



[www.csiro.au](http://www.csiro.au)

# SKA Astroinformatics

**Ray Norris**  
**June 2010**



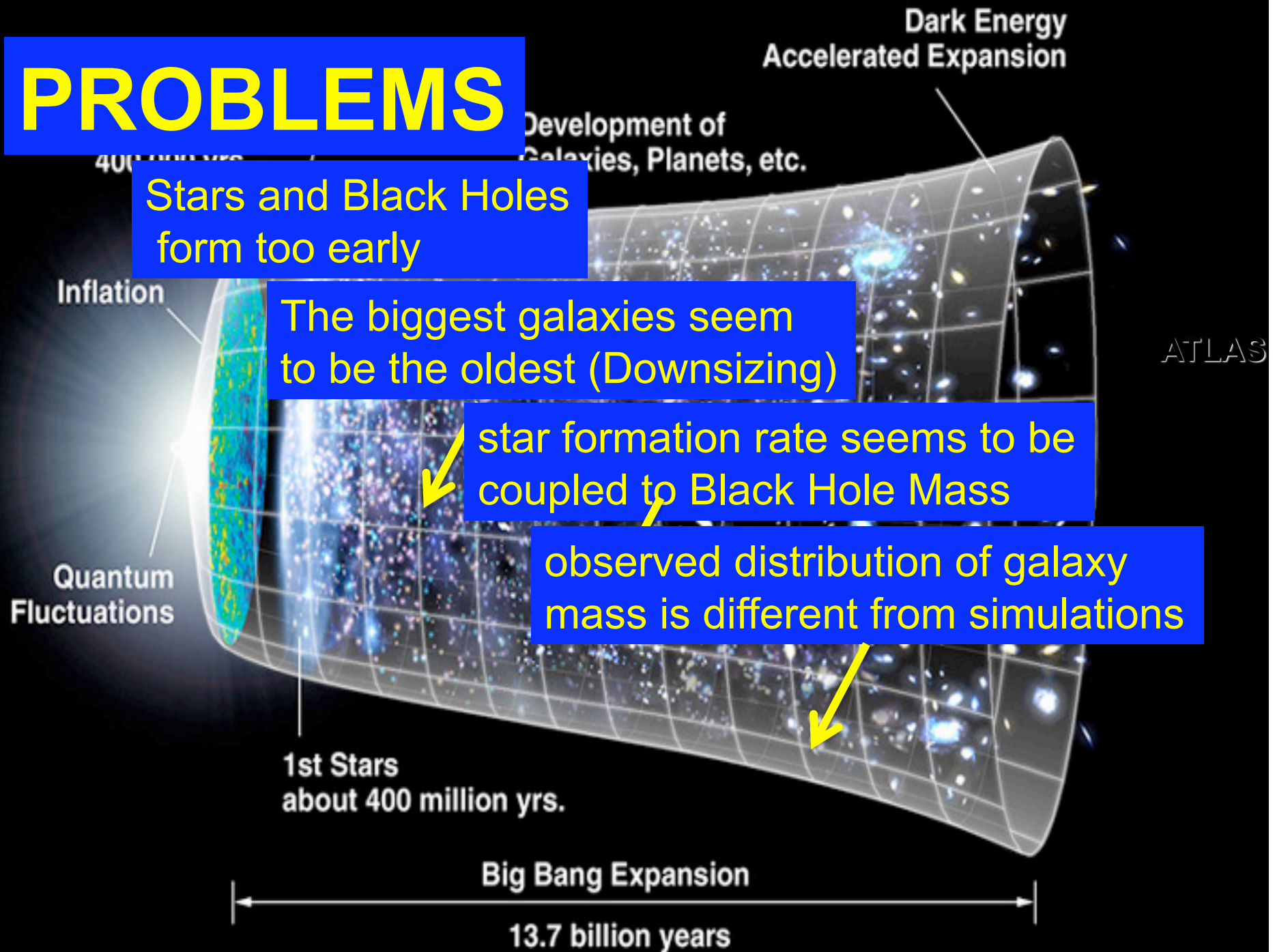
# PROBLEMS

Stars and Black Holes  
form too early

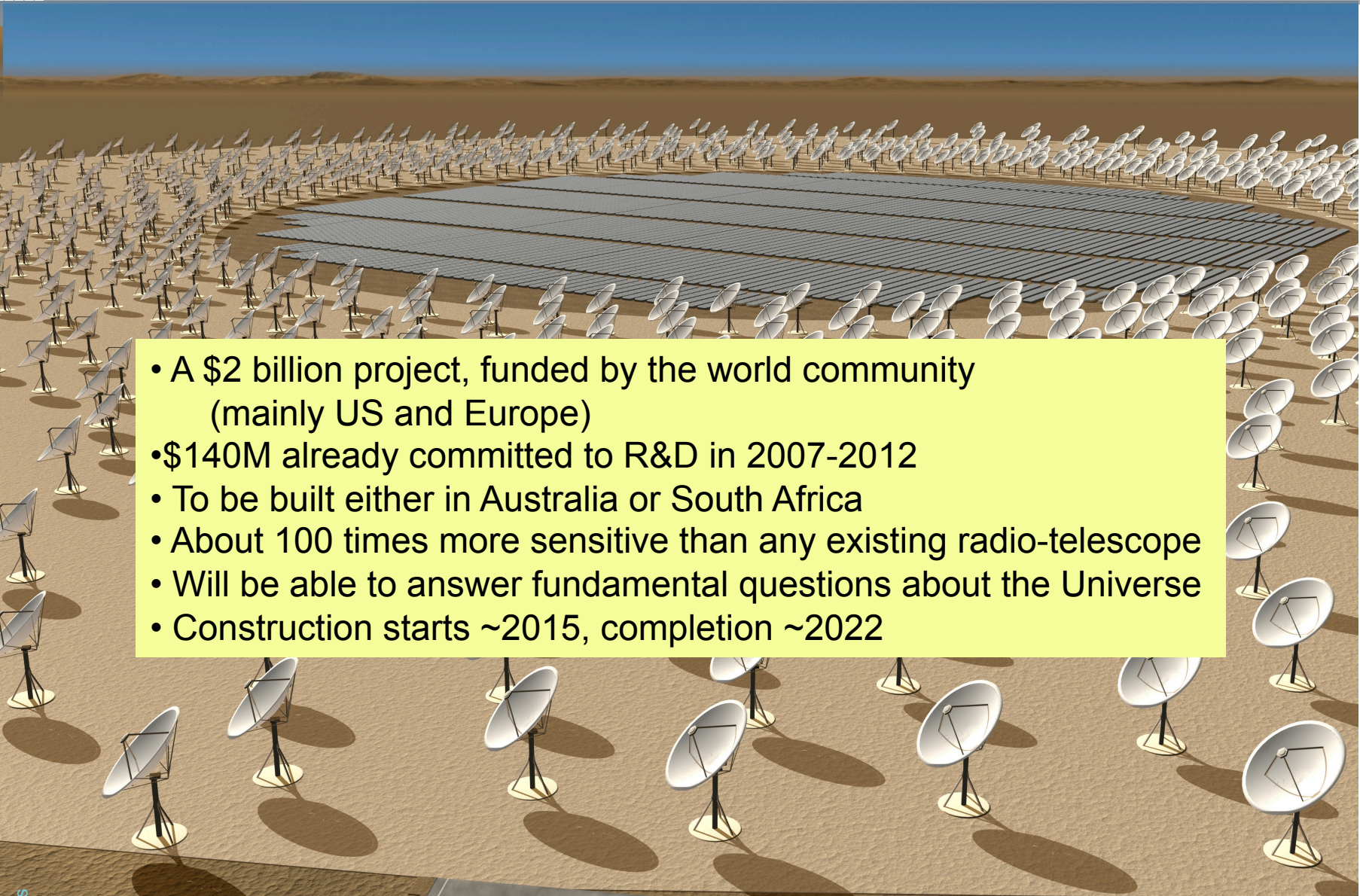
The biggest galaxies seem  
to be the oldest (Downsizing)

star formation rate seems to be  
coupled to Black Hole Mass

observed distribution of galaxy  
mass is different from simulations



# Square Kilometre Array (SKA)

- 
- A \$2 billion project, funded by the world community (mainly US and Europe)
  - \$140M already committed to R&D in 2007-2012
  - To be built either in Australia or South Africa
  - About 100 times more sensitive than any existing radio-telescope
  - Will be able to answer fundamental questions about the Universe
  - Construction starts ~2015, completion ~2022

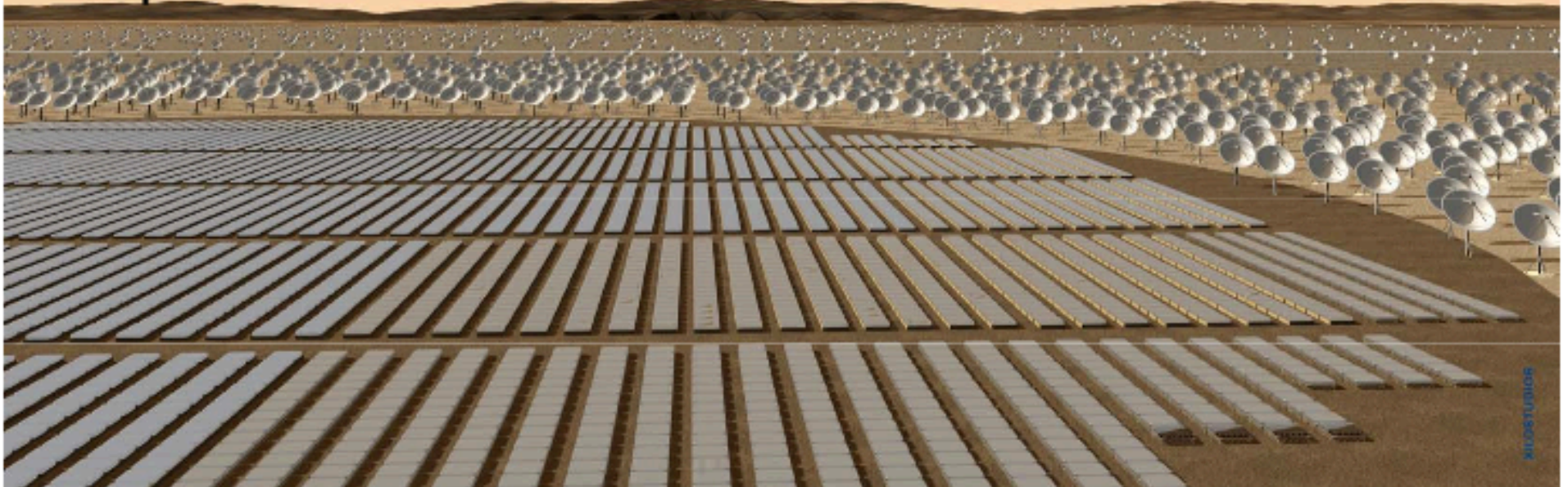
# The Square Kilometre Array (SKA)

1000- 1500 dishes (15m) in the central 5 km  
2000-3000 total

+

dense and/or sparse aperture arrays

connected to a massive data processor by an  
optical fibre network

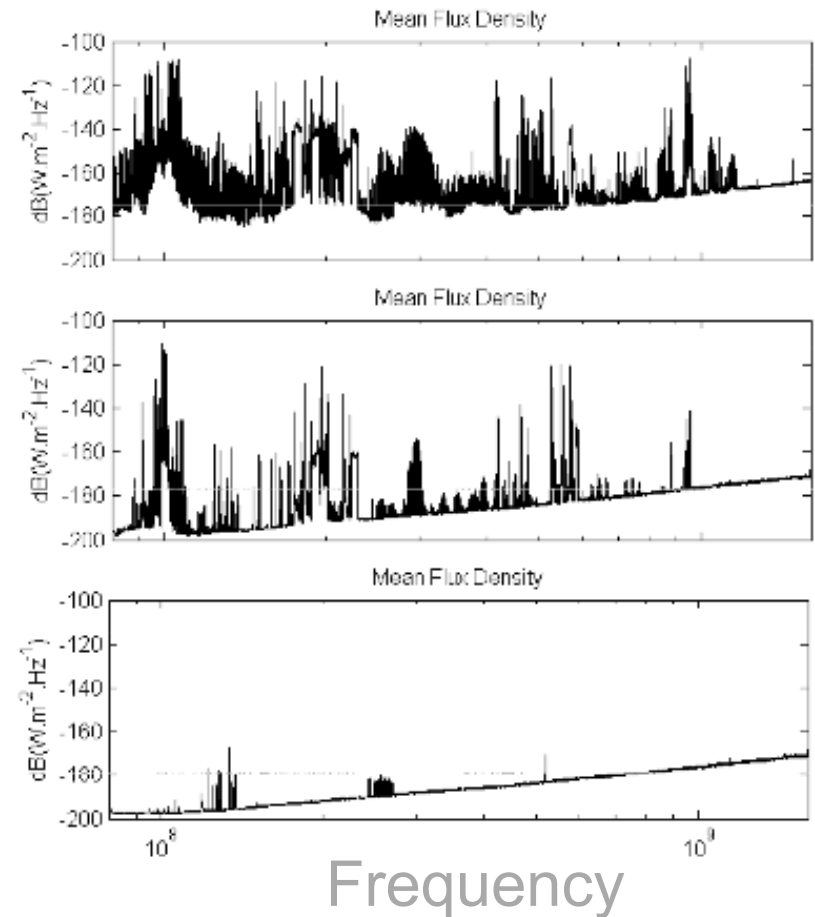


# Timeline

- **2005-2012 Pathfinder construction**
- **2013 → Pathfinder science**
- **2008-2012 SKA Telescope design**
- **2012 Site selection**
- **2012-2016 Phase 1 construction**
- **2016 → early Phase 1 science**
- **2017-2021 construction of full array**

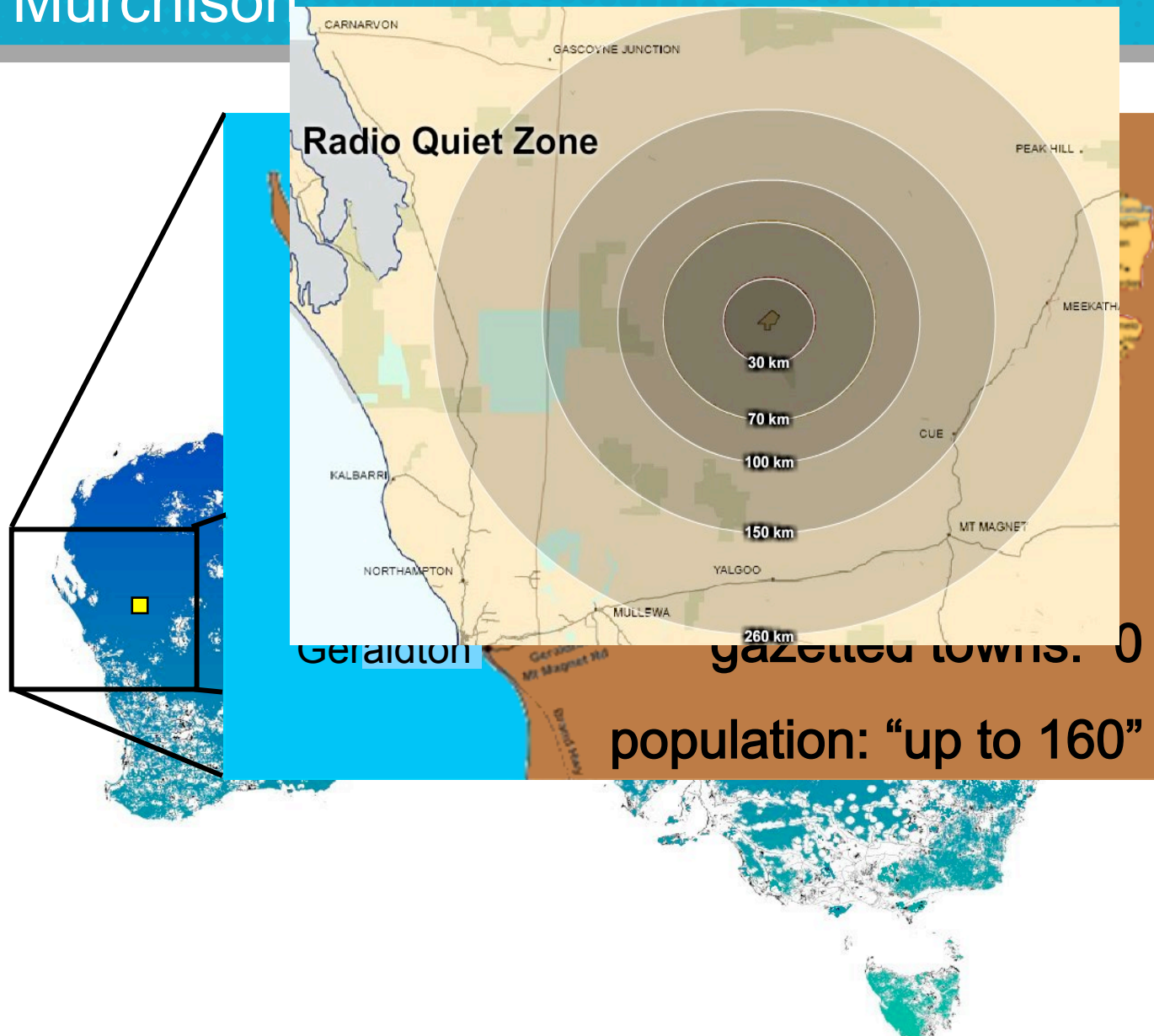
# Site Selection: where is the “Atacama” for radioastronomy?

- City (e.g LA)
- Town (e.g. Socorro)
- Candidate SKA site



Site selection in 2012:  
Australia or South Africa

# Murchison





# SKA Pathfinders

- **Pioneer potential SKA technology**
- **Are powerful science-driven observatories in their own right**
- **Examples:**
  - ASKAP (Australia)
  - MeerKAT (South Africa)
  - LOFAR (Netherlands)
  - Aperitif (Netherlands)
  - ATA (USA)
  - MWA (Australia)
  - EVLA (USA)
  - eMERLIN (UK)
  - eEVN (Europe)

# Current major 20cm surveys

ASKAP/EMU (Australia)



EMU  
75% of sky  
 $\text{rms}=10\mu\text{Jy}$   
70 million galaxies

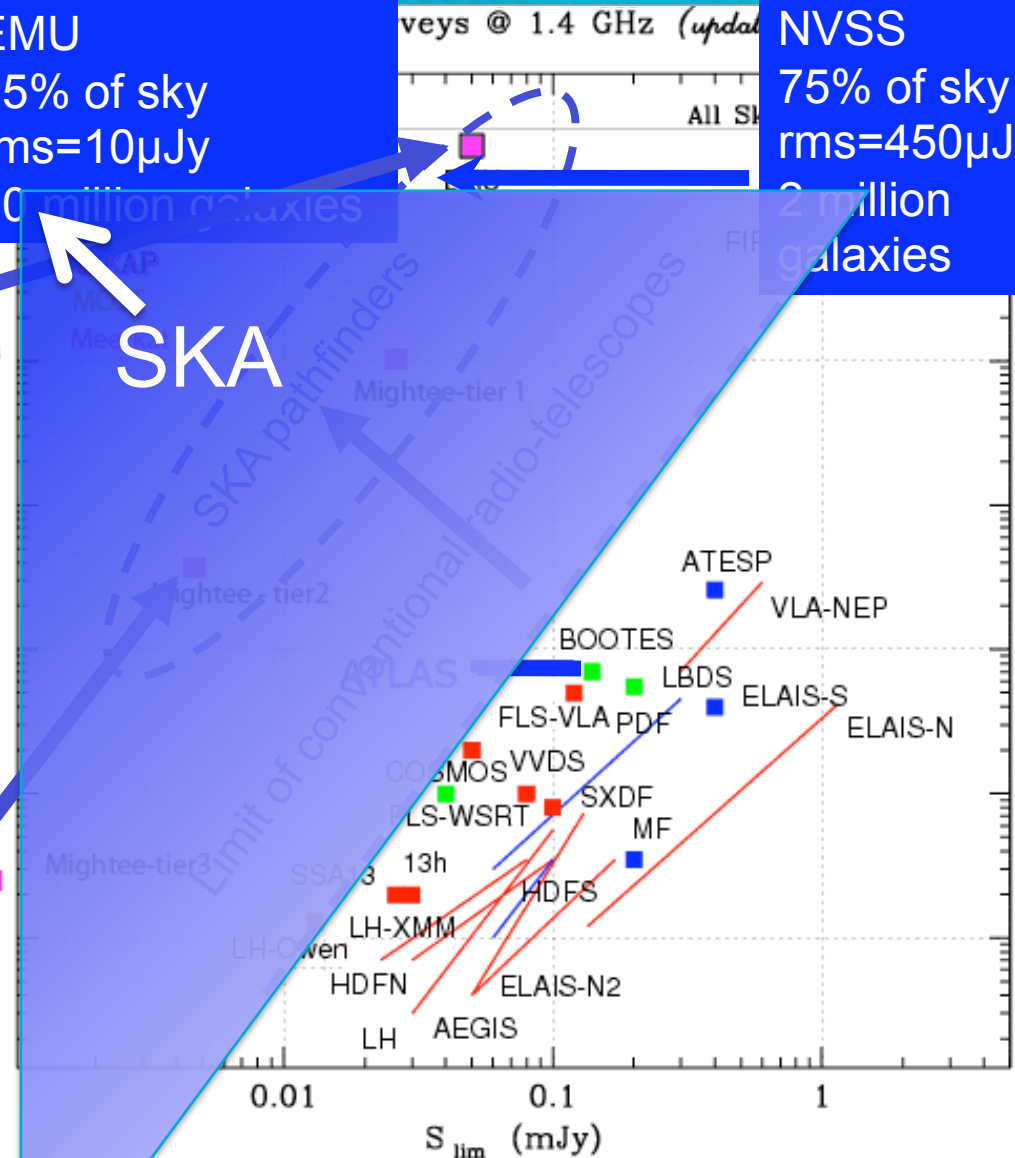
NVSS  
75% of sky  
 $\text{rms}=450\mu\text{Jy}$   
2 million galaxies

SKA

Meerkat/Mightee (South Africa)

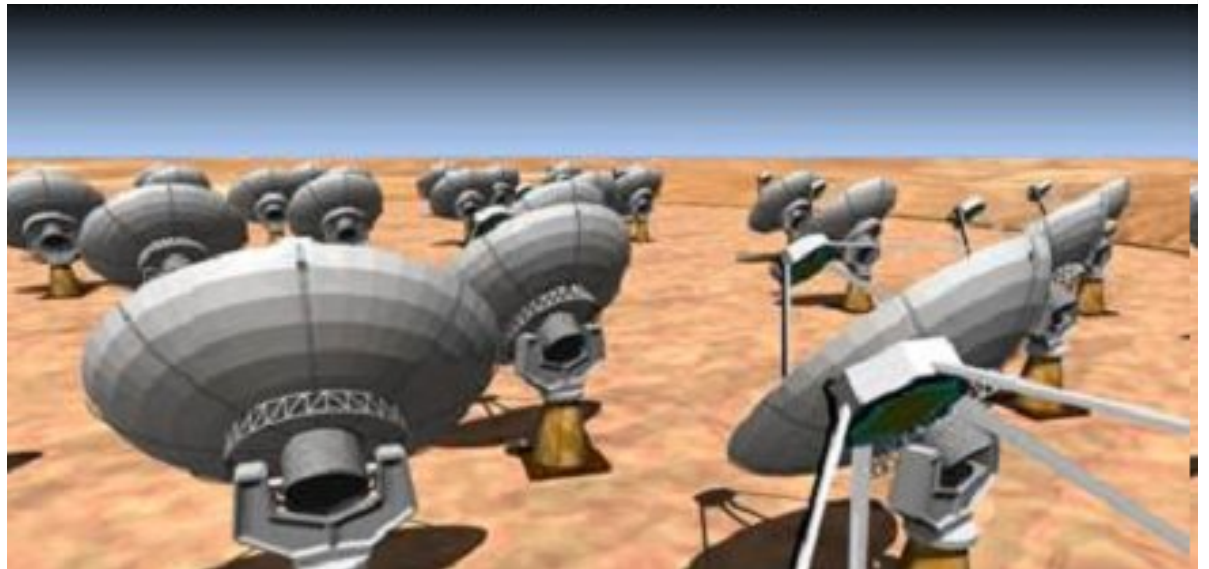


Area ( $\text{deg}^2$ )



# ASKAP=Australian SKA Pathfinder

- \$100m radio-telescope now under construction in Western Australia
- Completion late 2012
- 36\*12m antennas
- Antennas have a 100-pixel phased array feed
- 30 sq. deg FOV!



THE DESIGN CYCLE FORTNIGHTLY [WWW.NEWELECTRONICS.CO.UK](http://WWW.NEWELECTRONICS.CO.UK)

# New Electronics

23 SEPTEMBER 2008



## A better view of the skies

The Square Kilometre Array is set to provide astronomers with unprecedented views of what's out there – and opportunities for UK electronics.

This  
optic

when  
single-

# Data Challenges for SKA Pathfinder

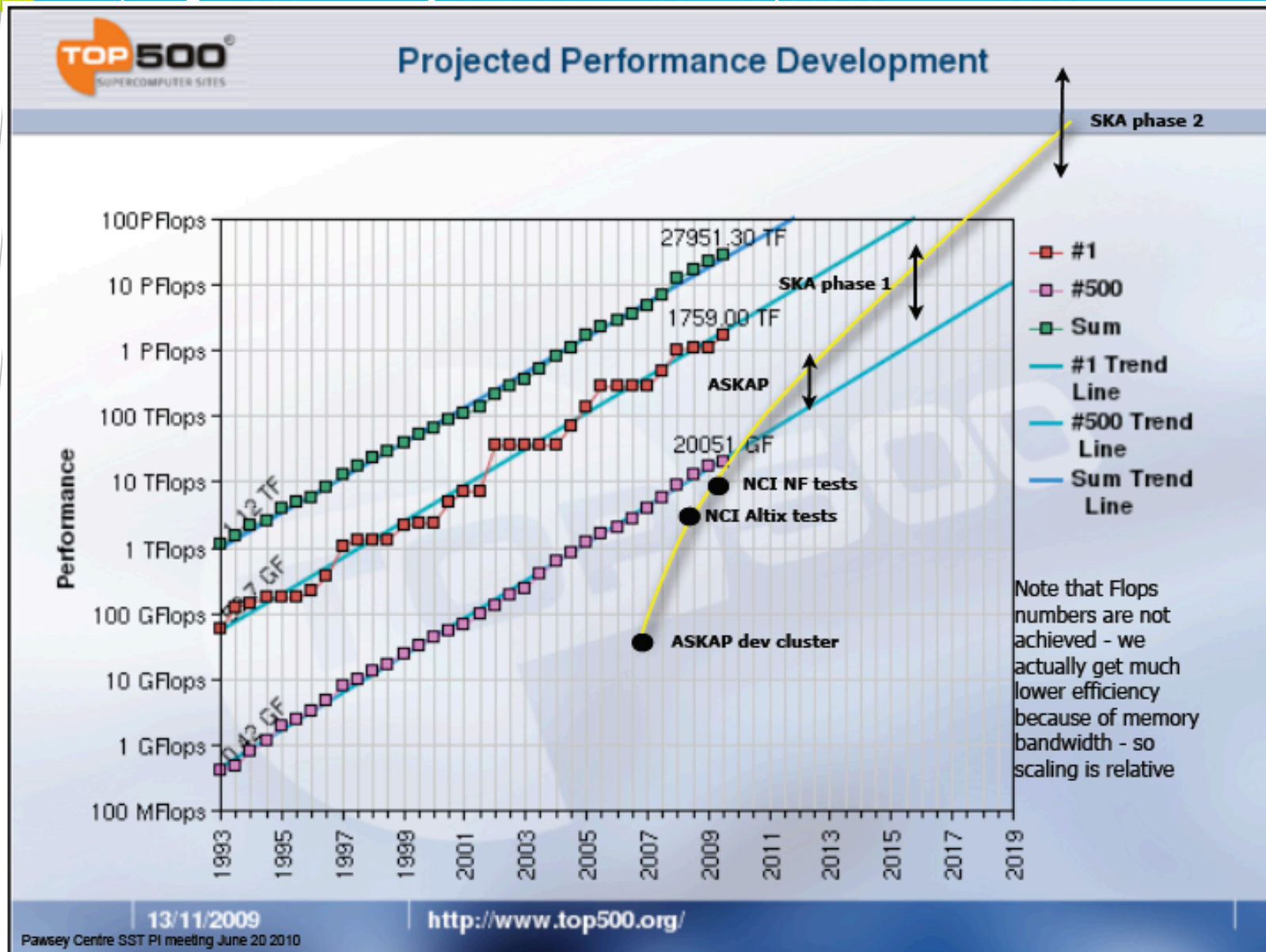
(numbers below are real budgeted numbers for ASKAP)

- **Large bandwidth from telescope to processor**
  - ~10 Tb/s from antennas to correlator (< 6 km)
  - 40 Gb/s from correlator to processor (~ 600 m)
- **Large processing power**
  - 750 Tflop/s expected/budgeted, 1 Pflop/s required
- **Power consumption of processing**
  - 1 MW at site, 10 MW for processing -> power bill ~\$3M p.a.
- **Pipeline processing**
  - including data reduction, source extraction, cross-identification, etc
- **Storage and management of data**
  - 70 PB/yr of products are kept
  - 5 PB/yr in current funding
  - write 12 h of data to disk at 10GB/s
- **Availability of data by users (all data in public domain)**
  - using VO tools & services
- **Data-intensive research on data**

**Multiply all these numbers by ~100 - 1000 for SKA in 2022**

# Climbing Mount Exaflop

(diagram courtesy of Time Cornwell)



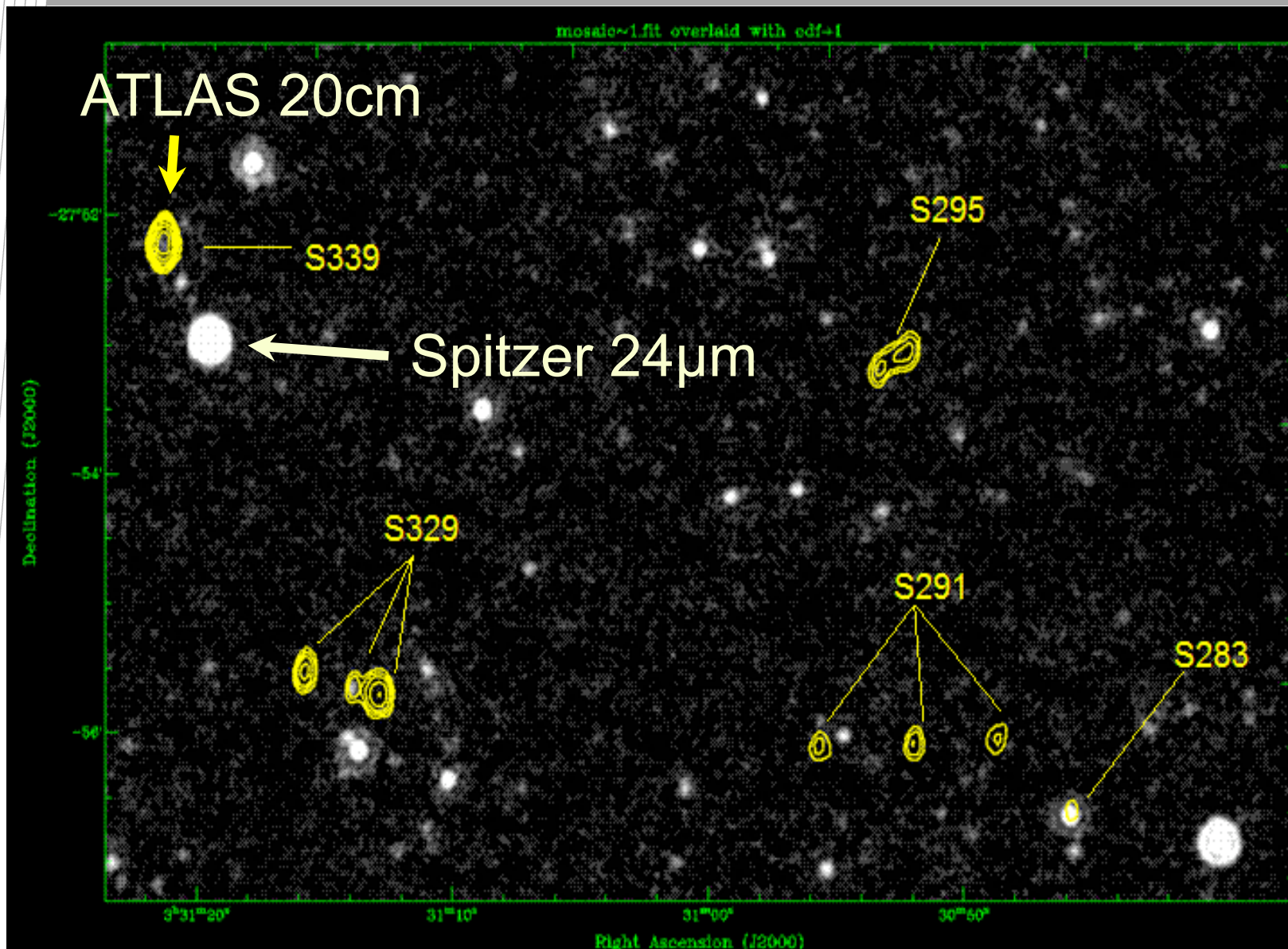
# Data-Intensive Research

- **Data mining**
  - correlation between radio properties and other multi-wavelength properties
- **Stacking**
  - measure the radio luminosity distribution of radio-quiet quasars
- **Cross-correlation**
  - radio positions vs. CMB brightness to measure Integrated Sachs-Wolf Effect -> Evolution of Dark Energy
  - field galaxies vs. radio brightness -> diffuse synchrotron emission from cosmic web
- **Etc...**
- **May require as much (more?) cpu time as main processor,**
- **Must either**
  - be geographically co-located with main data archive
  - or (better) have very high bandwidth links (e.g. 50GB/s sustainable) plus 100's of TB of disk

## An observing run in 2020...

- **Select what sources (or parameters) you want to observe**
- **Observe them by mining survey databases**
- **Perhaps you think of a really clever way of exploring a new part of observational phase-space using the database**
- **New discoveries are made by combining data from different wavelengths/regions of parameter space**
- **In some cases you need to go deeper, or observe a different part of parameter space, so...**
- **(in a few cases) back to the telescope**

# Cross-identification with other wavelengths



# Cross-Identification

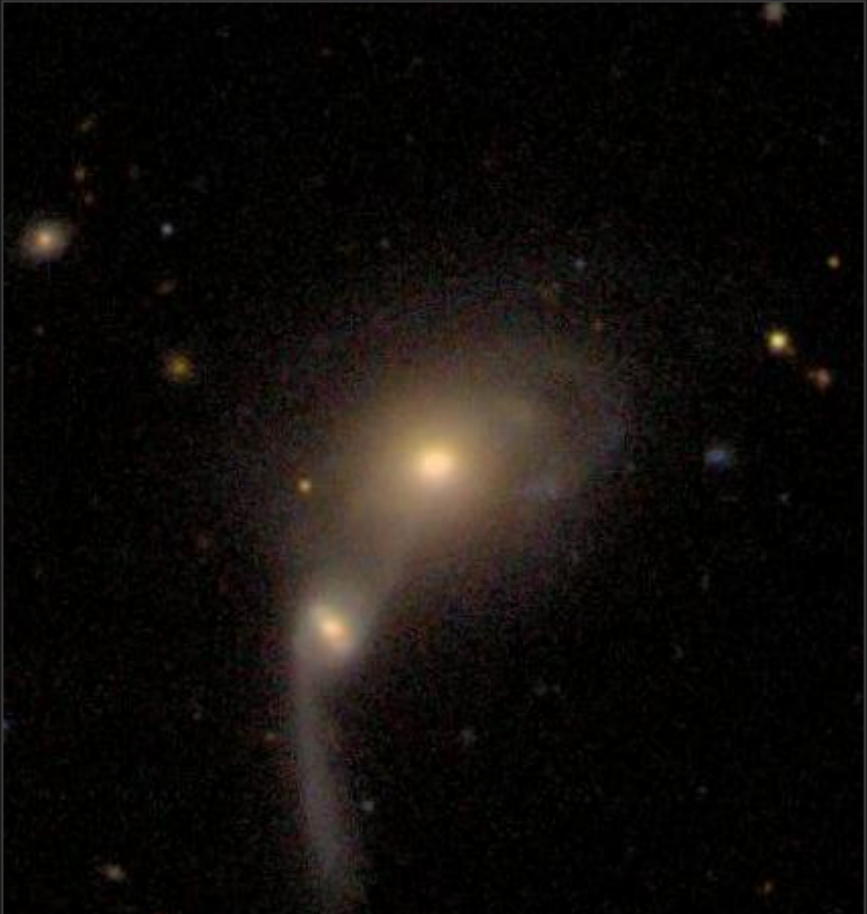
- **For ASKAP-EMU, we plan to develop a pipeline to automate cross-IDS with optical/IR surveys (Pan-STARRS, LSST, WISE, SDSS, SkyMapper, etc)**
  - using intelligent criteria
  - not simple nearest-neighbour
  - working closely with other survey groups
- **Expect to be able to cross-ID 70% of the 70 million galaxies that we discover**
- **20% of them won't have a cross-ID**
- **What about the remaining 10% (21 million galaxies)?**

# Brainpower-intensive research (aka “citizen science”)

# GALAXY ZOO

2

[Home](#) [How To Take Part](#) [My Galaxies](#) [Contact Us](#) [Register](#) [Log In](#)



## Classify galaxies

Answer the question below using the buttons provided.

**Is the galaxy simply smooth and rounded, with no sign of a disk?**



Smooth



Features or disk

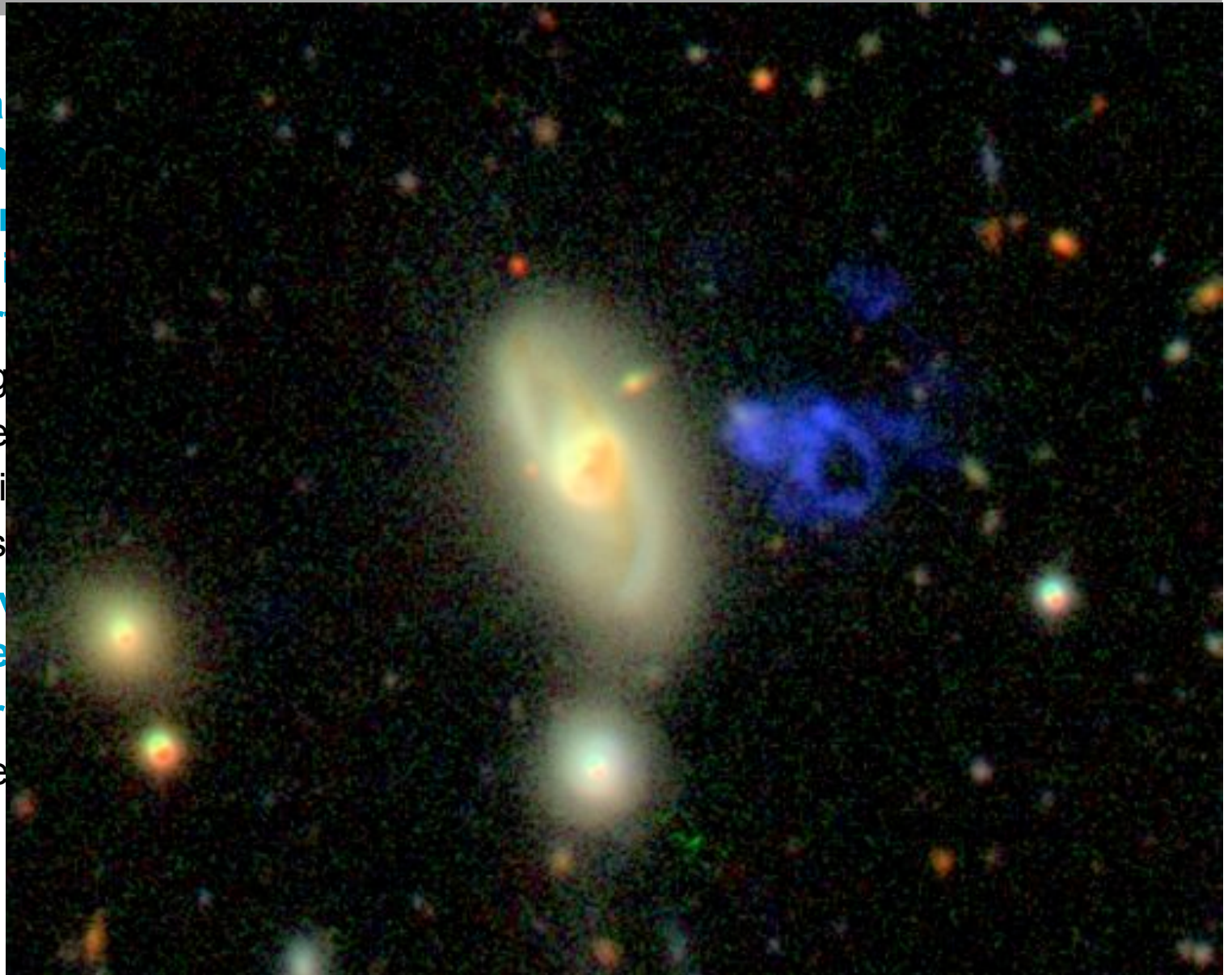


Star or artifact

[Need help?](#) [?](#)

# “Citizen Science”

- Enable  
ava
- Is an  
indi  
par
- g
- e
- fi
- S
- Prov  
dee
- Incr
- e



# Data Analysis

- Can we go even further than GalaxyZoo?
- Can pose astrophysical problems in a way that people without an astrophysics background can tackle them?
- Anyone with a good internet connection can potentially do front-line astronomy
- Can we tap in to brilliant minds that are currently unfulfilled because of lack of access to education and opportunities?
- Use Web2.0/3.0 techniques so people can interact intelligently with the data. (Semantic web?)

**CHALLENGE!**

**YOU ARE NOW LEAVING THE  
MURCHISON RADIO-ASTRONOMY  
OBSERVATORY  
THANK YOU FOR BEING RADIO QUIET**

