



REGIONAL  
CENTRE  
NETWORK

# SKA Regional Centres

Dr Rosie Bolton

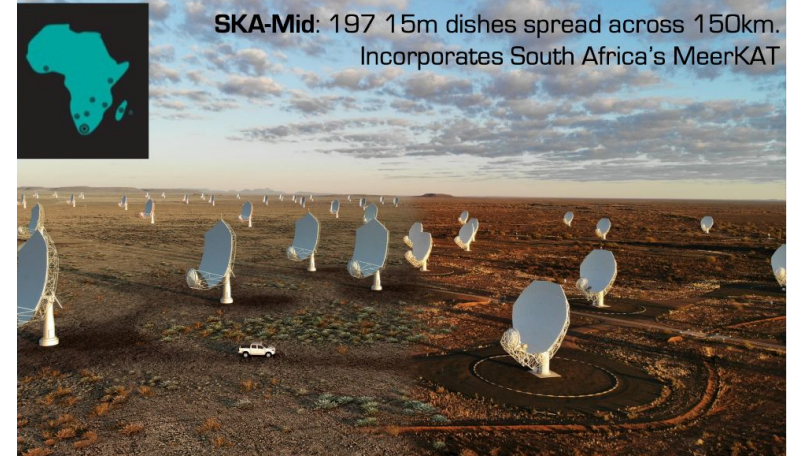
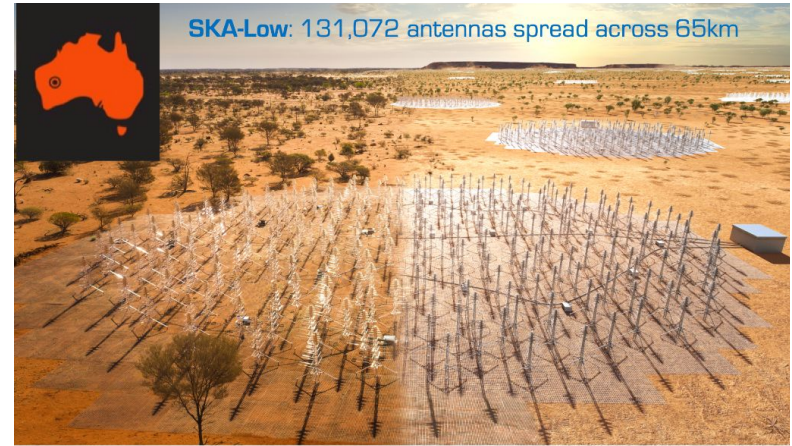
Head of Data Operations group, SKAO

7/11/2022



# SKA Observatory context

- Global collaboration of 16 countries to build and operate next-generation radio astronomy observatory
- SKAO Inter-Governmental Organisation governed by treaty (d.o.b. 4/2/2021)
- 7-8 year construction schedule. Cost  $\sim$ €2B (2021 euros) for first 10 years
- Internal data rates up to 2Pb/s
- Output data product rate  $\sim$ 100Gb/s per site:  $\sim$ 700PB per year



# The SRC Vision

**"To ensure that scientists can access SKA data products and use them to make discoveries"**

*(not part of WLCG! Here because we have common challenges to overcome)*



# What are we trying to achieve?

Globally distributed science teams  
Diverse skill base

Know who user is  
Understand group membership  
Respect proprietary periods  
Support public access

**"To ensure that scientists can access SKA data products and use them to make discoveries"**

get data products from Observatory  
allow inclusion of user-generated data products  
ensure long term reliability  
know where copies are

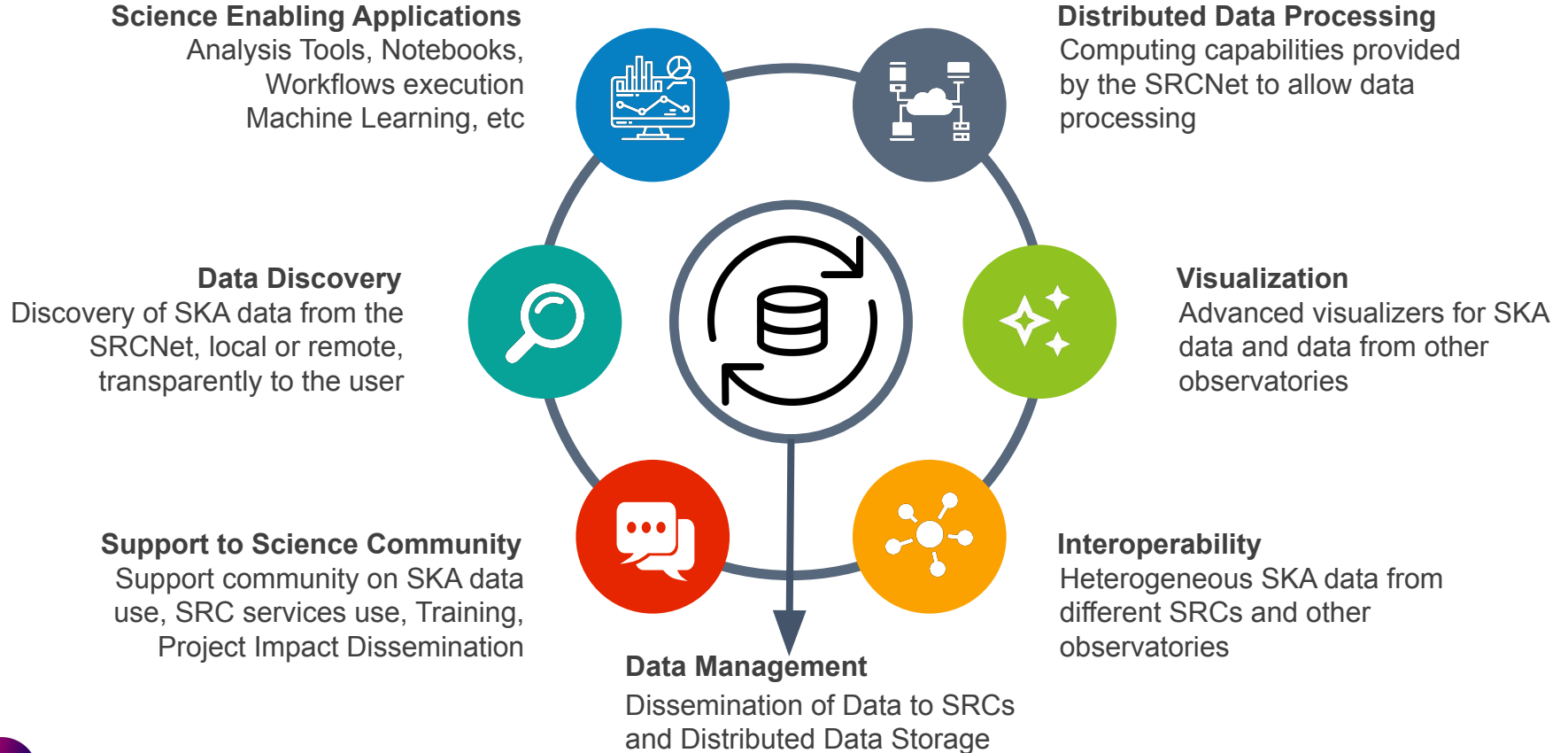
many and varied user interaction and workflow needs



# **SRC Capabilities**

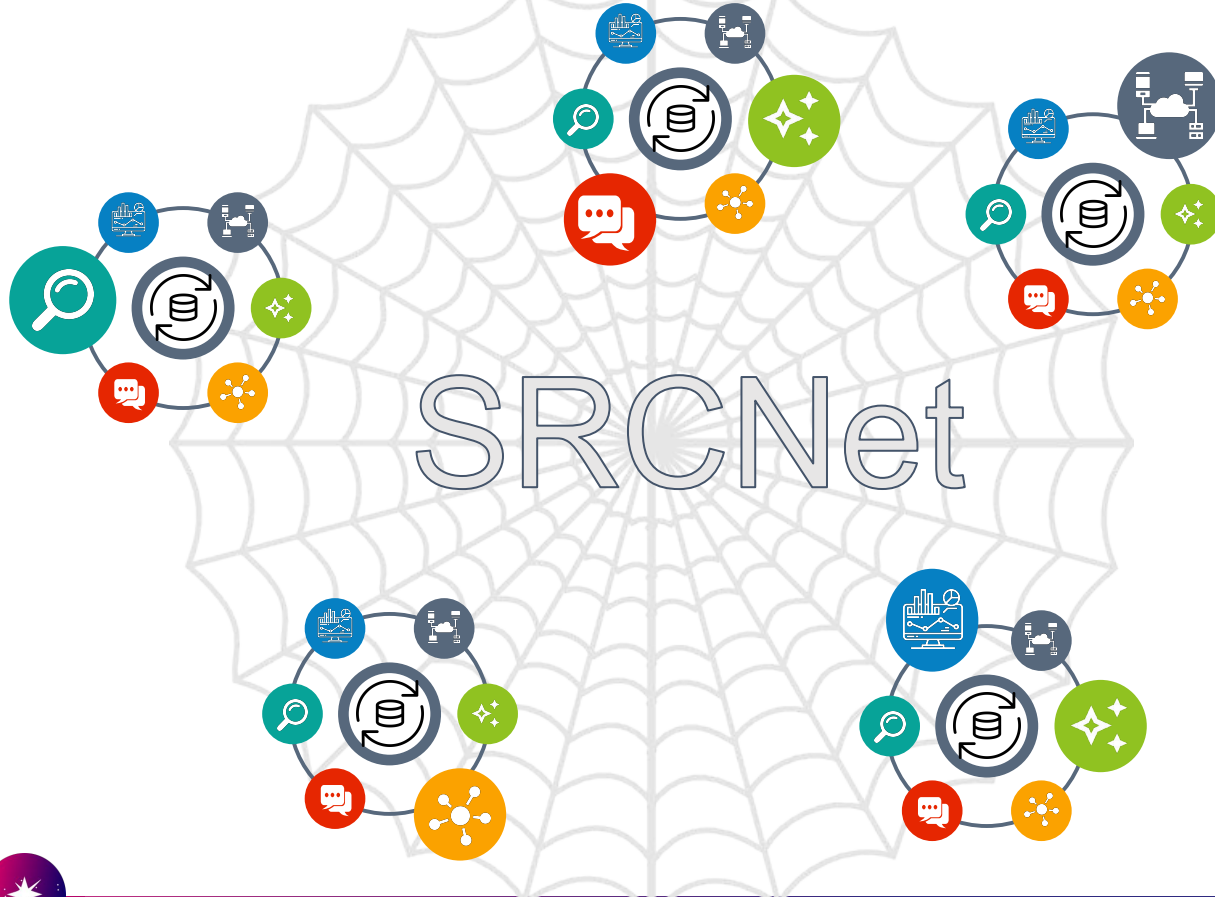


# SKA Regional Centre Capabilities





# SRC Network global capabilities



Collectively meet the needs of the global community of SKA users

Anticipate heterogeneous SRCs, with different strengths

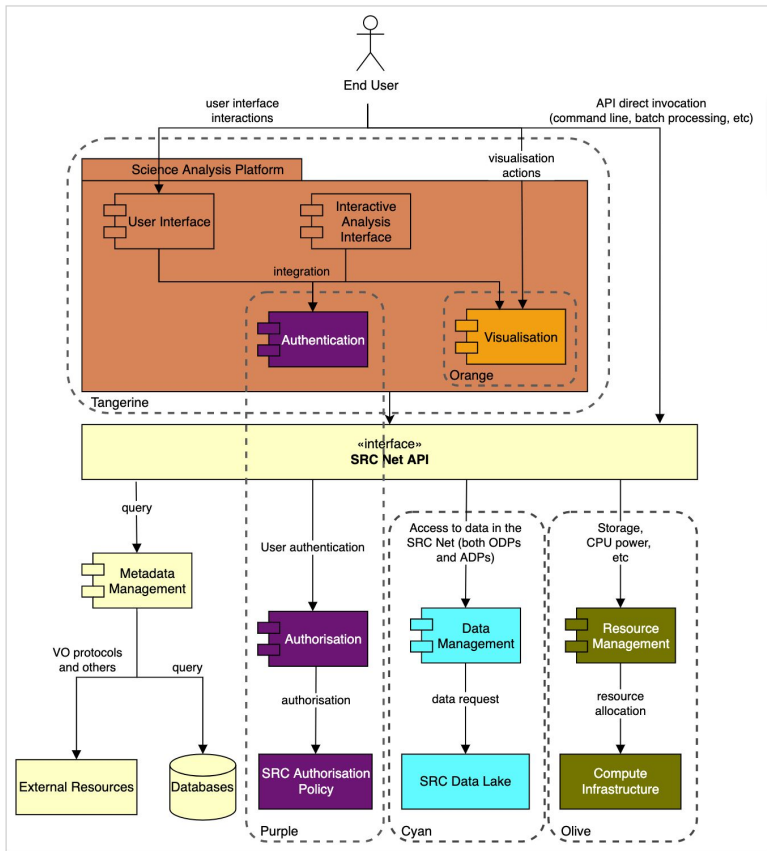


**How do we hope to deliver this?**





# Defining an architecture (WIP)



Users



Science Platform



Interface (API layer)



Metadata query - Science Data Discovery



Authentication Who? Permissions?



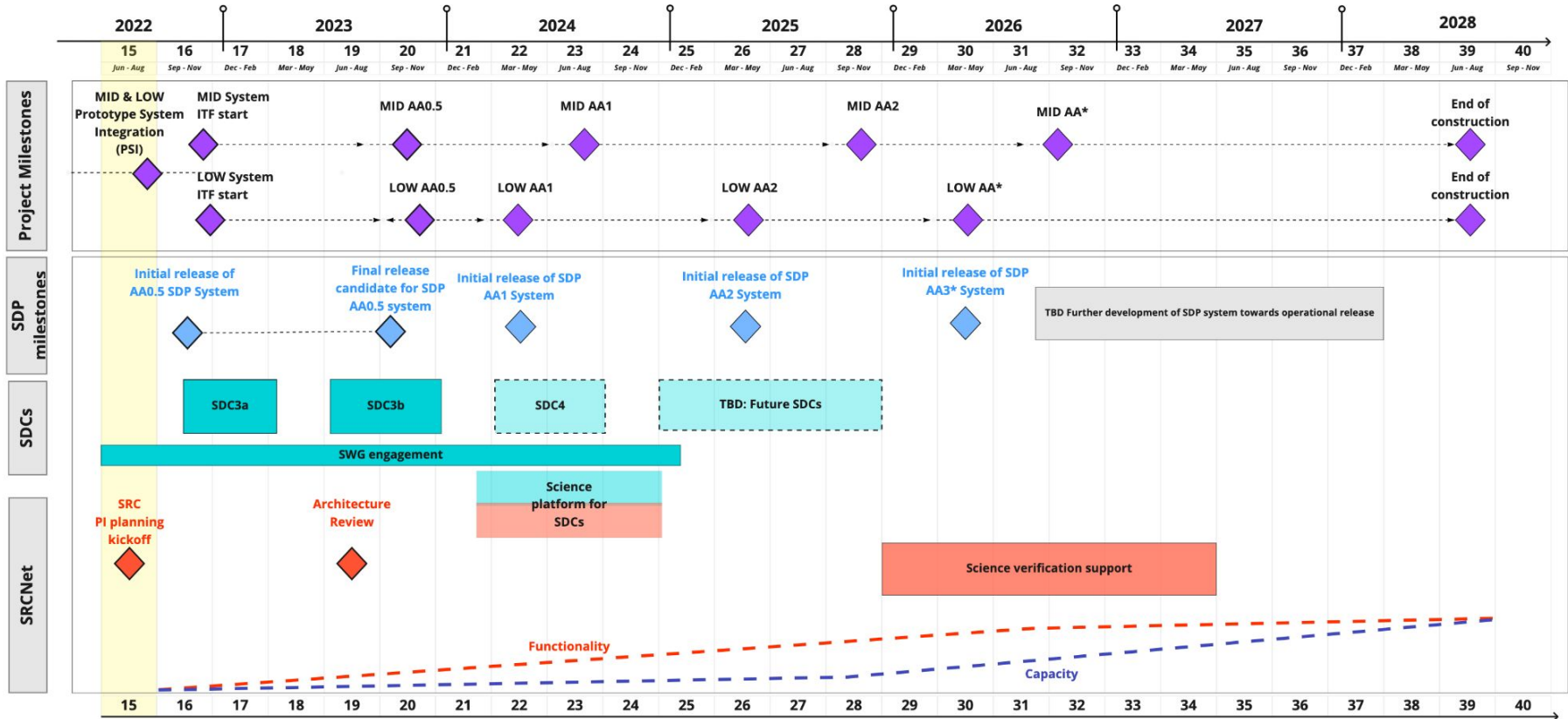
Data Logistics Globally distributed storage sites



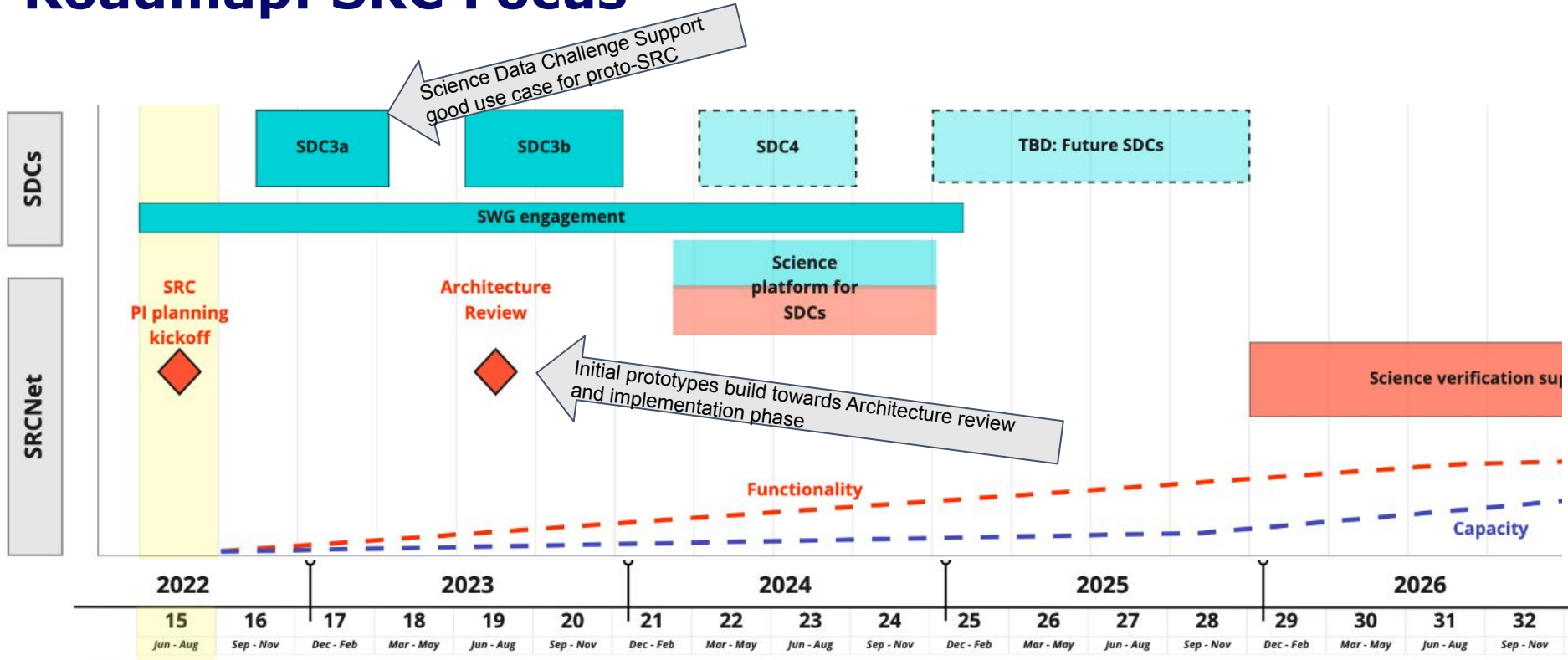
Compute Resource Management Work sharing



# Roadmap - SKA construction



# Roadmap: SRC Focus



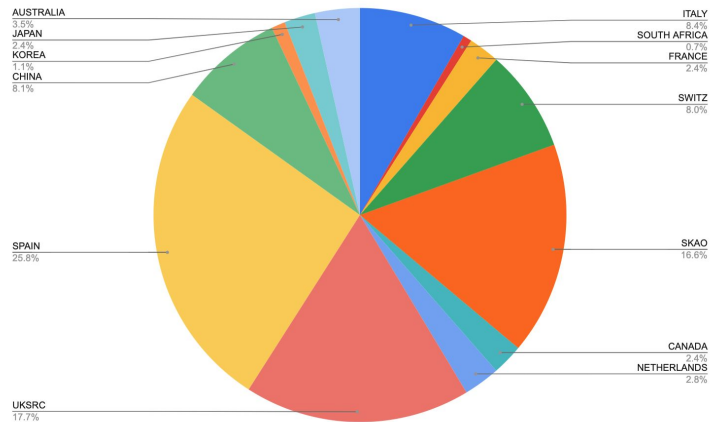
# SRC Development



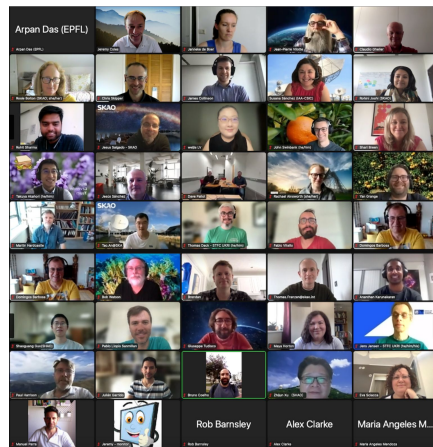
## SRC Steering Committee

- member from each SRC country (SKAO members or observers) plus SKAO
- Responsible for governance and policy planning
- Able to add SW development resources to build SRCNet, and later HW resources

Agreed to set up of a global SW development project "SRCNet". Funded in-kind from across the membership.



SRCNet developer-effort contributions September 2022



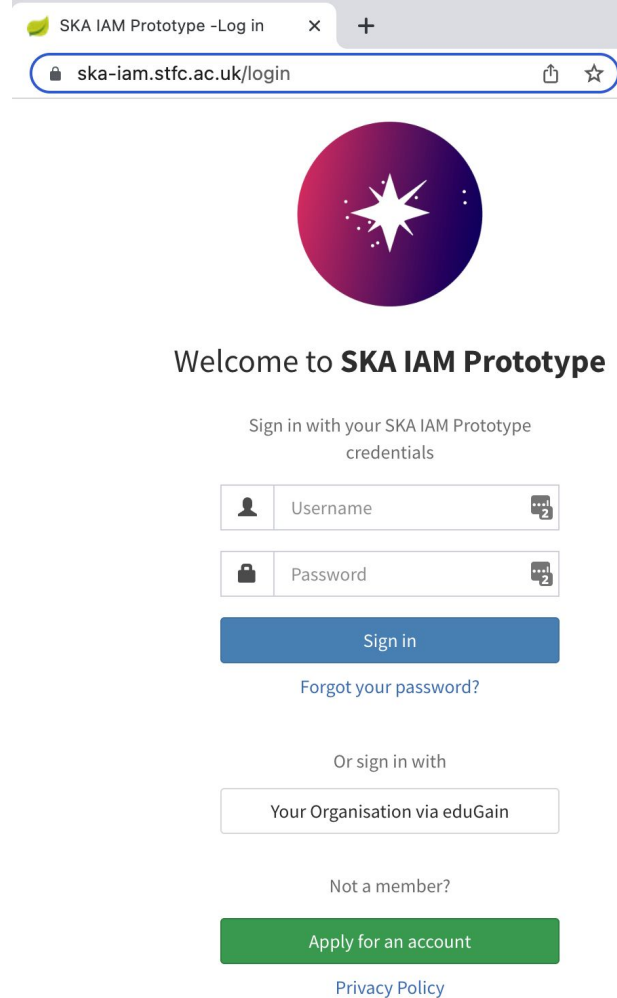
- Prototyping started June 2022
- Team members from 12 countries plus SKAO about 1000 developer-days per 3-months
- Anticipate growing as national funding to develop SRC nodes is available



# Progress - Identity management (Authentication and Authorisation)


- Instance of the IAM service deployed in UK (Rutherford Appleton Lab) to support SRC Development activities.
- Knowledge sharing within SRCNet team - supporting onboarding of new services connecting to IAM (e.g. grafana dashboard, test at AUSSRC science platform underway this month, hope to switch Rucio over from ESCAPE IAM by early 2023)
- Provides a way to probe site policy - can we get all SRC project sites to deal with this level of identity federation?
- Landscape report created ready for wider sharing and feedback within our community

*Shout out to Tom Dack, Jens Jensen and Ian Collier!*



SKA IAM Prototype -Log in

ska-iam.stfc.ac.uk/login



Welcome to **SKA IAM Prototype**

Sign in with your SKA IAM Prototype credentials

Username

Password

Sign in

[Forgot your password?](#)

Or sign in with

Your Organisation via eduGain

[Not a member? Apply for an account](#)

[Privacy Policy](#)



# Progress - SRC Rucio prototype

RSEs in SKA Rucio (SToRM webdav or dCache. S3 under investigation)



- *Not a production instance! Development and testing playground only!*
- Rucio is well suited to centralised Operations model for data management
- Performed long-haul transfers, Rucio stress tests, subscriptions (via our automated test framework)
- Integrating storage from national SRC efforts to increase understanding and inform assessment
- Continuing to assess suitability to SRC use cases. We've moved to a fully token-based system!

*shout to Xavi Espinal and the ESCAPE project for giving us the skills and confidence to bring our own system to life!*





# Rucio metadata interaction



- Work done on metadata filtering, now possible to search metadata using a range of operators
  - e.g. ranges, logical OR, compound inequalities
- Existing metadata plugin system extended to support external postgres (RDBMS) and mongodb (NoSQL) backend technologies
- Using this, some **promising exploratory work** done on exposing IVOA TAP services with data taken from an external postgres instance



*shout to Rob Barnsley leading Rucio metadata SIG, and Dave Morris: see Rucio Community Workshop later...*



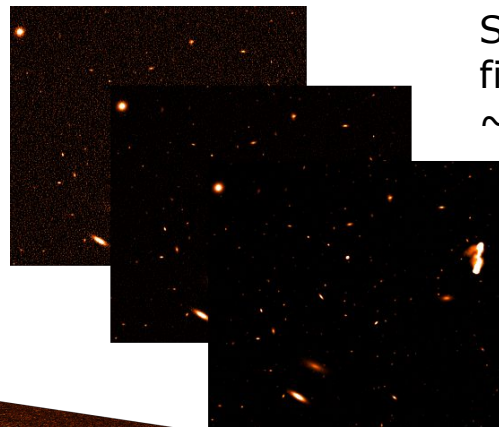


# SKA Science Data Challenges

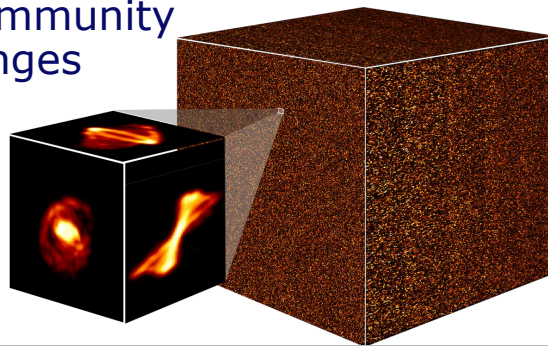
Science Data Challenges are a great way to involve science community with simulated SKA data products.

Successive challenges planned increasing in complexity and accuracy

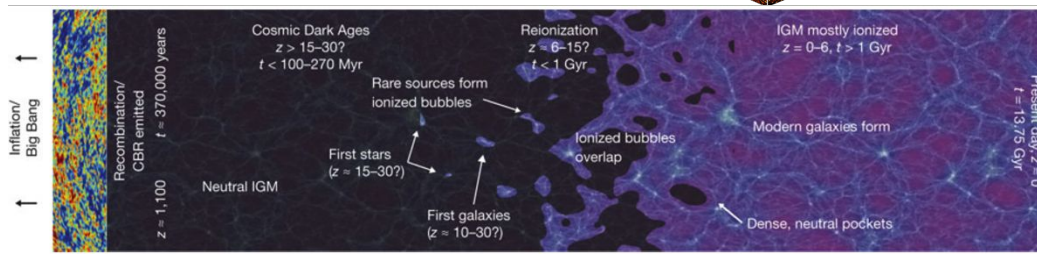
Moving to a model where SKA community working groups define the challenges



SDC1 - each file  $\sim 4\text{GB}$ ,  $\sim 50\text{GB}$  total



SDC2 - 1TByte simulated image cube



SDC3a -  $\sim 6\text{TByte}$  visibility data set



# Science Data Challenge Resources

## SKAO Science Data Challenge 2



MAP OF WORLDWIDE PARTICIPATION



Powerful vehicle to engage user and provider community

Anticipate that many providers will become, or relate to national SRC projects



# Summary

- SRCs are essential for connecting users with SKA data
- Governance and policy work still to be done
- Prototyping ideas for components now to drive global architectural decisions
- Implementation from end 2023, to keep pace with SKA Telescope development
- Our global team-of-teams is taking shape. Folks are learning now to interact with each other, with our SW development framework (SAFe™) and with the technologies of interest



**End - Questions!?**

