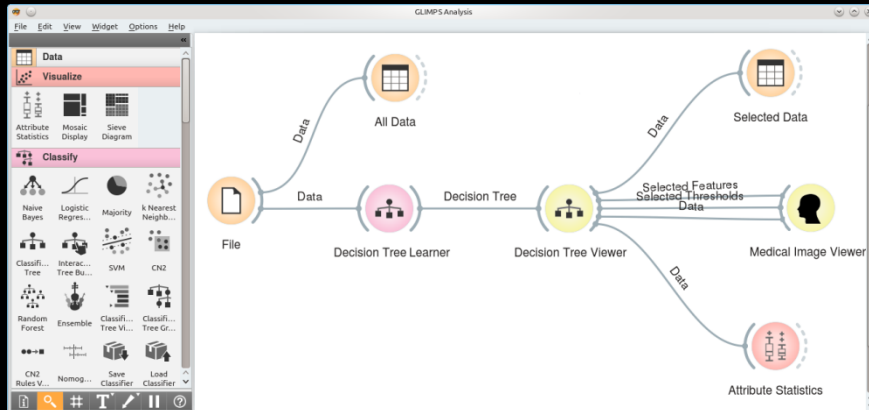


Interconnected



# GLIMPS - Visualization of decision trees

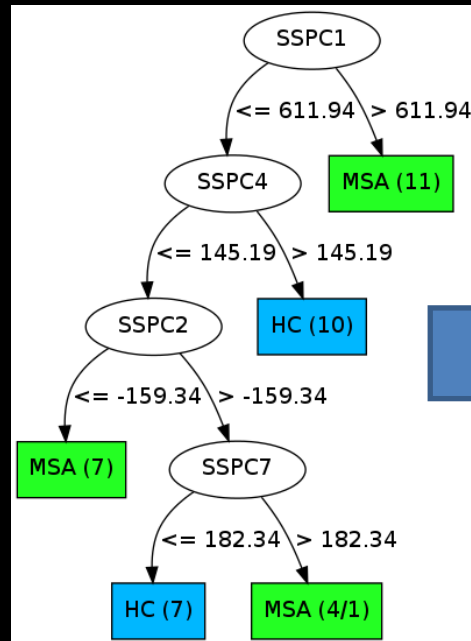
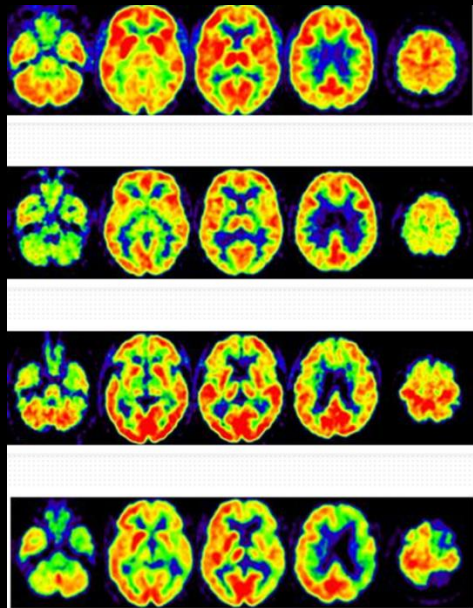
## Orange - Interactive



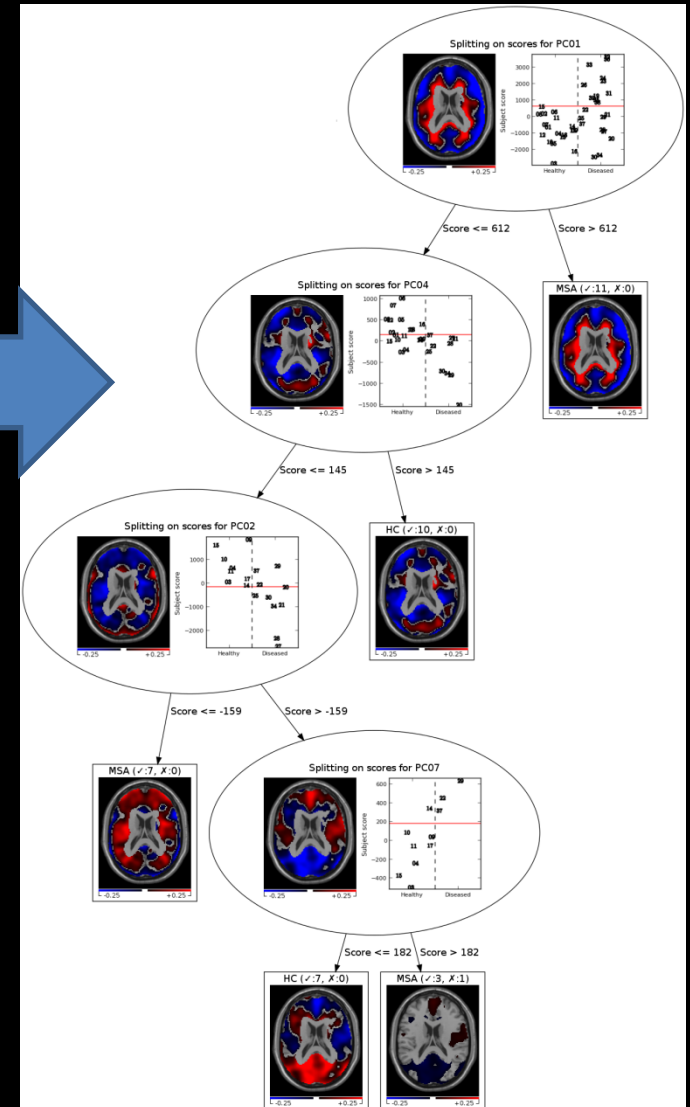
# GLIMPS - Visualization of decision trees

Static

J. Roerdink- D. Williams



- Decision trees are used to classify healthy vs. diseased subjects.
- The tree is augmented with scatter plots and thumbnails to better understand the data.

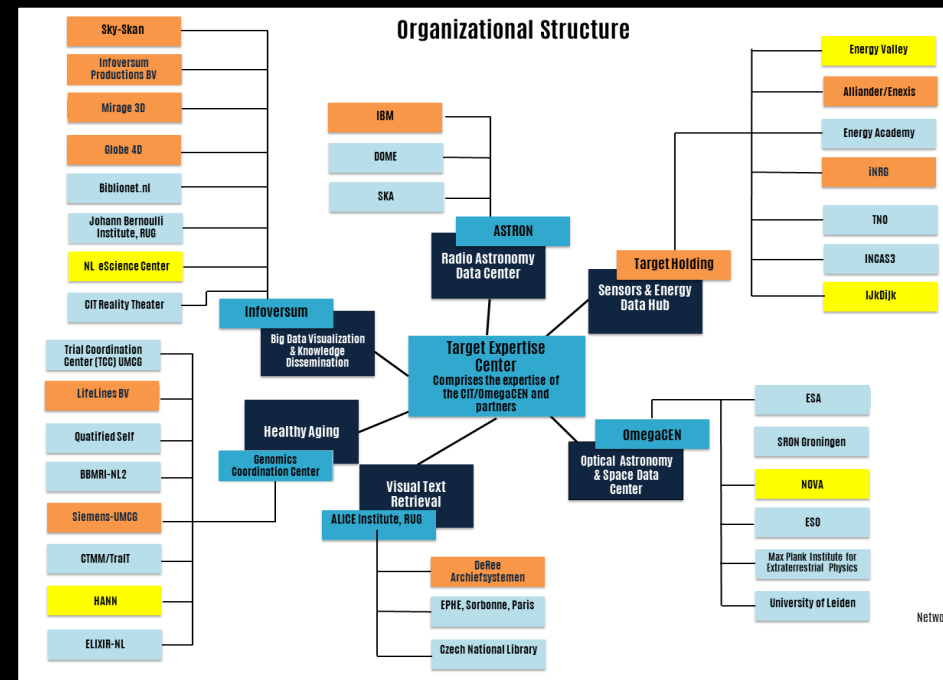




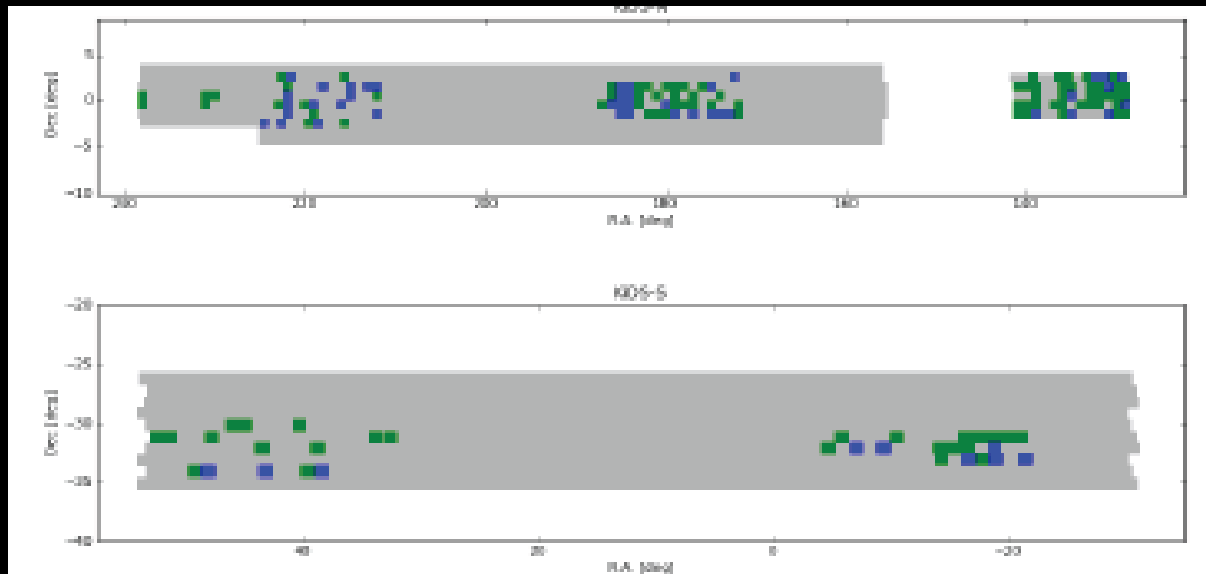
> 2015

# Data federations

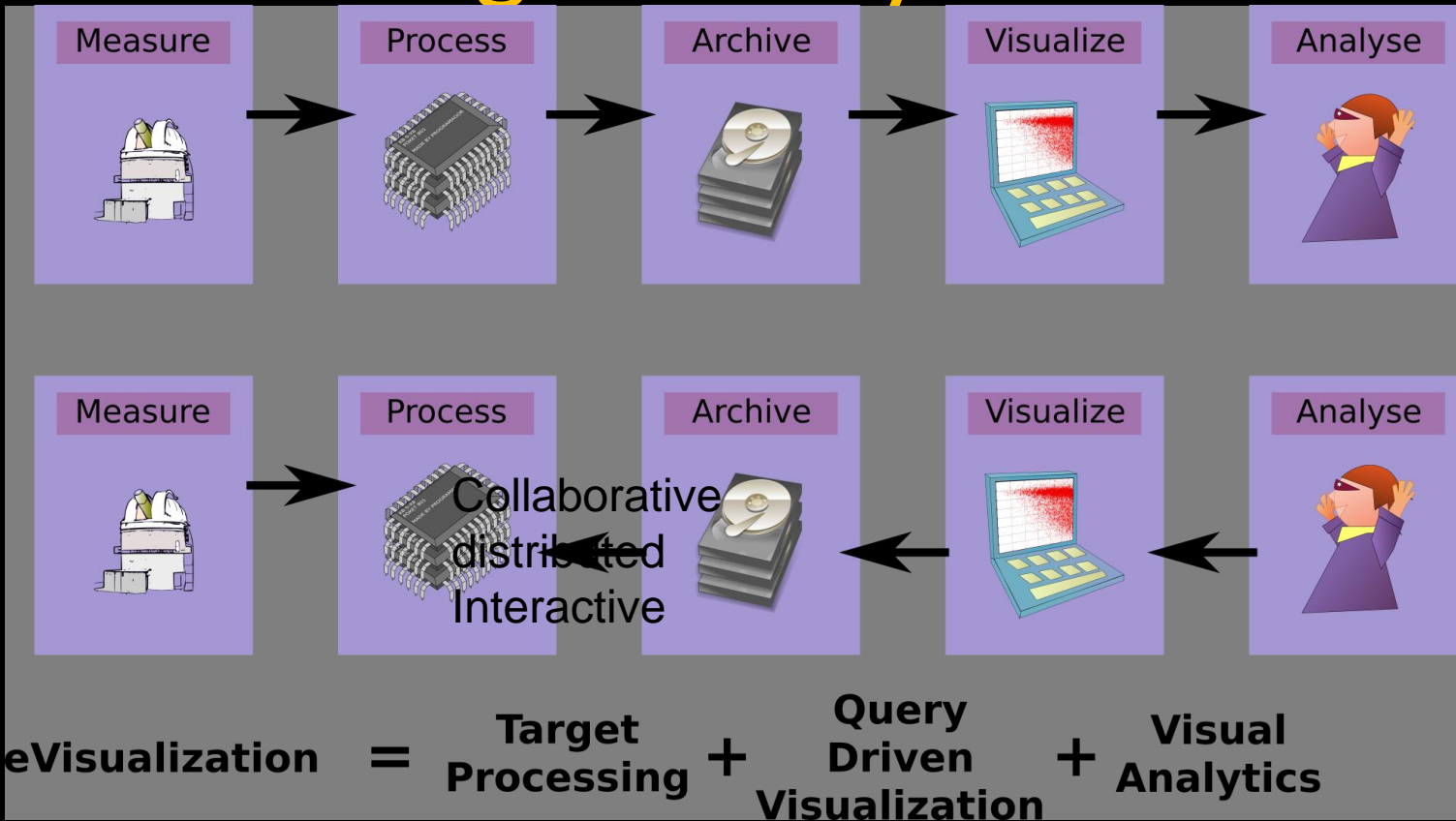
- Standards
- Project management
- Sociology
- Security
- Focus on projects – disciplines - domains
  - Share common infra
  - Share common expertise- datamanagement
- Stay away from “heal the world” type solutions



# KIDS DR1 DR2



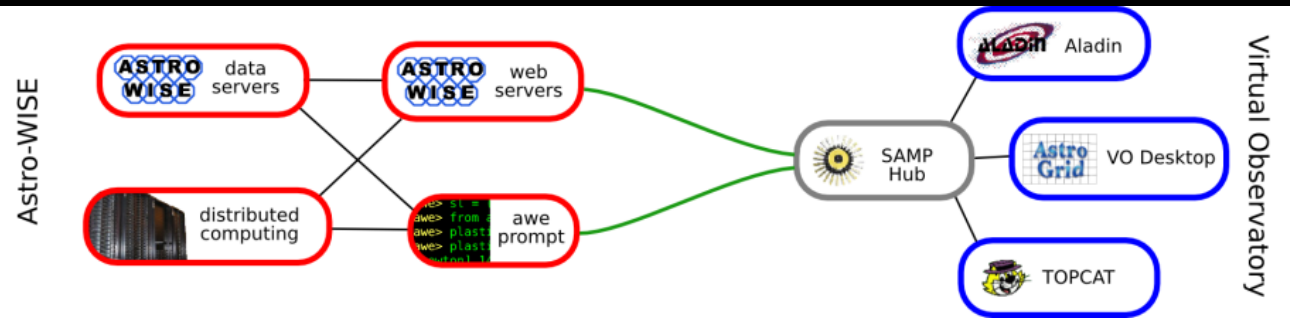
# Data Pulling - Query driven Visualization



Pushing

Pulling

Pushing to  $>10^9$  rows



# Query driven visualisation

**TOPCAT**

File Views Graphics Joins Windows YO Interop Help

Table List  
1: SourceList-135651

Current Table Properties  
 Label: SourceList-135651  
 Location: SL-135651-2df\_R\_17.votable  
 Name:  
 Rows: 4787  
 Columns: 35  
 Sort Order:   
 Row Subset: All   
 Activation Action: (no action)  Broadcast Row

SAMP  
Messages: Clients: AW

60 / 485 M

**Aladin v6.0** \*\*\* PROTOTYPE VERSION (based on v6.000) \*\*\*

File Edit Image Catalog Overlay Tool View Interop Help

Location:  ICRS

file:localhost:net:awton:datas:users:buddelmeijer:each:tmp:Sci-GalKEM

2 superimposed objects

BackGr	ERRA_IMAGE	SID	MAGERR_IS
5.44149848878E-13	8678.15332031	2521	1202.79
3.09843269748E-13	9033.68652344	2550	1760.5
1.46398917295E-11	16169.5732422	2604	853.52
1.96915183812E-13	4169.22070312	2740	1987.10
2.30313733579E-13	7846.08886719	2828	1228.8

TIP: Write a script command directly into the "Command/Location" field

33 sel / 4787 src 112Mb

**Scatter Plot**

File Export Plot Axes Subsets Errors Marker Style Error Style Help

FLUXERR\_ISO x10<sup>-3</sup>

MAG\_ISO

Main  
 Data  
 Table: 1: SourceList-135651  
 X Axis: MAG\_ISO  Log  Flip  
 Y Axis: FLUXERR\_ISO  Log  Flip  
 Row Subsets  
 All   
 bright

Potential: 4,787 Included: 4,787 Visible: 4,787 Position:

**SAMP Control**

File Connect Help

Sent Messages  
Received Messages

Clients  
 Aladin   
 Hub   
 topcat    
 Astro-WISE

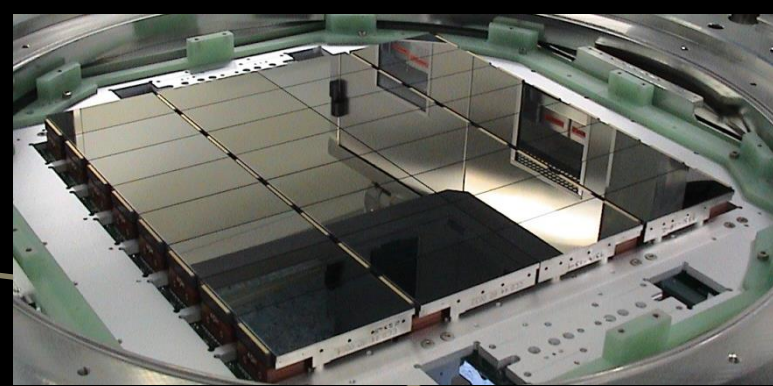
```

Current profile:
- username : AWHBUDELMEIJER
- database : db.astro.rug.astro-wise.org
- project : ALL
- current privileges : 1 (MyDB)

awe> sl = (SourceList.SLID == 135651)[0]
awe> from astro.services.samp.Samp import Samp
awe> samp = Samp()
awe> samp.broadcast(sl.frame)
awe> samp.broadcast(sl)
awe> s.highlighted(sl)
2474
awe>
    
```



# Target



provincie Drenthe



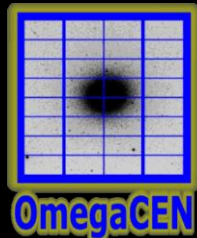
Ministerie van Economische Zaken



Dit project wordt medegefinancierd door het Europees Fonds voor Regionale Ontwikkeling en door het ministerie van EZ, Pieken in de Delta. Het project wordt mede mogelijk gemaakt door de gemeente Groningen, de provincie Groningen, de provincie Drenthe en SNN en staat onder auspiciën van Sensor Universe.

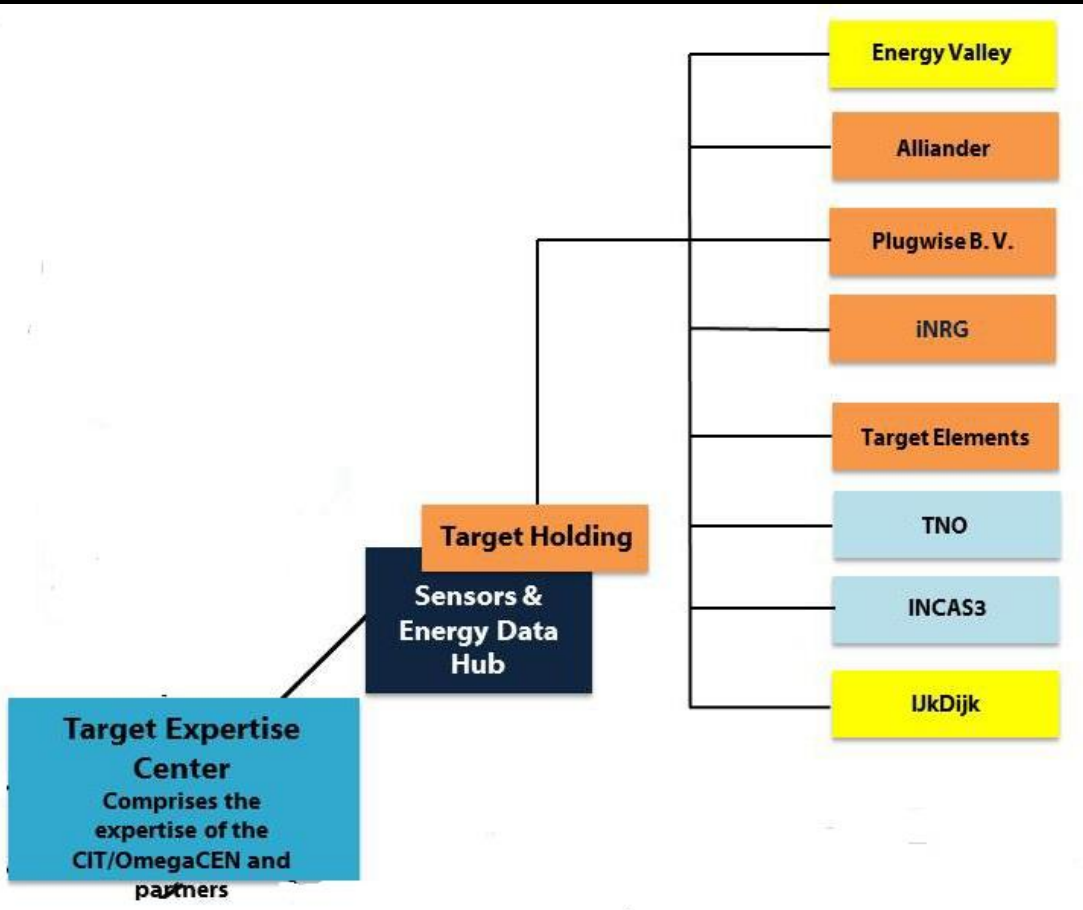
kunstmatige intelligentie

donald smits centrum voor informatie technologie





# Target Holding



SMART HOME



## I-CE Energie Management Systeem

- ✓ Realtime inzicht in je stroom- en gasverbruik
- ✓ Realtime inzicht in de opbrengst van je zonnepanelen
- ✓ Lokale energiedata en optimale privacy

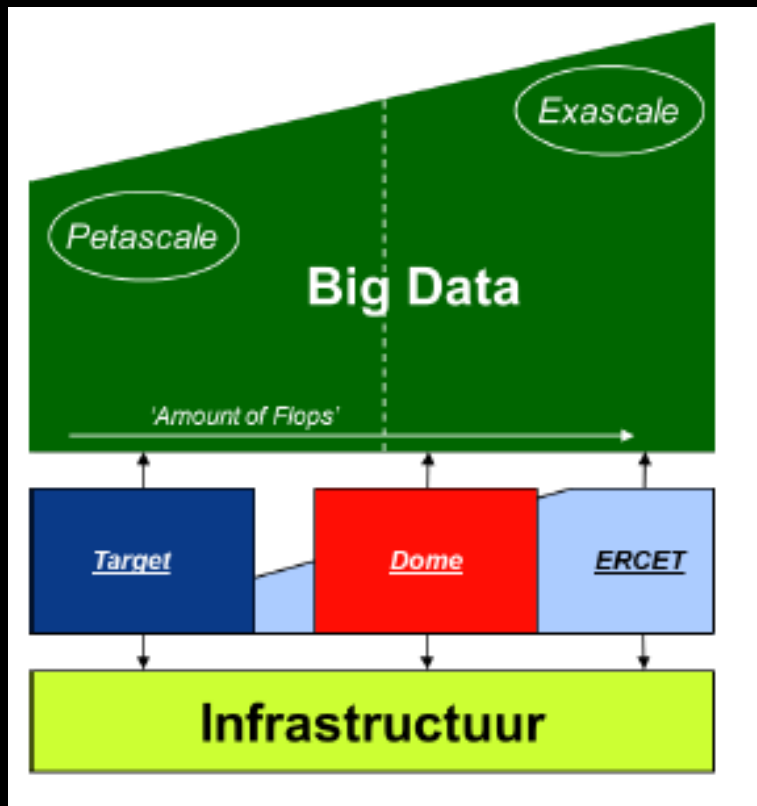
# Belastingdienst – NL tax office

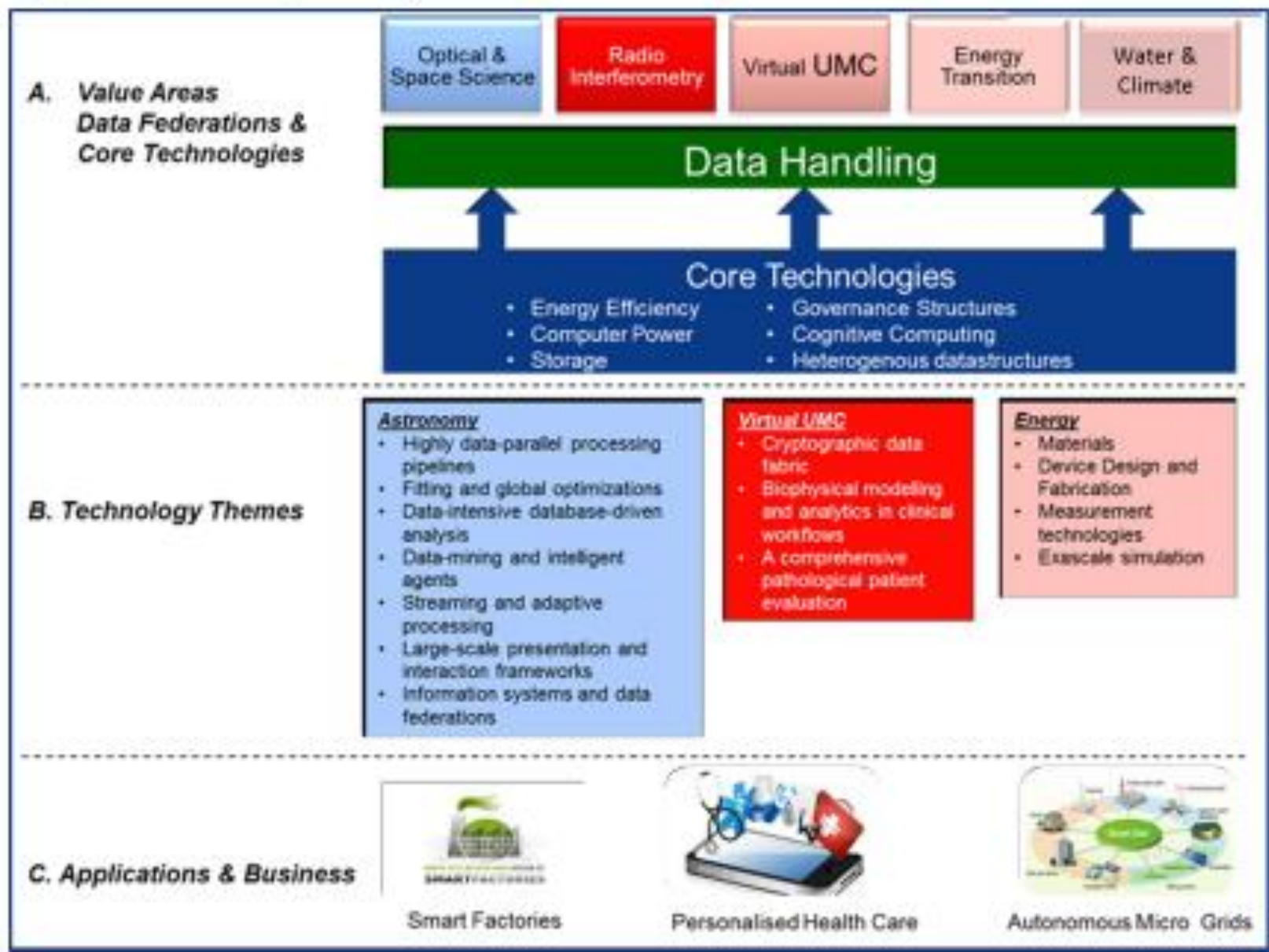
- 250M Euro holy grail system  
ETM system 260 M transactions/year
- federated i/f db

schreef directeur Belastingen Hans Blokpoel dat er in 2010 al “serieuze zorgen” over het systeem waren. Na onderzoek bleek eind 2013 dat het nooit goed zou komen met het systeem. Blokpoel schrijft:

“De hoeveelheid maatwerk die nodig zou zijn om het ETPM in gebruik te nemen, bleek enorm en niet beheersbaar te maken.”

Thank you !







# 'THE INFORMATION UNIVERSE' CONFERENCE

October 7-9, 2015  
Groningen, The Netherlands



[WWW.INFORMATIONUNIVERSE.RUG.NL](http://WWW.INFORMATIONUNIVERSE.RUG.NL)

## Keynote Speakers

**Erik Verlinde**, University of Amsterdam  
**Alex Szalay**, The Johns Hopkins University  
**Gregory Chaitin**, Federal University of Rio de Janeiro  
**Gerard 't Hooft**, University of Utrecht  
**Charley Lineweaver**, Australian National University  
**Lude Franke**, UMCG, University of Groningen

## Topics

- Is the universe one big information processing machine?
- Is there a deeper physical description of the world based on information?
- What is the significance of the Black hole information paradox?
- What is the role of information in highly organized complex life systems?
- Where will quantum information take us?
- What's the distinction between our numerical simulations (in vitro) and real natural system (in vivo)?
- Is information the road to understanding dark matter, dark energy?

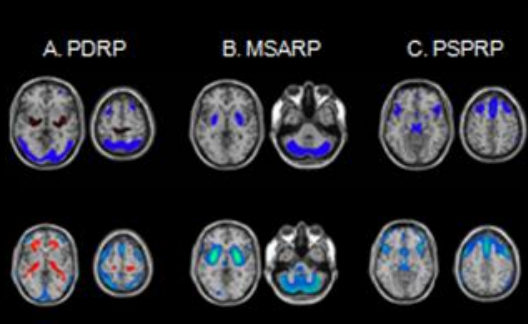
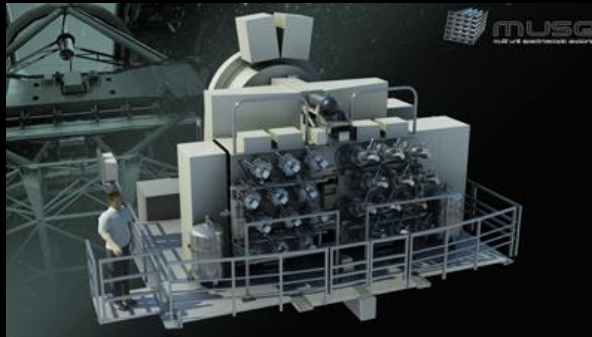
## Invited Speakers

**Tamara Davis**, University of Queensland, Australia  
**Alessandra Silvestri**, Lorentz Institute, Leiden University, Netherlands  
**Paul Vitanyi**, Center for Mathematics & Informatics (CWI), Amsterdam  
**Renate Loll**, IMAPP, Radboud University, Netherlands

# Current Development: Science

WISE concepts have been deployed in a diverse range of scientific projects (astronomy and non-astronomy).

- incorporate new technologies
- new projects provide an engine for improvement

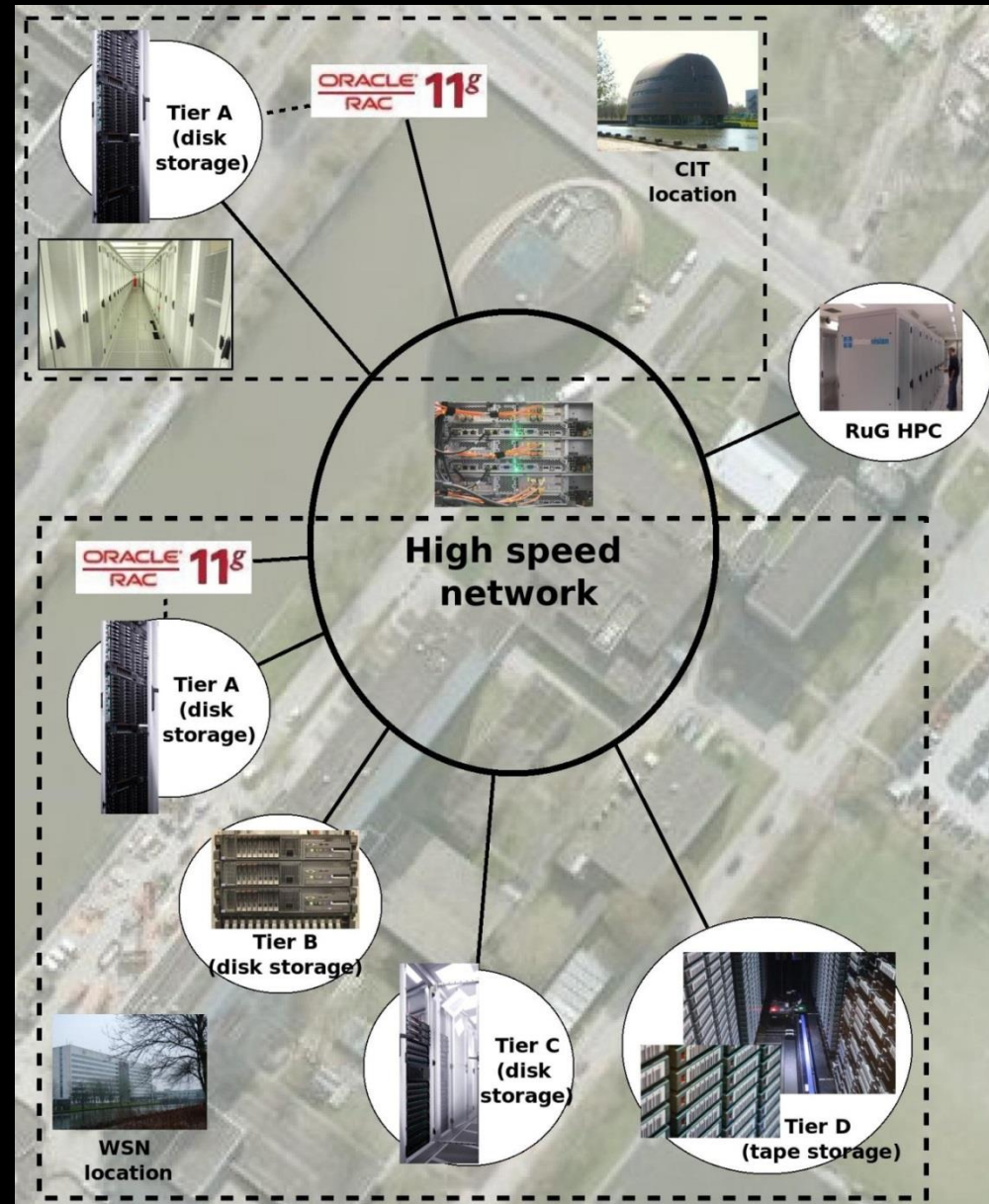
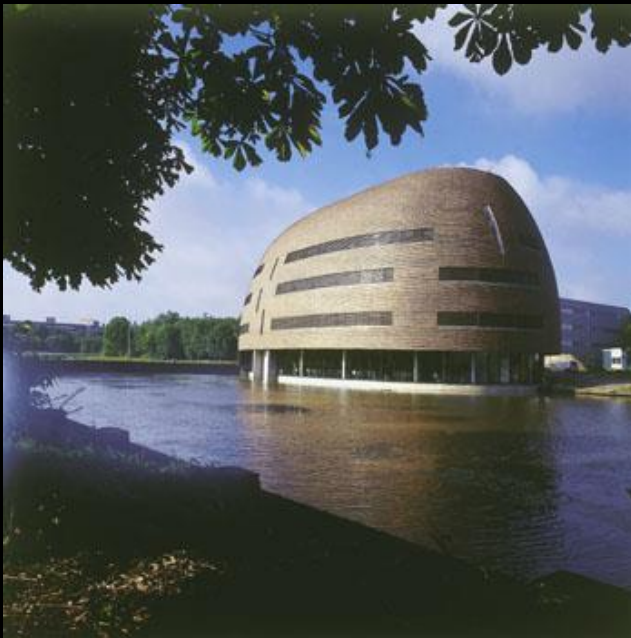




# WISE/Target Infrastructure

The WISE hub in Groningen consists of the following:

- HPC Cluster (4000 core)
- Grid CE (NL Big Grid)
- GPFS File System (10 PB)
- Oracle 11g RAC



# Most important

First datacentric information system in Astronomy

Full data lineage

Data model flexibility

Modular approach to pipelining

Infrastructural flexibility

Reproduction of any data product

Multi-user environment

Scalability

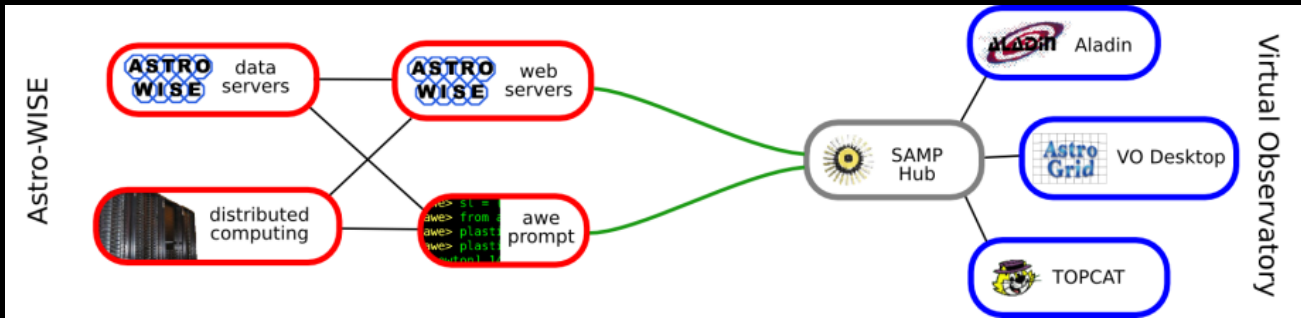
**Astro-WISE Online**

**Overall storage and user statistics**  
Online storage: 1.6 PB (=1600 TB)  
Number of files stored: 18696755  
Database accounts: 188  
Total queries<sup>1</sup>: 93605965  
<sup>1</sup>sum for all databases since their last restart

**Status of services at Astro-WISE nodes**

- [Bonn](#) ●●
- [Groningen](#) ●●
- [Leiden](#) ●●
- [Nijmegen](#) ●●
- [München](#) ●●
- [Napoli](#) ●●

[poli details](#)  
Updated: 02 Jul 2012 10:42:02



## Pushing it further catalogues $>10^9$ rows

- Extreme data lineage / back-wards chaining end – to – end
- Query driven visualization

