



CSIRO ASKAP Science Data Archive

Minh Huynh (CSIRO, ICRAR/UWA)

James Dempsey, Matt Whiting and the CASDA team

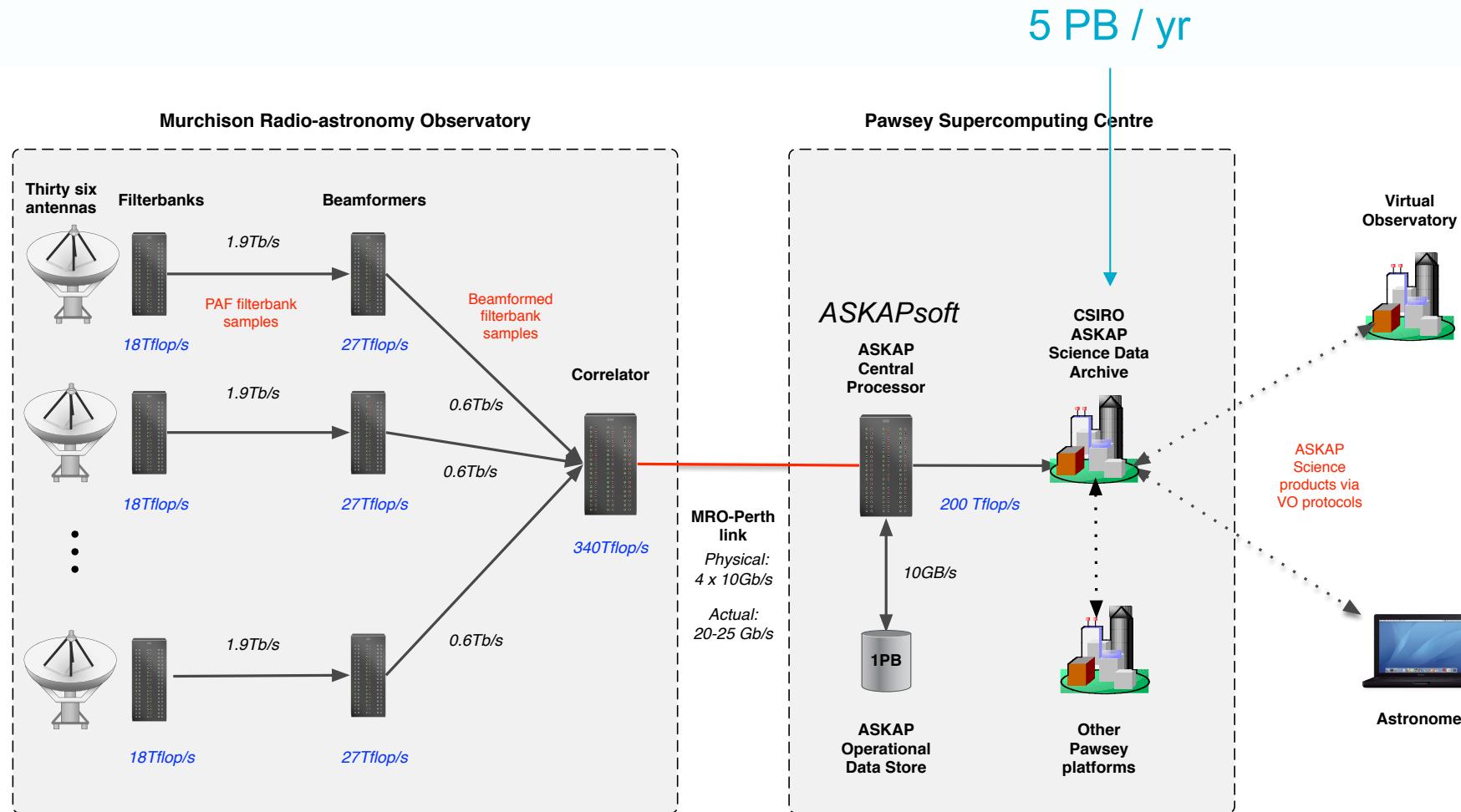
ASKAP – A next generation radio telescope

Australian SKA Pathfinder (ASKAP):

- 36 element interferometer
- Steerable parabolic dishes, 12m in diameter.
- Operating frequency: 700-1800 MHz.
- Wide instantaneous BW: 300 MHz.
- **High Spectral resolution:** 16200 channels (~18.5kHz)
- **Phased Array Feeds: Very Wide Field of view** (~30 sq deg)
- Roll-axis: High dynamic range imaging
- **High survey speeds**
- **Also very High Data Volume/rates**



ASKAP and all its wonderful data



CASDA Deployment

PAWSEY SUPERCOMPUTING CENTRE		CANBERRA
Location	Perth, WA	Canberra, ACT
Functions	<ul style="list-style-type: none">• Deposit ASKAP data products• Data access• Virtual Observatory	<ul style="list-style-type: none">• Interactive search• Collections• Authentication/Authorisation• Data Validation/Release• User interface
Facilities (Prod)	<ul style="list-style-type: none">• 2 dedicated servers• 2 x 10 PB tape library (shared)• Lustre filesystem (shared)• 256 TB disk (initial allocation)	<ul style="list-style-type: none">• 5 virtual machines• Integrated with CSIRO Data Access Portal
Facilities (Dev, Test, AT)	<ul style="list-style-type: none">• 5 dedicated servers• Lustre filesystem (shared)• 384 TB disk (initial allocation)	<ul style="list-style-type: none">• 15 virtual machines• Integrated with CSIRO Data Access Portal

CASDA Software Environment

Java



Tomcat



PostgreSQL



NGAS

SLURM



Elastic Stack (formally ELK)



Open source libraries

CASDA Functionality

CASDA provides long-term archiving of and access to the large scientific datasets taken by ASKAP.

Functionality includes:

- Long term storage of ASKAP science data products
- Searches and data access via web (CSIRO Data Access Portal) and Virtual Observatory services
- Validation of ASKAP observations
- Upload of value-added science catalogues and image cubes by users
- Digital Object Identification (DOI) for all datasets
- Archive administration, inc. team member access to unreleased data

Approach:

- Agile Scrum software development

CASDA Data Products

CASDA stores the following data products:

Calibrated visibilities

- Archived long-term for continuum data only (frequency averaged)

Images and Image cubes, Spectra and Moment Maps

- Basic image products produced by the pipelines
- Derived image products, such as extracted spectra for detected sources, moment maps, etc

Catalogues

- Detected sources and their parameters

Project and Observation information

Data quality information

- Used to evaluate the quality of an observation
- Evaluation metric files

CSIRO Data Access Portal

<https://data.csiro.au/dap/specificSearch>

The screenshot shows the CSIRO Data Access Portal homepage at https://data.csiro.au/dap/specificSearch. At the top, there is a message: "We are working on improving your DAP experience. Please contact researchdatasupport@csiro.au if you would like to be involved." Below this is a "Close" button. The header includes the CSIRO logo, a "Data Access Portal" title, and links for "Contact Us", "Help", and "API". On the right side, there is a "Registered Users" section with a "Login Using" dropdown set to "Nexus", and fields for "Username" and "Password" with a "LOGIN" button. The main content area features a "SEARCH" tab, a "BROWSE" tab, and a "DOMAIN SEARCH" tab with a search icon. To the left, there is a sidebar with text about the Domain Search Tools and a list of facilities:

- AAHL MICROSCOPY SEARCH**: Query microscopy images from Australian Animal Health Laboratory. [About the Australian Animal Health Laboratory...](#)
- ATNF PULSAR OBSERVATION SEARCH**: Query pulsar observations taken at the Parkes radio telescope. [About the Australia Telescope National Facility...](#)
- CASDA OBSERVATION SEARCH**: Query the CSIRO ASKAP Science Data Archive (CASDA) for observations taken by the ASKAP radio telescope. [About the Australian Square Kilometre Array Pathfinder \(ASKAP\)...](#)
- CASDA SKYMAP SEARCH**: Query the CSIRO ASKAP Science Data Archive (CASDA) using an interactive sky map. [About the Australian Square Kilometre Array Pathfinder \(ASKAP\)...](#)
- SENSOR NETWORK SEARCH**: Query CSIRO's Sensors and Sensor Networks empirical data about the natural world. Sensors measure rainfall, temperature and many other phenomena across Australia. [About the Sensor and Sensor Networks Capability](#)

At the bottom, there is a copyright notice: "Copyright (c) 2010-2017 CSIRO Australia. All Rights Reserved. Server Name: Production. Build Number: 2.19.1489 (04 Oct 2017)". There are also links for Accessibility, Acknowledgements, Legal Notice and Disclaimer, Privacy, and Copyright.

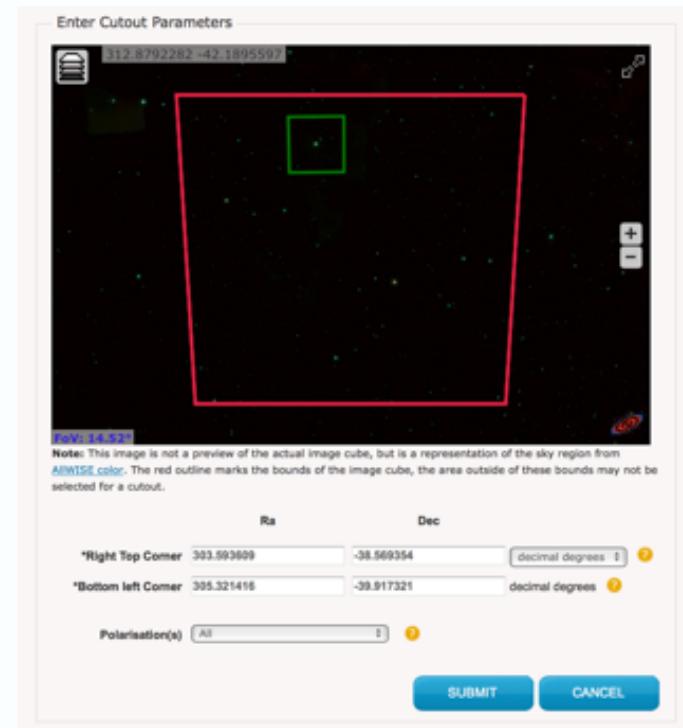
Data Access Portal query interface

<https://data.csiro.au/dap/public/casda/casdaSearch.zul>

The screenshot shows the Data Access Portal's CASDA Observation Search page. At the top, there's a navigation bar with links for Contact Us, Help, and API. On the right, there's a 'Registered Users' section for logging in via Nexus or other methods. Below the header, there are three main search tabs: SEARCH, BROWSE, and DOMAIN SEARCH. The DOMAIN SEARCH tab is selected. The main content area is titled 'CASDA OBSERVATION SEARCH'. It includes an 'INSTRUCTIONS' section with detailed guidance on how to use the search interface. The search form is divided into two main sections: 'CONE SEARCH' and 'Observation / Project'. The 'CONE SEARCH' section contains fields for 'Object Name', 'Right Ascension', 'Declination', and 'Search Radius'. The 'Observation / Project' section contains a 'FILE' section with a 'Filename' field. There are also 'hide' and 'hide_all' buttons for each section.

Data Access Portal Capabilities

- Search for images/cubes and catalogues via
 - Coordinates, cone search
 - Object name
 - Project or Scheduling Block ID
 - filename
- Generate image/cube cutouts via Aladin Lite tool
- Data validation and release
- Upload of User Generated Catalogues, Images and Cubes



Data Access Portal skymap search with Aladin Lite

<https://data.csiro.au/dap/public/casda/casdaSkymapSearch.zul>

The screenshot shows the Data Access Portal's CASDA Skymap Search page. At the top, there is a navigation bar with links for Contact Us, Help, API, and a 'Registered Users' section for logging in via Nexus. Below the navigation is a search bar and a header titled 'Data Access Portal'. The main content area has tabs for SEARCH, BROWSE, and DOMAIN SEARCH. The URL in the address bar is https://data.csiro.au/dap/public/casda/casdaSkymapSearch.zul. The page title is 'CASDA SKYMAP SEARCH'. On the left, there is a 'TARGET' field containing 'NGC 7232' with a 'Go' button, and a 'CATALOGUE TYPE' section with radio buttons for Continuum Component, Continuum Island (which is selected), Spectral Line - Absorption, Spectral Line - Emission, Polarisation, and None. In the center, there is a map showing numerous red square markers representing continuum islands. A specific marker is highlighted with a cyan border and labeled with coordinates: J2000 22:19:37.983 -45:51:0.04. Below the map, a 'FoV: 2°' indicator is visible. To the right of the map, there is a 'Selected Item Details' table and an 'Item Preview' image. The 'Selected Item Details' table contains the following data:

Name:	J221428-453305
RA (J2000):	333.608398
Dec (J2000):	-45.551559
Peak Flux:	55.835 mJy
Integrated Flux:	3.314 mJy
Major Axis:	43.17 arcsec
Minor Axis:	16.76 arcsec
Position Angle:	179.44 degrees

On the far right, there is a CSIRO logo.

CASDA and Virtual Observatory

CASDA work is continuing to build CSIRO's skills, knowledge and international reputation in VO protocols and services.

Main Protocols

- Table Access Protocol (TAP)
- Simple Cone Search
- Simple Image Access Protocol (SIA v2)
- Server Operations for Data Access (SODA)
- Simple Spectral Access Protocol (SSAP)

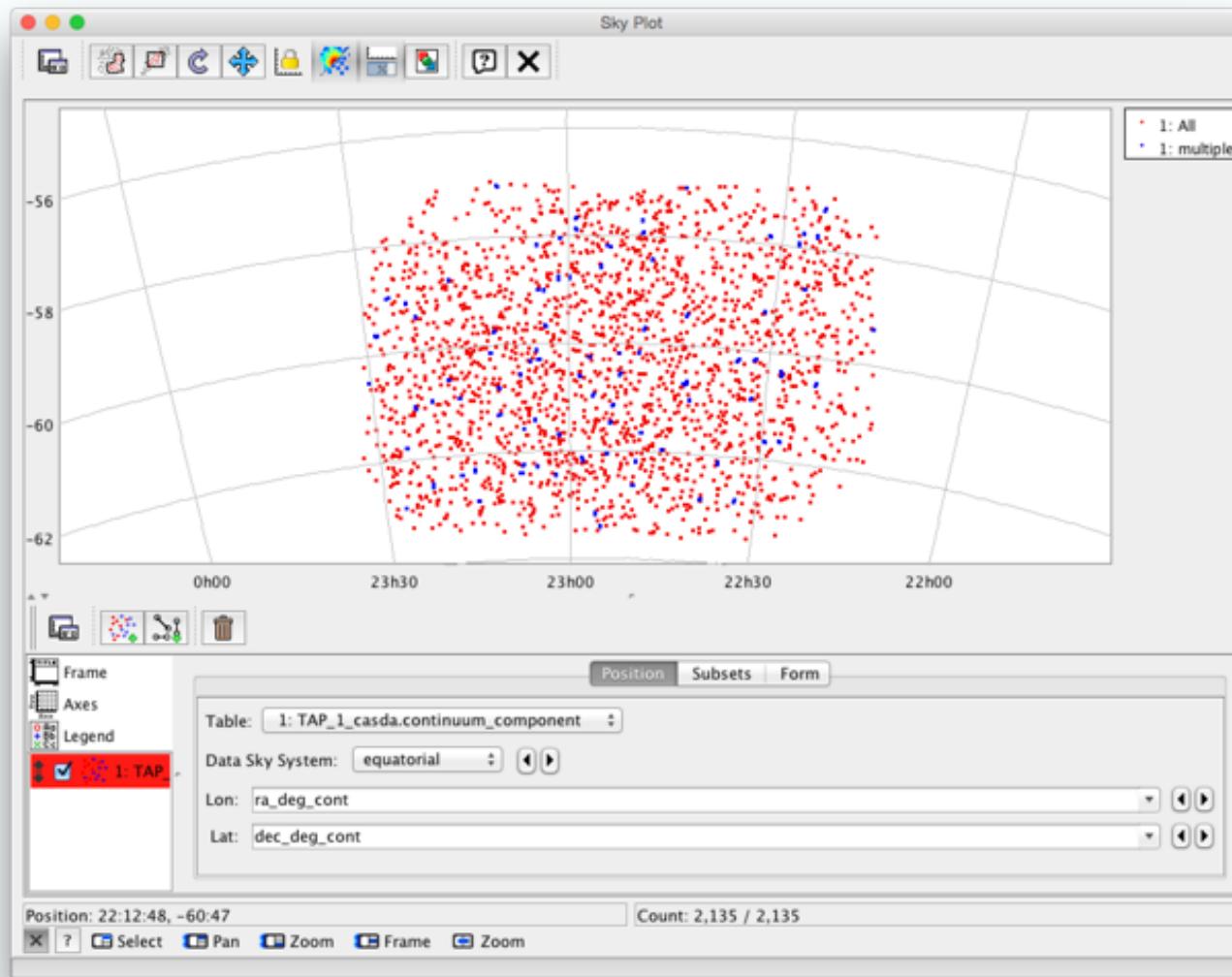
Automated data discovery and retrieval via scripts

Also used as a service layer:

- Cutouts
- File access
- Catalogue Retrieval

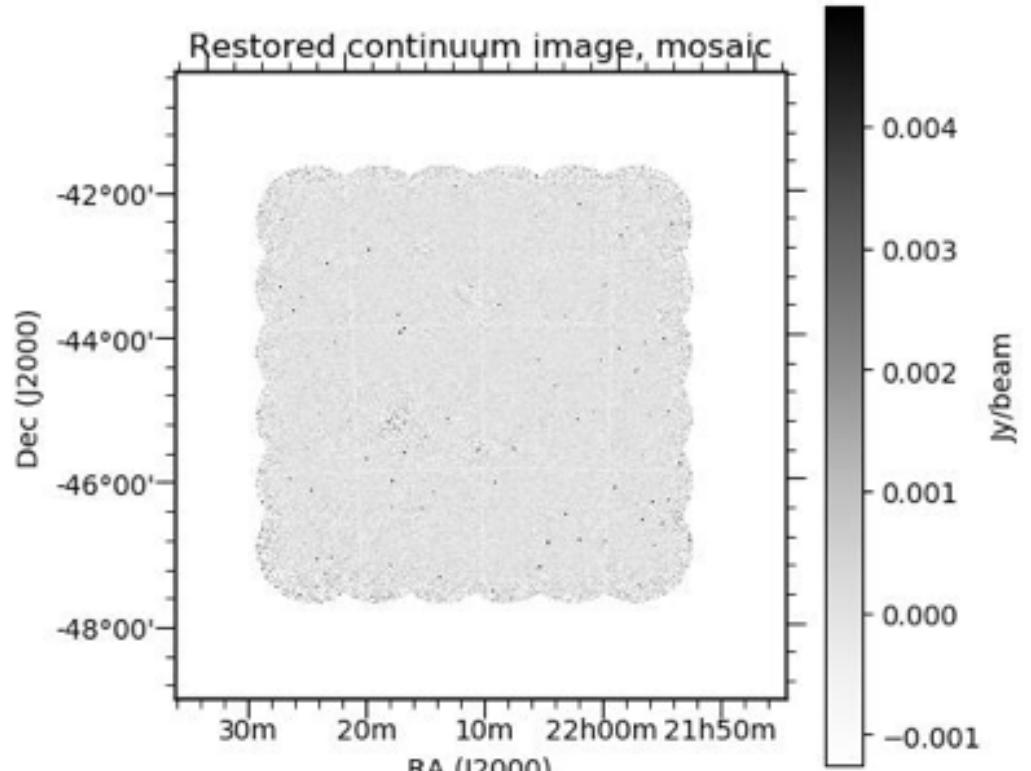
CASDA Virtual Observatory Tools available: <http://doi.org/10.4225/08/59dc56fda2997>

CASDA VO via TOPCAT



First ASKAP Early Science CASDA Data Release

- Observations of NGC 7232 and surrounding area
 - 12 MkII PAFs, 48 MHz bandwidth, 36 beams
- Eleven scheduling blocks (or days)
 - not combined, this is left to the user
- Data released:
 - ASKAPsoft continuum images
 - Selavy continuum source catalogues
 - Calibrated visibilities (freq binned) as CASA measurement sets



Summary

ASKAP will produce ~5 PB a year of processed data products

CSIRO ASKAP Science Data Archive (CASDA) provides:

- Long term storage for science-ready data products - images, cubes, catalogues, calibrated visibilities (continuum), spectra and moment maps
- Search and download via web (Data Access Portal) and Virtual Observatory services

ASKAP Early Science has started, first dataset now available.

- Expect more datasets soon!

