# Open Science with EOSC in SKA Regional Centres

Sebastián Luna*; Michael Jones; Susana Sánchez-Expósito; Julián Garrido; Javier Moldón; Lourdes Verdes-Montenegro

Instituto de Astrofísica de Andalucía (CSIC)

* sluna@iaa.es ; http://amiga.iaa.es

SKA is an international project to build the largest and most sensitive radio telescope ever conceived. A worldwide distributed network of SKA Regional Centres (SRCs) will host and provide access to the SKA data, to the analysis tools and processing power. The SRCs will be at the core of the exploitation of SKA data, being the place where the science will be done. The Institute of Astrophysics of Andalusia-CSIC (IAA-CSIC) is both developing an SRC that will be aligned with the Open Science Principles and contributing to build EOSC through our participation in the H2020 ESCAPE project and the Spanish Thematic Network in Open e-Science.

In this poster, we show our work to make an astrophysics paper reproducible (Jones et al. (2019)) using existing technologies and tools such as git/GitHub, Conda, Docker, Jupyter notebooks and EOSC resources (the main repository for the project is at https://github.com/AMIGA-IAA/hcg-16). EOSC played a central role in the progress of our work by enabling convenient access to compute (EGI Federated Cloud) and data (EUDAT B2SHARE) resources, forming a collaborative environment in which to develop and test software for the project. We will share our experience with EOSC resources from the end user point of view so other researchers can identify how to benefit from EOSC. This work represents a step forward to bring Open Science into the SRCs and to pave the way in which astronomers and scientists approach the challenge of reproducible science.

SKA

SKA Fact sheets. August 2018. skatelescope.org
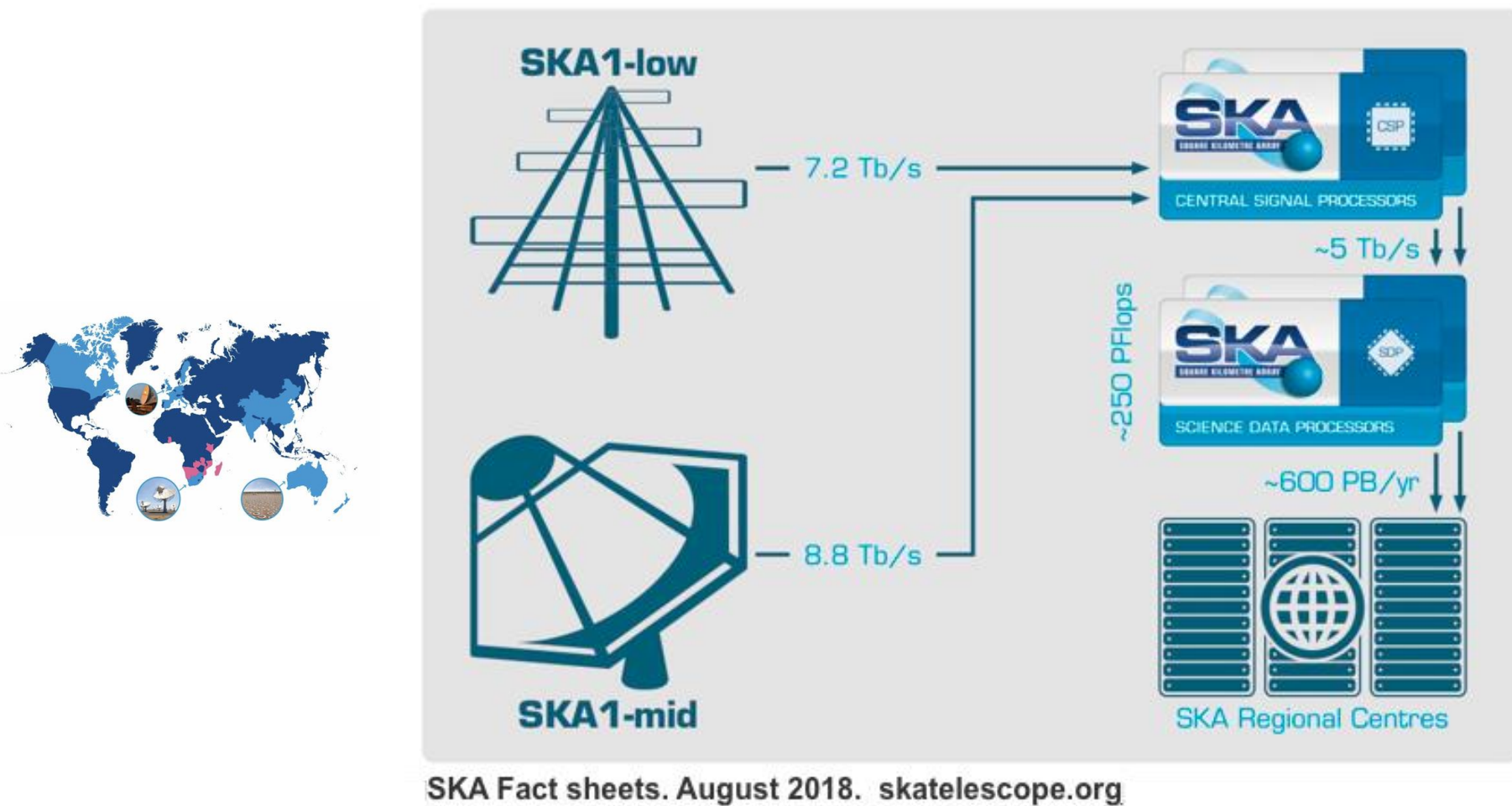
SKA is an international project to build in this decade the largest and most sensitive **radio telescope** ever conceived, being the greatest public data research project, once complete. It will be composed of thousands of antennas on Africa and Australia and it will generate a copious amount of data (estimated 600 PB per year).

SKA Regional Centre prototype at IAA-CSIC
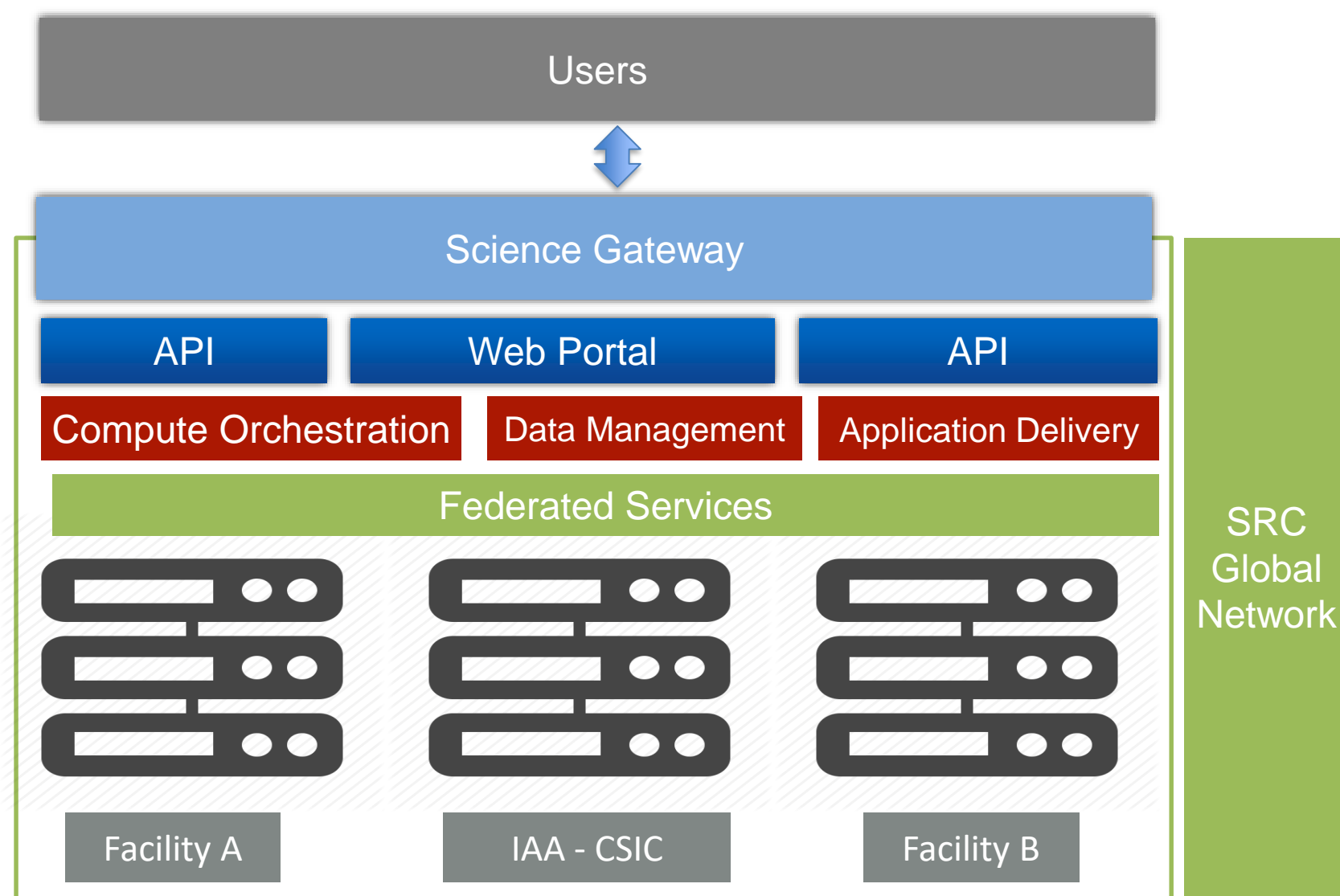
Science Analysis Platform
- Identify technical specs from the use case requirements
- Set-up the associated IAA computing/storage resources
- Partnership with national computing facilities
- Collaborations with e-Infrastructures and other SRCs prototype initiatives
- Deploy a Science Gateway to provide unified access to research services across the SRC Global Network

Case study: HCG-16
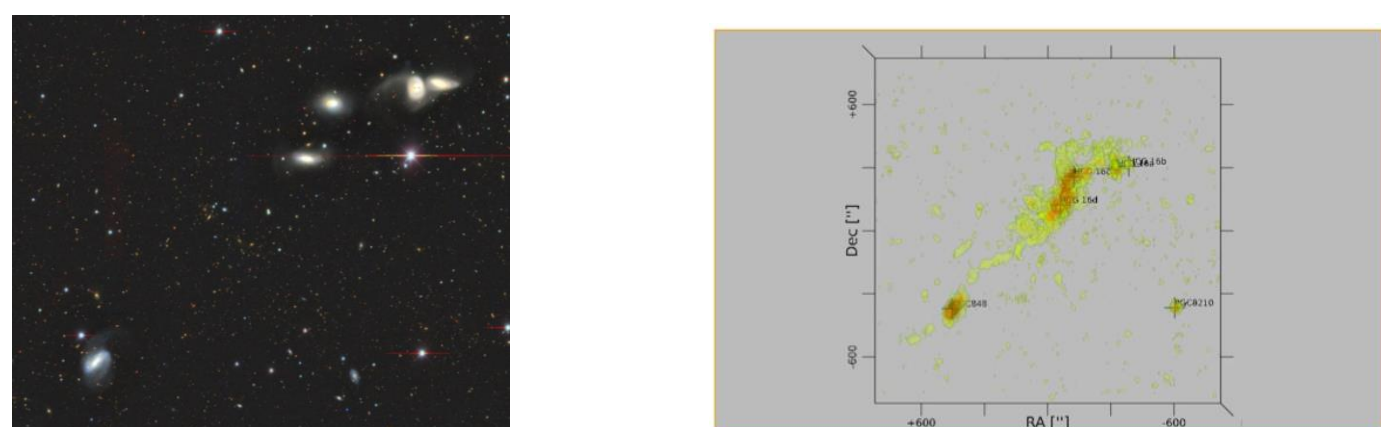
HCG-16 is a complex compact group of galaxies suffering strong interactions and displaying complex structures like tidal tails that can be revealed by radio interferometric observations. We have studied the HI content of the group and the on-going processes that are causing it to change (Astronomy & Astrophysics, Volume 632, A78).

End User experience with EOSC resources

Through the EOSC marketplace we requested compute and storage services from EGI Cloud Compute and EUDAT B2SHARE respectively.

The EGI Cloud Compute service offers a convenient interface to create test and collaborative environments for scientific analyses. However, we have noticed that getting virtual machines running is easier in some sites than others in the fedcloud Virtual Organisation.
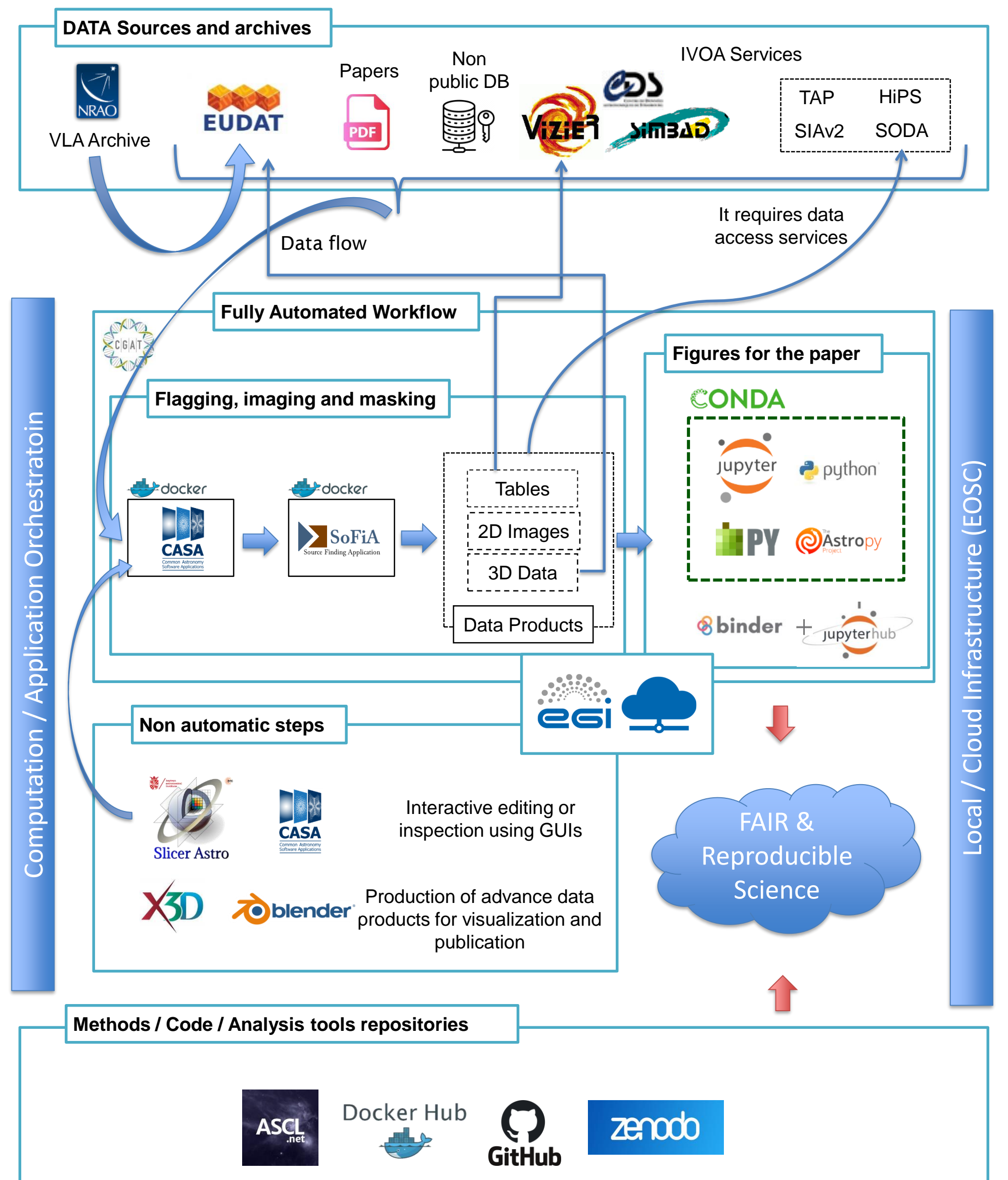
The EUDAT B2SHARE storage service makes it extremely easy to store and share scientific data automatically assigning a DOI and PID to each upload. It also allows adding a scientific discipline which helps with discoverability of related data. However, we find it difficult to locate under what terms of use our data is stored (eg. data retention).

Fostering SRCs as Open Science Hubs

It will not be feasible to move data volumes in the order of PB per observation around. Instead, a worldwide distributed network of SRCs will provide access to the data, analysis tools and processing power. Therefore, the SRCs will be at the core of the exploitation of SKA data, being the place where the science will be done.

An Open Science implementation will facilitate sharing data, resources and tools across the global SKA community through the SRCs. Users will be able to verify, reuse and repurpose methods, accelerating discovery and the transfer of knowledge.

Towards Open Science with existing tools

Instituto de Astrofísica de Andalucía, IAA-CSIC

EXCELENCIA SEVERO OCHOA