Science Platforms and the IVOA
The SKA Regional Centres Network Use Case

Jesús Salgado
SKA Regional Centres Network Architect

And the SRCNet members

SKA1-low – the SKA's low-frequency instrument

The Square Kilometre Array (SKA) is a next-generation radio astronomy facility that will revolutionise our understanding of the Universe. It will have a uniquely distributed character: **one** observatory operating **two** telescopes on **three** continents. Construction of the SKA will be phased and work is currently focused on the first phase named SKA1, corresponding to a fraction of the full SKA. SKA1 will include two instruments – SKA1-mid and SKA1-low – observing the Universe at different frequencies.

SKA SQUARE KILOMETRE ARRAY

Location: Australia

Frequency range: **50 MHz** to **350 MHz**

~131,000 antennas spread between **500 stations**

Total collecting area: **0.4km²**

Maximum distance between stations: **>65km**

SKA1-low

Data transfer rate: **7.2 Terabits** per second

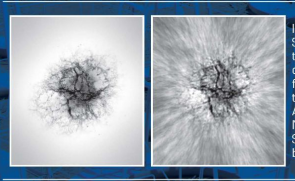


Image quality of SKA1-low (left) versus the best current facility operating in the same frequency range, the LOw Frequency ARray (LOFAR), in the Netherlands (right). SKA-low's resolution will be similar to LOFAR.

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

25% better resolution

8x more sensitive

135x the survey speed

www.skatelescope.org @SKA_telescope SKAtelescope ska_telescope Square Kilometre Array ska-organisation

SKA1-mid – the SKA's mid-frequency instrument

The Square Kilometre Array (SKA) is a next-generation radio astronomy facility that will revolutionise our understanding of the Universe. It will have a uniquely distributed character: **one** observatory operating **two** telescopes on **three** continents. Construction of the SKA will be phased and work is currently focused on the first phase named SKA1, corresponding to a fraction of the full SKA. SKA1 will include two instruments – SKA1-mid and SKA1-low – observing the Universe at different frequencies.

SKA SQUARE KILOMETRE ARRAY

Location: South Africa

Frequency range: **350 MHz** to **15.3 GHz** with a goal of 24 GHz

197 dishes (including 64 MeerKAT dishes)

Total collecting area: **33,000m²**

or **126 tennis courts**

Maximum distance between dishes: **150km**

SKA1-mid

Data transfer rate: **8.8 Terabits** per second

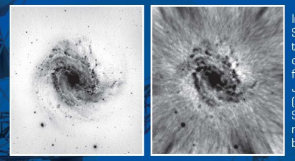


Image quality of SKA1-mid (left) versus the best current facility operating in the same frequency range, the Jansky Very Large Array (JVLA) in the United States (right). SKA-mid's resolution will be 4x better than JVLA.

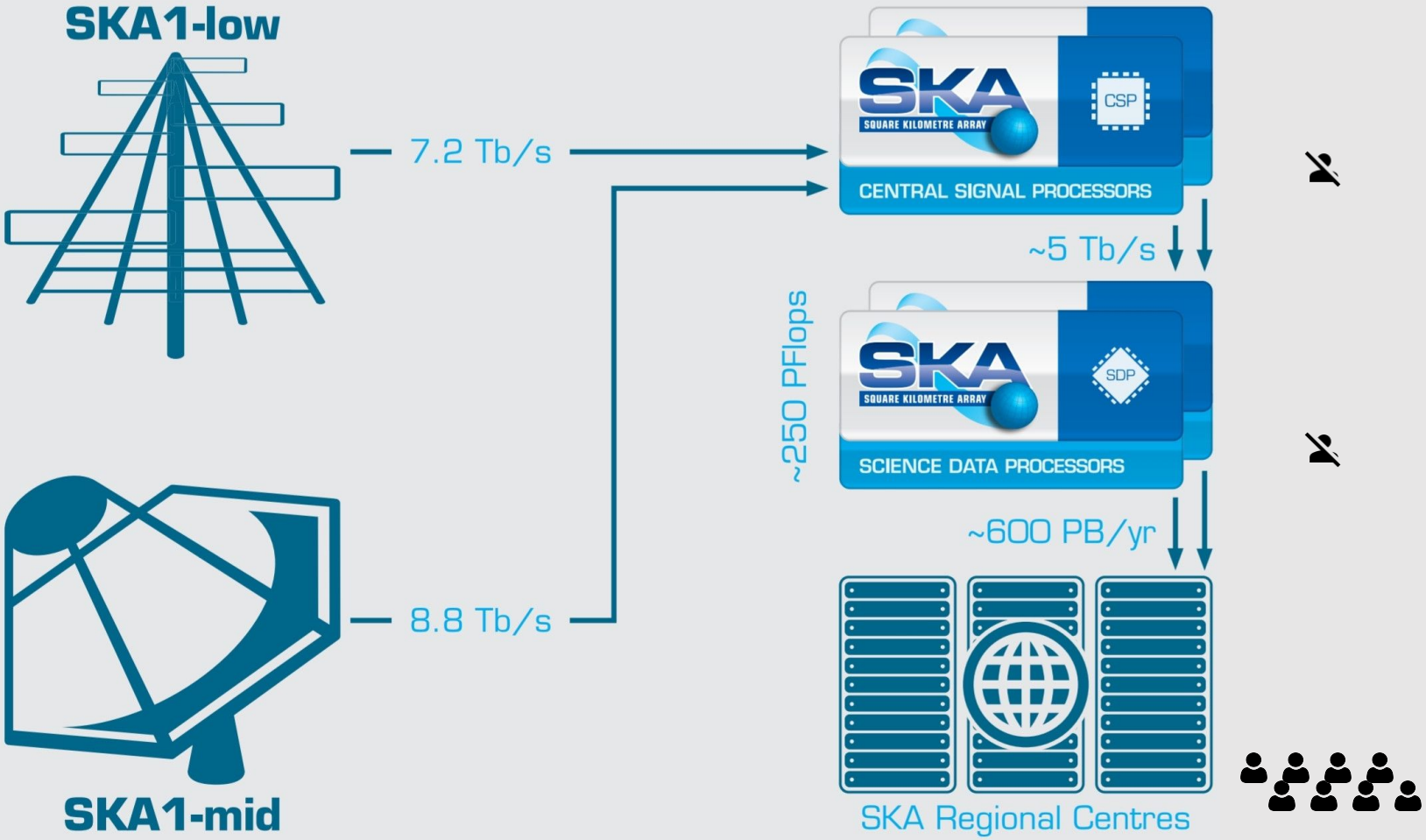
Compared to the JVLA, the current best similar instrument in the world:

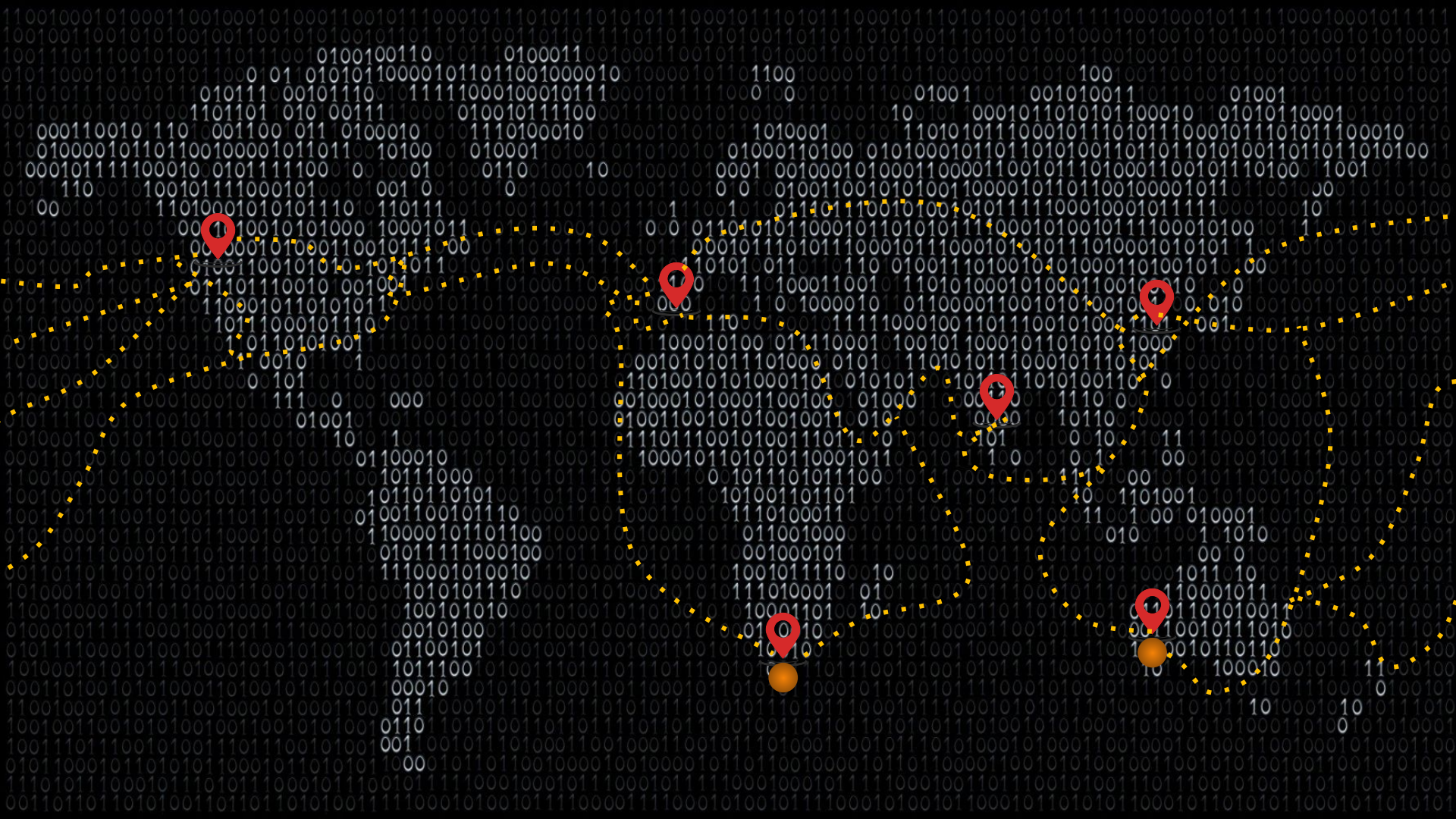
4x the resolution

5x more sensitive

60x the survey speed

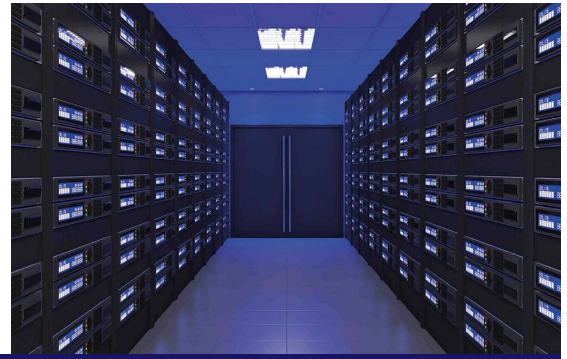
www.skatelescope.org @SKA_telescope SKAtelescope ska_telescope Square Kilometre Array ska-organisation


SKA1-low
7.2 Tb/s
SKA1-mid
8.8 Tb/s
SKA
SQUARE KILOMETRE ARRAY
CSP
CENTRAL SIGNAL PROCESSORS
~5 Tb/s
~250 PFlops
SKA
SQUARE KILOMETRE ARRAY
SDP
SCIENCE DATA PROCESSORS
~600 PB/yr
SKA Regional Centres

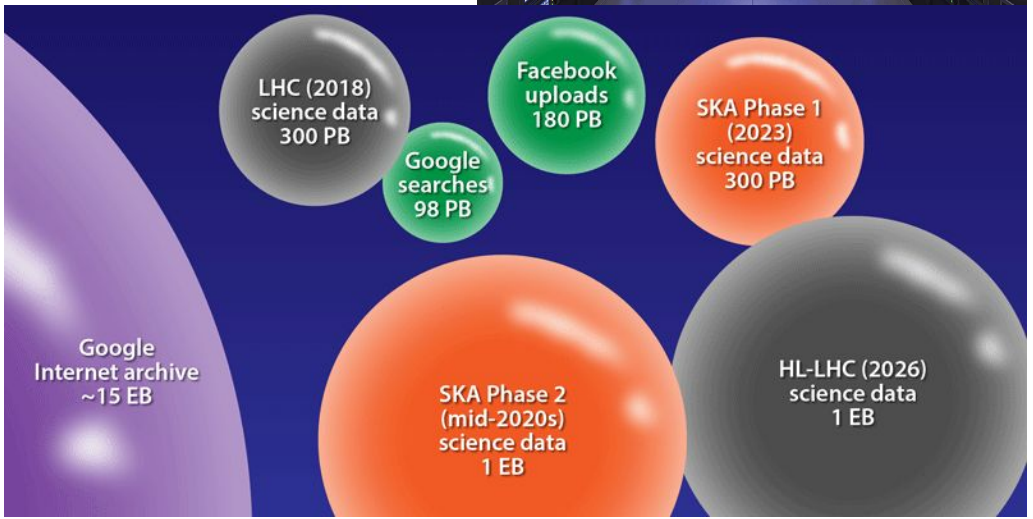


SKA Regional Centres (SRC) Network in Numbers

- ~ 600 PB/year of Scientific Data
- 16 countries involved
- Up to 100 FTEs during development phase



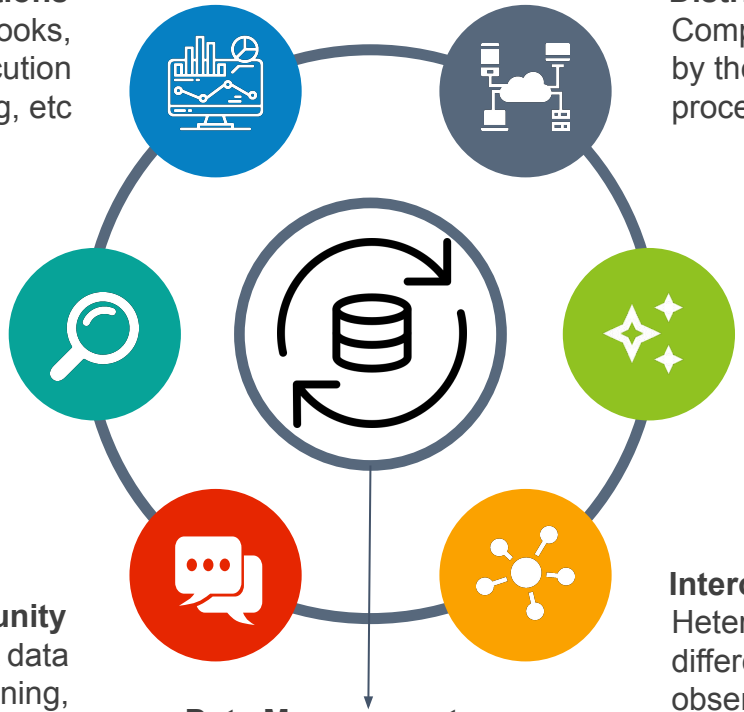
- Collaboration agreements with CERN, GEANT, CTAO
- Collaborations with CNRS, Vera Rubin and others



SKA Regional Centre Capabilities Blueprint

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

SRC Network global capabilities

SRCNet Principles

1 Data Management

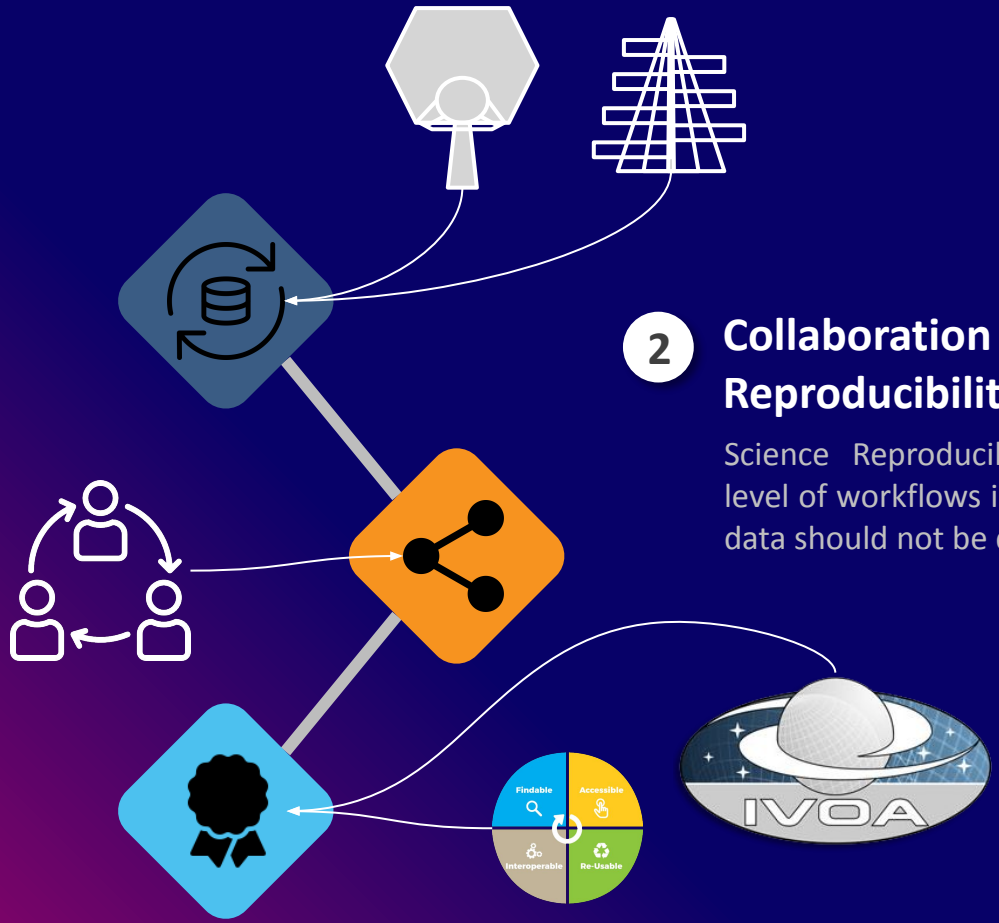
Avoid unnecessary duplication and transfers
Roughly 5-10 million dollars per year in new data, for one copy

2 Collaboration and Reproducibility

Science Reproducibility at the level of workflows is essential as data should not be downloaded

3 Use of Standards

Build SKA science archive around FAIR and IVOA standards



The IVOA Context

Democratic Science and AI

Harmonisation
Transparent Data Access
Combined Computing Resources

Science Enabling Applications

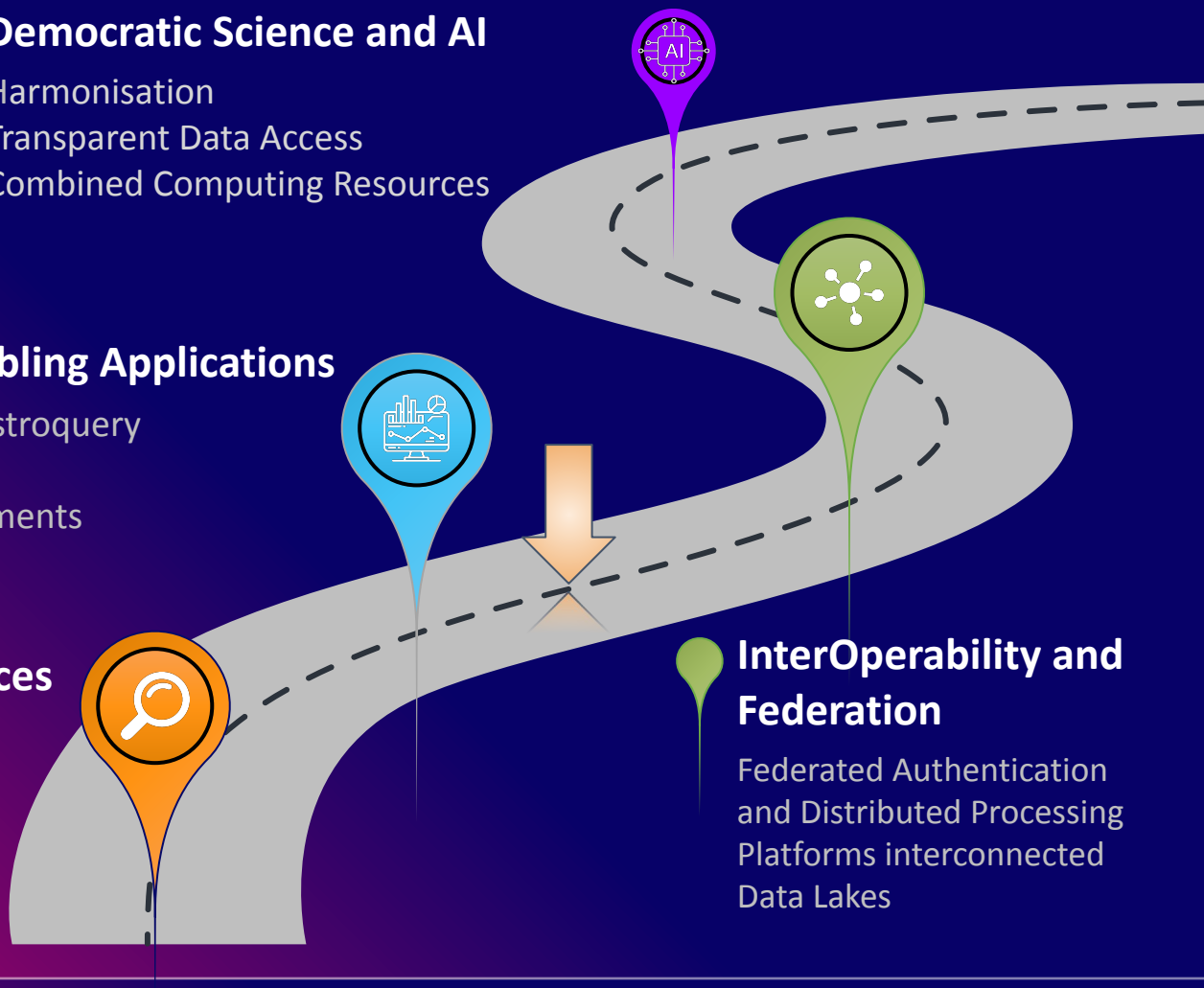
Astropy and Astroquery
Notebooks
Users environments

Discovery And Access Services

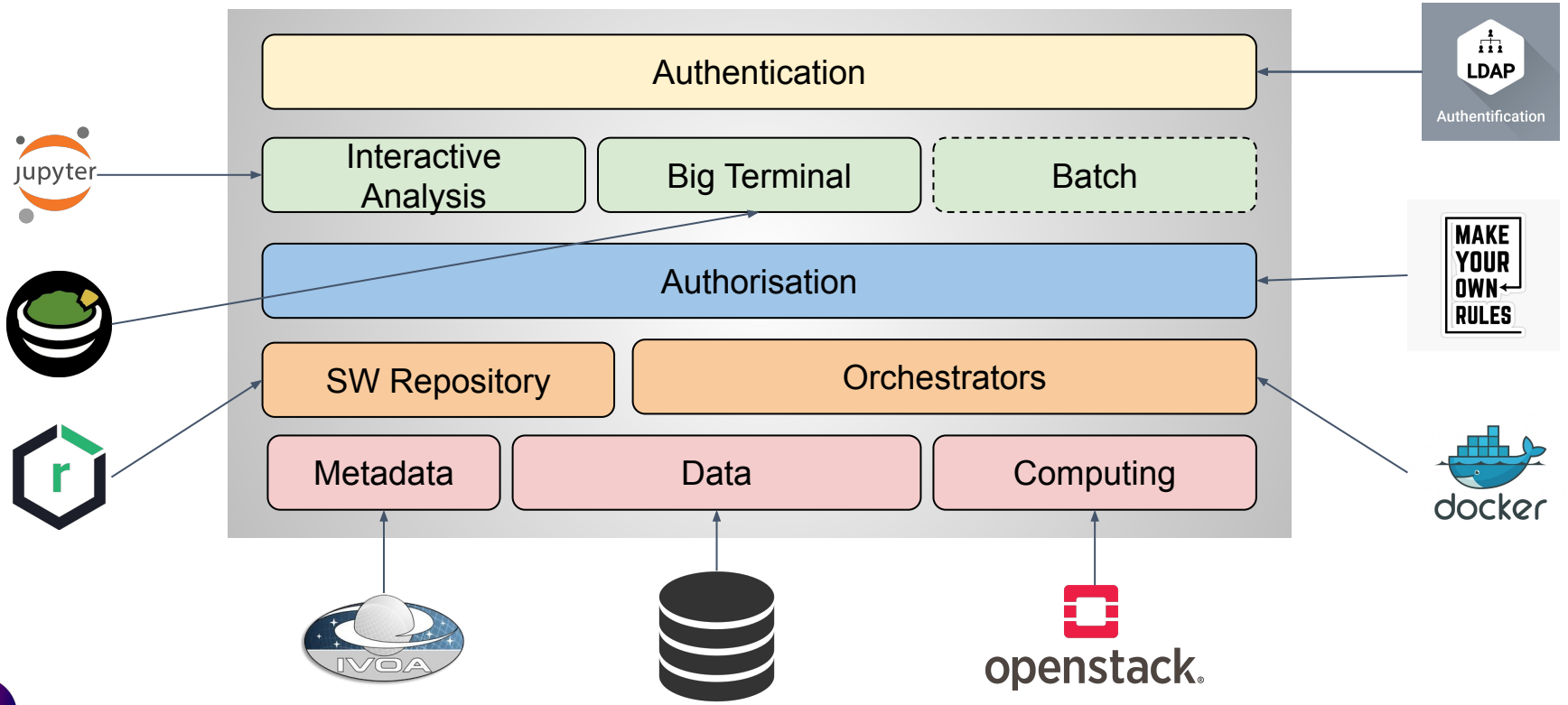
Cone Search
SSAP, SIAP
TAP

InterOperability and Federation

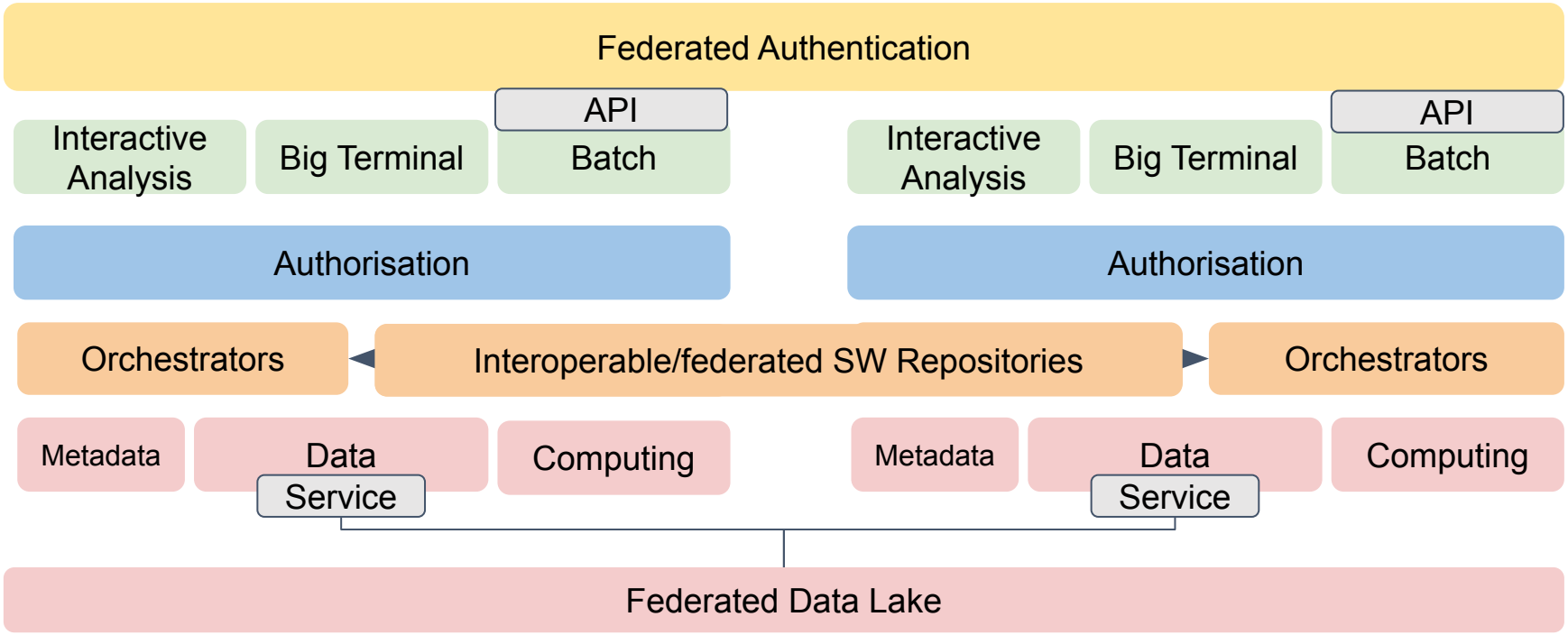
Federated Authentication and Distributed Processing Platforms interconnected Data Lakes



Science platforms



Science Platforms Interoperability

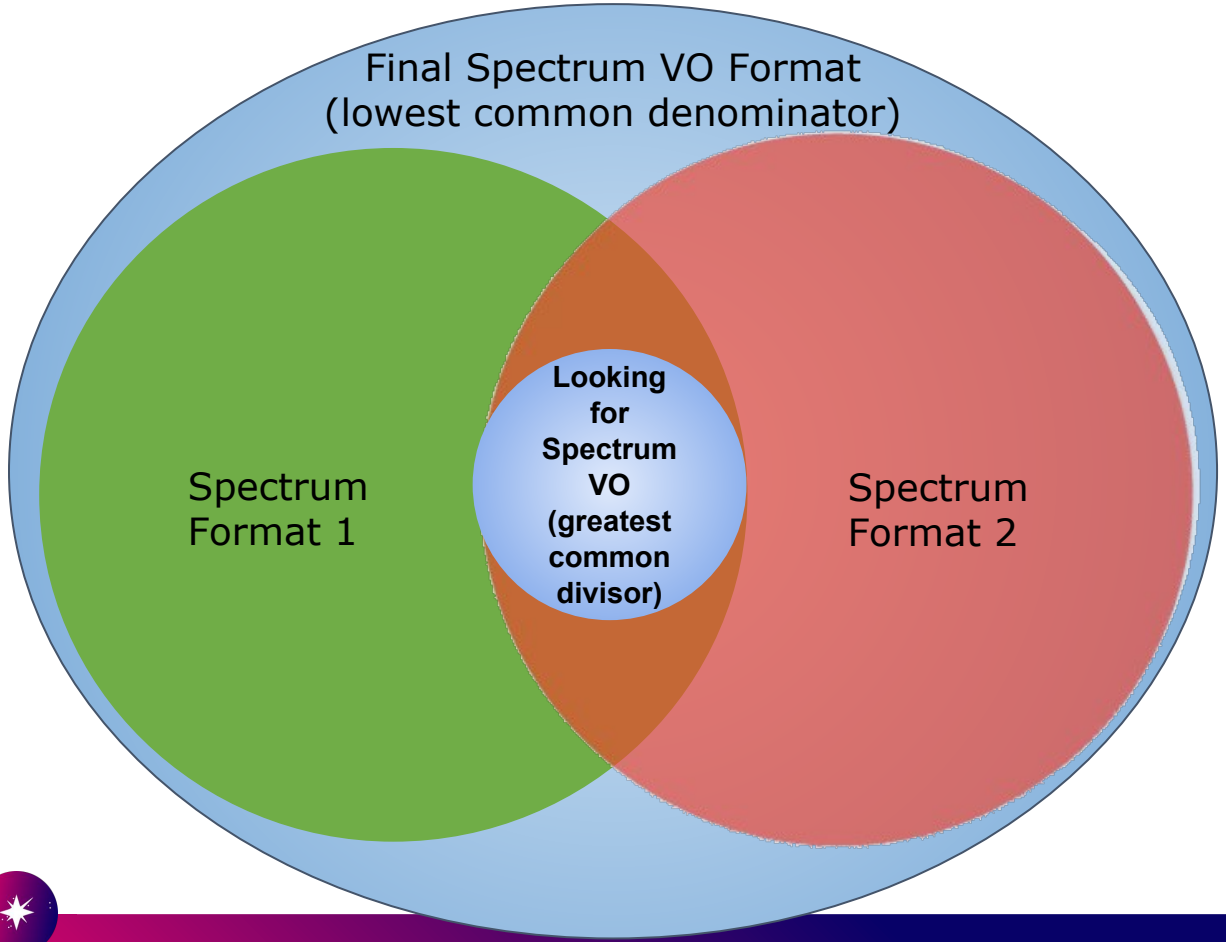


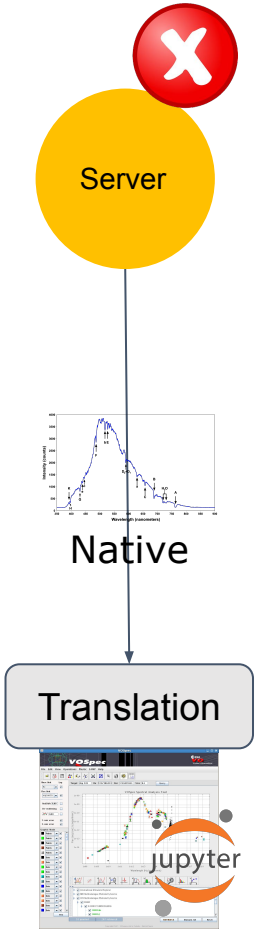
Some possible data mesh services

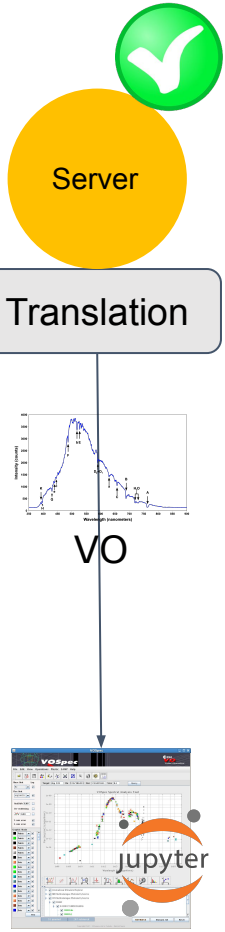
Data Type	Operation	Input	Output
Any Type	Get Stream	ID	Input Stream
Data Cube	Cut-out	ra, dec, size, resolution	Data Cube
Data Cube	Get Spectra	ra, dec, size	Spectrum
Data Cube	Get Time Series	ra, dec, size	Time Series
Data Cube	Get Slice	wavelength	Image
Image	Change Resolution	ra, dec, size, resolution	Image (FITS to HiPS)
Image	Source Extraction	ID, algorithm params	Source Catalogue
Spectrum native	Convert to VO	ID	Spectrum VO
Source Catalogue	Similar Source	Source ID	Source Catalogue

Extend SODA for Remote (Data Atomic) Operations APIs

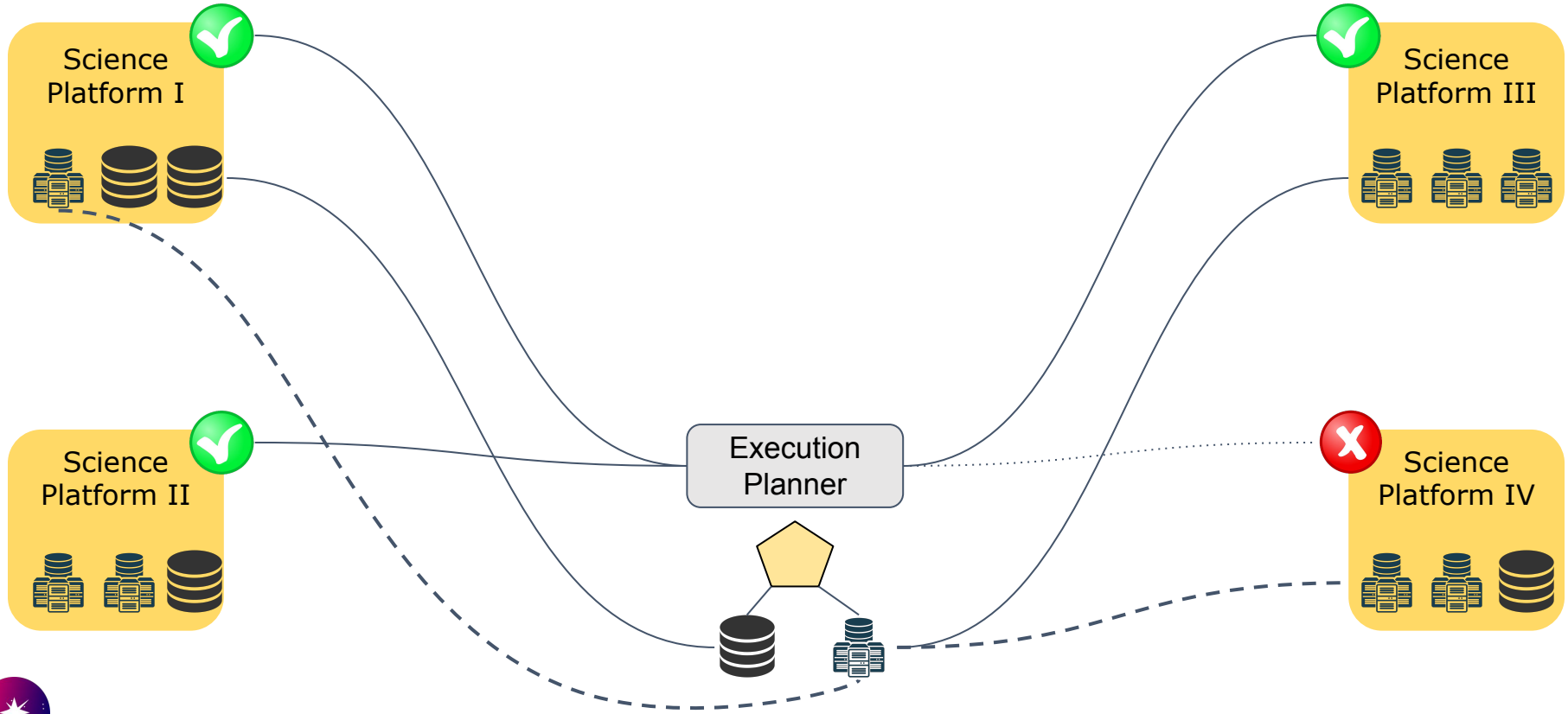
The problem of the formats



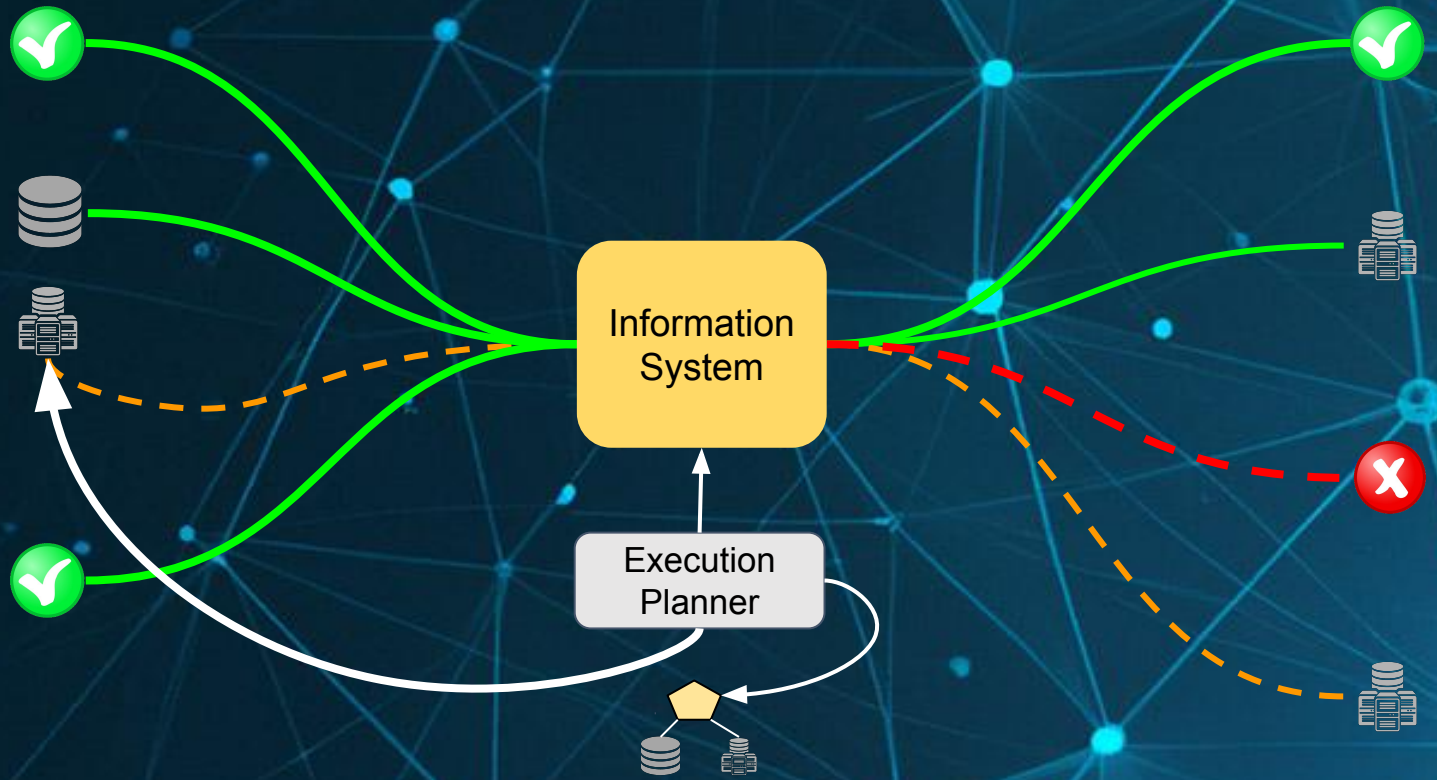




Execution Planner



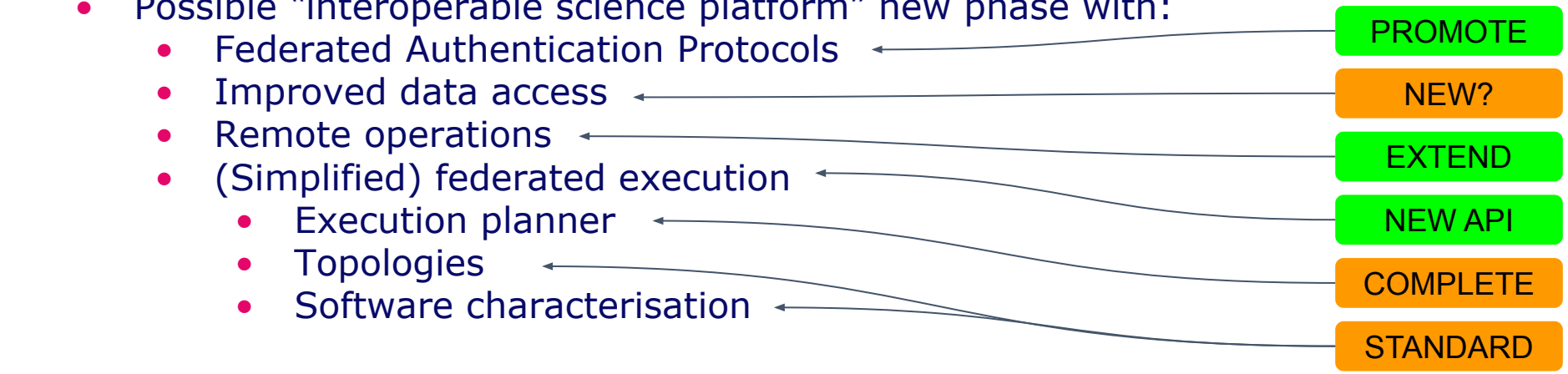
Solving the Topology



Summary

- IVOA provides discovery and access protocols for most of the astronomical data
  - Standards, Integration with scripting languages, Easy publication and collaboration environments
- Many astronomical use cases are enabled due to IVOA standards

- Possible “interoperable science platform” new phase with:
  - Federated Authentication Protocols
  - Improved data access
  - Remote operations
  - (Simplified) federated execution
    - Execution planner
    - Topologies
    - Software characterisation



Thanks for your attention