The Virtual Observatory in Australia

Connecting to International Initiatives

Peter Lamb

CSIRO Mathematical & Information Sciences

The Grid & eScience

- Convergence of high-performance computing, huge data stores and high-bandwidth computing to:
  - share resources (data/computation/instruments)
  - enhance collaboration
  - allow for new ways of conducting research

International Virtual Observatories Alliance

- Global electronic access to astronomical data archives of space and ground-based observatories & sky surveys
- Coordination data analysis using common standards, high-bandwidth networking & state-of-the-art analysis tools

International Virtual Observatories Alliance

- UK – AstroGrid
- Australia – Aus-VO
- EU – AVO
- China – ChinaVO
- Canada – CVO
- France – VOFrance
- Germany – GAVO
- Italy – DRACO
- Japan – JVO
- USA – NVO
- Russia – RVO
- India – VO-I
- Korea – KVO

http://www.ivoa.net/

Aus-VO Projects

- HI Parkes All-Sky Survey
- 2QZ – spectroscopic quasar survey
- MACHO archive
- ATCA archive
- Machine learning
- SUMSS 843MHz southern sky survey
- RAVE radial velocity survey
- Remote Visualisation

http://www.aus-vo.org/

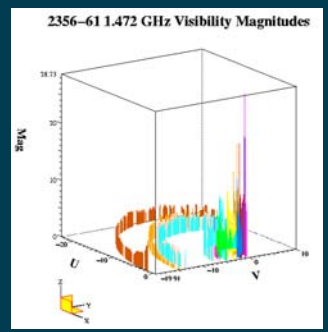
IT Challenges

- Provide a uniform model of observational, derived and modelled data for an entire scientific discipline.
- Provide a model of computation for tools used to reduce the data.
- Allow for automatic, assisted and convenient manual assembly of data sources and tools.

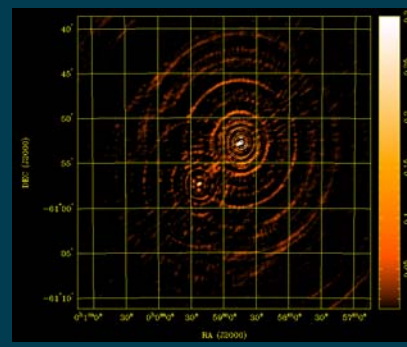
Radio Interferometry Imaging

Raw data to source extraction

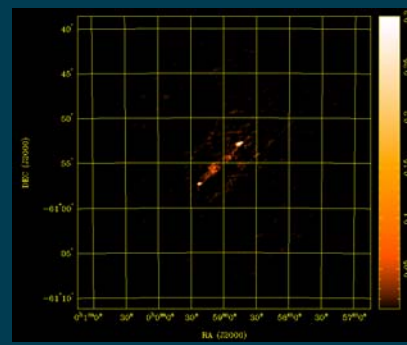
Amplify, filter, correlate



Calibrate, Inverse Fourier Transform

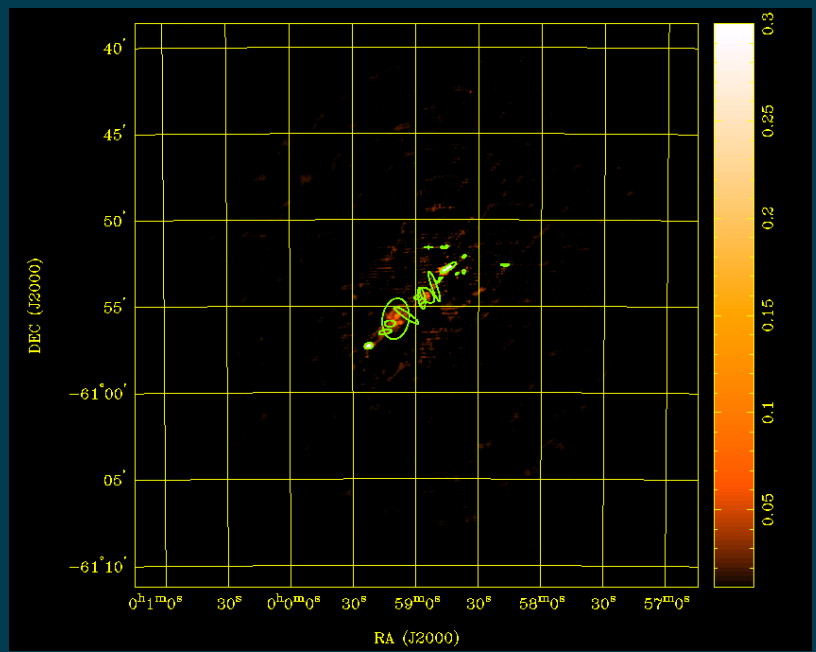


Deconvolve



Source extraction

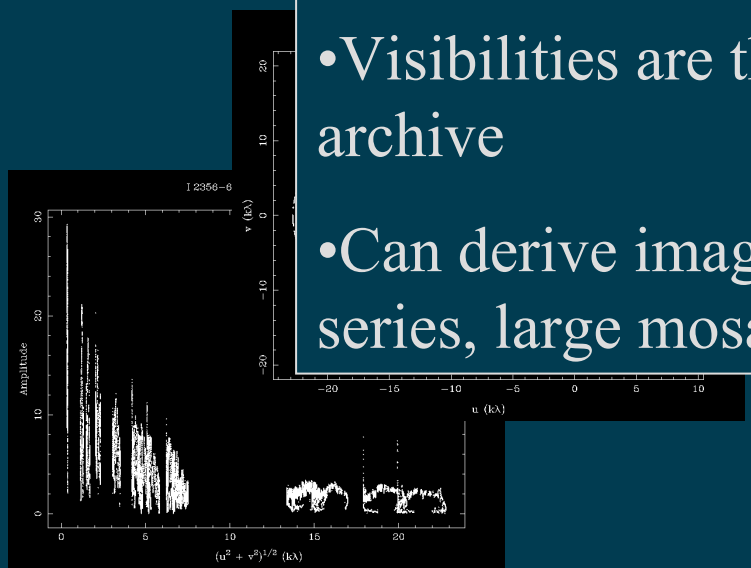
Many details omitted!

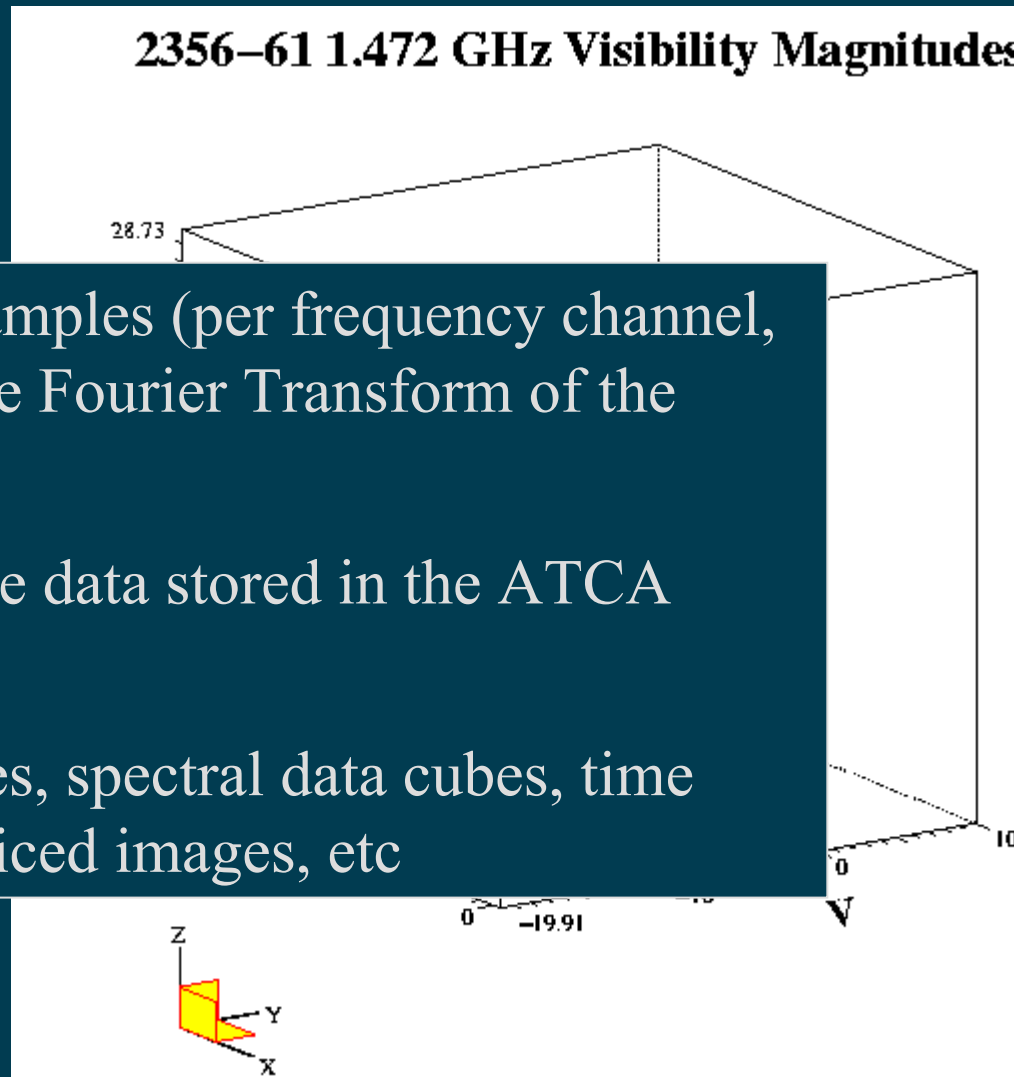


Raw data - visibilities

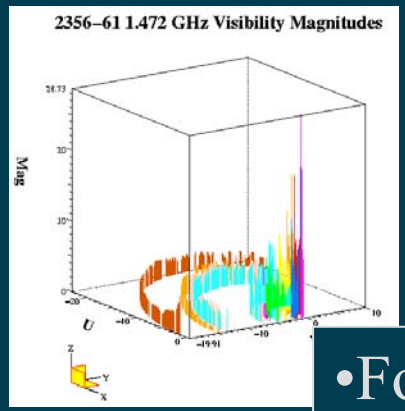
Amplify, filter

- Visibilities are samples (per frequency channel, polarisation) of the Fourier Transform of the image
- Visibilities are the data stored in the ATCA archive
- Can derive images, spectral data cubes, time series, large mosaiced images, etc

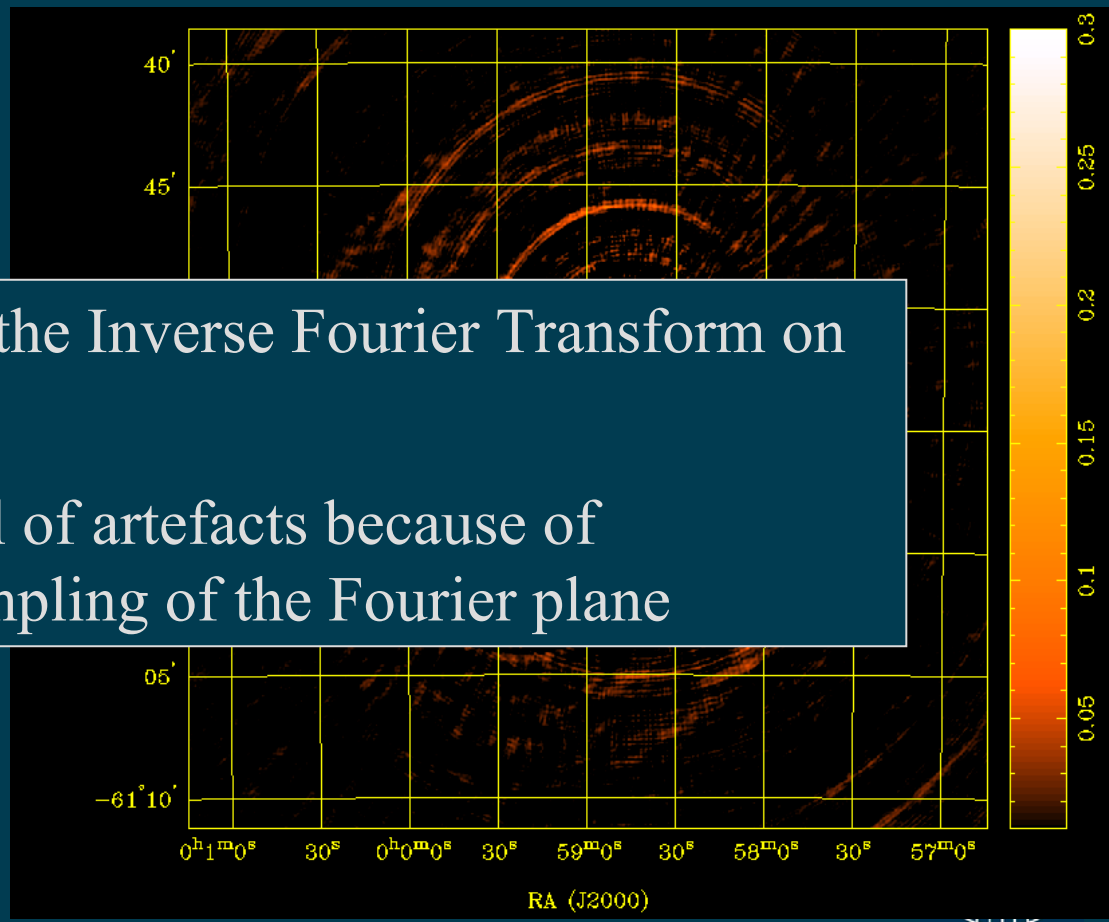




“Dirty” Image



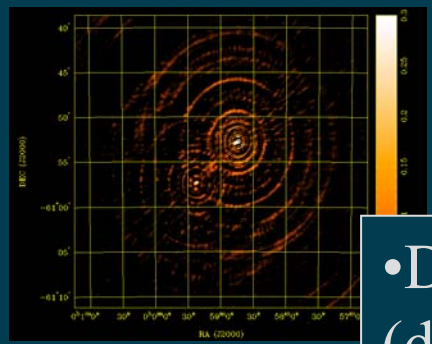
Calibrate, Inverse Fourier



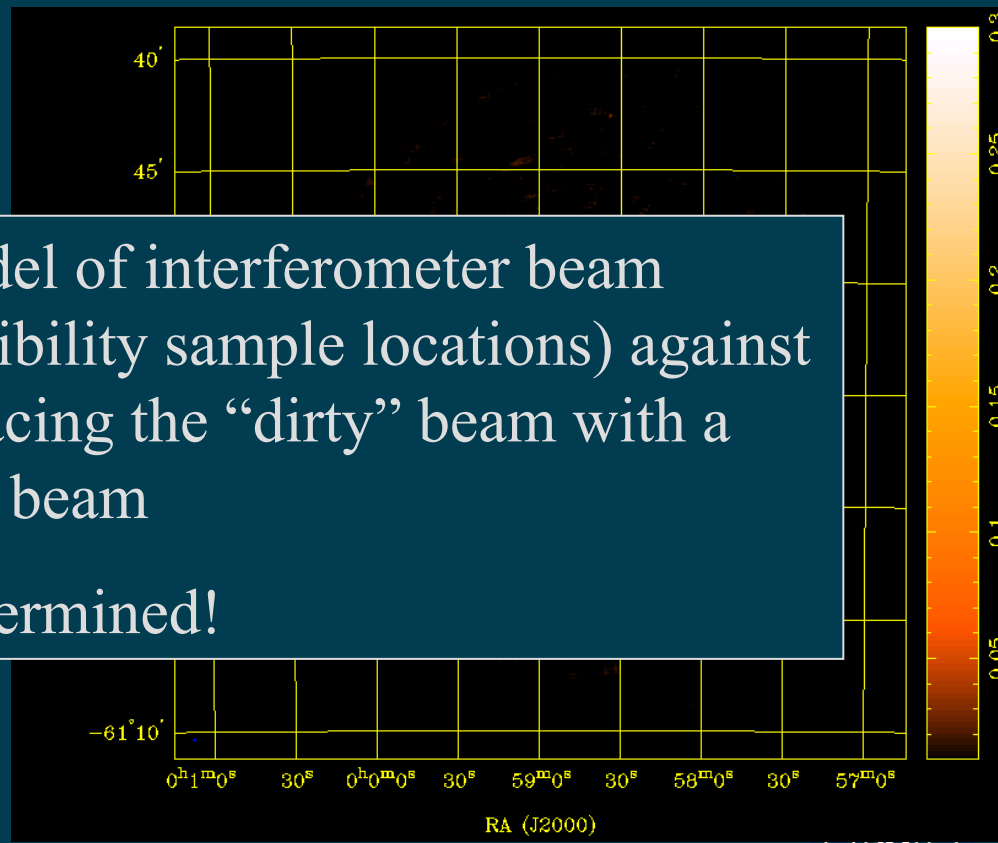
- Formed from the Inverse Fourier Transform on the visibilities
- Has high level of artefacts because of incomplete sampling of the Fourier plane

CSIRO

“Clean” image

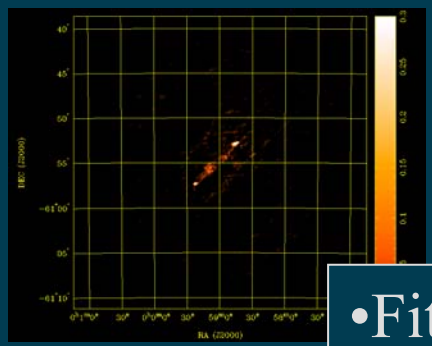


Deconvolve

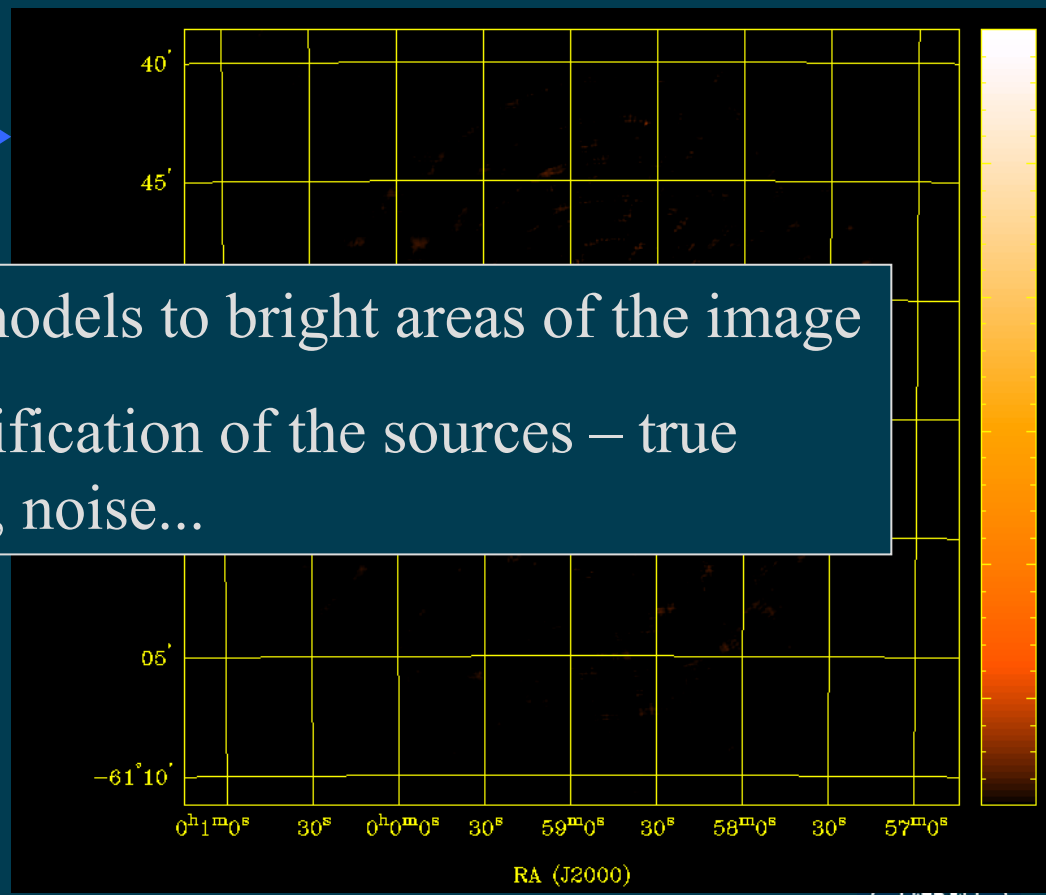


- Deconvolve model of interferometer beam (derived from visibility sample locations) against dirty image, replacing the “dirty” beam with a “clean” Gaussian beam
- Highly underdetermined!

Source Extraction



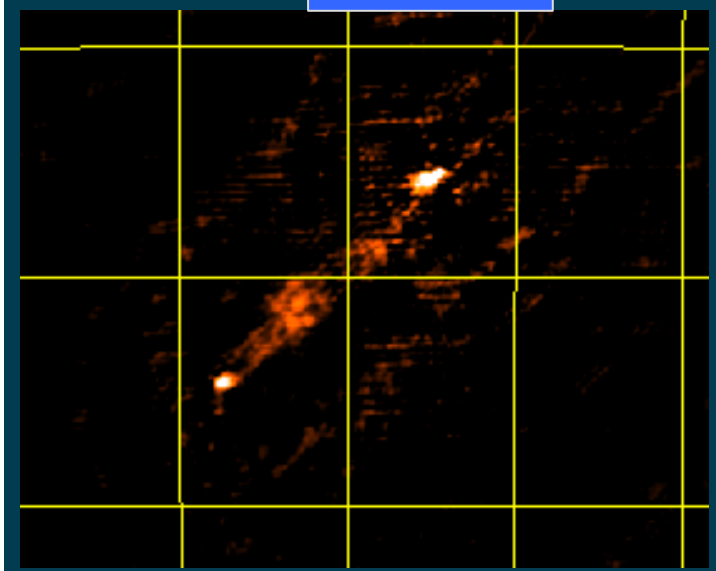
Source extraction

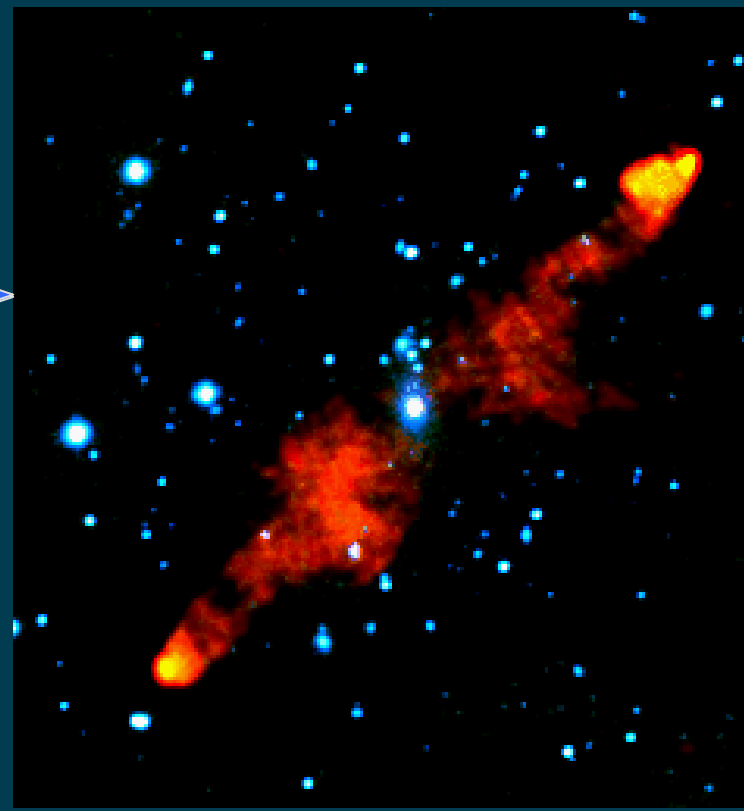


- Fit Gaussian models to bright areas of the image
- Requires classification of the sources – true source, artefact, noise...

Merge with other data

Search optical VO data



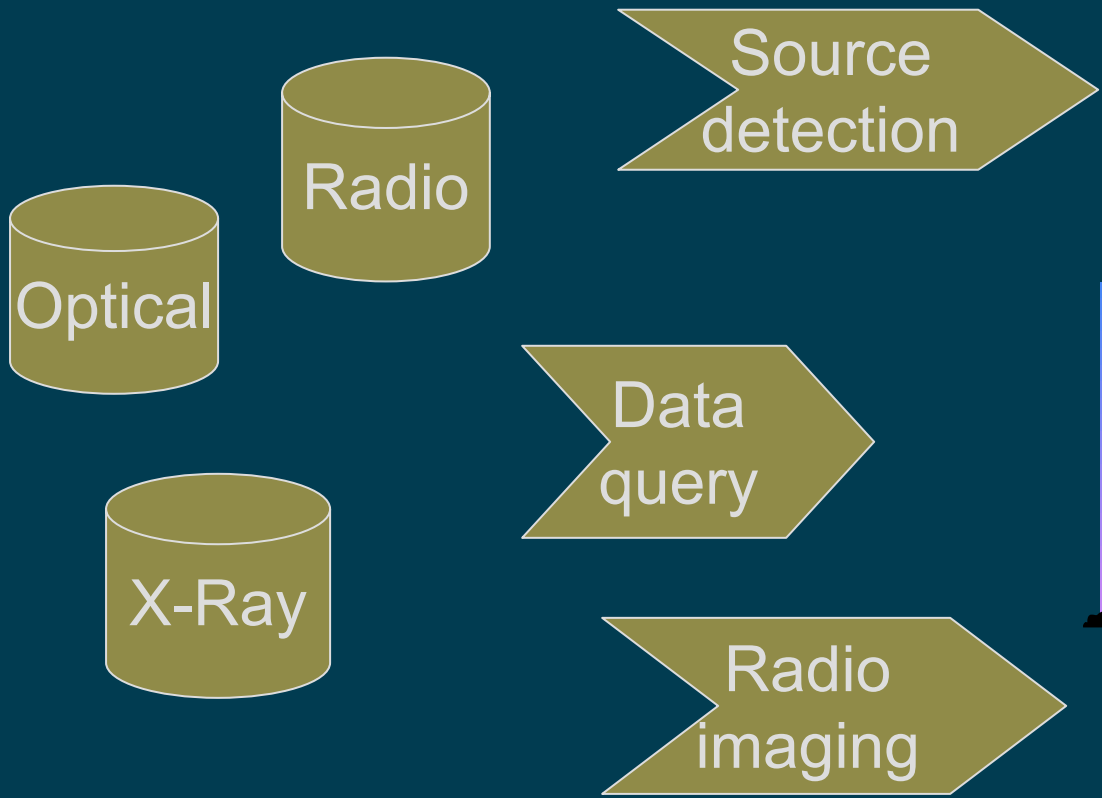


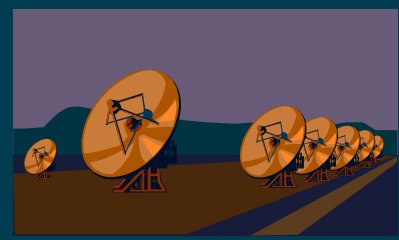
IT Demands

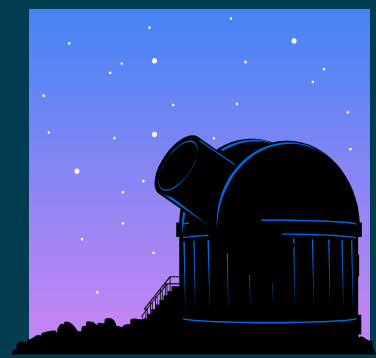
- Large distributed database (ATCA data alone is ~1.8TB)
- Need common data model
- Need to
  - describe
  - compose
  - execute

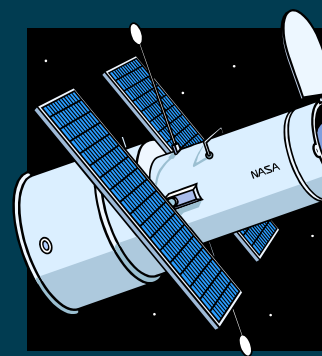
  data queries and processing

Service description



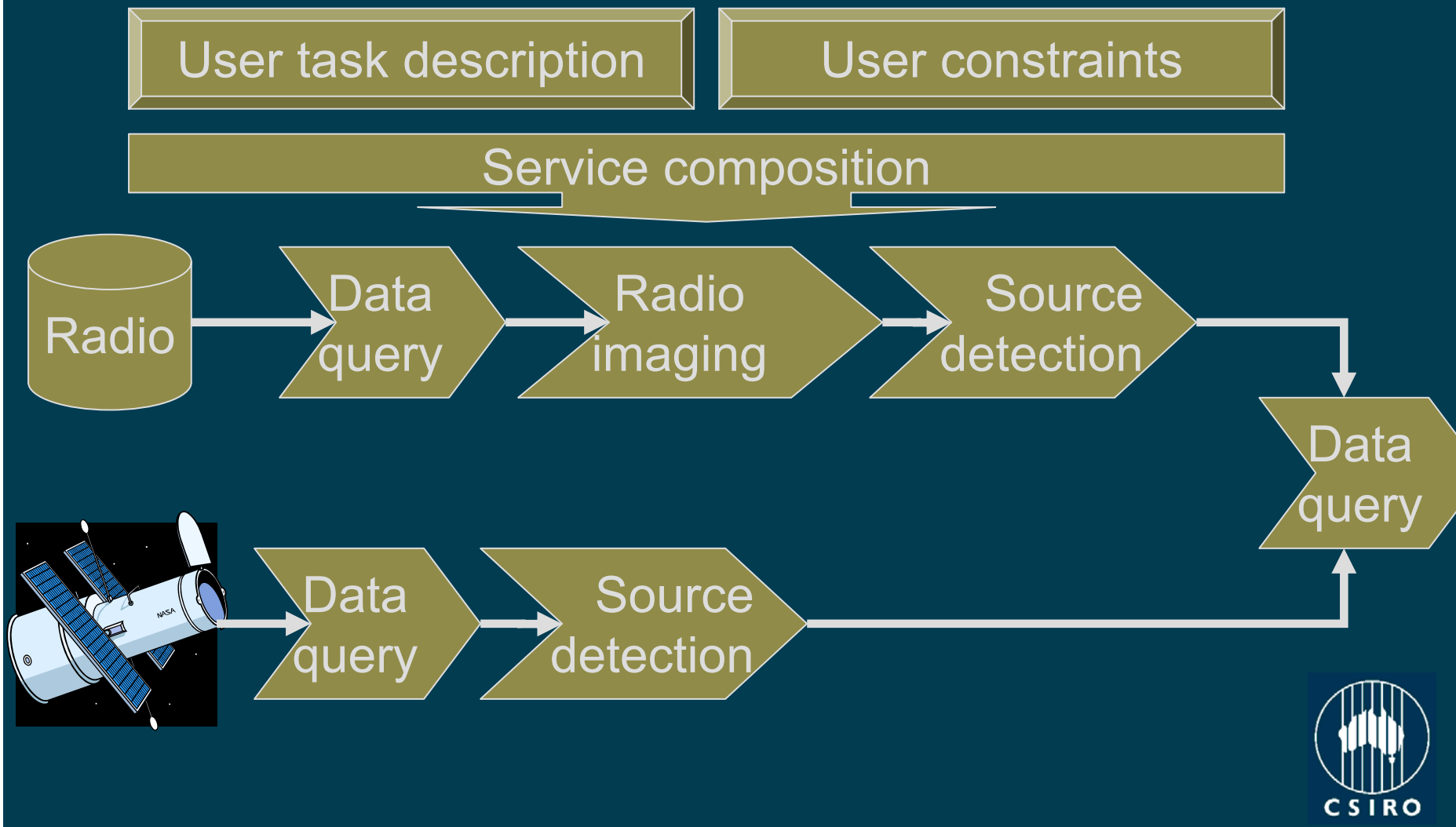






CSIRO

Service composition



Where we are

- Select an area of interest in the sky
- Assemble or acquire the data sets (does the data I need exist already)?
- Process/reduce
- Combine
- ✓Scientific results

Where we want to be

- State the data and processing *requirements*
- Receive the data processed as needed
- Check the hypothesis
- Double-check the data and the processing
- ✓ Scientific result; sooner & with less effort

What's in the way

- We have archives , not data access nodes
- Many kinds of data: radio, optical, X-ray; images, spectra, polarisation maps
- Much data processing/reduction needed to make inspection, search & fusion possible
- Data treatment is a craft

Questions?