



*Advanced European Network of E-infrastructures  
for Astronomy with the SKA*



# AENEAS: An SKA Regional Centre for Europe

Michael Wise

ASTRON (Netherlands Institute for Radio Astronomy)

SKA-Link Kickoff, April 03, 2017





JVLA



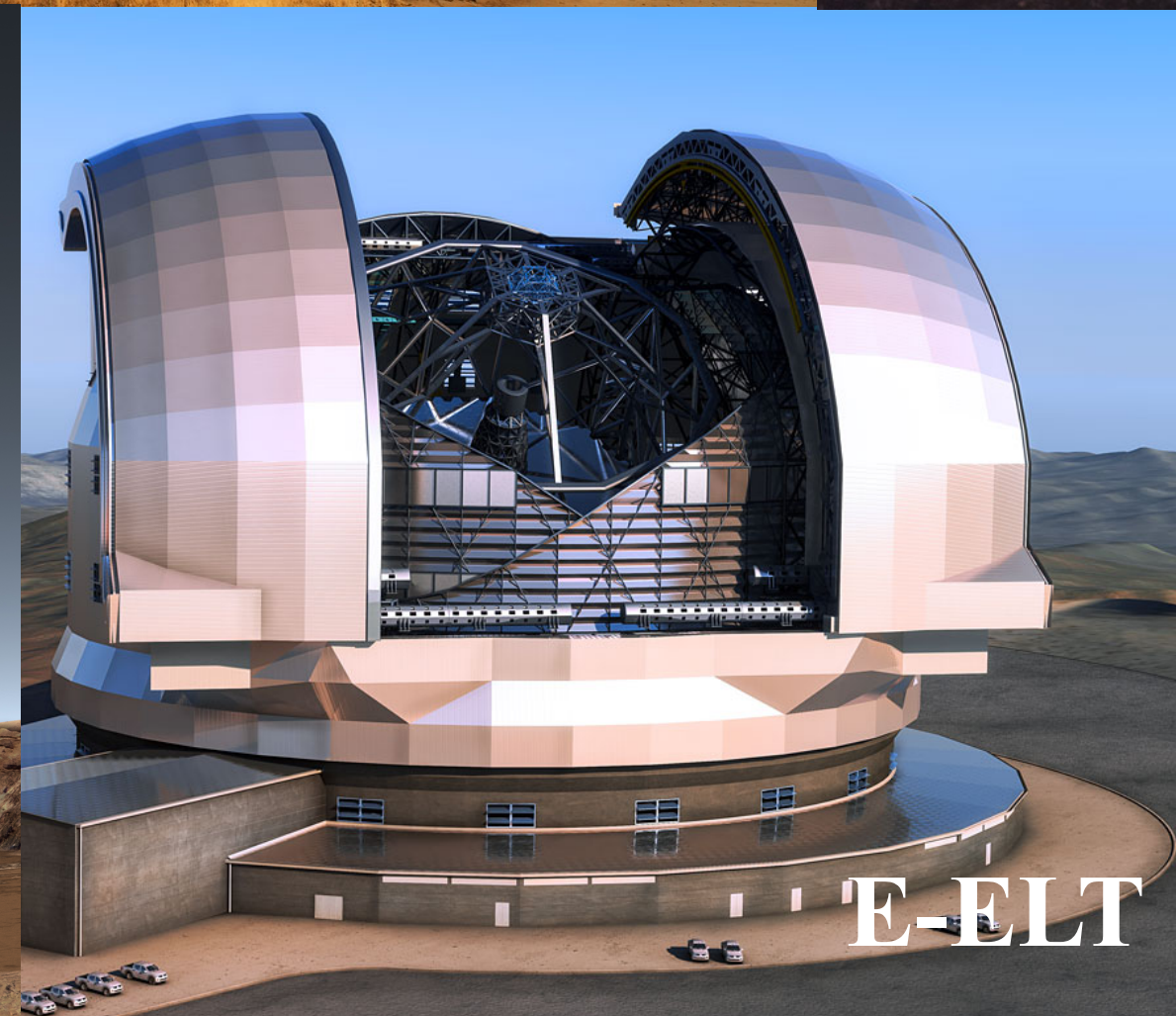
ALMA



LOFAR



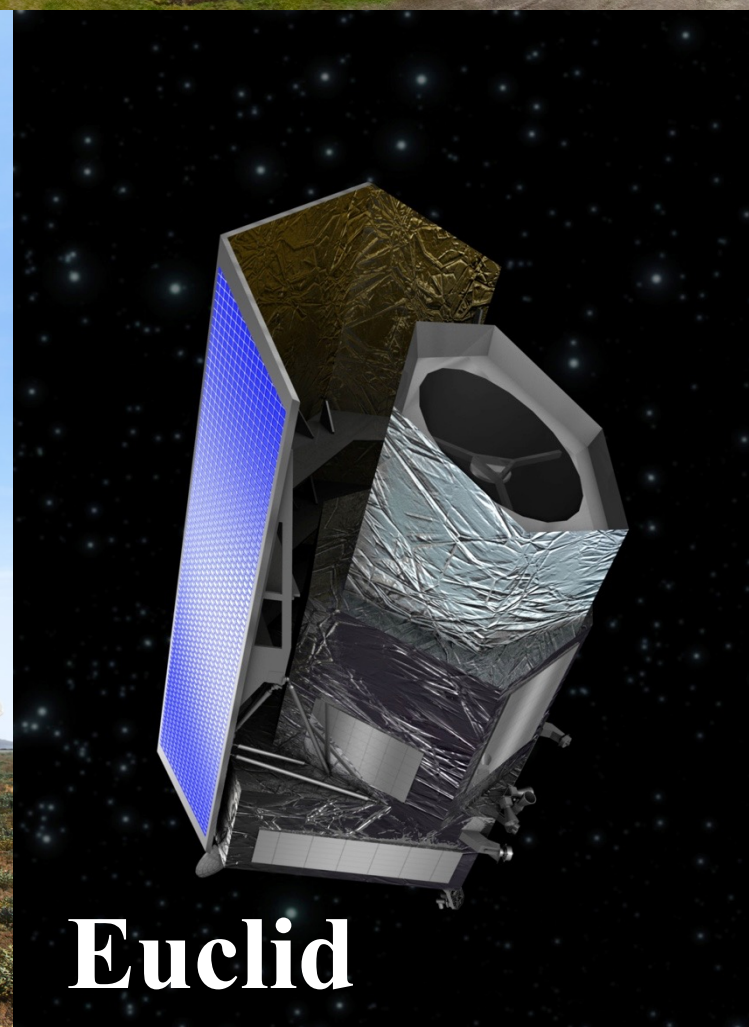
TMT



E-ELT



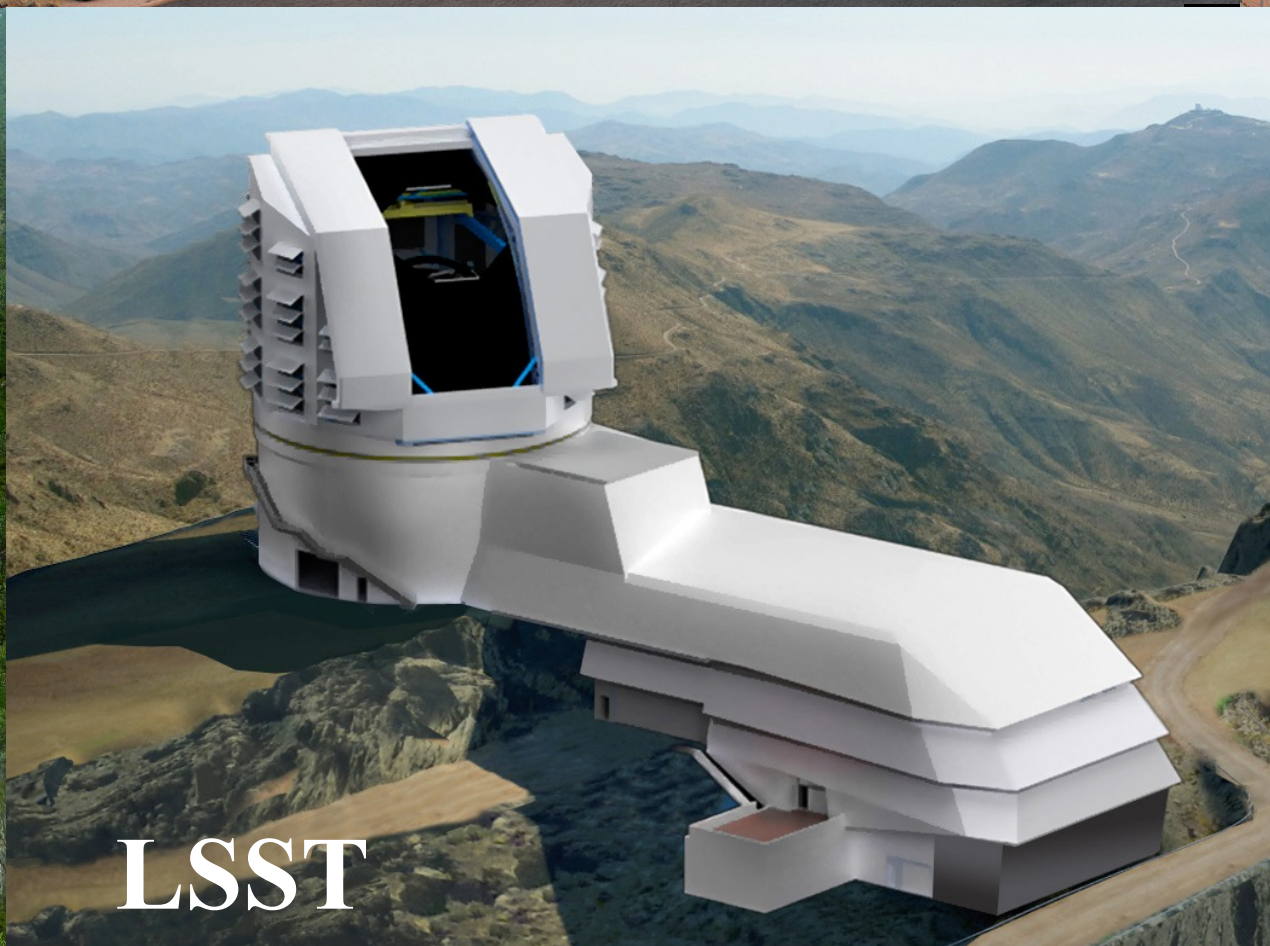
SKA



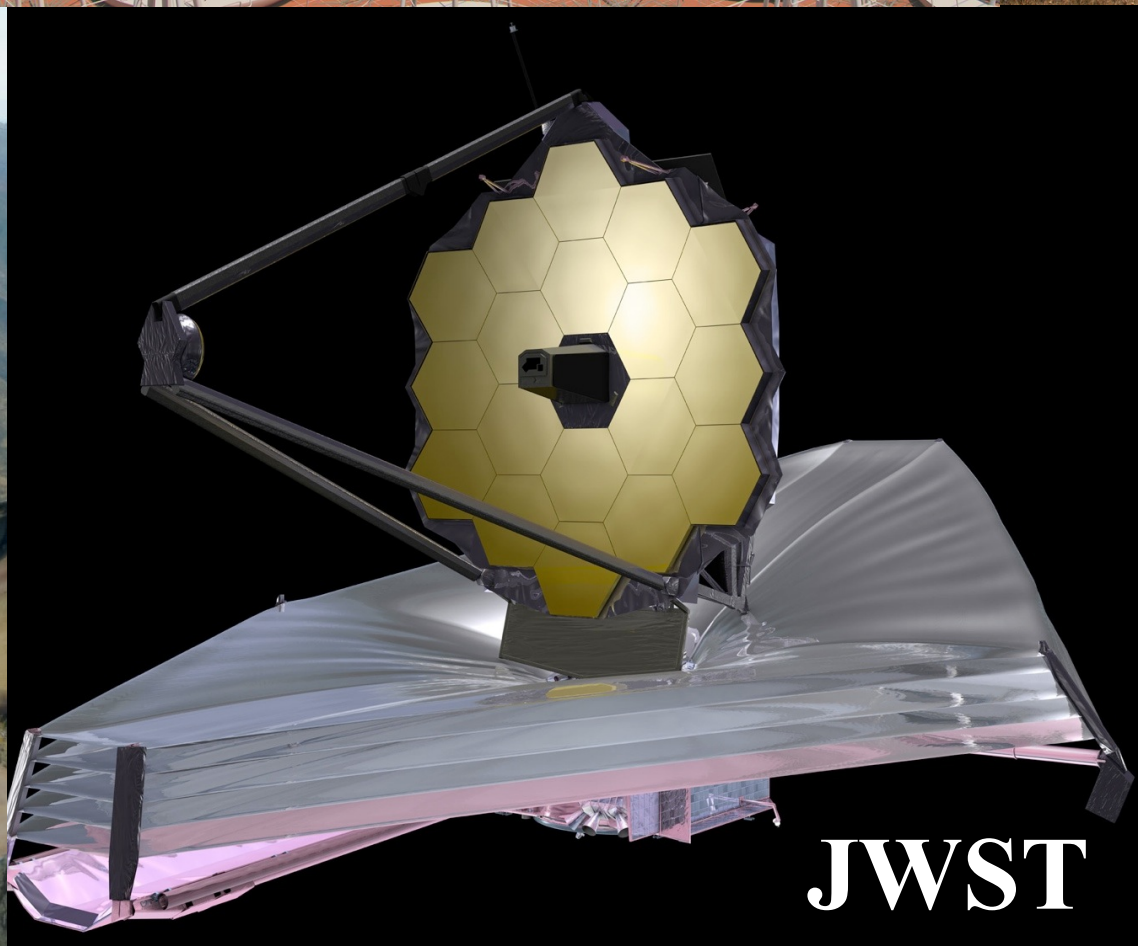
Euclid



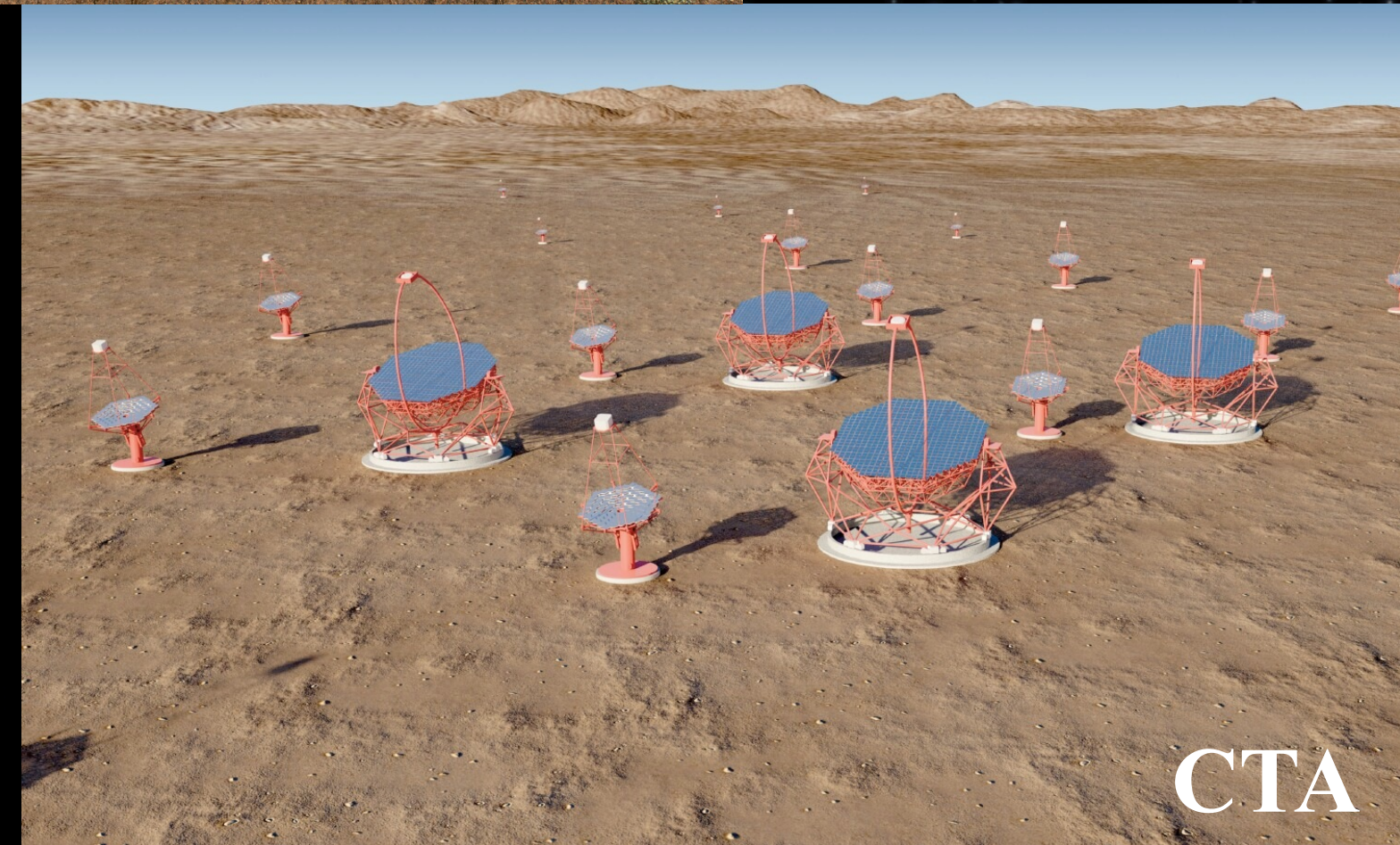
FAST



LSST



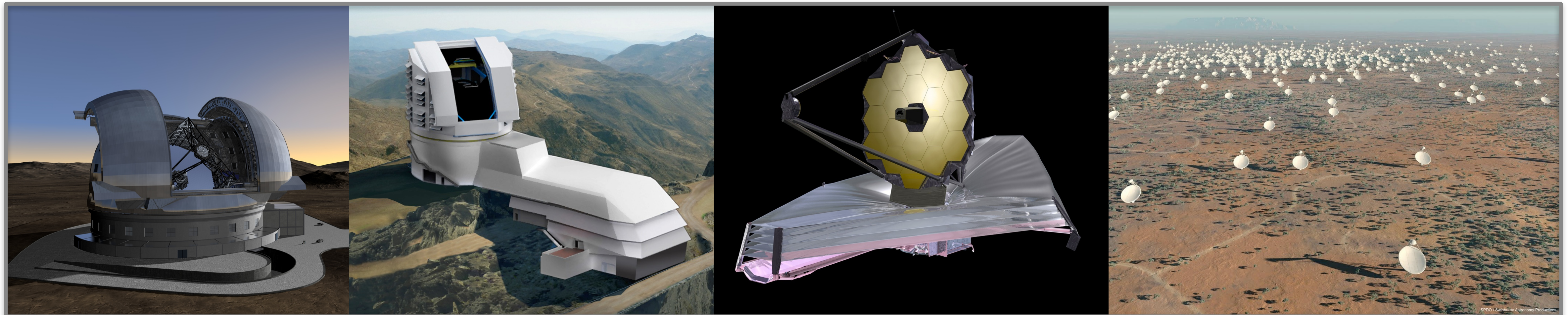
JWST



CTA



# Data Intensive Astronomy



- Science is increasingly driven by large data sets
- Massive data collections and large scientific collaborations
- Most science extraction is based on the archived data
- Current instruments already producing petascale datasets

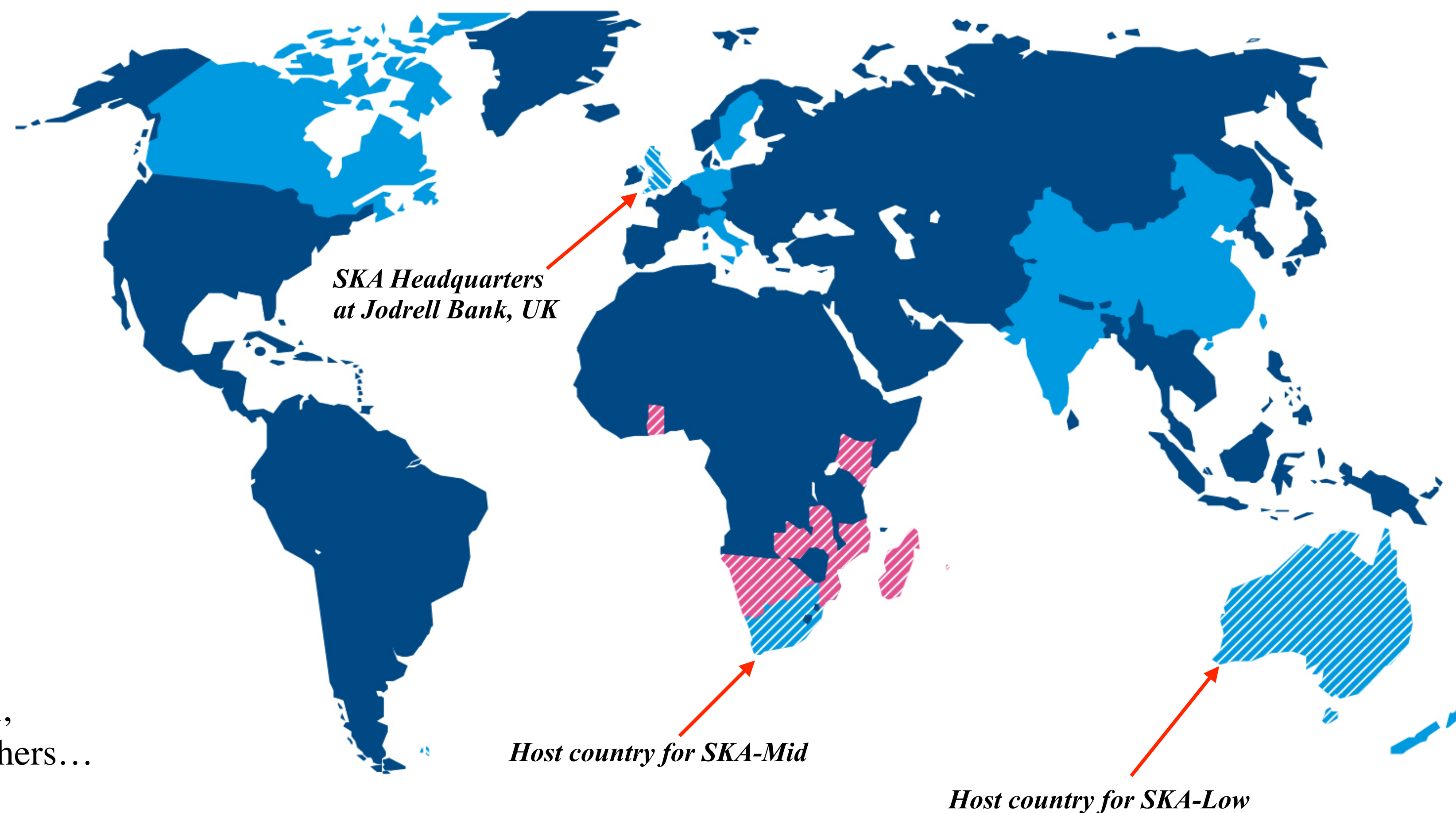
***New science  
infrastructures  
will produce  
exascale data!***



# The Square Kilometre Array

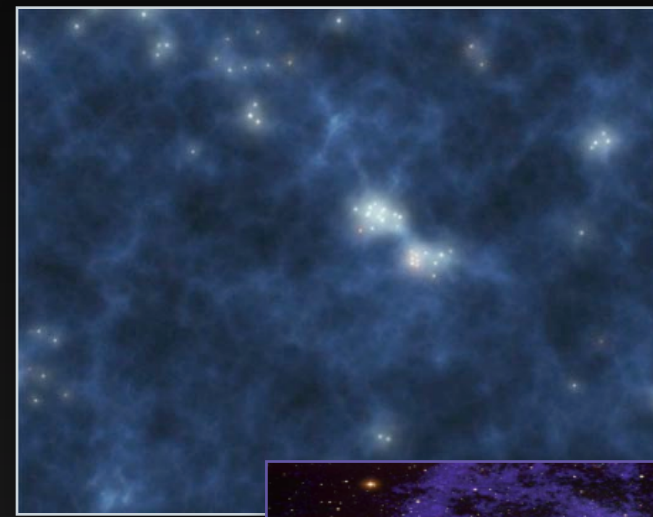
- Australia
- Canada
- China
- India
- Italy
- Netherlands
- New Zealand
- South Africa
- Sweden
- UK

Potential new members: Spain,  
Portugal, Germany, France, others...





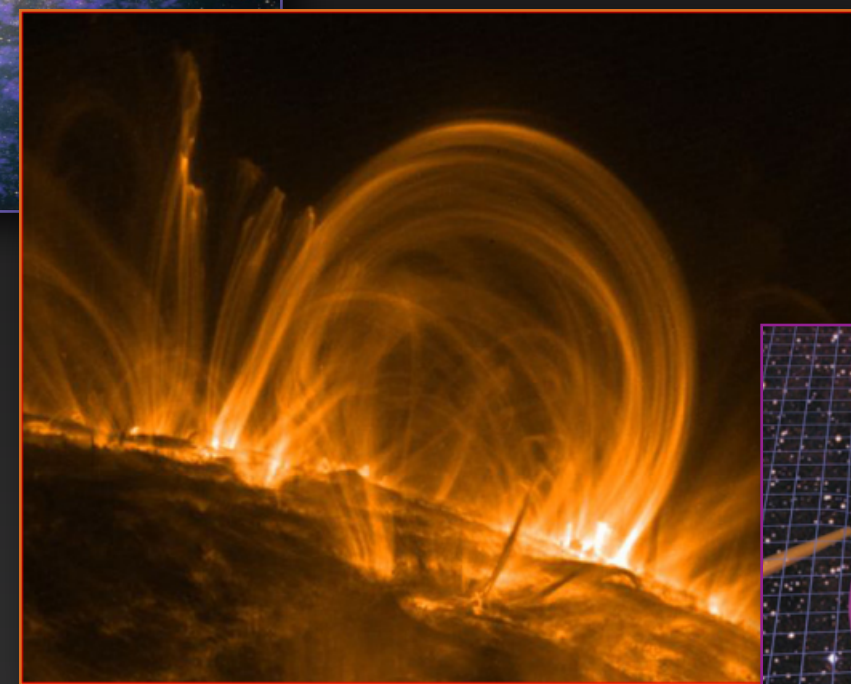
# SKA Science Drivers



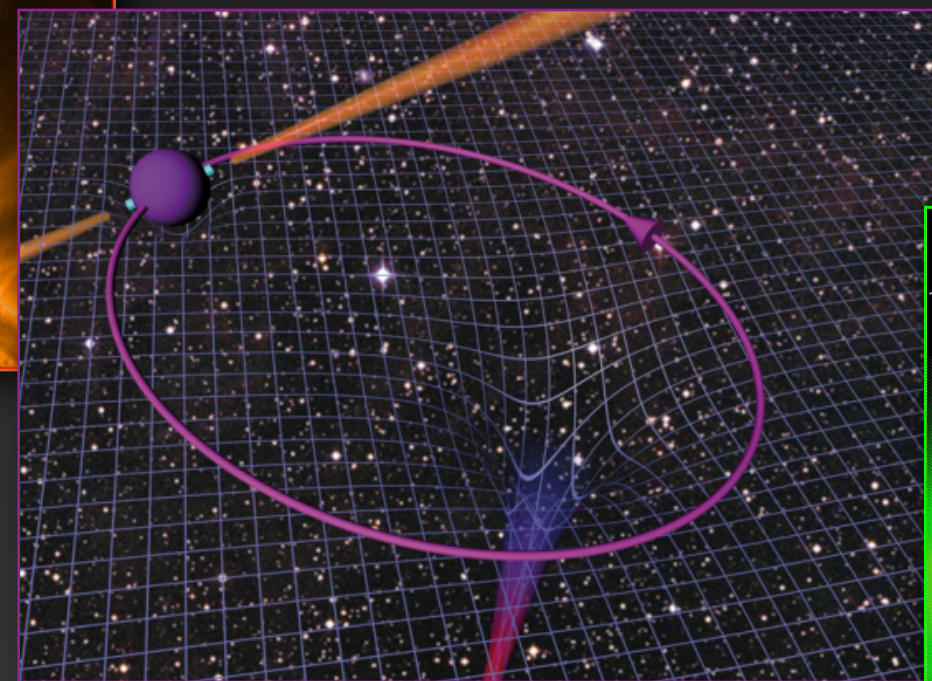
**The First Stars**



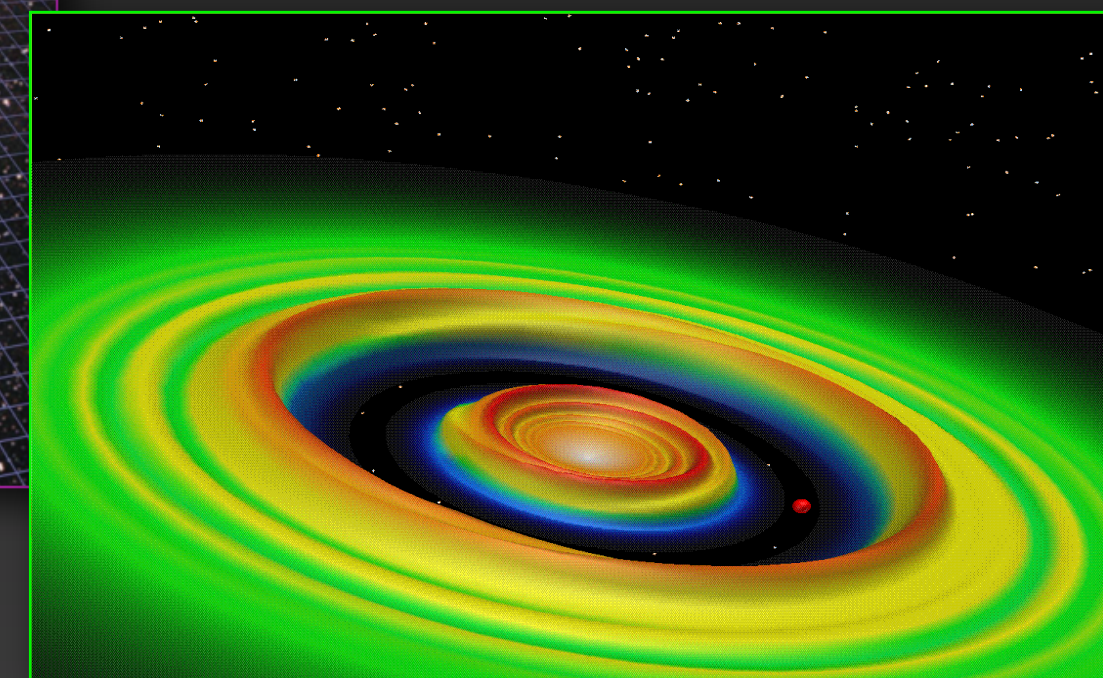
**Cosmic Evolution**



**Cosmic Magnetism**



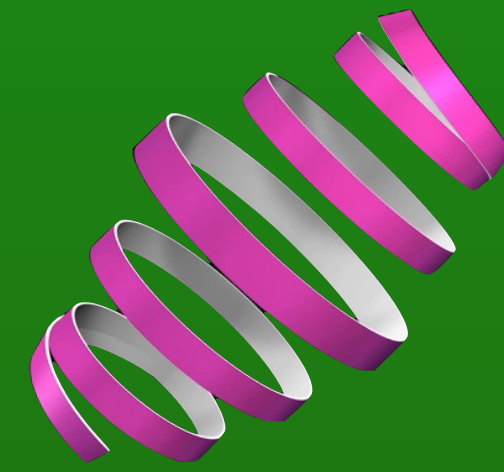
**Gravitational Physics**



**Origins of Life**



# Current Radio Astronomy Archives



**LOFAR**

Long Term Archive

25PB



6PB

United States  
**Census**  
Bureau

4PB

**NASDAQ**

3PB



LIBRARY OF  
CONGRESS

5PB

1 Petabyte (PB)



# Future SKA Science Archive

searches on  
**Google**  
98PB

uploads to  
**facebook**<sup>®</sup>  
180PB

**YouTube**  
15PB

**CERN**  
15PB

 **LOFAR**  
Long Term Archive

25PB

**IBRA**  
6PB

**United States  
Census**  
4PB

**NASDAQ**  
3PB

**LIBRARY OF  
CONGRESS**  
5PB

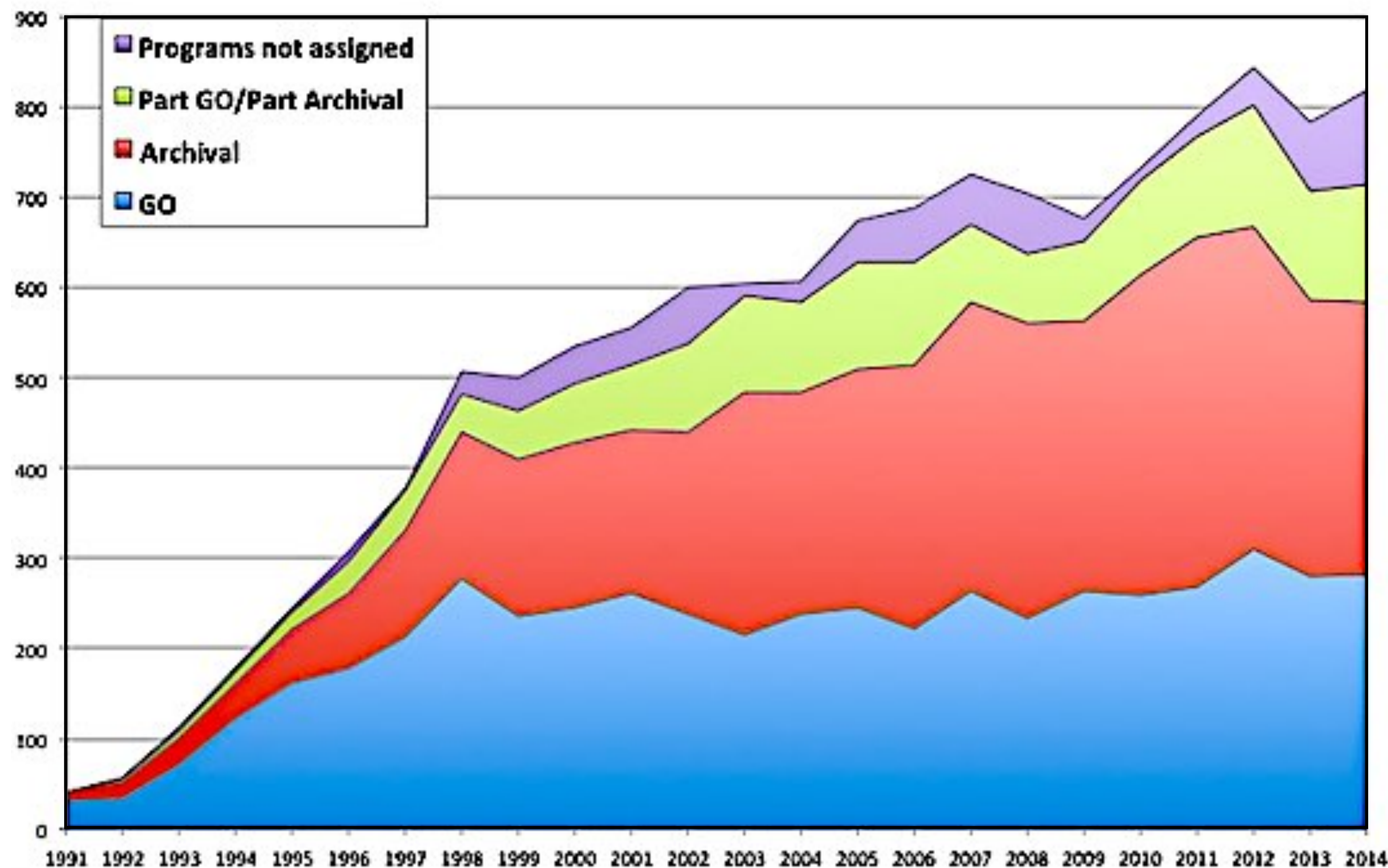
PER YEAR  
1 Petabyte

**SKA**  
Phase1 Science Archive

300PB



# Impact of Science Archives



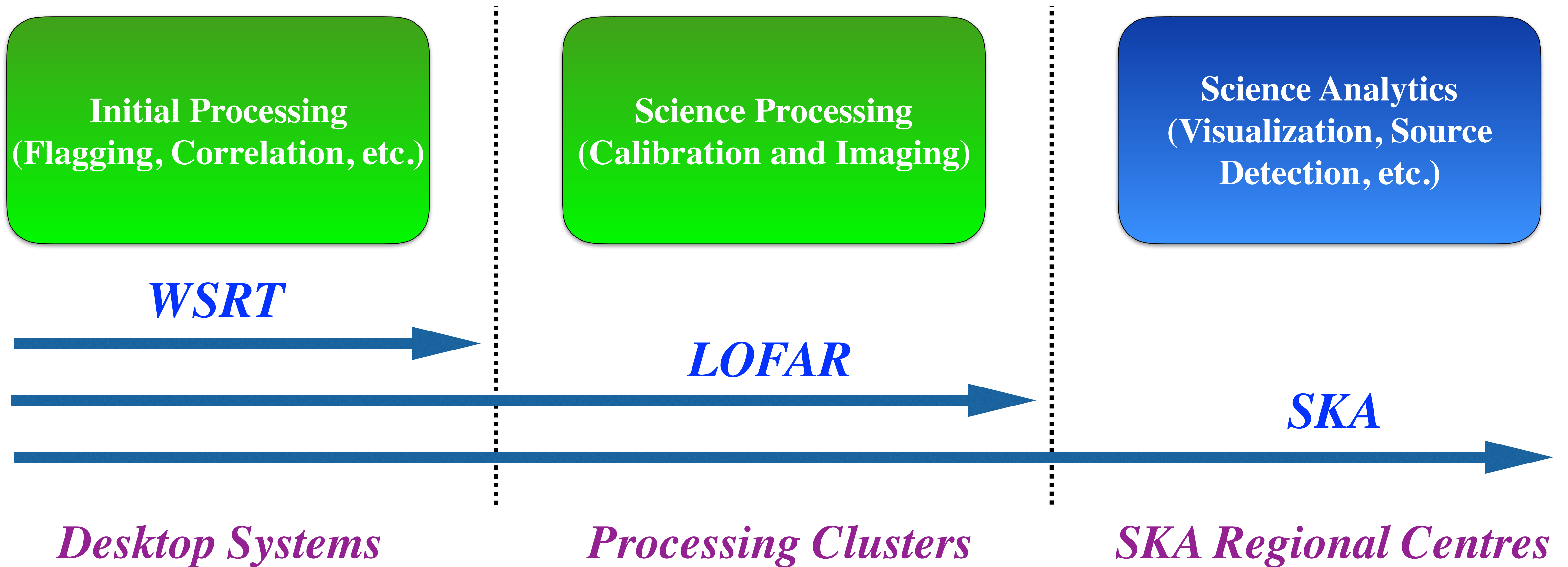
*Hubble Space Telescope Publication Rate*

*Science archives increase the total science output!*

- Assumes the archives are persistent
- Assumes archival data is accessible
- Assumes data products stored are appropriate for general use
- Assumes users have sufficient resources to fully extract science



# Full Cost of Science Extraction





# What is an SKA Regional Centre?

**Astronomers**

*Place to find their data  
Place to analyze their data  
Place to find support*

**Funding Agencies**

*Way to fund science  
Way to maximize investment  
Way to spur innovation*

**Observatories**

*Place to curate their data  
Place to develop new capabilities  
Place to support their users*

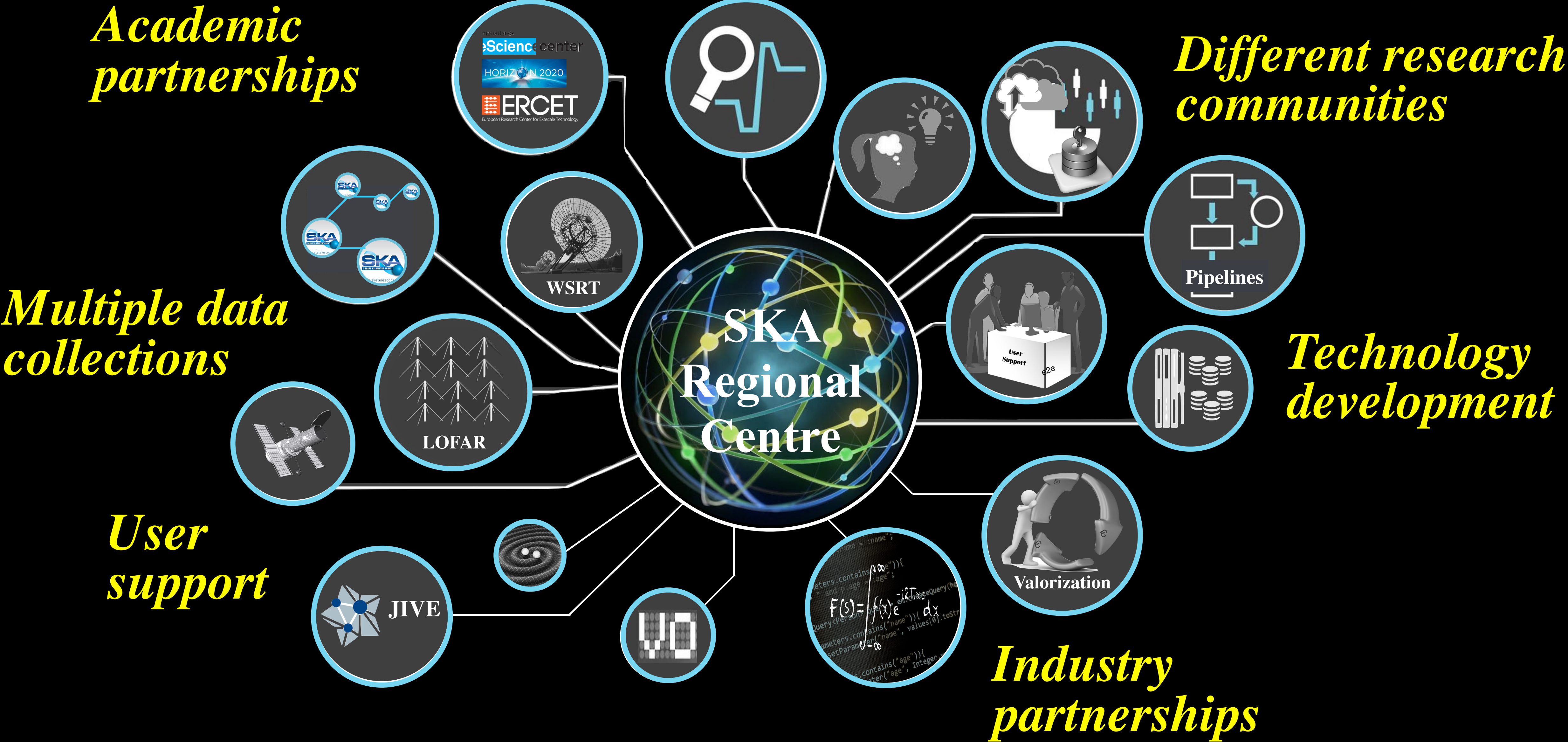
**Industry**

*Place to find new challenges  
Place to collaborate with academia  
Place to test new technologies*





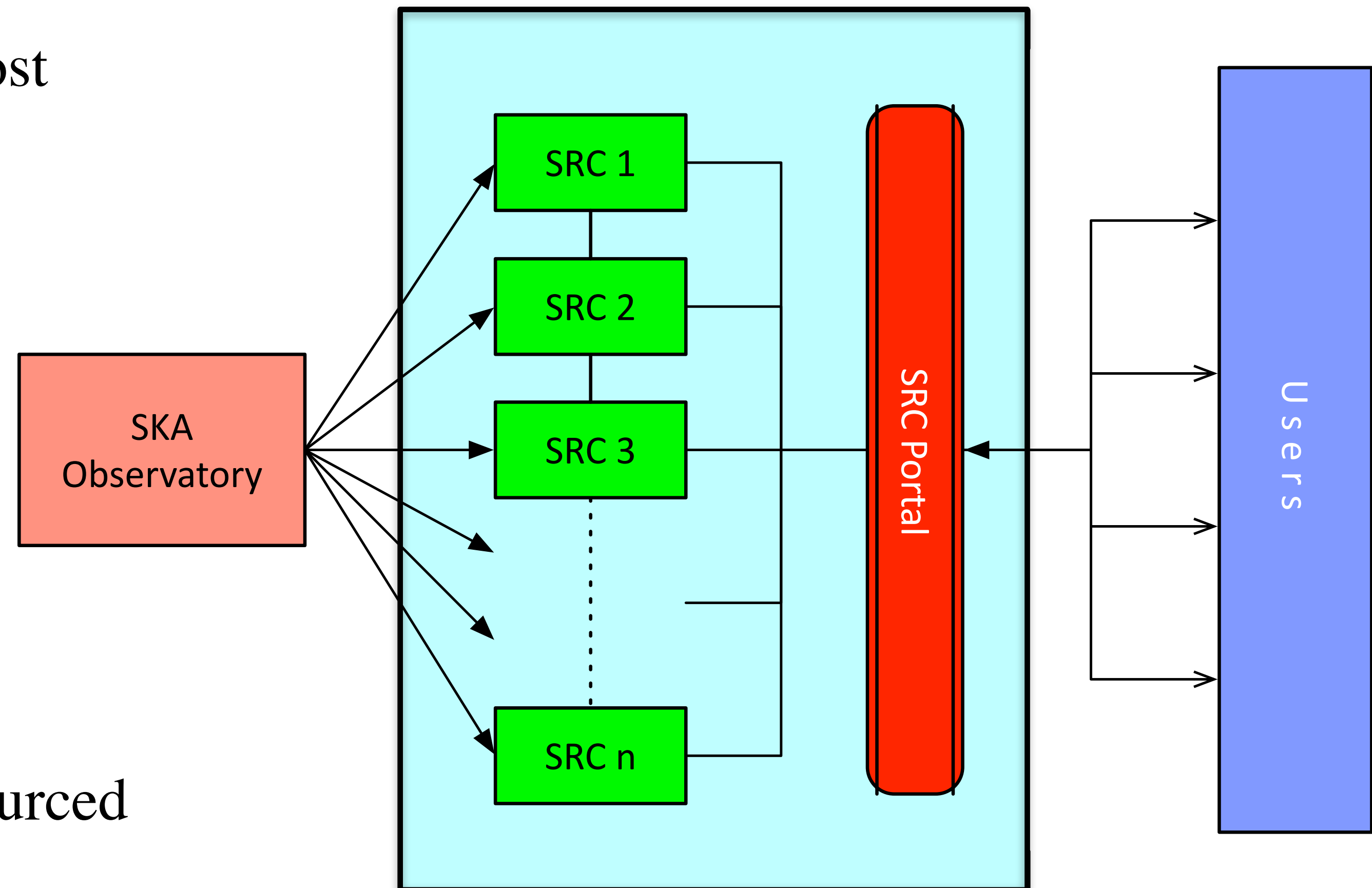
# Connections Beyond Astronomy





# SKA Regional Centres

- Science Data Centres (SDCs) will host the SKA science archive
- Provide access and distribute data products to users
- Provide access to compute and storage resources for users
- Provide analysis capabilities
- Provide user support
- Multiple regional SRCs, locally resourced





# Global Network of Centres

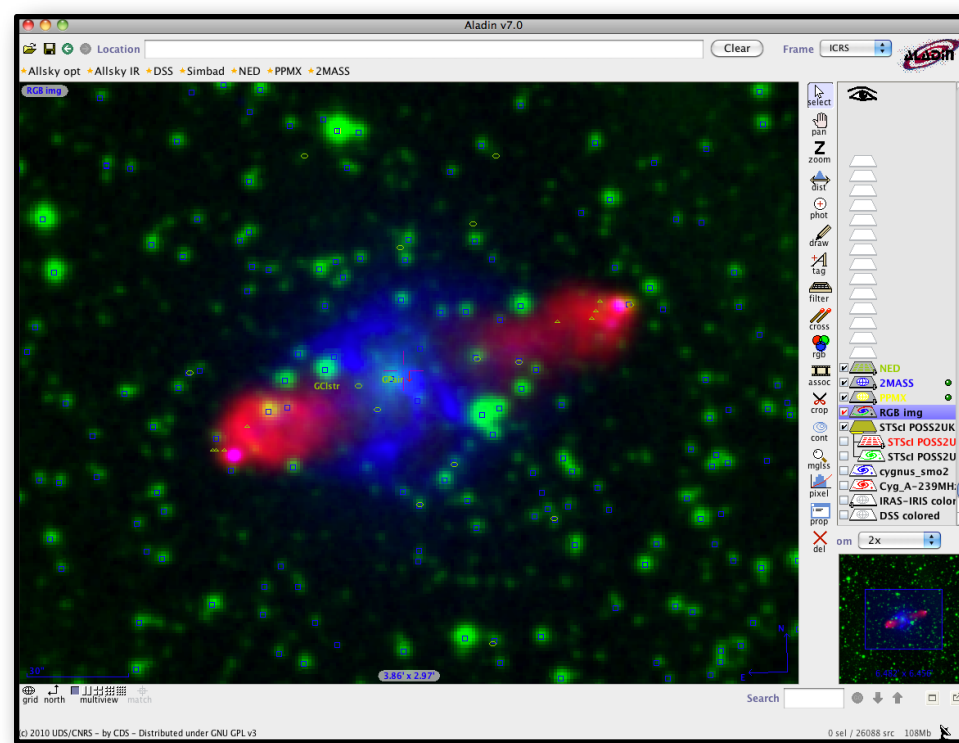




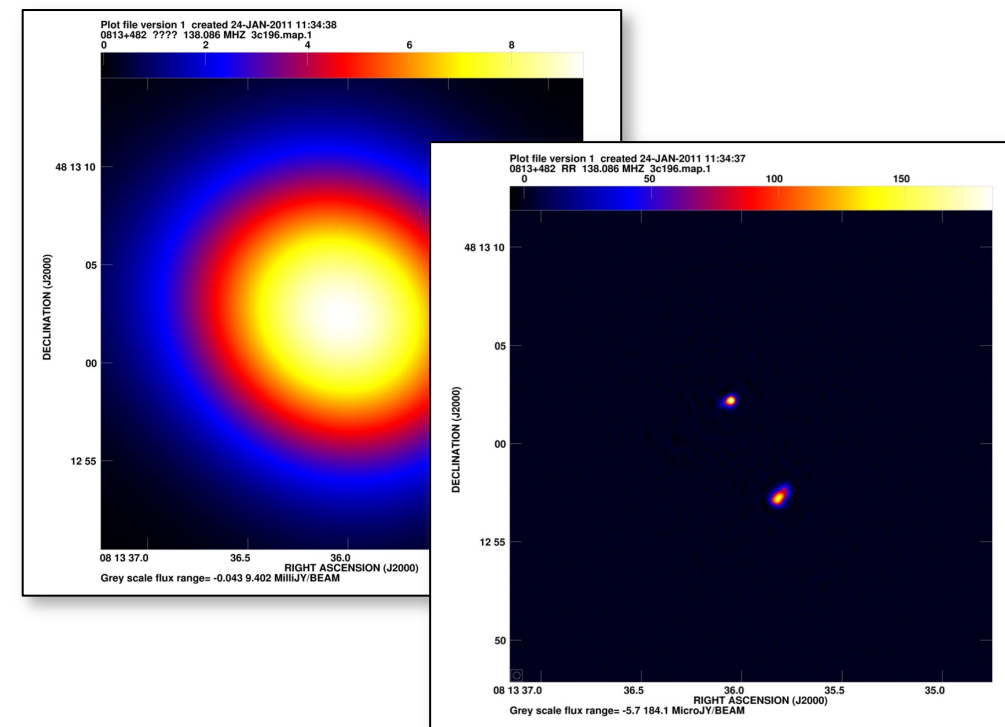
# Regional Centre Functionality

## Data Discovery

- Observation database
- Quick-look data products
- Flexible catalog queries
- Integration with VO tools
- Publish data to VO



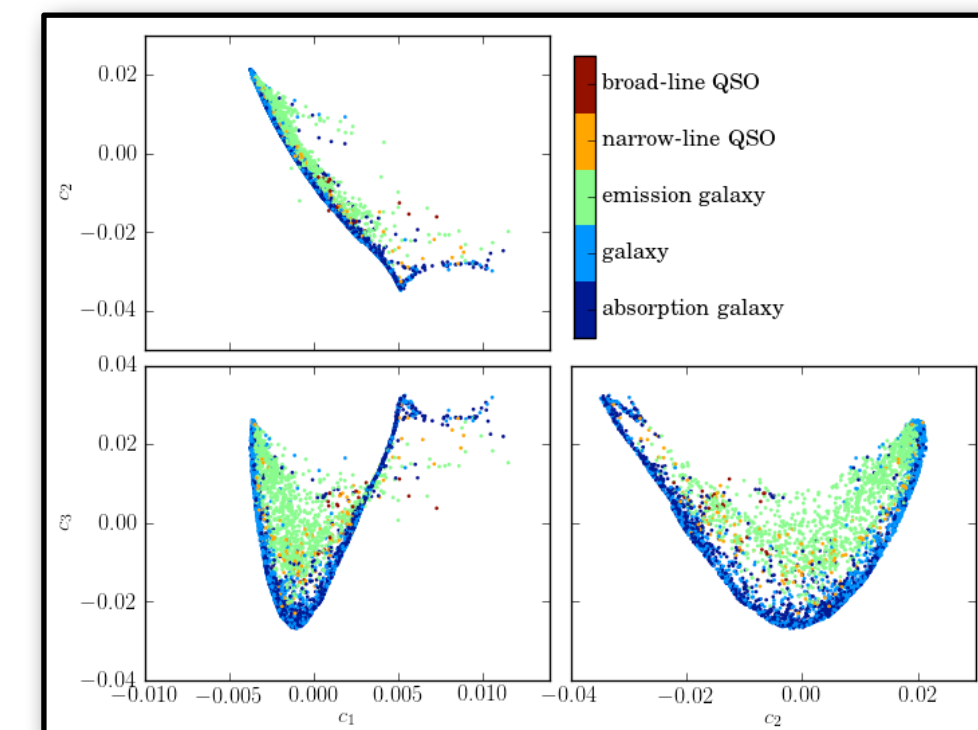
## Data Processing



- Reprocessing
- Calibration and imaging
- Source extraction
- Catalog (re-)creation
- DM searches

## Data Mining

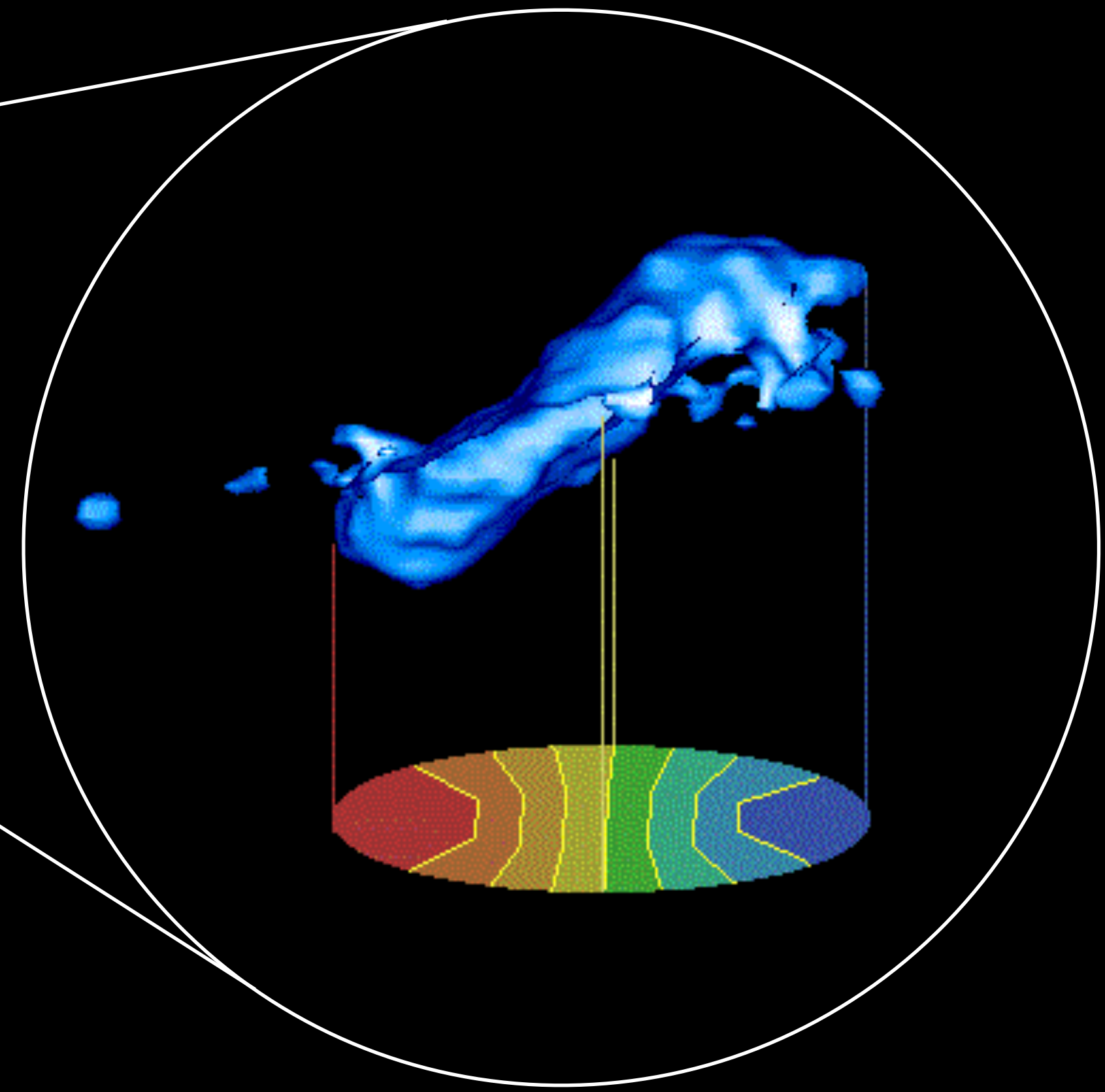
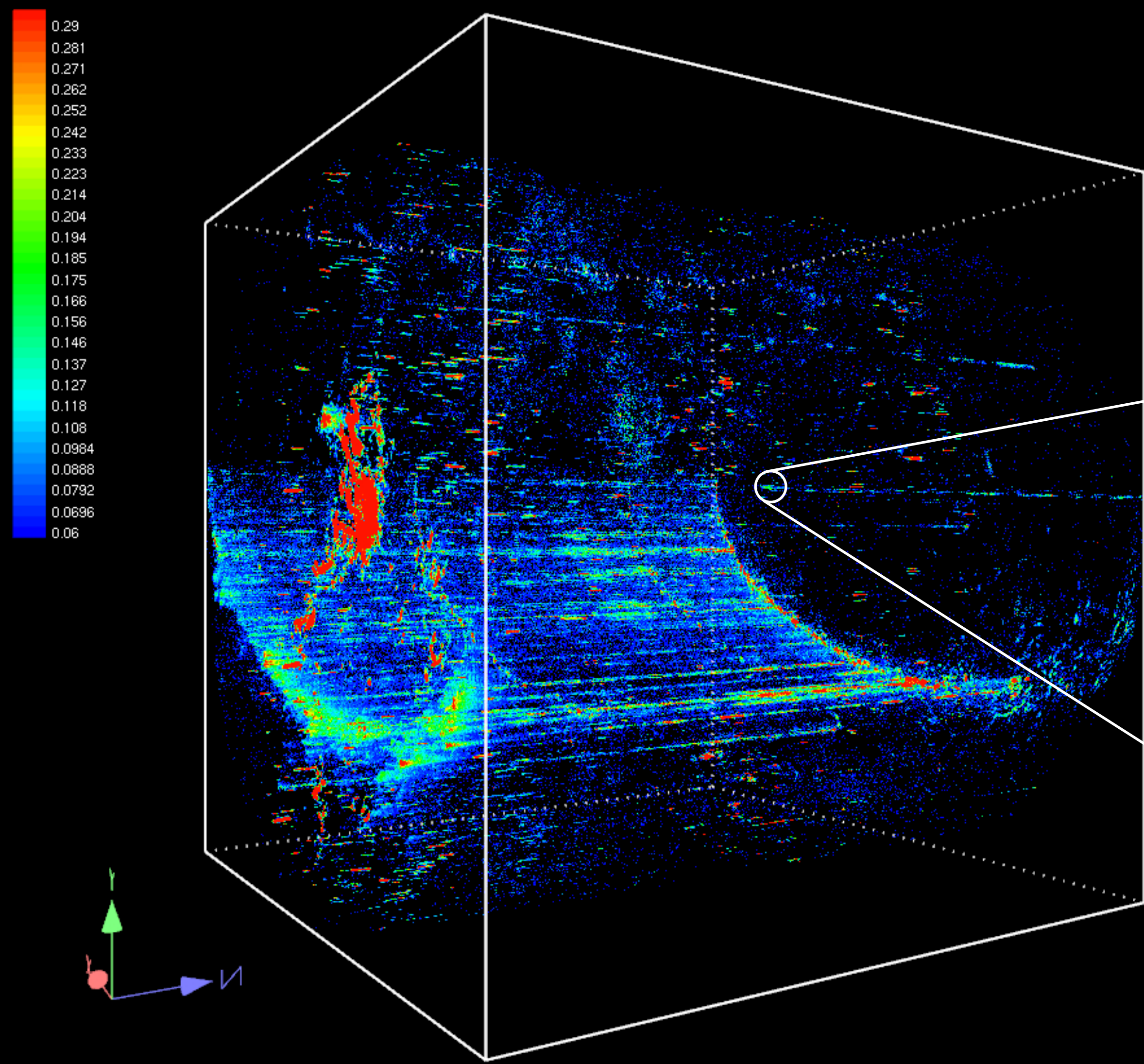
- Multi-wavelength studies
- Catalog cross-matching
- Transient classification
- Feature detection
- Visualization





# Visualization, Classification, Inference

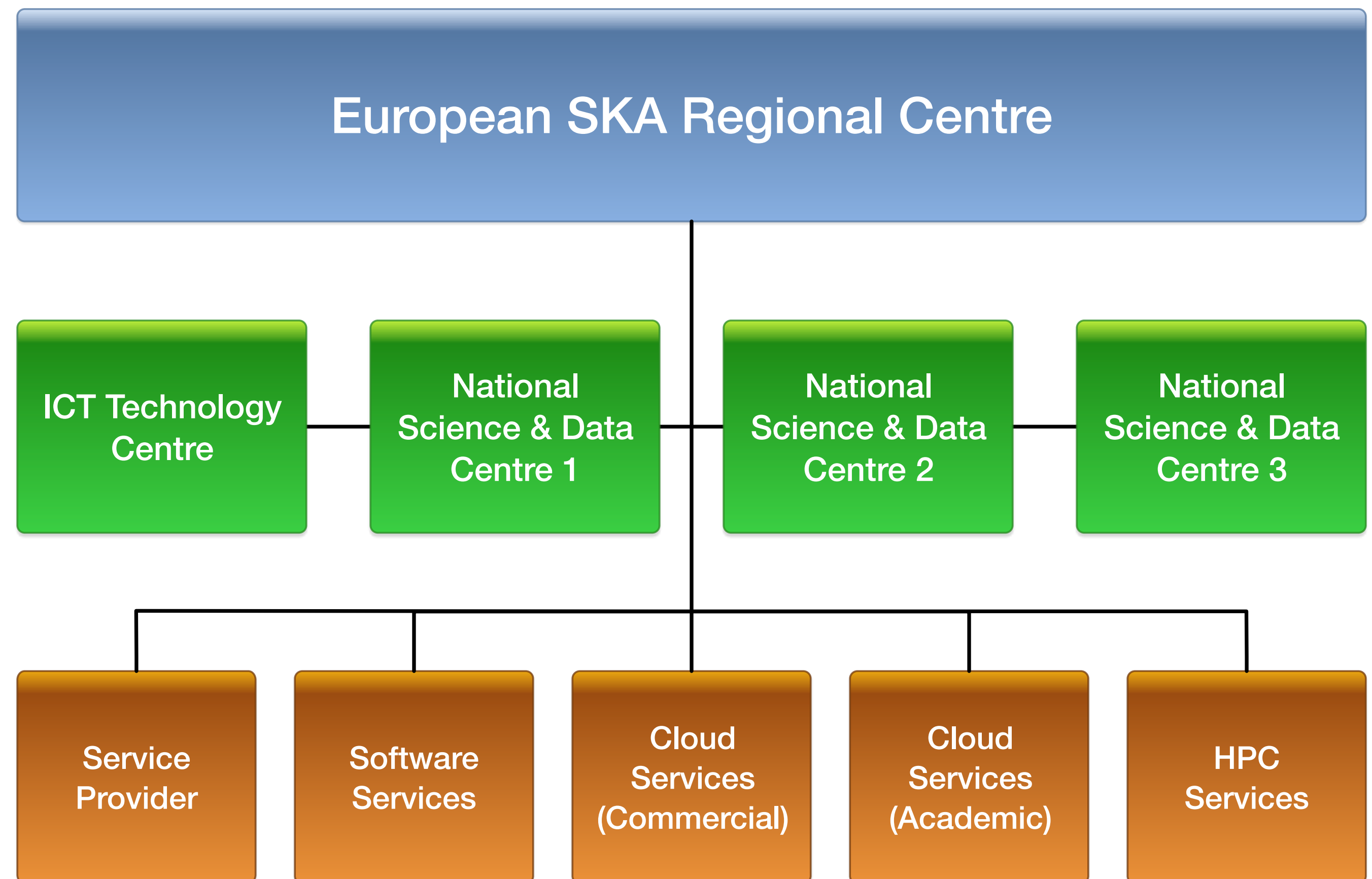
*Automated detection and calculation of galaxy rotation curves in HI surveys*





# European SKA Regional Centre

- Create a European-scale, federated Regional Centre for the SKA
- Provide resources for SKA science extraction to users
- Coordination with ICT communities, industry, and service providers
- Facilitate shared development, interoperability, and innovation
- European counterpart for engagement with other SRCs internationally





# Open Questions

Where will the SKA science archive data be hosted?

How will that data be transported from the sites to Europe?

How can we take optimal advantage of existing infrastructure?

What are the processing requirements and technologies to consider?

What interfaces, tools, and techniques will users need for analysis?

How do we setup and operate an international network of SRCs?





*Design and specification of a distributed, European SKA  
Regional Centre to support the pan-European astronomical  
community in achieving the scientific goals of the SKA*

*EC Horizon 2020 (€3 million)*

*13 countries, 28 partners, SKAO, host countries,  
e-infrastructures (EGI, GÉANT, RDA), NREN's*

*Three year project (2017-2019)*

Advanced European Network of E-infrastructures  
for Astronomy with the SKA

- WP1: Project Management
- WP2: Governance Structure and Business Models
- WP3: Computing and Processing Requirements
- WP4: Data Transport and Optimal European Storage Topologies
- WP5: Data Access and Knowledge Creation
- WP6: User Services





# Summary

- Regional Centres will be essential for the scientific success of the SKA
- SRCs will be the primary interface for SKA science extraction
- An international network of SRCs will be required to serve the SKA community
- A European SKA Regional Centre will be a key part of that network
- The AENEAS project is a first step in establishing such an European SRC