The ASTRON Science Data Centre

John Swinbank — swinbank@astron.nl — SDC Program Manager

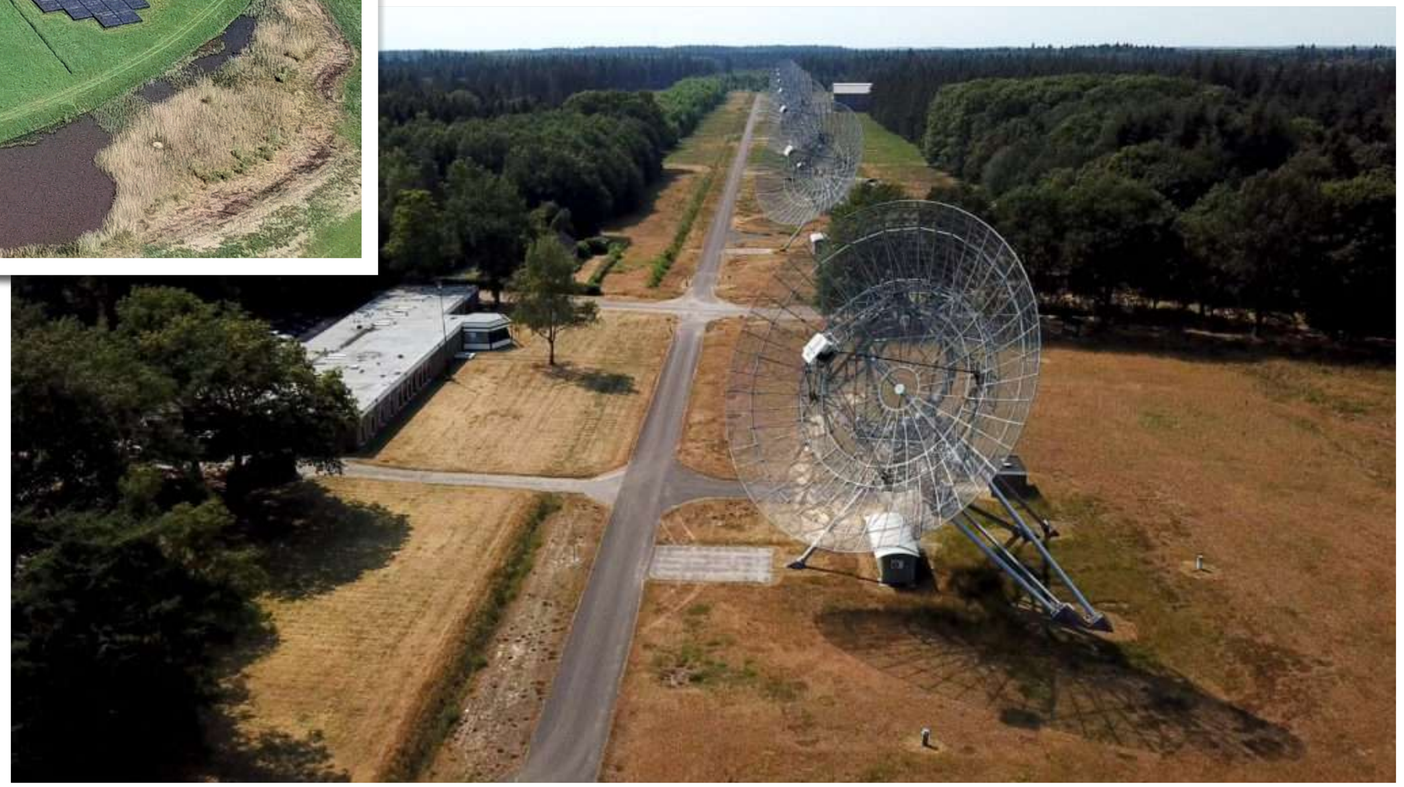
Yan Grange, Michiel van Haarlem, Hanno Holties, Mattia Mancini,
Zheng Meyer-Zhao, Roberto Pizzo, Nico Vermass

ASTRON
Netherlands Institute for Radio Astronomy

2021-06-28

WSRT-APERTIF

14 steerable 25 metre dishes (12 with PAFs)
1.1 – 1.4 GHz
Survey operations to end 2021
~7.5 PB
https://alta.astron.nl



LOFAR

100,000 dipoles across Europe
30 – 240 MHz
Major upgrade underway (“LOFAR 2.0”)
~50+ PB
https://lta.lofar.eu

The Mission

The ASTRON Science Data Centre will enable fundamental scientific discoveries by ***providing access to tools****; instrumental & science-ready* ***data products****; and* ***data processing & analysis services*** *that enable astrophysicists of all levels of experience to work effectively with the extremely large and complex data products produced by current and future radio astronomy instrumentation, with a particular focus on* ***LOFAR****,* ***Apertif*** *and the* ***Square Kilometre Array****.*

The ASTRON SDC is (...or will be...)

An operational facility, available to the community

A provider of specialist software and pipelines

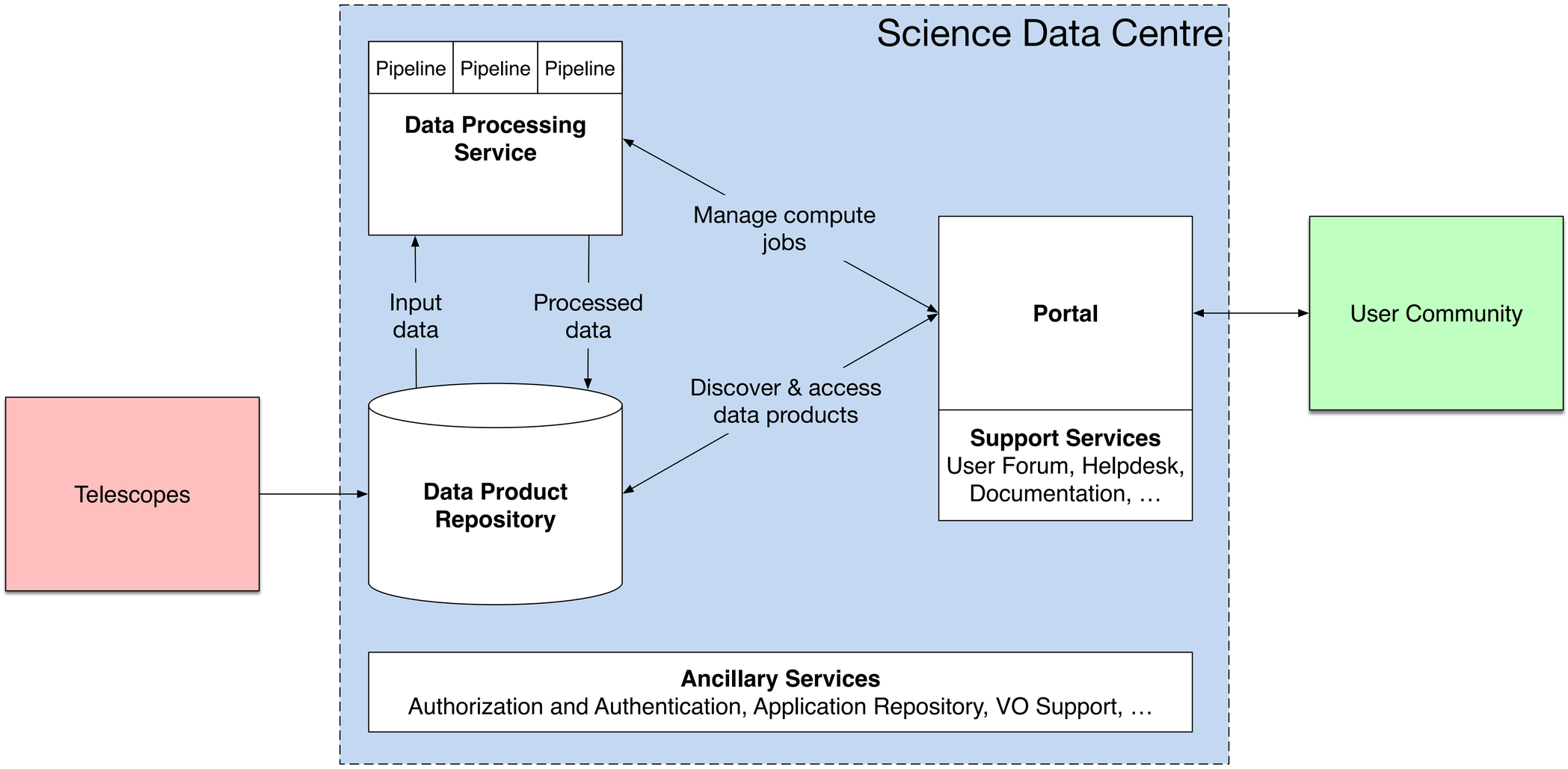
The home of support and expertise for LOFAR & Apertif

Contributing to, and learning from, international projects

Preparing to host SKA data in the Netherlands

The SDC Facility Provides

- A “**one stop shop**” for users of instrumentation operated by ASTRON.
  This is expected to include:
  - Proposal preparation and submission tools.
  - Access to instrumental data.
  - Generation of, and access to, standardized “science-ready” data products.
  - User-friendly interfaces to discover and access data.
  - On-line interactive analysis facilities (e.g. Jupyter notebooks).
  - Giving users the ability to define and run their custom pipelines in the SDC framework.
  - A curated library of tools designed to work with our data.
  - ...etc!
- The primary means of access to support, documentation, and expertise on instrumentation operated by ASTRON.


Science Data Centre
Pipeline
Pipeline
Pipeline
Data Processing Service
Manage compute jobs
Input data
Processed data
Portal
User Community
Discover & access data products
Telescopes
Data Product Repository
Support Services
User Forum, Helpdesk, Documentation, …
Ancillary Services
Authorization and Authentication, Application Repository, VO Support, …

Data Product Repository

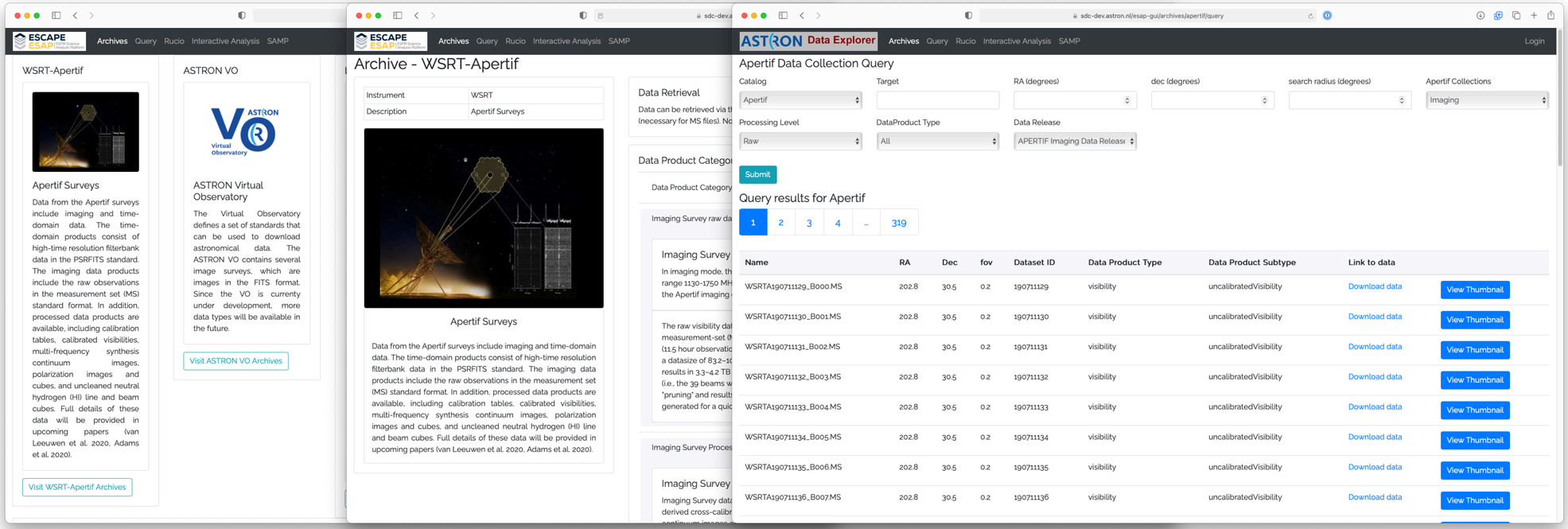
- Provides bulk storage of low-level (instrumental) and advanced (science-ready) data products, as well as structured data (e.g. source catalogues).
- Abstracted physical infrastructure: the repository may be spread over multiple sites, while providing end users with a logically coherent view of their data...
- ...but optimizing compute jobs to take account of data locality.
- Deep integration with the VO and other data sharing tools.

Pipelines & Processing

- The Data Processing Service will be generic with respect to the pipelines being executed: user-supplied pipelines will be executable if they are written to the appropriate API (and subject to policy, resource allocation, etc).
- Underlying infrastructure TBD, but we anticipate ultimately providing a mix of high- and low-memory-per-core systems as well as accelerators like GPUs.
- ASTRON will support a set of standard pipelines, initially focusing on LOFAR imaging.
  - ...these will be made available both for use inside the SDC, and released directly to the community.
- SDC staff will execute those pipelines and make available science-ready data products (subject to data rights!).

Portal

- Starting from the work on the ESFRI Science Analysis Platform, being developed as part of the ESCAPE project, and building from there.
- We hope to be rolling out upgrades to ASTRON's existing archive interfaces as early as this year.
- Future plans for bulk compute & interactive data analysis (Jupyter...).



ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

Where next?

- More coming in this session!
  - Yan Grange, *Science analysis platform and data lake for the ASTRON Science Data Centre*, in ~10 minutes.
  - Marco Iacobelli, *Valorising the multi PetaByte scale LOFAR data collection*, at 11:17 CEST.
  - André Offringa, *Next level processing of radio astronomy data: The Rapthor framework*, at 11:47 CEST.
- SDC services gradually rolling out across ASTRON over the coming months and years; watch www.astron.nl for further announcements.
- Feedback/requests/comments welcome: swinbank@astron.nl.