

## Connecting Scientists to Data in the Era of Global Astronomy

**Michael Wise**

*Head, ASTRON Astronomy Group*

**The Connecting Strength of Big Science Projects**  
**ASTRON, June 10, 2016**

ASTRON is part of the Netherlands Organisation for Scientific Research (NWO)







JVLA



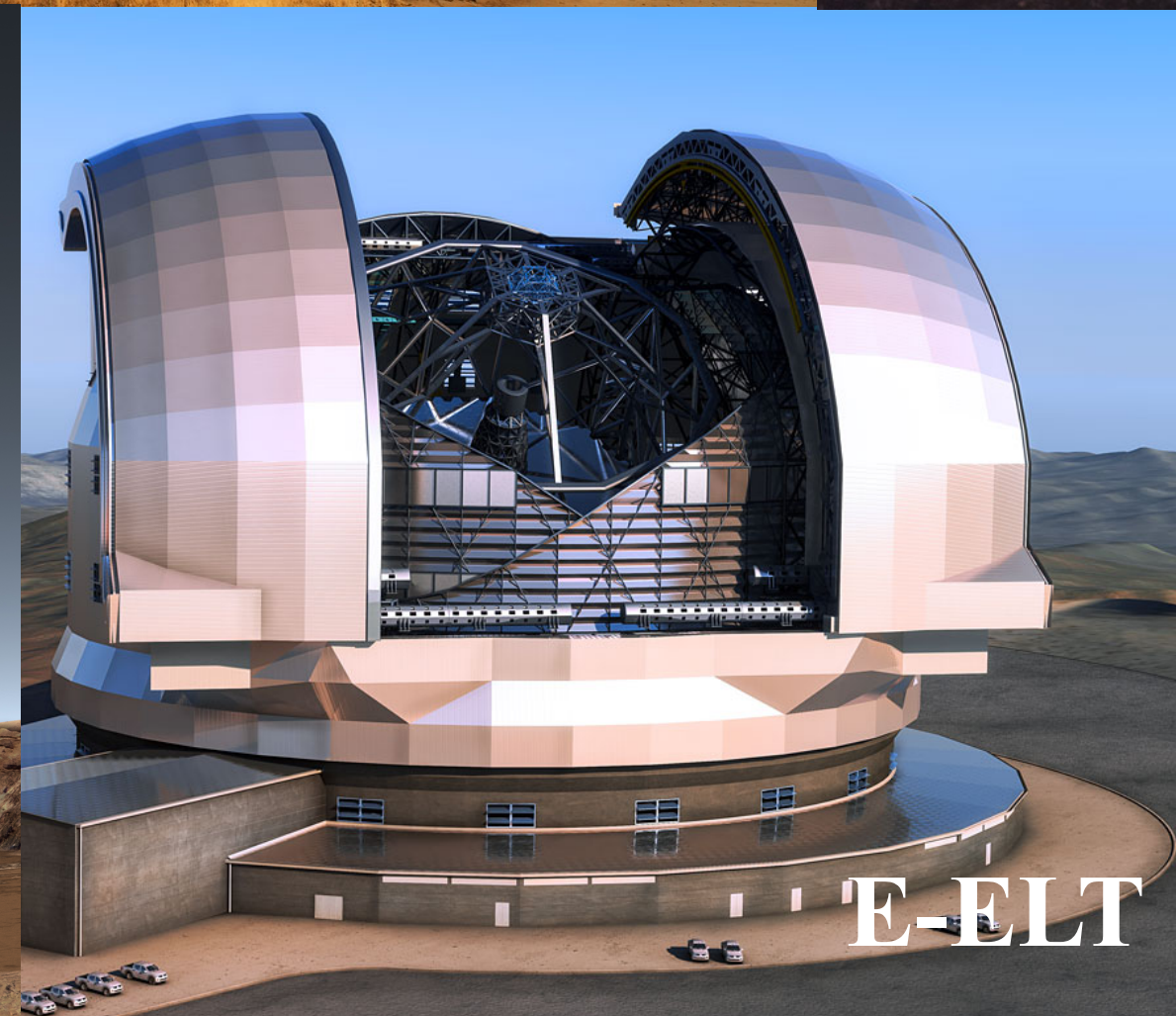
ALMA



LOFAR



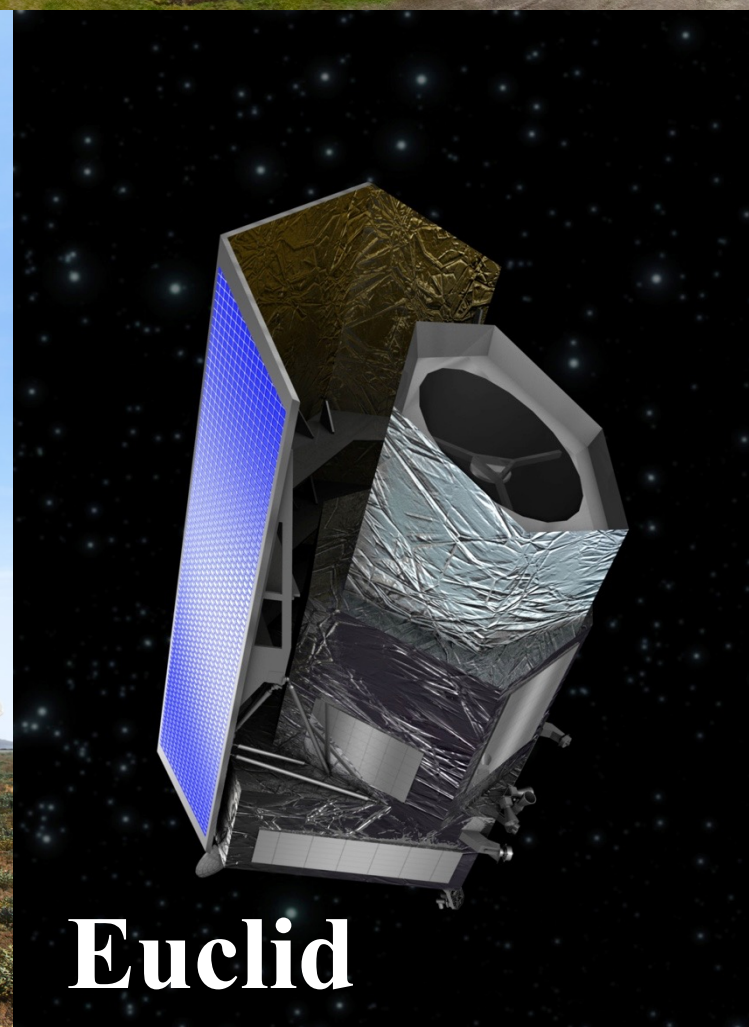
TMT



E-ELT



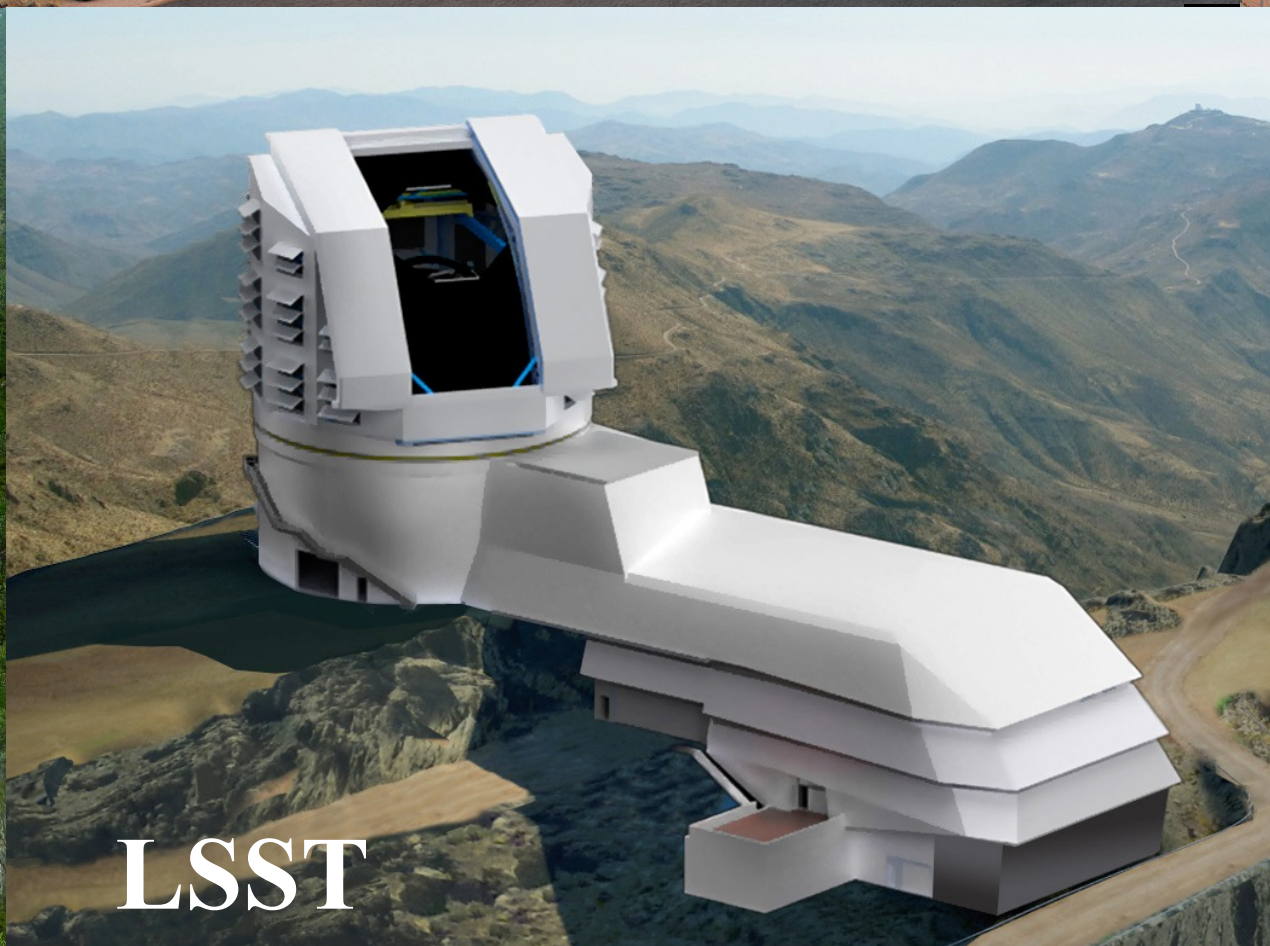
SKA



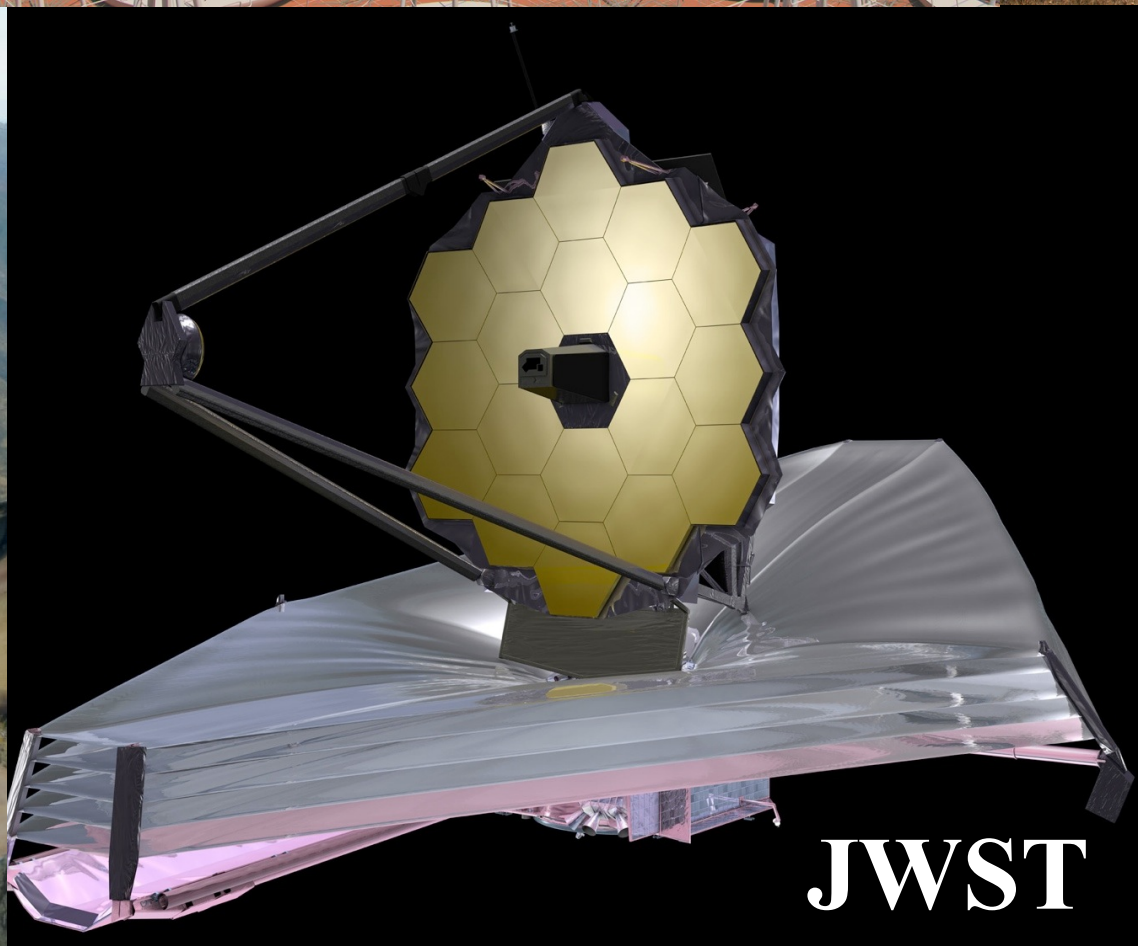
Euclid



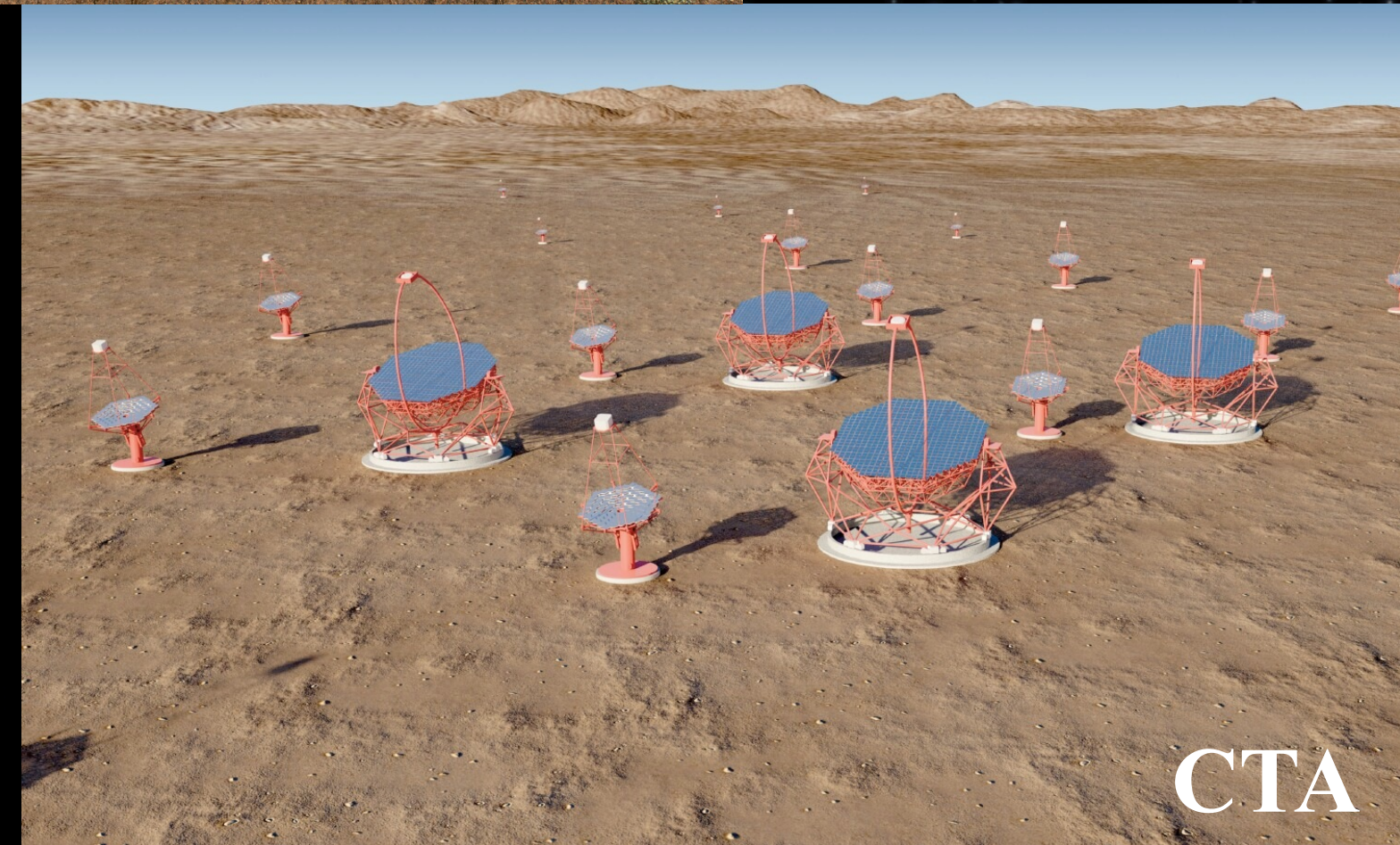
FAST



LSST

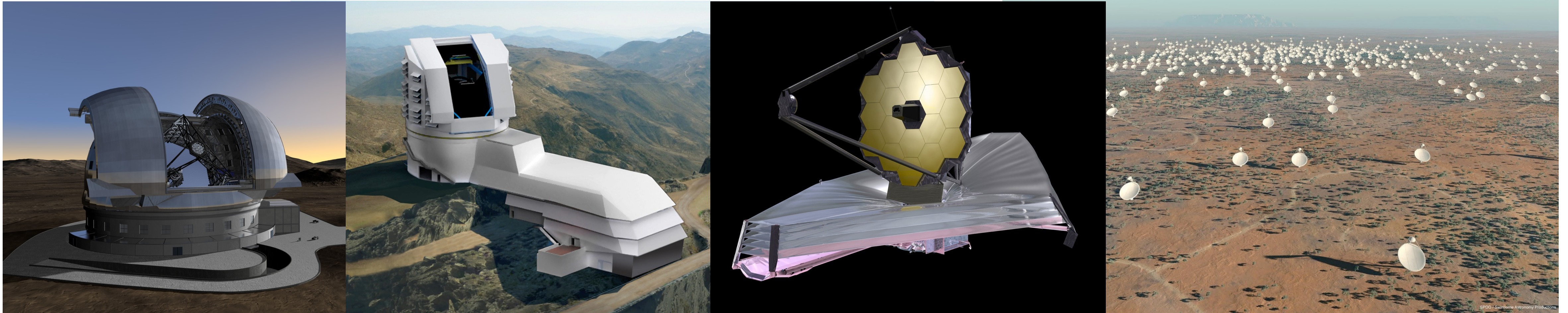


JWST



CTA





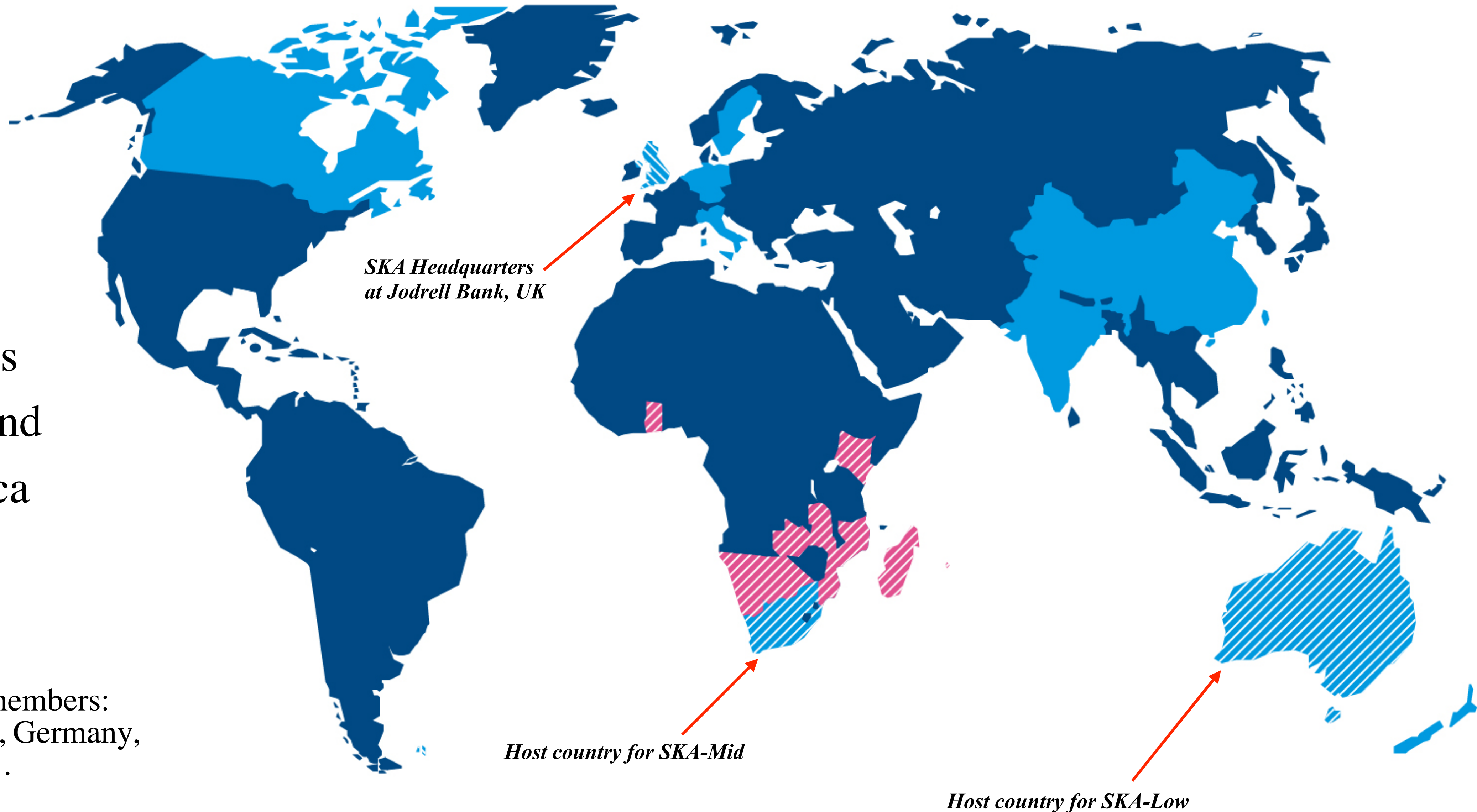
- Cutting edge science is increasingly data intensive
- Massive data collections and large scientific collaborations
- Most science extraction is based on the archived data
- Current instruments already producing petascale datasets

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

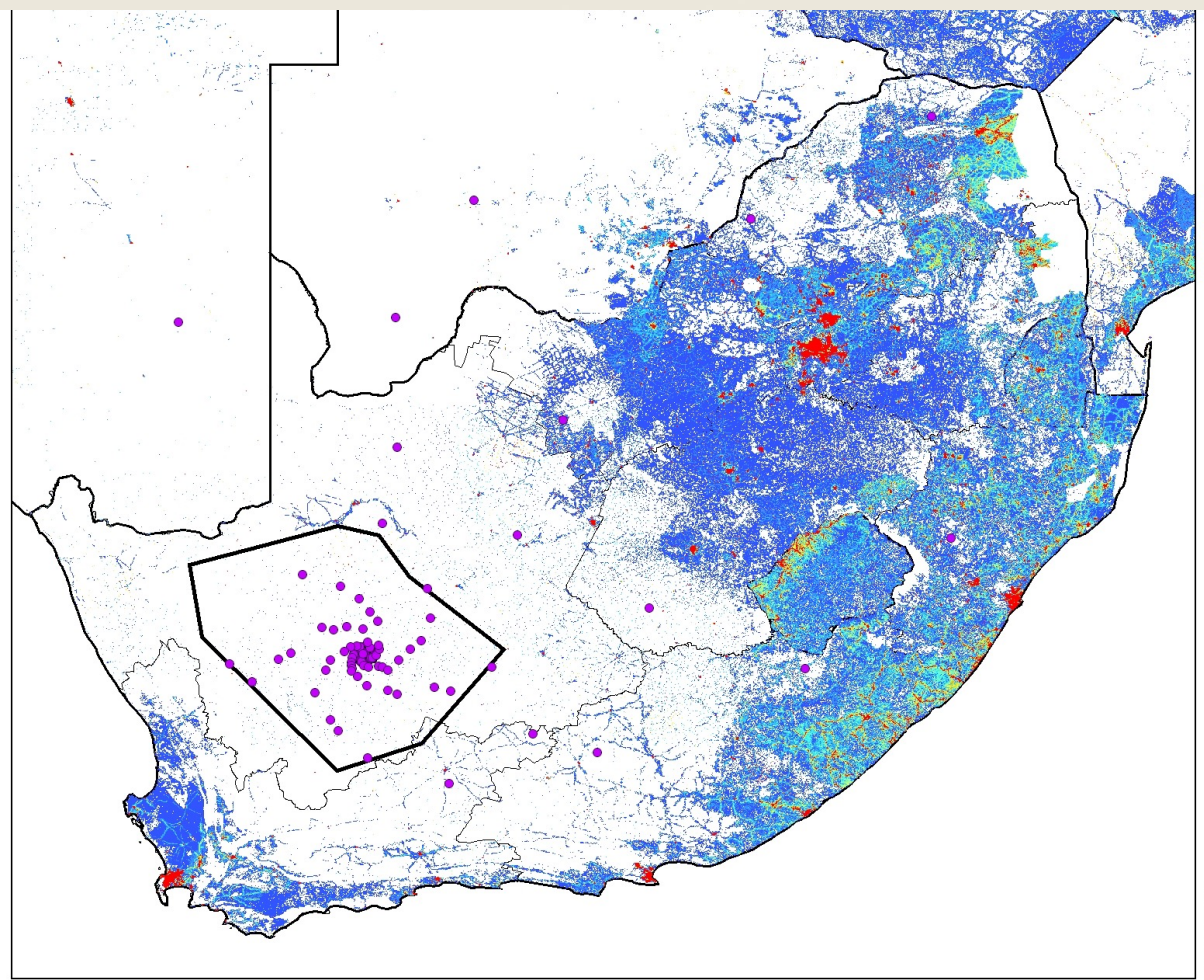


- Australia
- Canada
- China
- India
- Italy
- Netherlands
- New Zealand
- South Africa
- Sweden
- UK

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







## SKA1 MID - the SKA's mid-frequency instrument

The Square Kilometre Array (SKA) will be the world's largest radio telescope, revolutionising our understanding of the Universe. The SKA will be built in two phases - SKA1 and SKA2 - starting in 2018, with SKA1 representing a fraction of the full SKA. SKA1 will include two instruments - SKA1 MID and SKA1 LOW - observing the Universe at different frequencies.

Location: South Africa

Frequency range: **350 MHz** to **14 GHz**

~**200 dishes**  
(including 64 MeerKAT dishes)

Total collecting area: **33,000m<sup>2</sup>**  
or **126 tennis courts**

Maximum distance between dishes: **150km**

Total raw data output:

**2 terabytes** per second

**62 exabytes** per year

Enough to fill **340,000** average laptops with content **every day**

x**340,000**

Compared to the JVLA, the current best similar instrument in the world:

**4x** the resolution

**5x** more sensitive

**60x** the survey speed

# SKA1 MID in South Africa



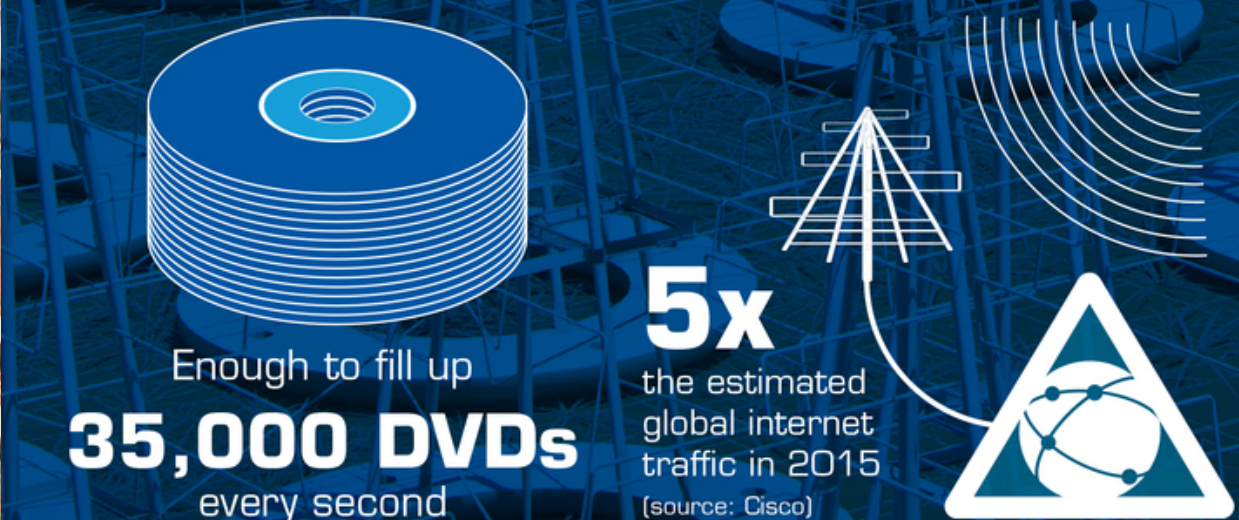
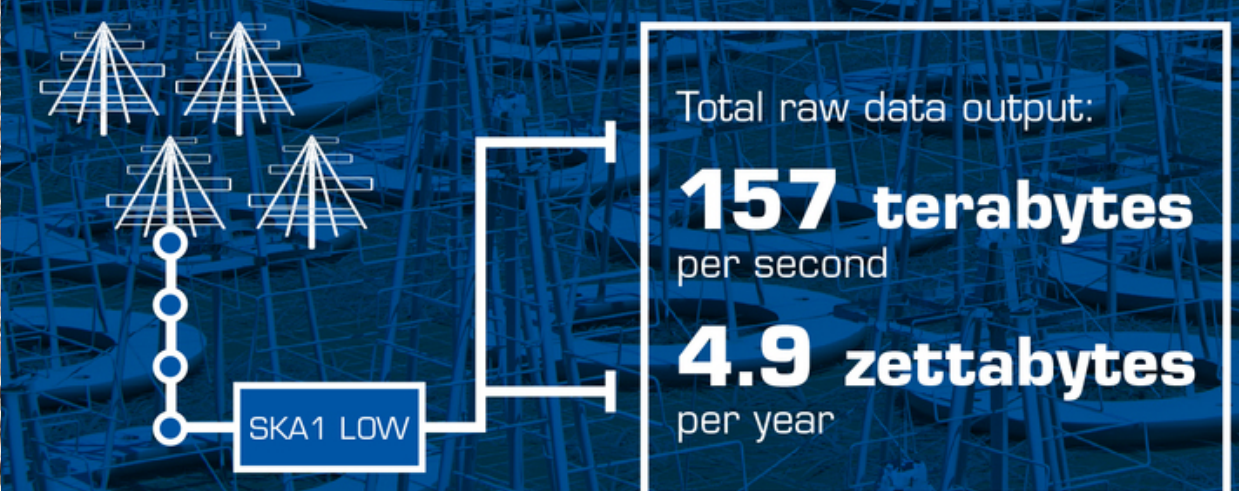
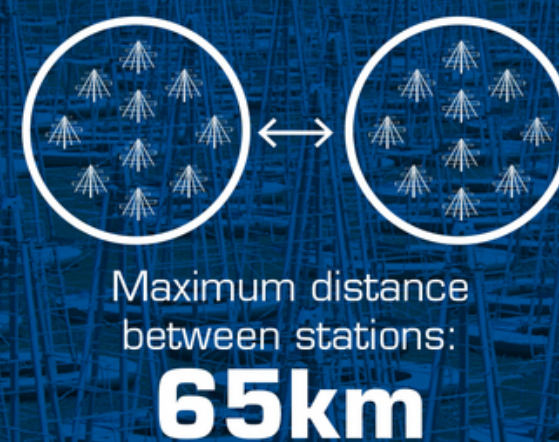
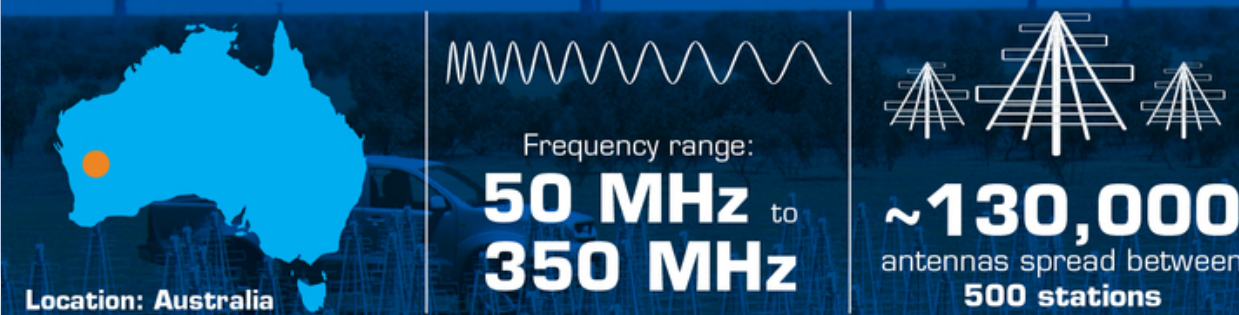


# SKA1 LOW in Australia



## SKA1 LOW - the SKA's low-frequency instrument

The Square Kilometre Array (SKA) will be the world's largest radio telescope, revolutionising our understanding of the Universe. The SKA will be built in two phases - SKA1 and SKA2 - starting in 2018, with SKA1 representing a fraction of the full SKA. SKA1 will include two instruments - SKA1 MID and SKA1 LOW - observing the Universe at different frequencies.

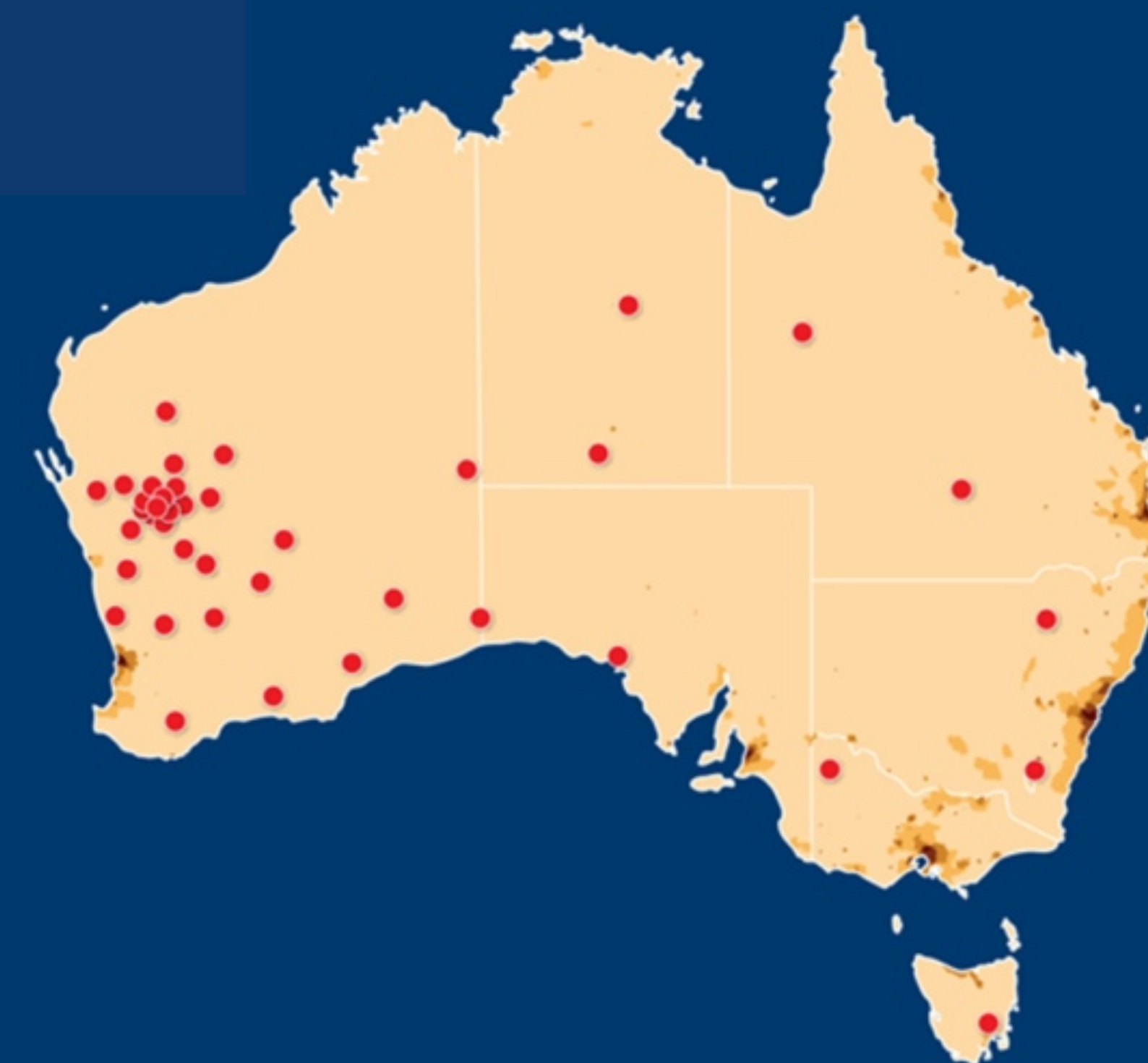
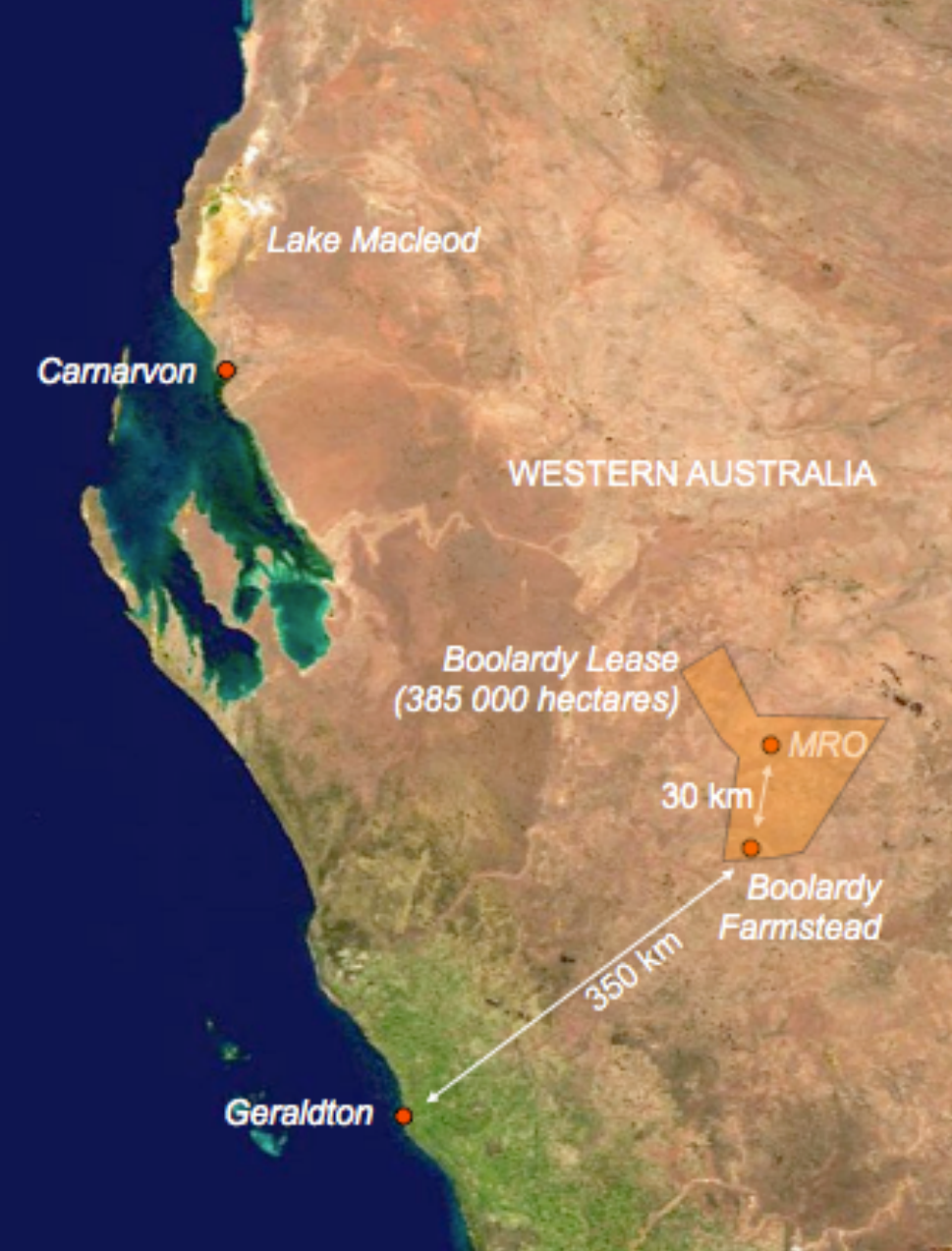


Compared to LOFAR Netherlands, the current best similar instrument in the world

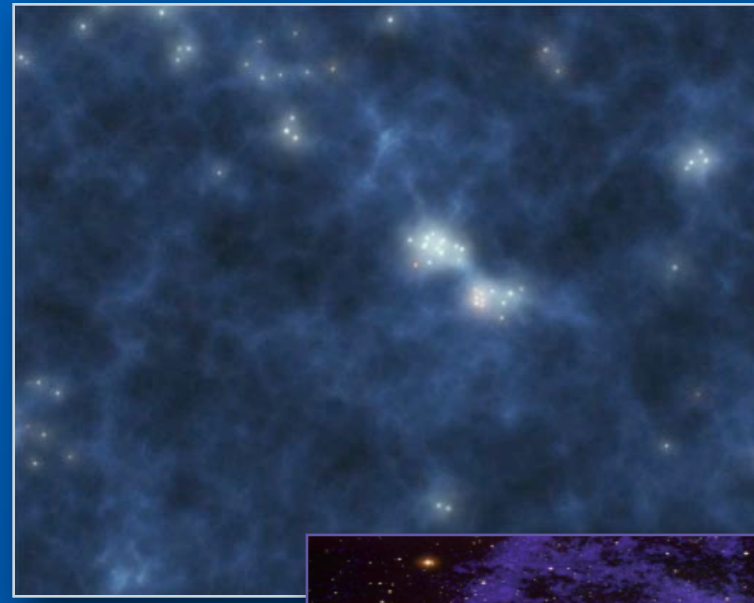


### Shire of Murchison:

- 50,000 km<sup>2</sup> Size of the Netherlands
- 0 gazetted towns
- 29 sheep/cattle stations
- 110 population





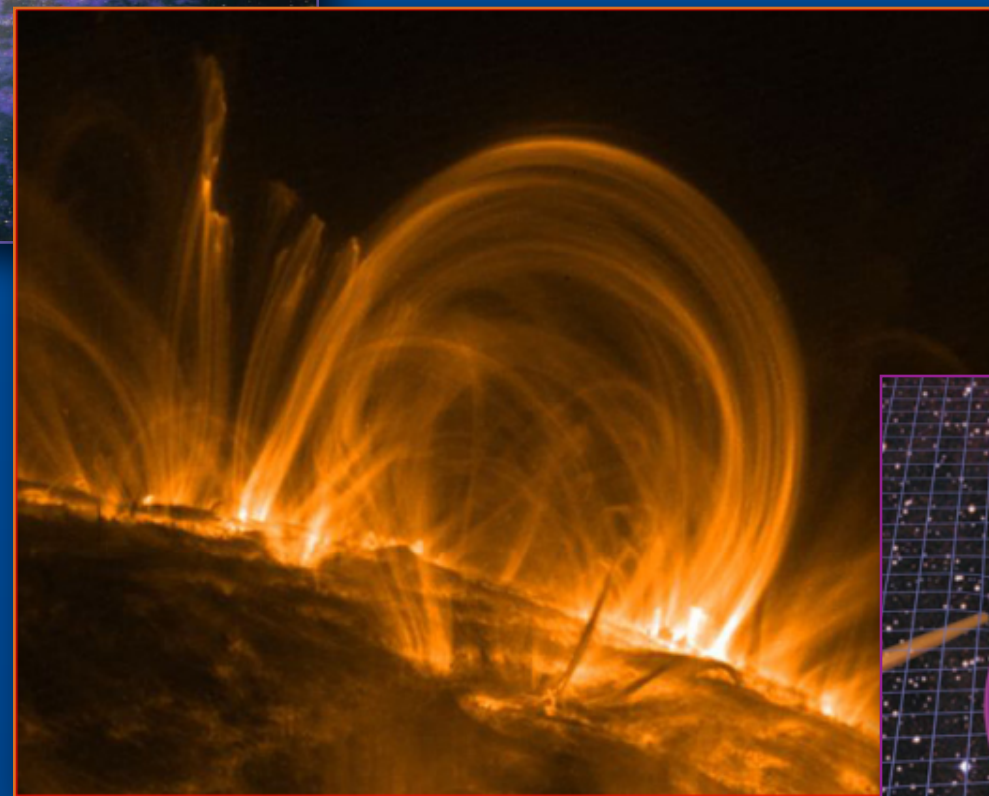


**The First Stars**

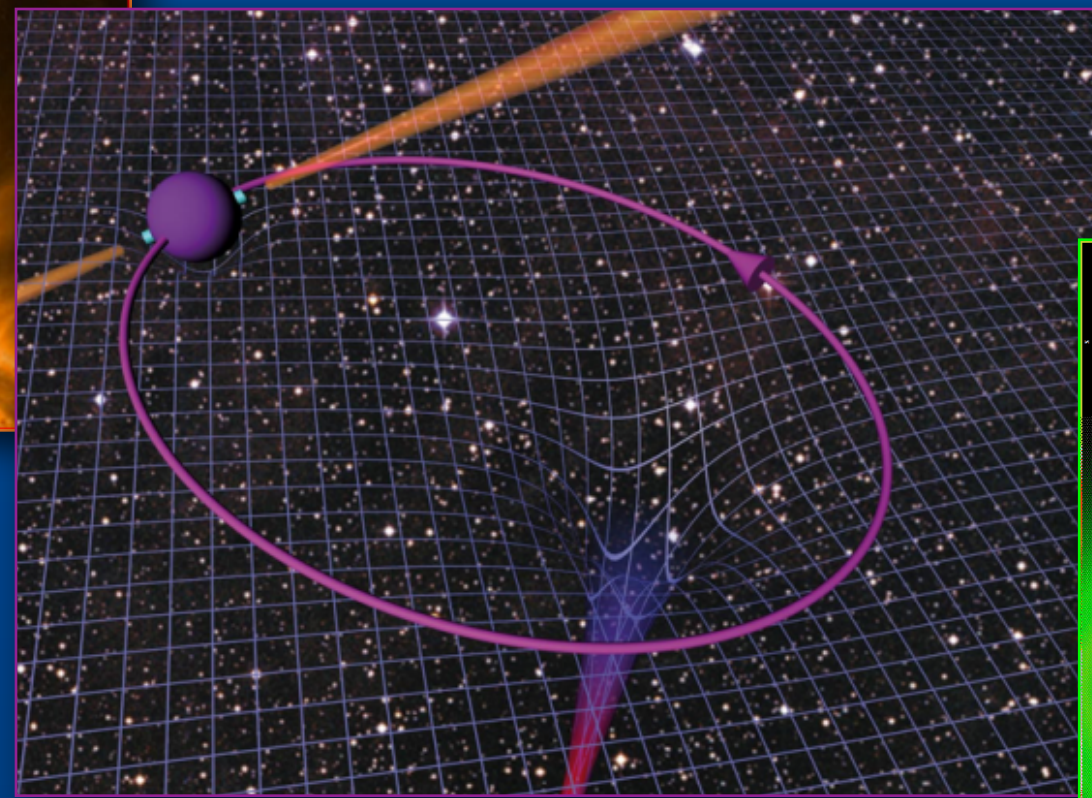
**Cosmic Evolution**



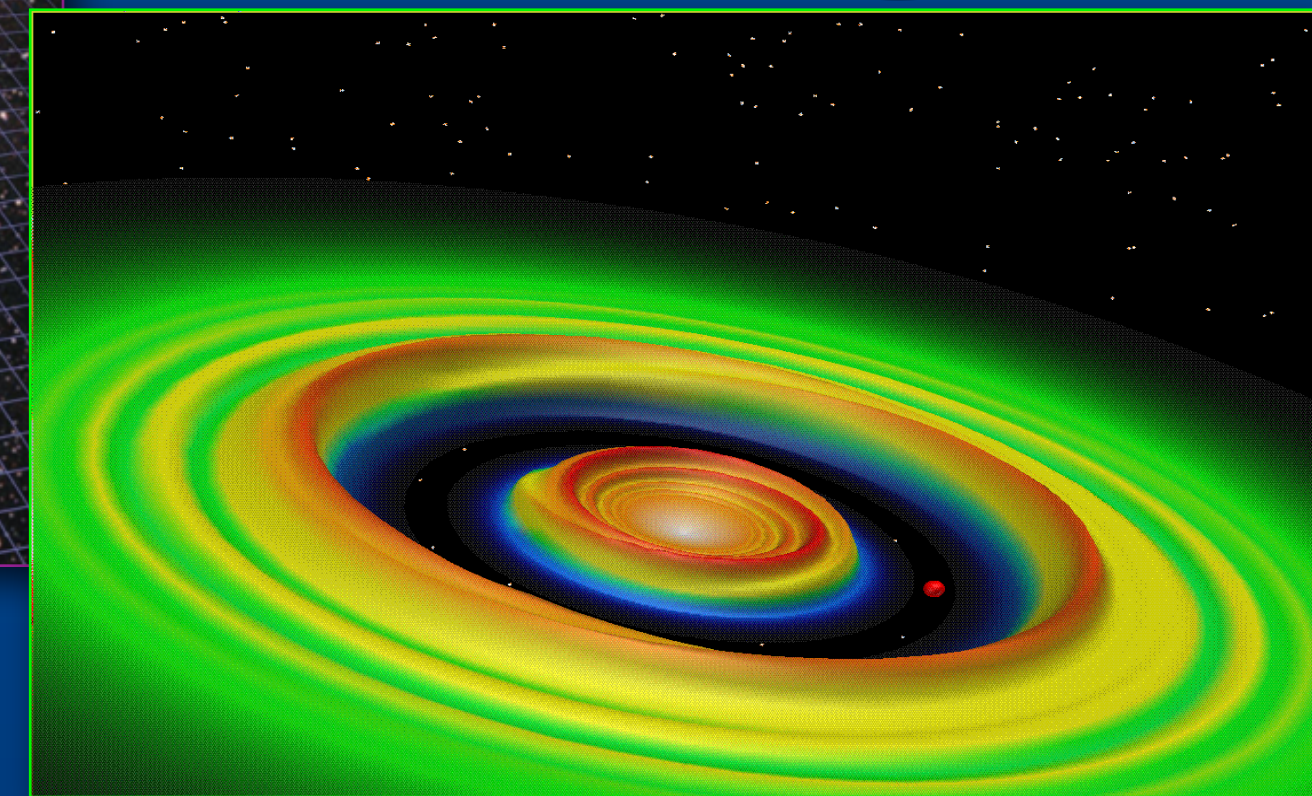
**Cosmic Magnetism**



**Gravitational Physics**

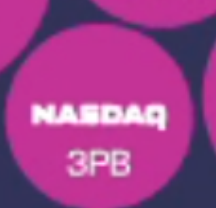
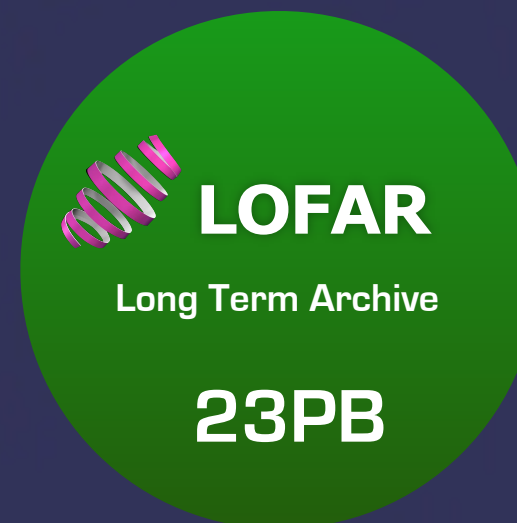


**Origins of Life**





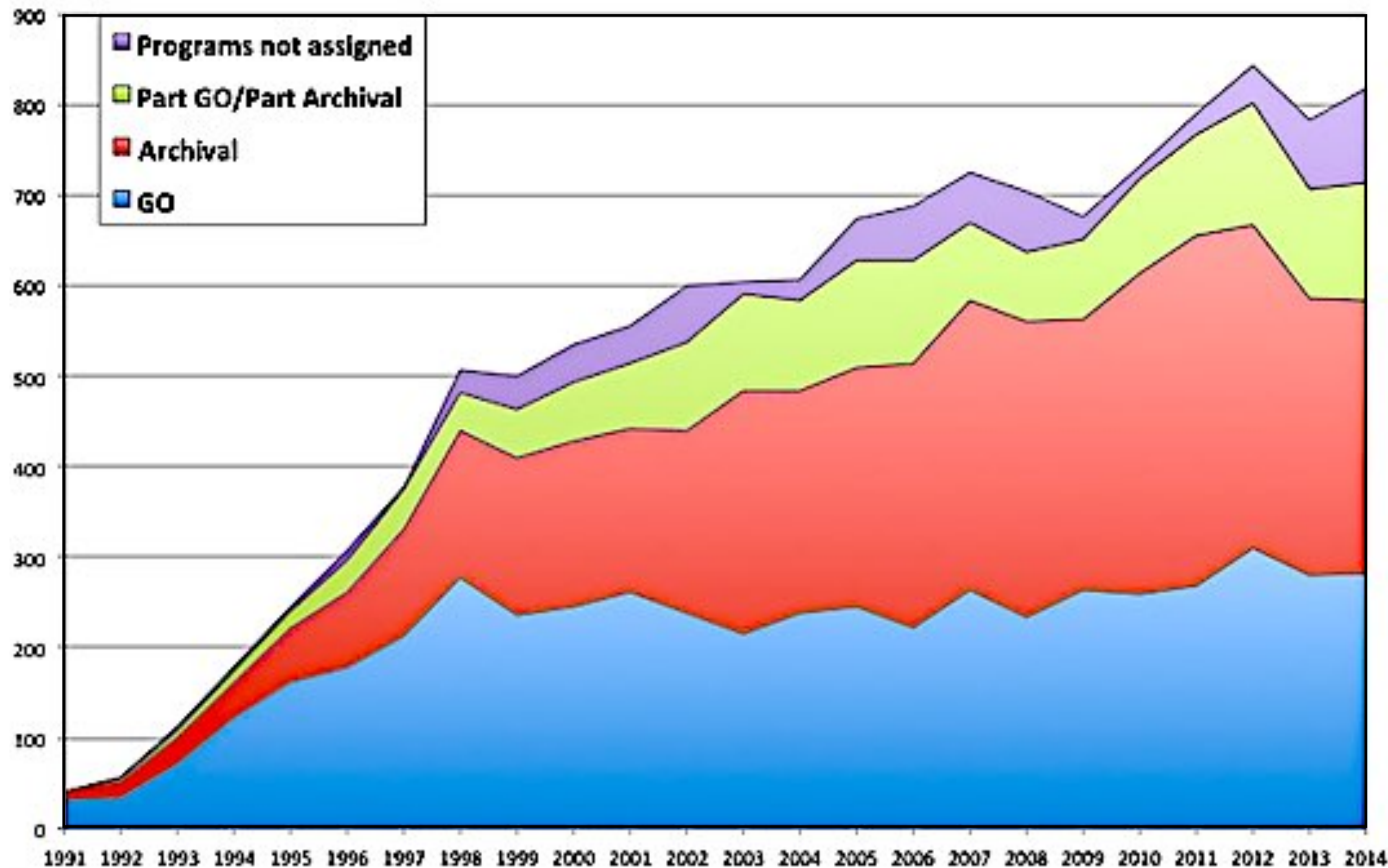
# SKA Science Archive



PER YEAR  
● 1 Petabyte



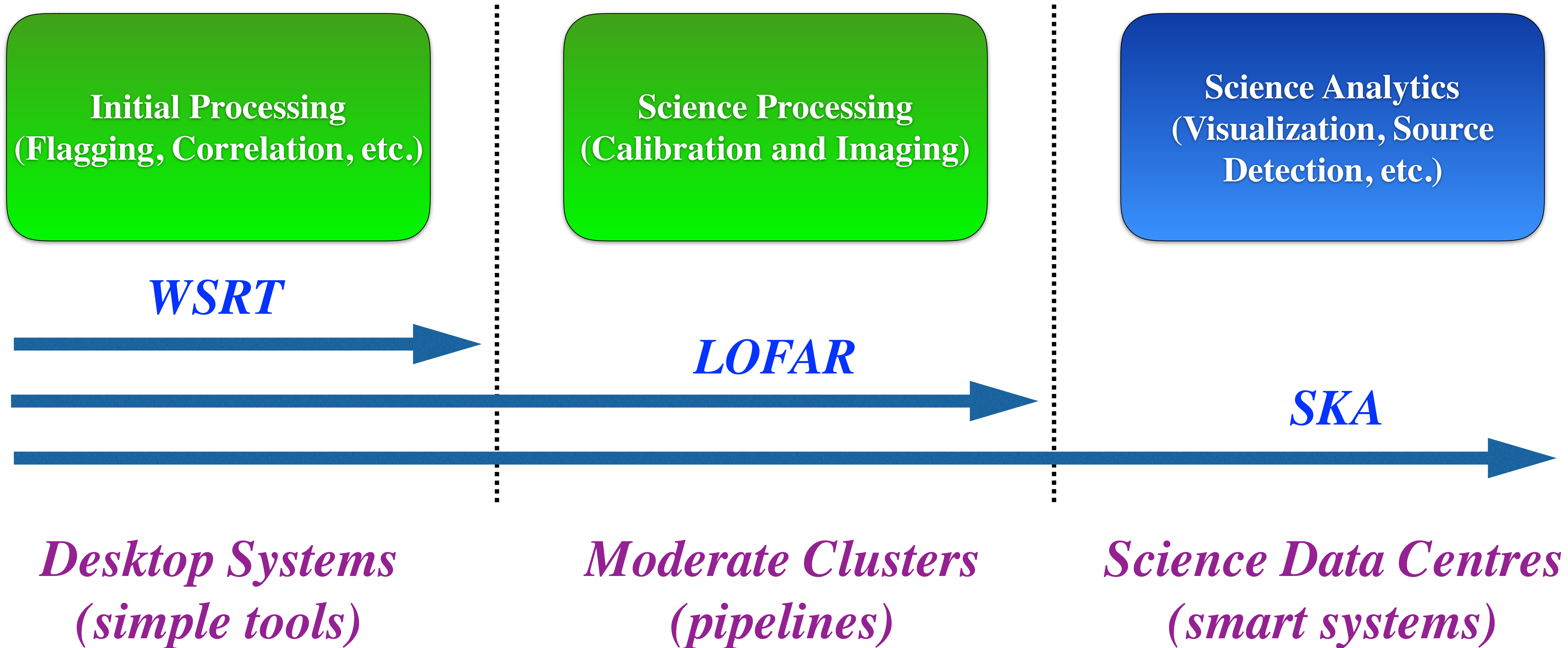
## *Science archives are a multiplier for total science output*



- Assumes the archives are persistent and maintained
- Assumes archival data is open and accessible
- Assumes data products stored are appropriate for general use
- Assumes users retrieving data have resources to process to a science result

## *Hubble Space Telescope Publication Rate*

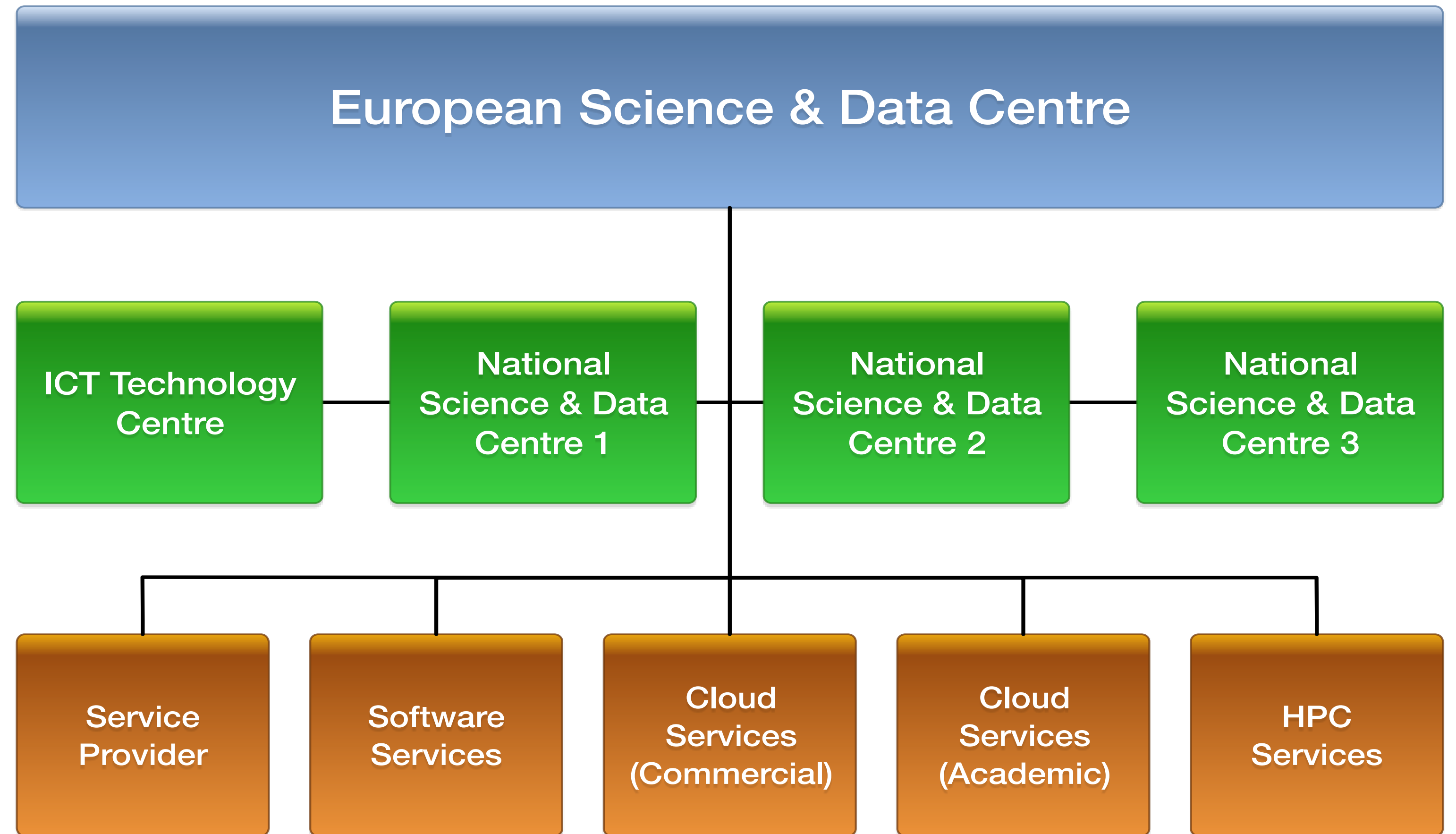






*Science Data Centres (SDCs) will host the SKA science archive, and provide access to infrastructure and expertise*

- Archive
- Access
- Analysis
- Analytics



*SDCs will be how scientists use the SKA!*









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

*Design and specification of a distributed, European Science Data Centre (ESDC) 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)*

- WP2: ESDC Governance Structure and Business Models
- WP3: ESDC Computing and Processing Requirements
- WP4: SKA Data Transport and Optimal European Storage Topologies
- WP5: User Data Access and Knowledge Creation







*Data is the subtle knife that separates fact  
from fiction, and answers from questions.*

*Thanks!*



# Micro Data Centers: lowering the energy-per-answer

## Prof. Dr. Ton Engbersen

Sci. Dir. ASTRON & IBM Center for Exascale Technology

Prof. Data Science Engineering, Rijksuniversiteit Groningen, Netherlands

Member IBM Academy of Technology

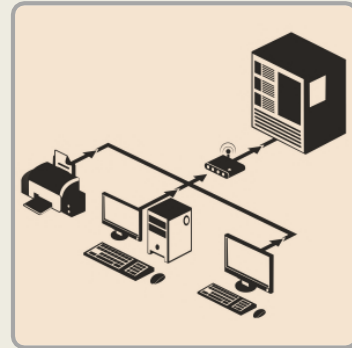
IBM Research GmbH - Switzerland



# Major Waves of Technology



**Back-Office Computing**



**Client-Server  
PC - 1981**



**World Wide Web  
and eBusiness**



**Confluence of Social, Mobile,  
Cloud, Big Data / Analytics**

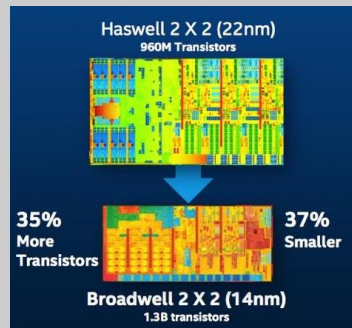




# Value in IT Industry is Shifting

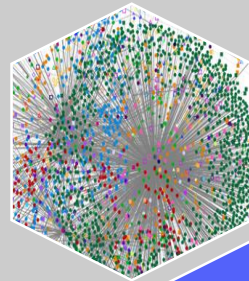
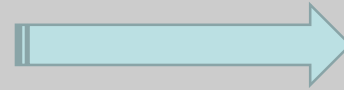
## Automating the World

Moore's Law



### Broadwell delivers:

- 2.2x increase in transistor density
- Up to 40% better 3D graphics perf<sup>1</sup>
- Enables <9mm fanless designs

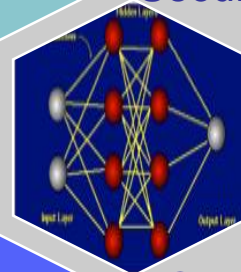


Graph Analytics

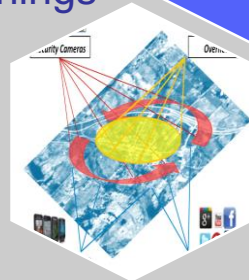
Genomics  
Social Networks  
Security

Healthcare  
Education  
Internet of Things

Machine Learning



Object recognition  
Complex video analytics  
Multi-model analytics



Video, Speech Analytics

## Understanding the World









# Micro Datacentre



5cm

