The SRC Development Timeline

Rosie Bolton

Head of Data Operations group, SKAO
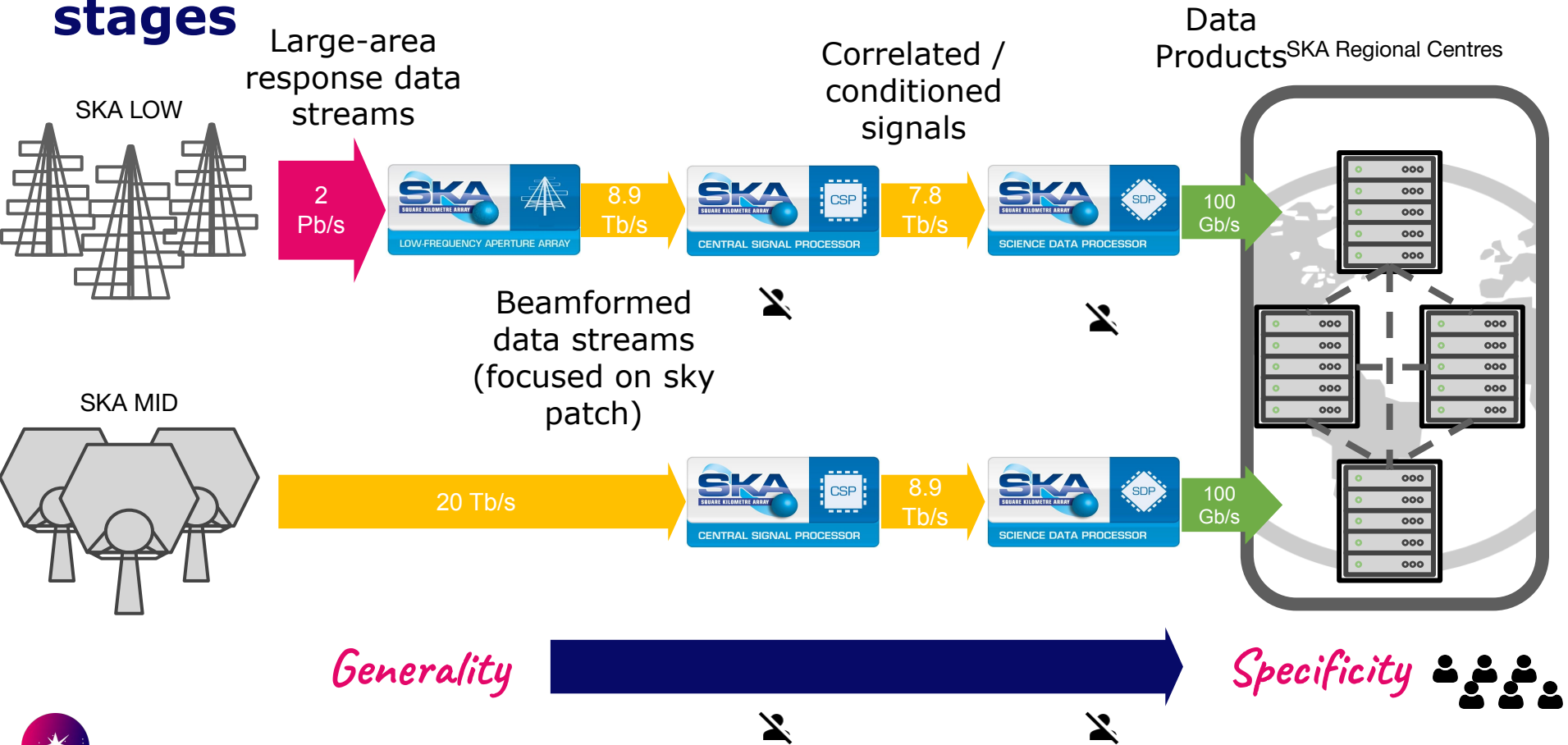SRCNet Product Manager

13/09/2022

With some slide material from Rohini Joshi and Jesús Salgado

The SRC Vision

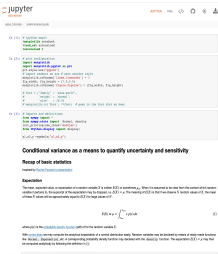"To ensure that scientists can access SKA data products and use them to make discoveries"

# SKA Regional Centres: SKAO data processing stages

Large-area response data streams

Correlated / conditioned signals

Data Products SKA Regional Centres

SKA LOW

2 Pb/s

8.9 Tb/s

7.8 Tb/s

100 Gb/s

LOW-FREQUENCY APERTURE ARRAY

CENTRAL SIGNAL PROCESSOR

SCIENCE DATA PROCESSOR

Beamformed data streams (focused on sky patch)

SKA MID

20 Tb/s

8.9 Tb/s

100 Gb/s

CENTRAL SIGNAL PROCESSOR

SCIENCE DATA PROCESSOR

Generality

Specificity

SRCs: Collaboration platform, Support data reuse

Workflows

Power spectra
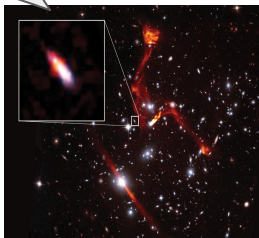
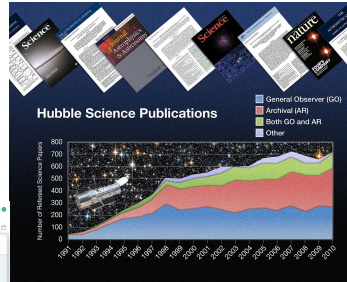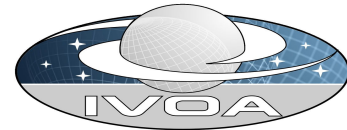Paul, Sourabh et al. (2016). ApJ 833. 10.3847/1538-4357/833/2/213.
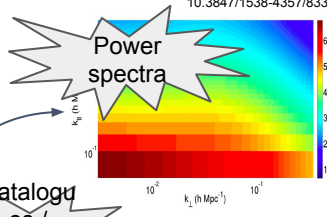
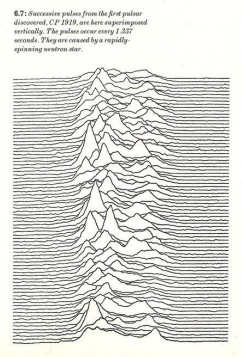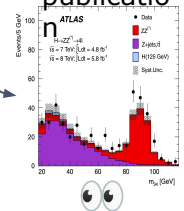Catalogues / Source List

Credit: Heywood et al.; Sophia Dagnello, NRAO/AUI/N

Plots for publication

Hubble Science Publications

Example from https://cartavis.org/

FAIR principles
(Find, Access, Interoperate, Re-use)

SKA Regional Centre Capabilities

Science Enabling Applications
Analysis Tools, Notebooks, Workflows execution Machine Learning, etc

Distributed Data Processing
Computing capabilities provided by the SRCNet to allow data processing

Data Discovery
Discovery of SKA data from the SRCNet, local or remote, transparently to the user

Visualization
Advanced visualizers for SKA data and data from other observatories

Support to Science Community
Support community on SKA data use, SRC services use, Training, Project Impact Dissemination

Interoperability
Heterogeneous SKA data from different SRCs and other observatories

Data Management
Dissemination of Data to SRCs and Distributed Data Storage
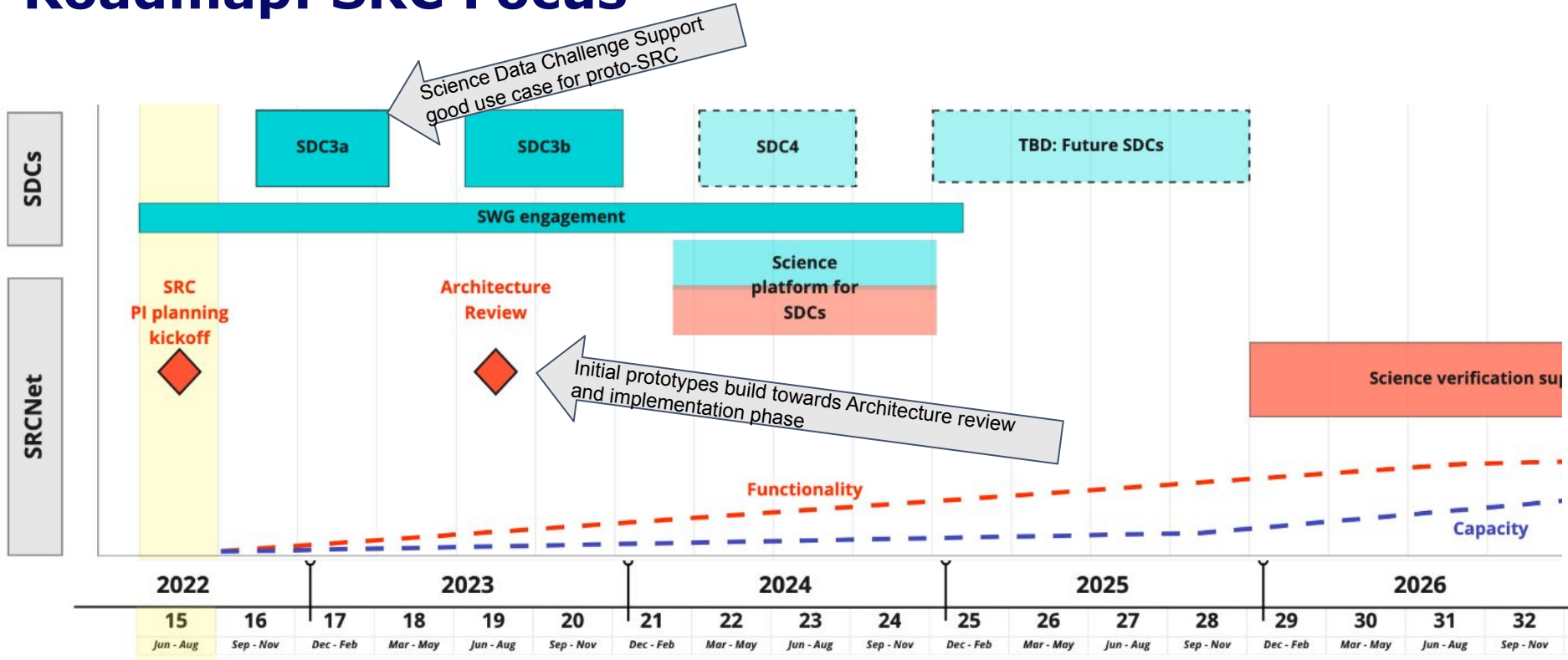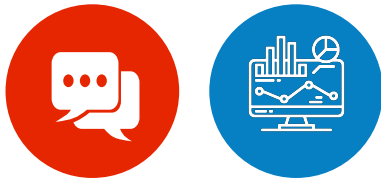
Roadmap

Roadmap: SRC Focus

First Prototyping Phase: 2022-2023
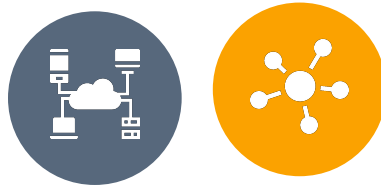
Data Management service: Replication, distribution, synchronisation of data products and location index

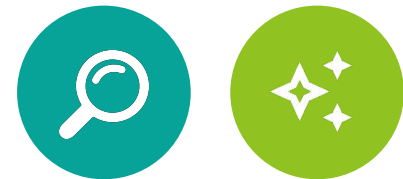**Data Analysis**: Science Extraction, Processing in Notebooks

Central Services and Software Distribution: SW infrastructure, compute provision

Federated **Authentication and Authorization**: identity management, compatible with SKAO

Data Visualisation and discovery - performance at SKA scale

Progress - Data Management

Transfer success site matrix					
Src\Dst	SPSRC_STORM	IMPERIAL	CNAF	AUSSRC_STORM	STFC_STORM
STFC_STORM	97%	89%	100%	99%	NO DATA
SPSRC_STORM	NO DATA	91%	100%	99%	27%
IMPERIAL	98%	NO DATA	100%	99%	39%
CNAF	98%	93%	NO DATA	99%	34%
AUSSRC_STORM	84%	91%	100%	NO DATA	31%

Hope for Swiss site in this next 3-month plan

Upgraded Rucio (global distributed data management service) instance to use tokens for Authentication

Successfully integrated storage in Spanish SRC (Granada) and Australian SRC (Perth) into existing system

Improved monitoring - rich (extensible) dashboard and supporting database and slack integration

Rucio: https://rucio.cern.ch/

Progress - Identity management (Authentication and Authorisation)

- "IAM"* service deployed in UK (Rutherford Appleton Lab)
https://ska-iam.stfc.ac.uk/login

- Landscape report created ready for wider sharing and feedback

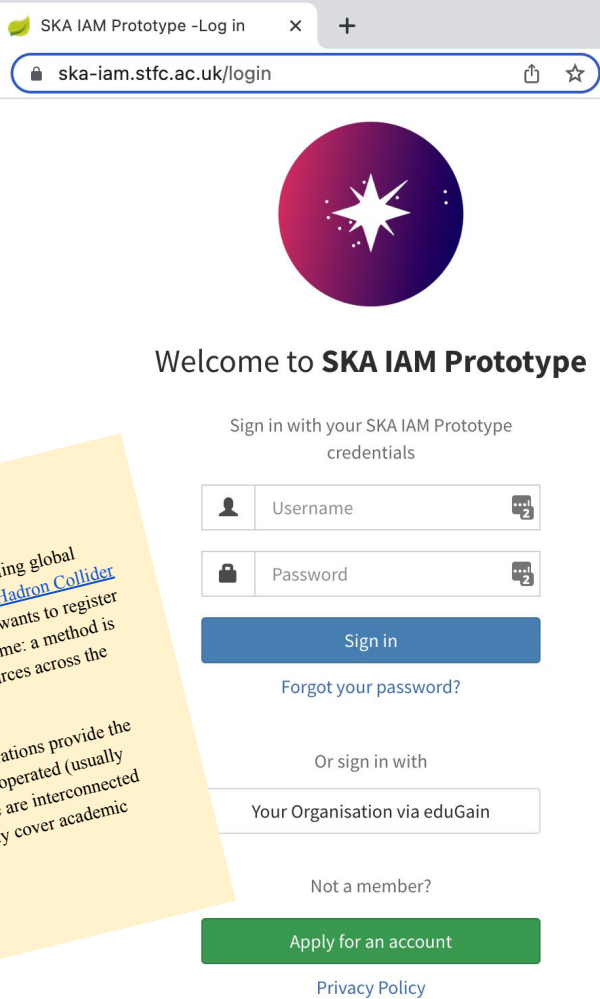*https://indigo-iam.github.io/docs/v/current/about.html

Executive Summary

Science has large scale research activities that require pooling global computing and data storage resources, such as the Large Hadron Collider (WLCG), climate research, and NGS. Naturally no one wants to register with every remote service with a new password every time: a method is needed to connect the user's existing identities to resources across the world, possibly in different trust domains.

Authentication and authorisation infrastructure federations provide the means to connect users to resources. NRENs have operated (usually national) identity federations for decades and these are interconnected through activities like eduGAIN, but (usually) only cover academic member organisations.

…

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 - Visualisation

- Comparative review of three tools where teams have expertise: CARTA, ALADIN and visIVO

- Collection of data products (with links) for use in demos

- Demo of these tools to the SRCNet team of teams

- All three tools deployed at CHINA SRC, and access given to external users (user pass, IP address whitelist)

- Next steps: INTERFACES - understand interfaces for these tools, towards developing a common API layer
  - Fit with SRCNet architecture
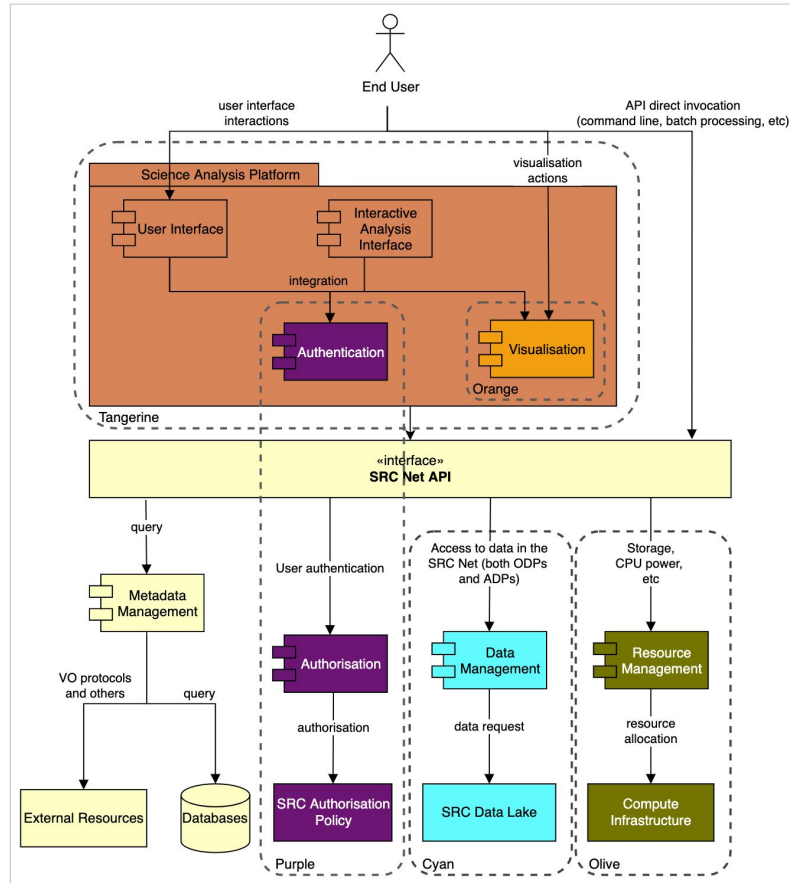 - Allow extension to other tools
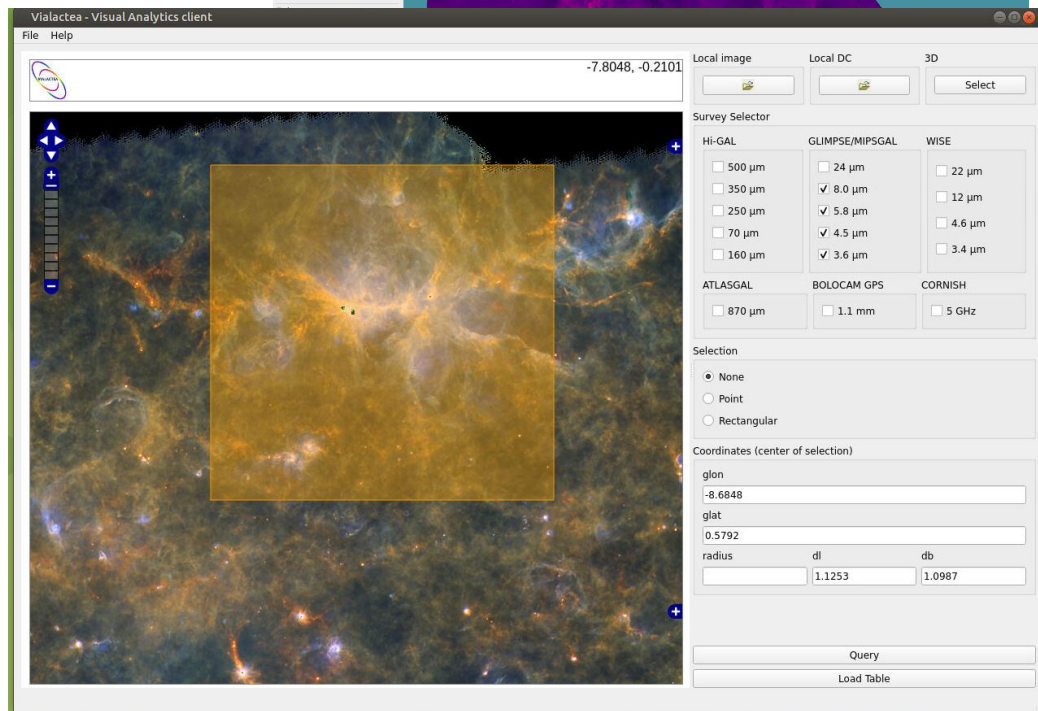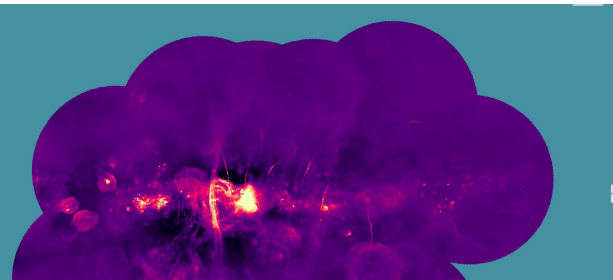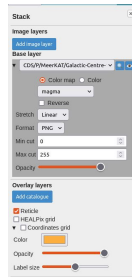
Progress - Science Platform

- Evaluation of (30) existing platforms considering many different criteria

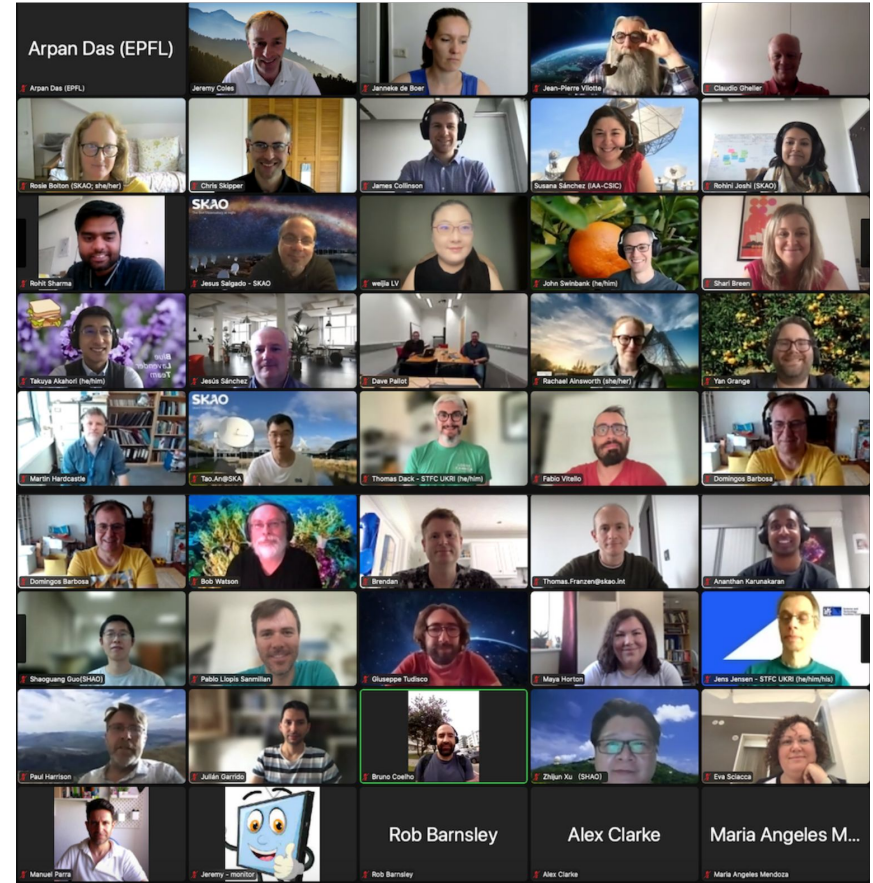Shortlist of 8 (or so) for further work

- Science Platform vision document

Next steps are to consider architectural aspects of science platform vision, update vision to incorporate this.
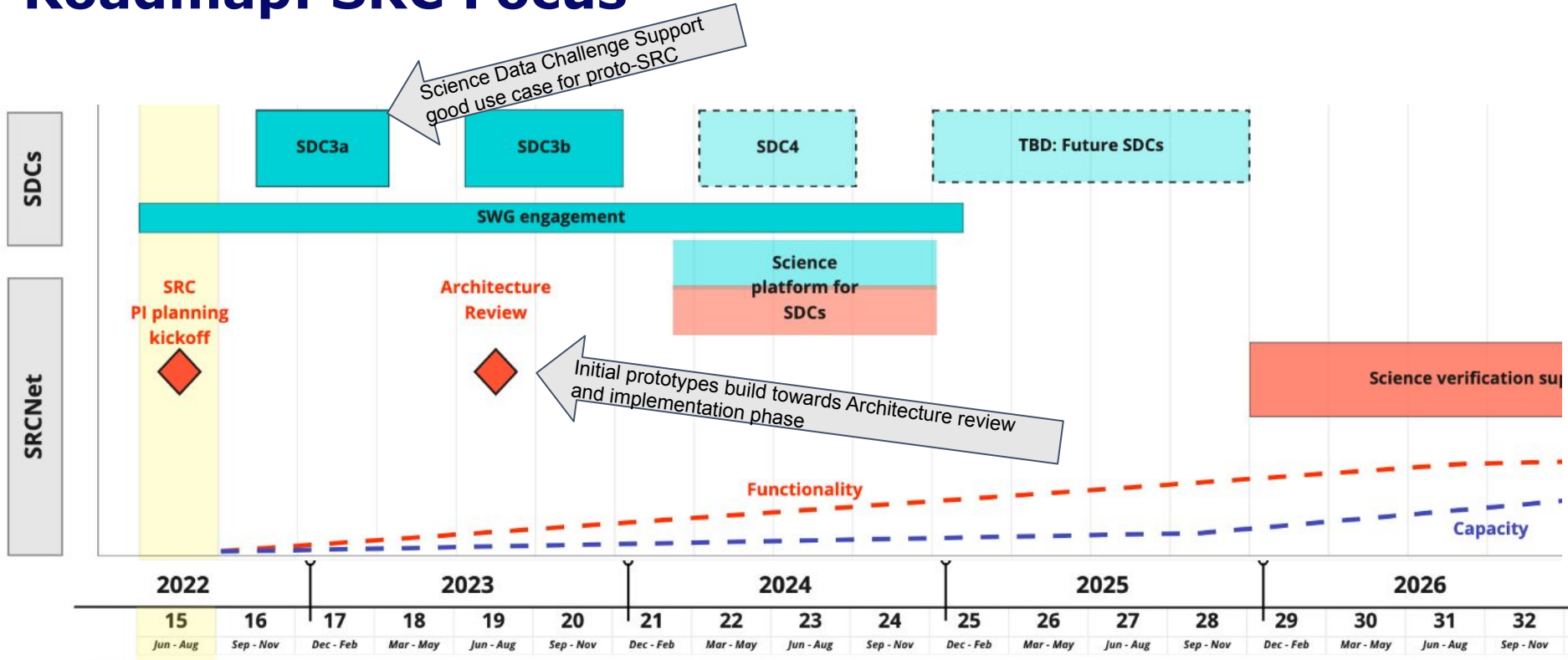
SRC Prototyping: A team of teams

- Adopting Scaled-Agile framework (SAFe)

- Teams are enabled to make their own decisions about work to commit to for 3 months. Members from 12 countries plus SKAO currently

- Program team guides roadmap and helps develop feature ideas - working with team leaders and stakeholders ("Business Owners")

- Same approach as used elsewhere in SKA SW development but we are able to adjust it to suit our different funding circumstances

Roadmap: SRC Focus

End