ASTRON

Netherlands Institute for Radio Astronomy

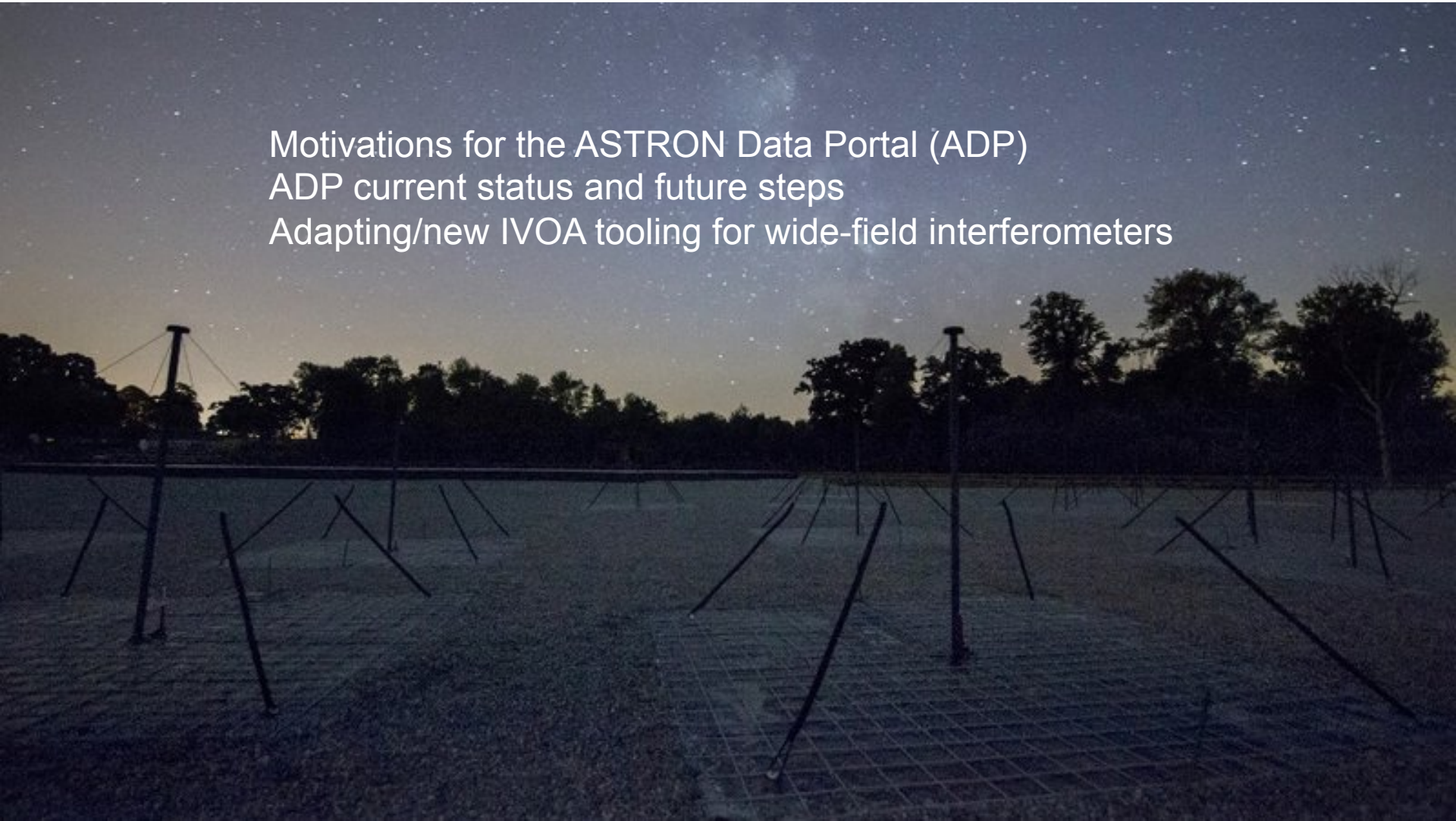
The ADP: enabling access and exploitation of radio data collections through the IVOA

Marco Iacobelli
LOFAR telescope scientist

Asterics European Data Provider Forum and Training Event
Heidelberg, June 28th 2018

I-LOFAR Intern Luis Alberto Canizares

Outline



Motivations for the ASTRON Data Portal (ADP)
ADP current status and future steps
Adapting/new IVOA tooling for wide-field interferometers

ASTRON data collections

A large and heterogeneous collection of scientific data product from different radio interferometers:

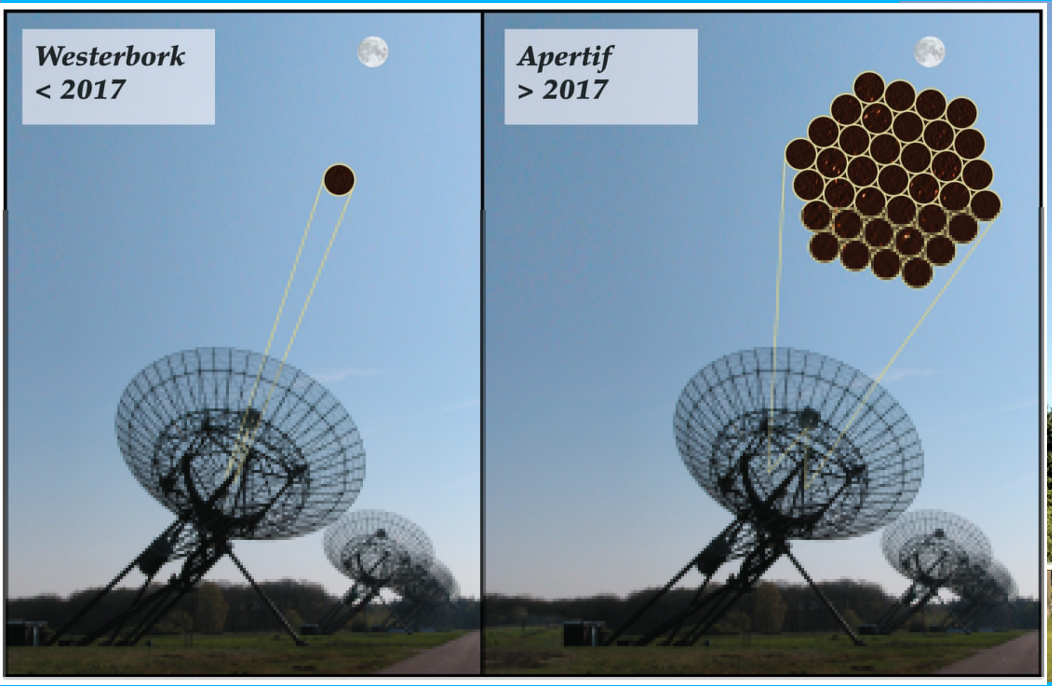
International LOFAR Telescope: 51 stations ➔ 38 NL+13 Europe

Data transport rates over the array of order 150 Gbit/s:

- new generation wide-field interferometer (SKA pathfinder)
- data storage challenges (35 Tb/h of raw, correlated visibilities)

ASTRON data collections

A large and heterogeneous collection of scientific data product from different radio interferometers:

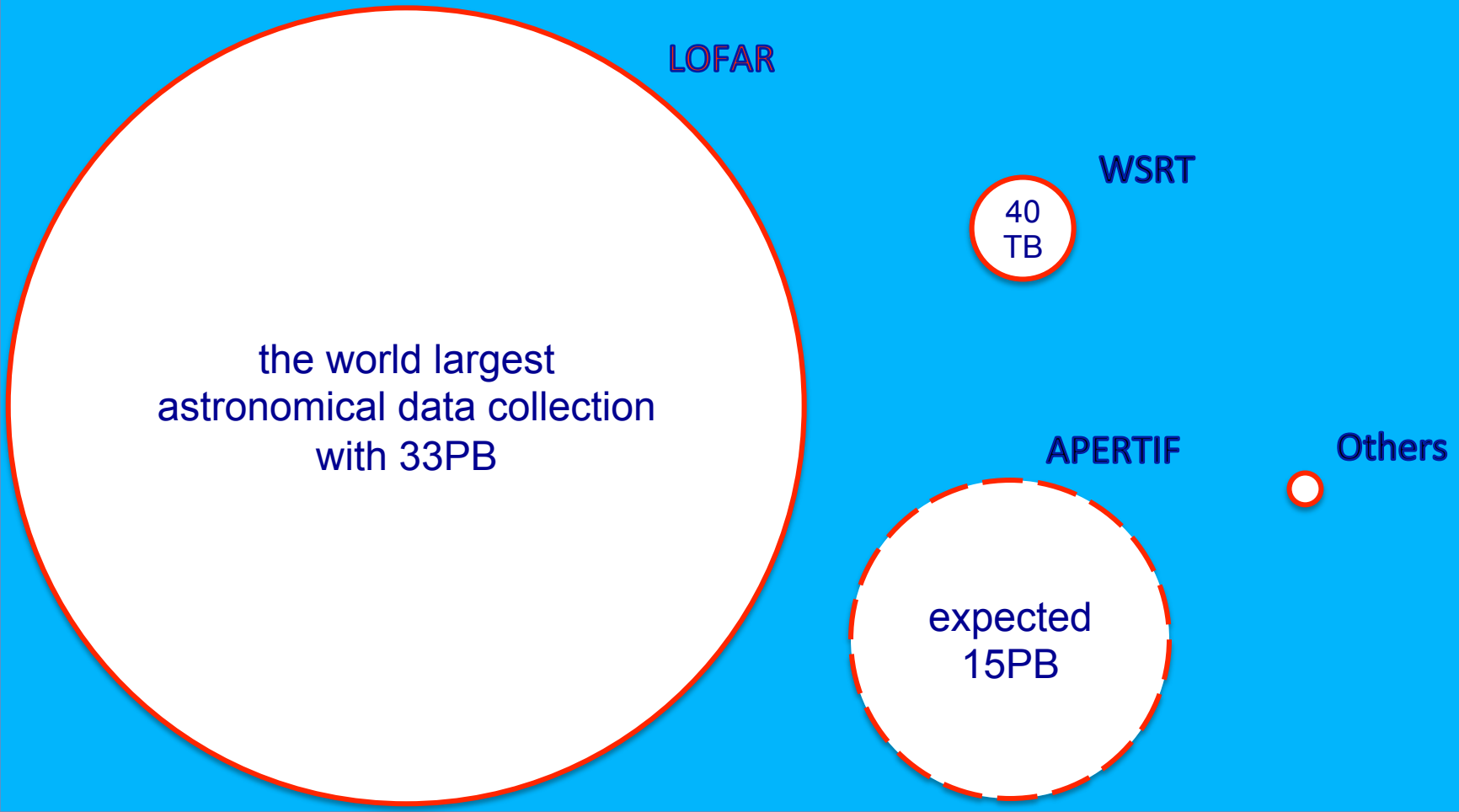


WSRT Telescope / APERTIF:

- Telescope operational since 1970
- 2nd life as wide-field interferometer due to new generation of receiver array

ASTRON data collections

A large and heterogeneous collection of scientific data product from different radio interferometers:



ASTRON data collections

A large and heterogeneous collection of scientific data product from different radio interferometers:

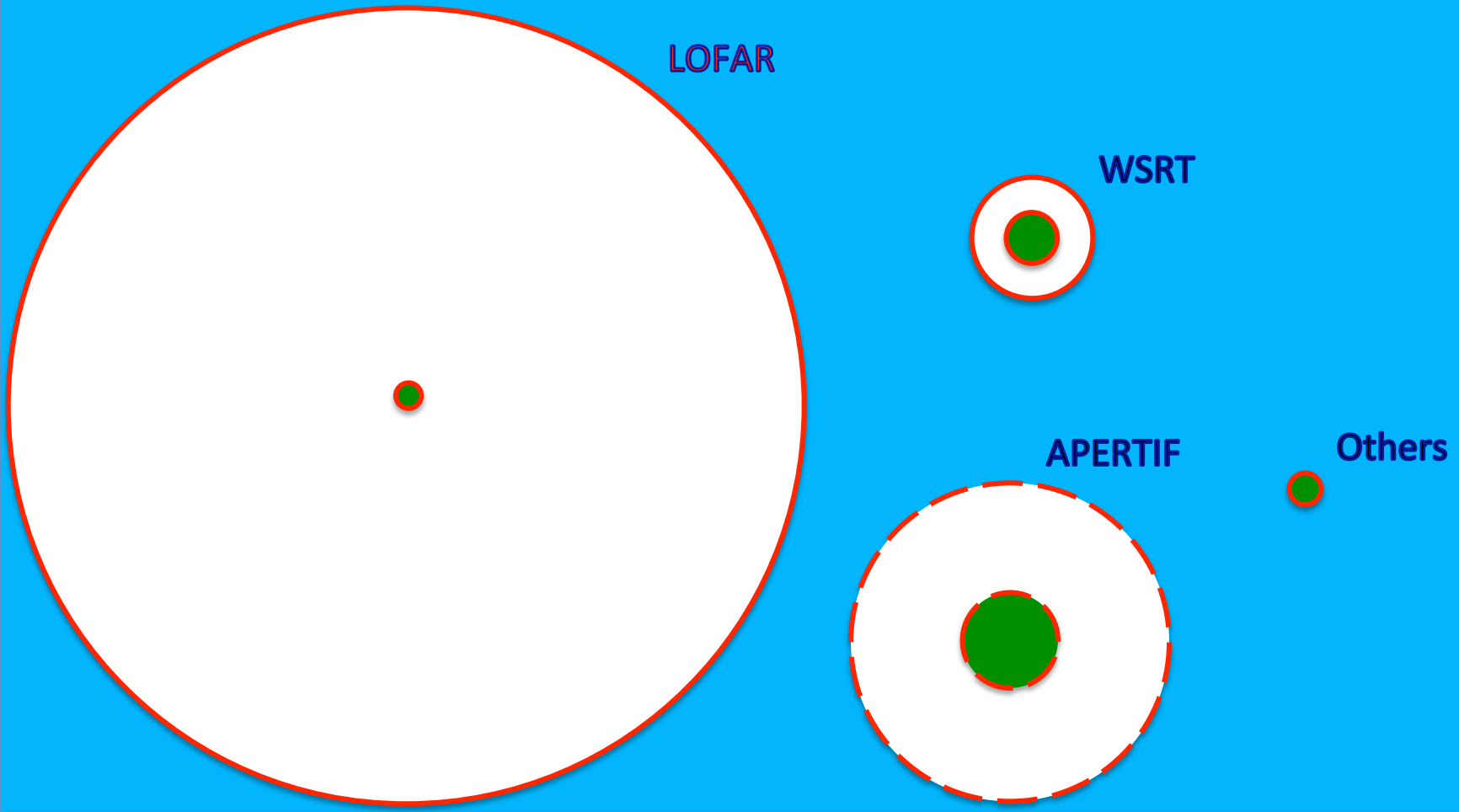
- Different data types (level 0 to 3)
- Different sizes 10^{-4} – 10^{2} TB

Domain	Data Type	NASA/EOS Terminology	Description
Instrument	Device monitoring data	N/A (Level-1)	Produced by instrumentation, typically not preprocessed, and typically not stored as part of raw data products. Useful for preparing trending data to detect emerging instrument failures, and providing operational responses to other failures to dynamically improve the potential for producing quality data products.
Instrument	Raw data	Level 0	Produced by instrumentation. May be subject to limited preprocessing in firmware (for example, autocorrelation spectrometers).
Instrument	Calibrated data	Level 1	Produced by removing instrumental and environmental effects. EOS breaks this down into Levels 1A (raw data appended with annotations and calibration information) and Level 1B (raw data processed to calibrated data).
Science	Derived data	Level 2	Produced by combining calibrated data with other calibrated data, or with other derived data, according to processes, techniques, or algorithms. Scientific analysis can take place at this level or any higher level.
Science	Assimilated data	Level 3	Produced by gridding, resampling, and/or changing the frame of reference for derived data.
Science	Model data	Level 4	Produced by applying one or more mathematical, physical, or stochastic models to collections of assimilated and derived data products.

©2006, ASQ

ASTRON data collections

A large and heterogeneous collection of scientific data product from different radio interferometers:



ASTRON data collections

A large and heterogeneous collection of scientific data product from different radio interferometers:

- Different data types (level 0 to 3)
- Different sizes 10^{-4} – 10^{2} TB
- Data sharing impacted by the network limitations, especially for level 0 data products.
- Processing requires large computing power, which can not be available to a generic user.
- RO and KSPs are exploring the option to move the processing directly to the archiving sites. A successful pilot project is ongoing at SARA. Currently being replicated at Julich.

Domain	Data Type	NASA/EOS Terminology	Description
Instrument	Device monitoring data	N/A (Level-1)	Produced by instrumentation, typically not preprocessed, and typically not stored as part of raw data products. Useful for preparing trending data to detect emerging instrument failures, and providing operational responses to other failures to dynamically improve the potential for producing quality data products.
Instrument	Raw data	Level 0	Produced by instrumentation. May be subject to limited preprocessing in firmware (for example, autocorrelation spectrometers).
Instrument	Calibrated data	Level 1	Produced by removing instrumental and environmental effects. EOS breaks this down into Levels 1A (raw data appended with annotations and calibration information) and Level 1B (raw data processed to calibrated data).
Science	Derived data	Level 2	Produced by combining calibrated data with other calibrated data, or with other derived data, according to processes, techniques, or algorithms. Scientific analysis can take place at this level or any higher level.
Science	Assimilated data	Level 3	Produced by gridding, resampling, and/or changing the frame of reference for derived data.
Science	Model data	Level 4	Produced by applying one or more mathematical, physical, or stochastic models to collections of assimilated and derived data products.

©2006, ASQ

ASTRON data collections

A large and heterogeneous collection of scientific data product from different radio interferometers:

- LOFAR: ASTROWISE ➔ need for improvements: chance for new algorithms
- WSRT: de-commissioned + WENSS
- APERTIF: ALTA ➔ under development
- Others (MSSS, LBCS, LOTSS, TGSS) ➔ VO@ASTRON

The ADP scopes

In view of an European Science and Data Centre, ASTRON has the ambition to set up a science data portal to serve a broad astronomical community to:

- support the exchange of data by scientists
- maximise scientific return from existing data collections.

Main goals:

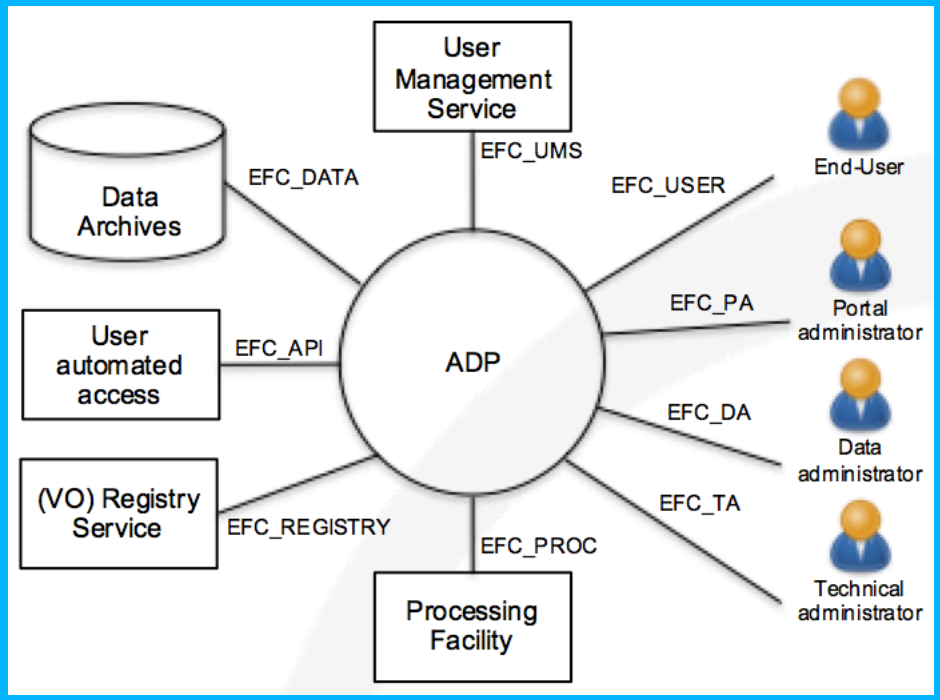
- exposing LOFAR, WSRT and APERTIF data collections, following principles of Findability, Accessibility, Interoperability, and Reusability
- added value services ➔ pipelines, analytics and visualization

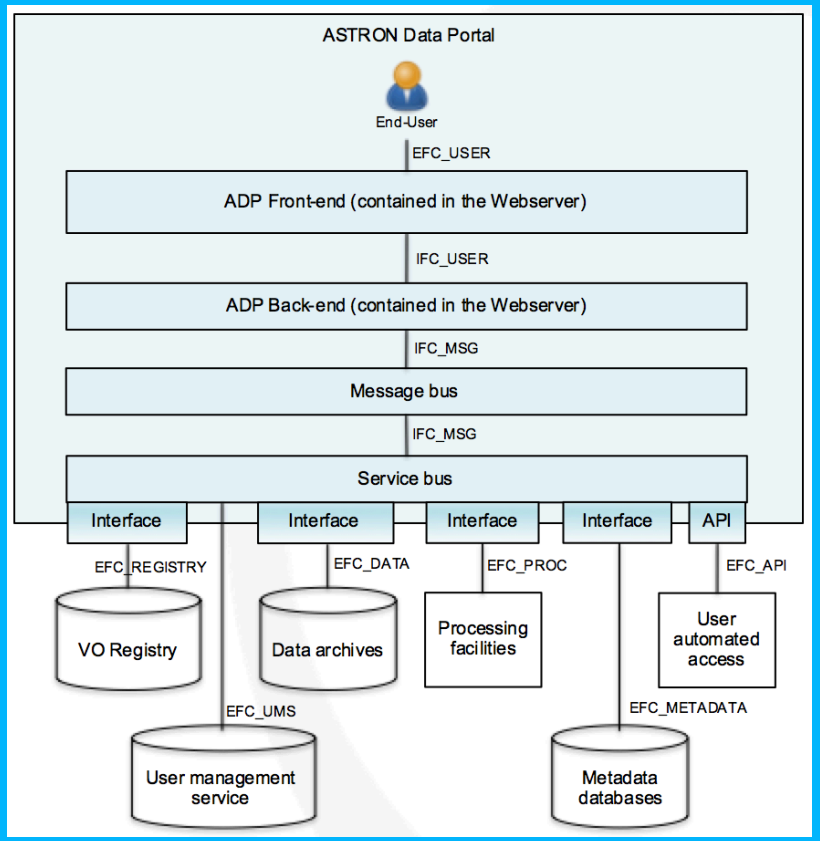
The portal can grow wider by including data collections of future instruments (MeerKat, NCLE, SKA).

The ADP overview

The ADP will consist of:

- a graphical user interface (frontend), publicly available on the internet.
- interfaces to the existing data archives that ASTRON uses (backend), to query the different data archives and to translate the results back to users.





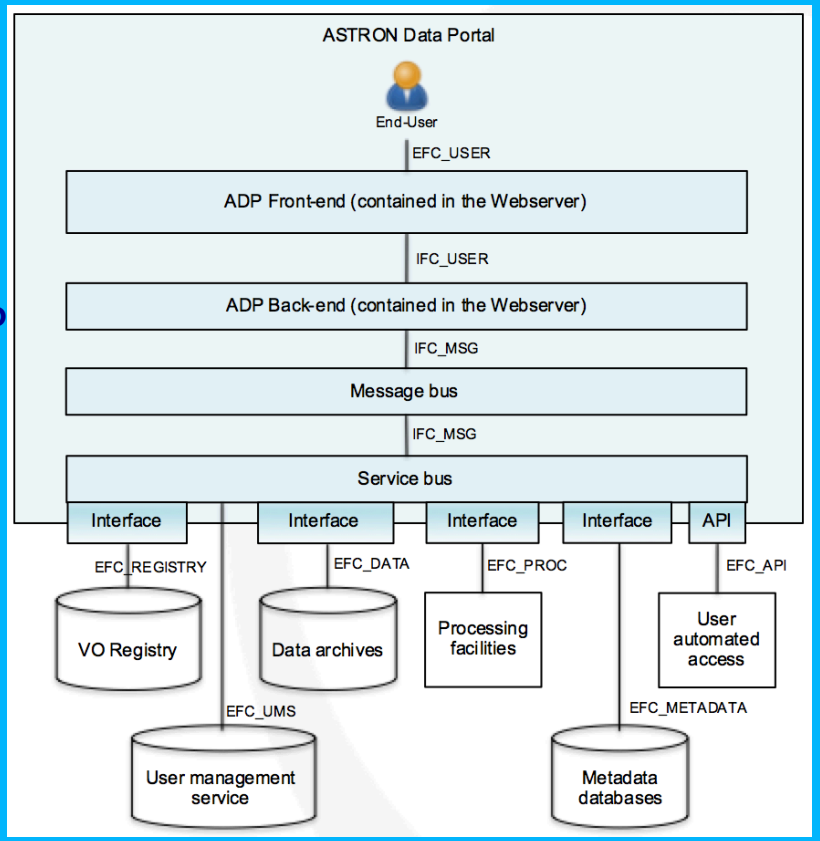
The ADP overview

The ADP will consist of:

- a graphical user interface (frontend), publicly available on the internet.
- interfaces to the existing data archives that ASTRON uses (backend), to query the different data archives and to translate the results back to users.

The ADP will adopt IVOA standards for, e.g.:

- visualization ➔ Aladin / HiPS
- protocols ➔ADQL for queries / TAP / SAMP
- data model ➔ ObsCore / Datalink



The ADP roadmap

Current strategy:

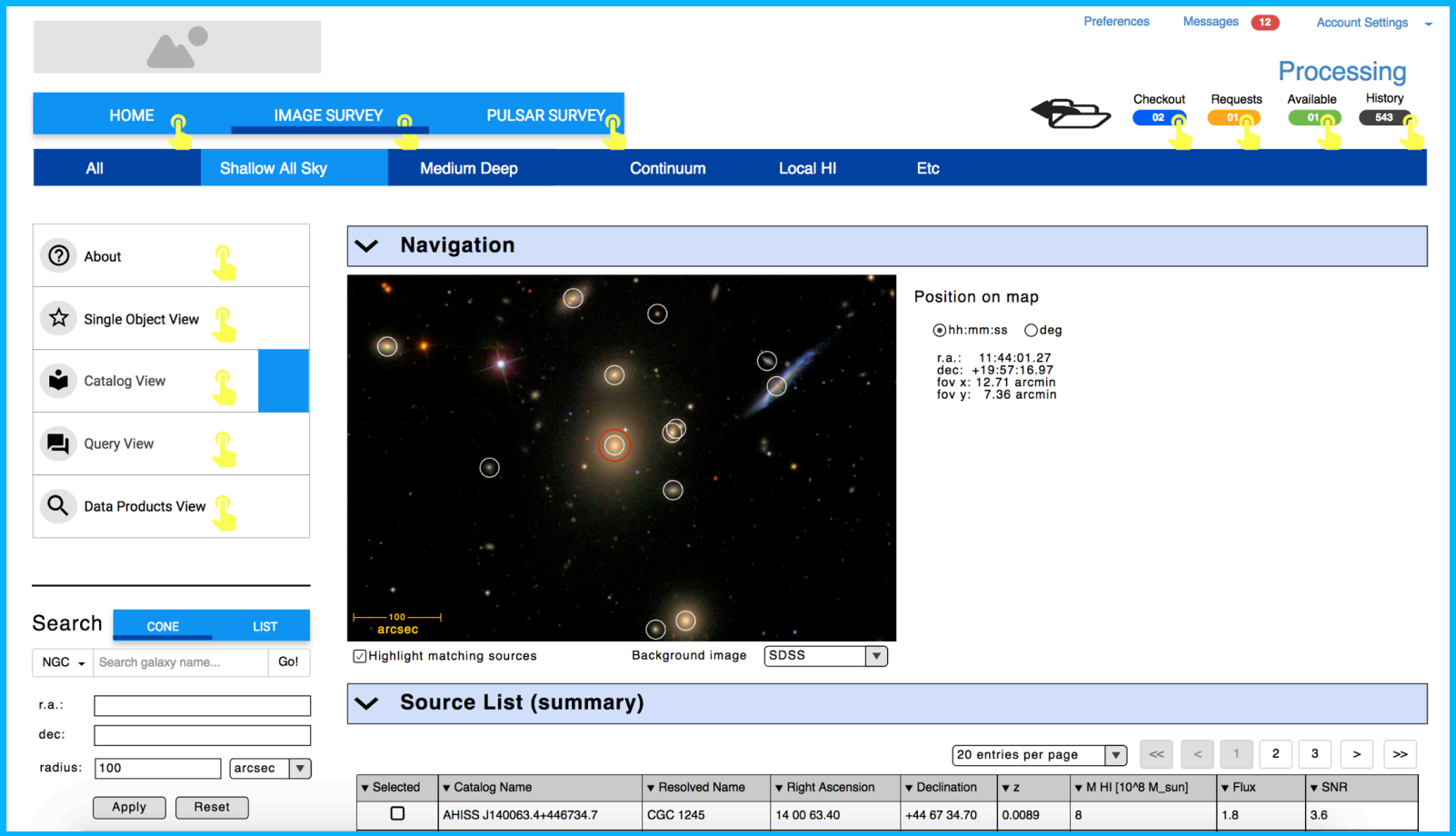
- long term goal: a platform for scientists to do science with ASTRON provided data collections, including services to process and analyse data
- short term goal: a simple but usable release of the ADP

Project status: concept system design / system requirements phase finalized

- the ADP will be developed iterative and scalable
- phase 1 ➔ development & implementation in 1 year
  - search & retrieve interface
  - user management (& user area)

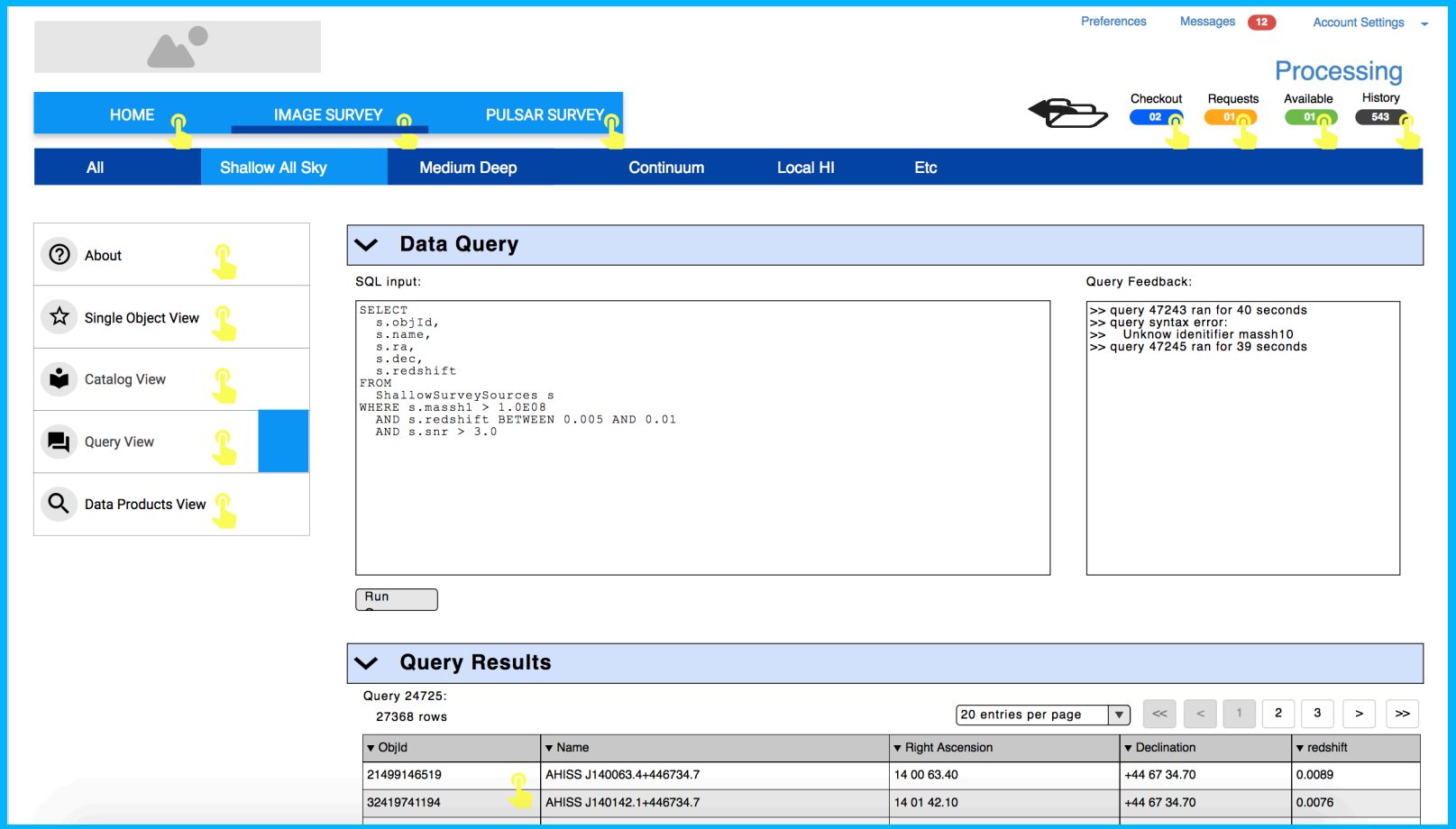
The ADP design

Design inspired by the ALTA (in progress) and ESA Sky 2.x portals



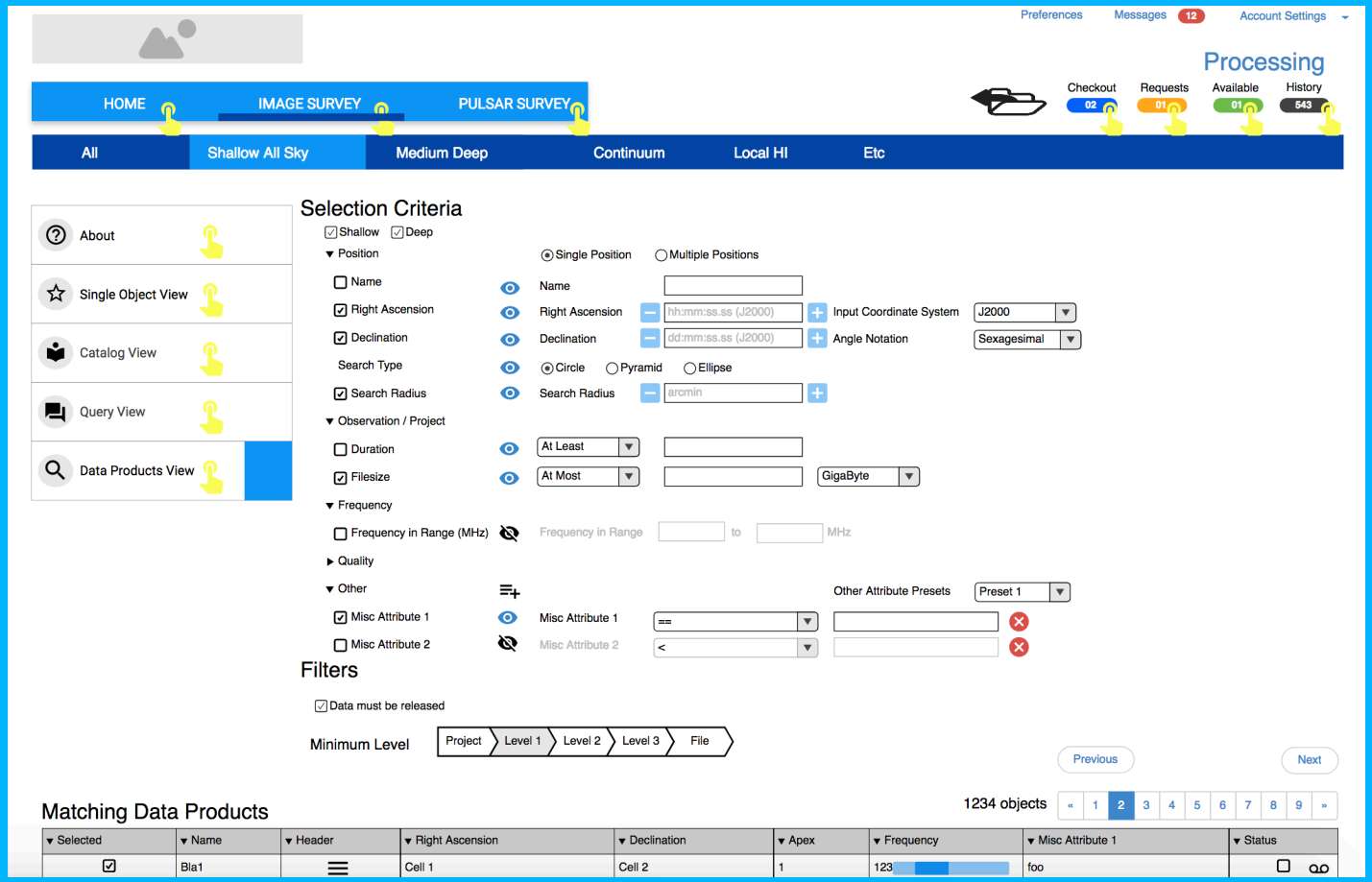
The ADP design

Design inspired by the ALTA (in progress) and ESA Sky 2.x portals



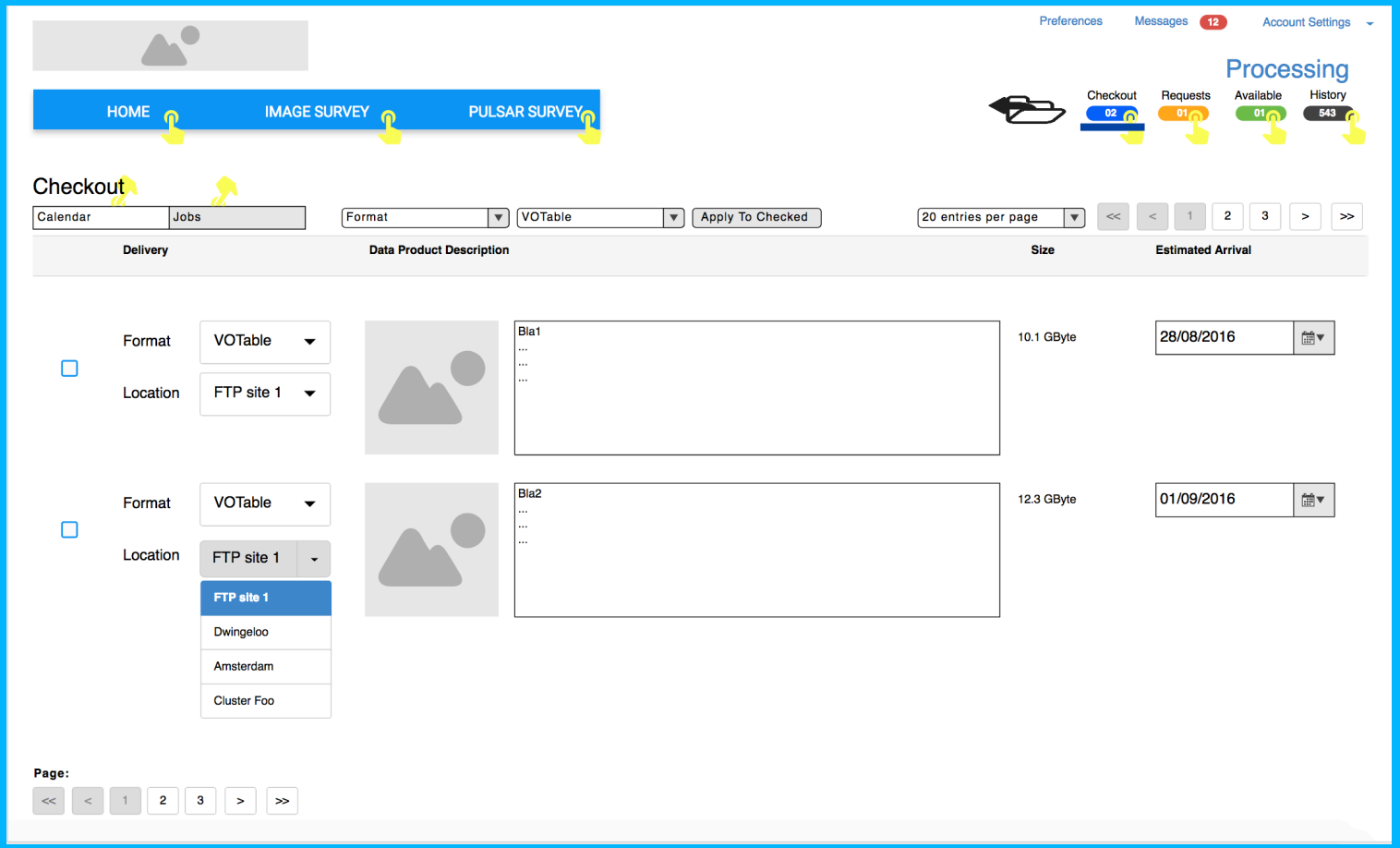
The ADP design

Design inspired by the ALTA (in progress) and ESA Sky 2.x portals



The ADP design

Design inspired by the ALTA (in progress) and ESA Sky 2.x portals



The ADP challenges

Users management & private area (authentication)

- authentication vs federation ➔ federated identity management

Data modeling

- see A. Renting talk

(Meta-)Data quality

- which standards and visualization / flagging tools ?

Processing (including ingest) & analysis area

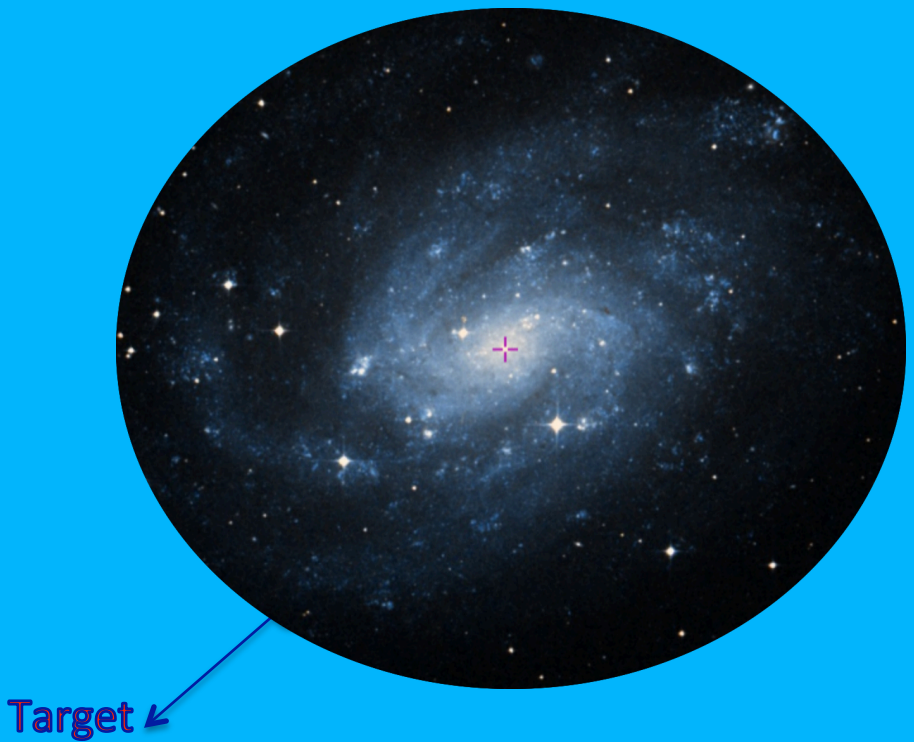
- allocation and monitoring of resources crucial for operations

The ADP & new tooling

Characterization tools needed when dealing with level 0/1 data products from wide-field interferometers ➔ accounting for beam-shapes of different modes for searches and selection but also for processing / analysis

The ADP & new tooling

Characterization tools needed when dealing with level 0/1 data products from wide-field interferometers ➔ accounting for beam-shapes of different modes for searches and selection but also for processing / analysis

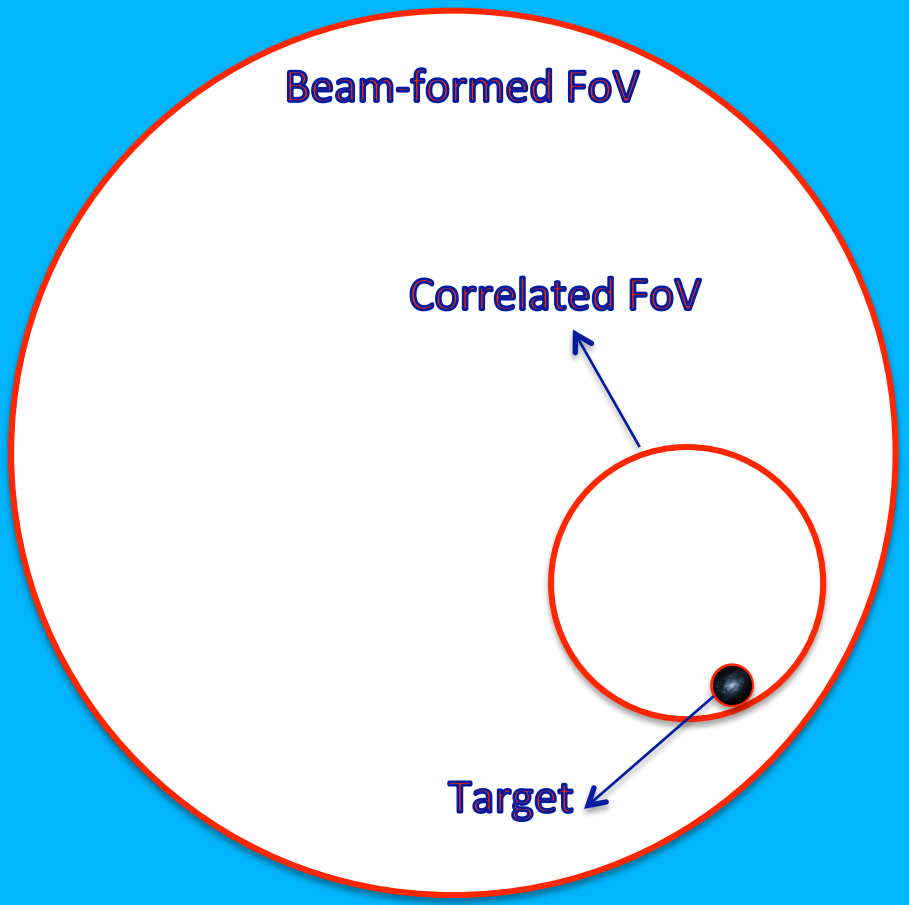


The ADP & new tooling

Characterization tools needed when dealing with level 0/1 data products from wide-field interferometers ➔ accounting for beam-shapes of different modes for searches and selection but also for processing / analysis

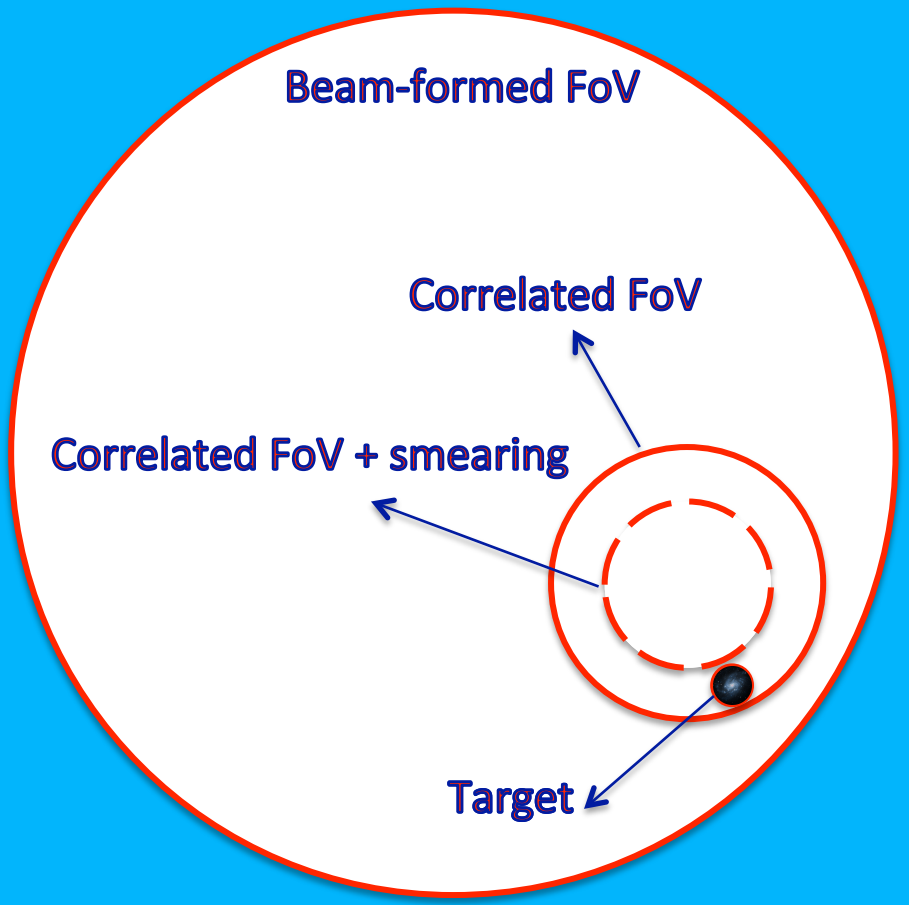
The ADP & new tooling

Characterization tools needed when dealing with level 0/1 data products from wide-field interferometers ➔ accounting for beam-shapes of different modes for searches and selection but also for processing / analysis



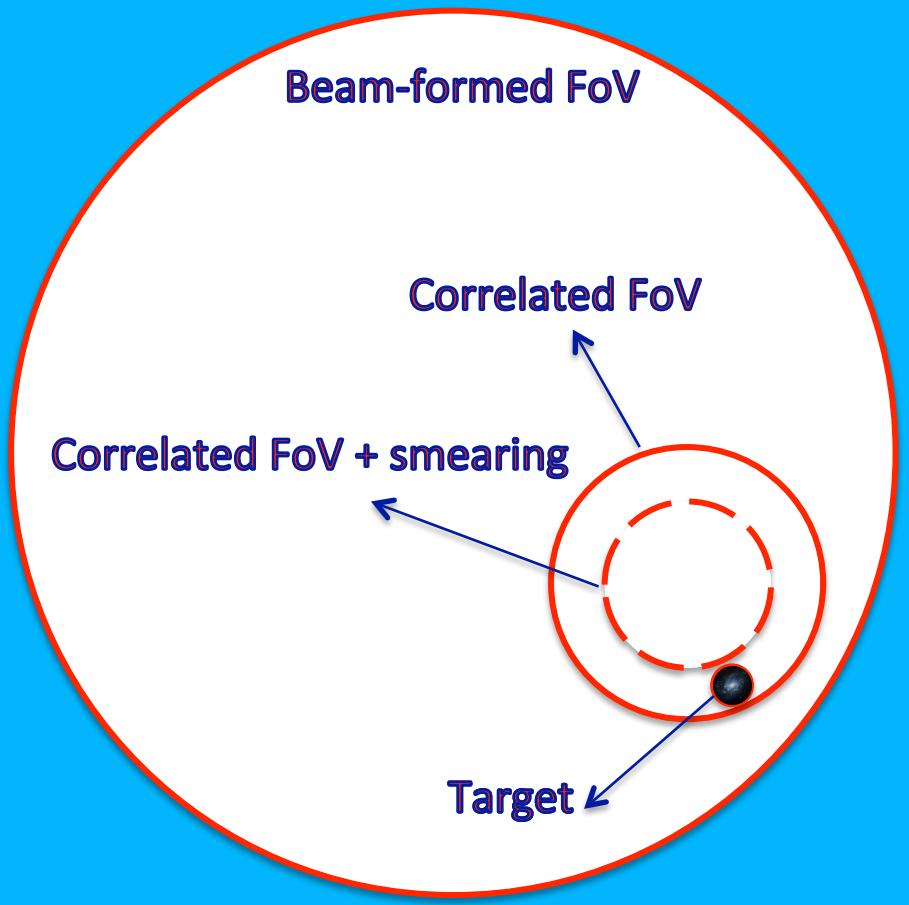
The ADP & new tooling

Characterization tools needed when dealing with level 0/1 data products from wide-field interferometers ➔ accounting for beam-shapes of different modes for searches and selection but also for processing / analysis



The ADP & new tooling

Characterization tools needed when dealing with level 0/1 data products from wide-field interferometers ➔ accounting for beam-shapes of different modes for searches and selection but also for processing / analysis



Observed / processed data characterization in terms of:

- uv-coverage
- beam pattern (FoV & sensitivity)

Accurate reconstruction:

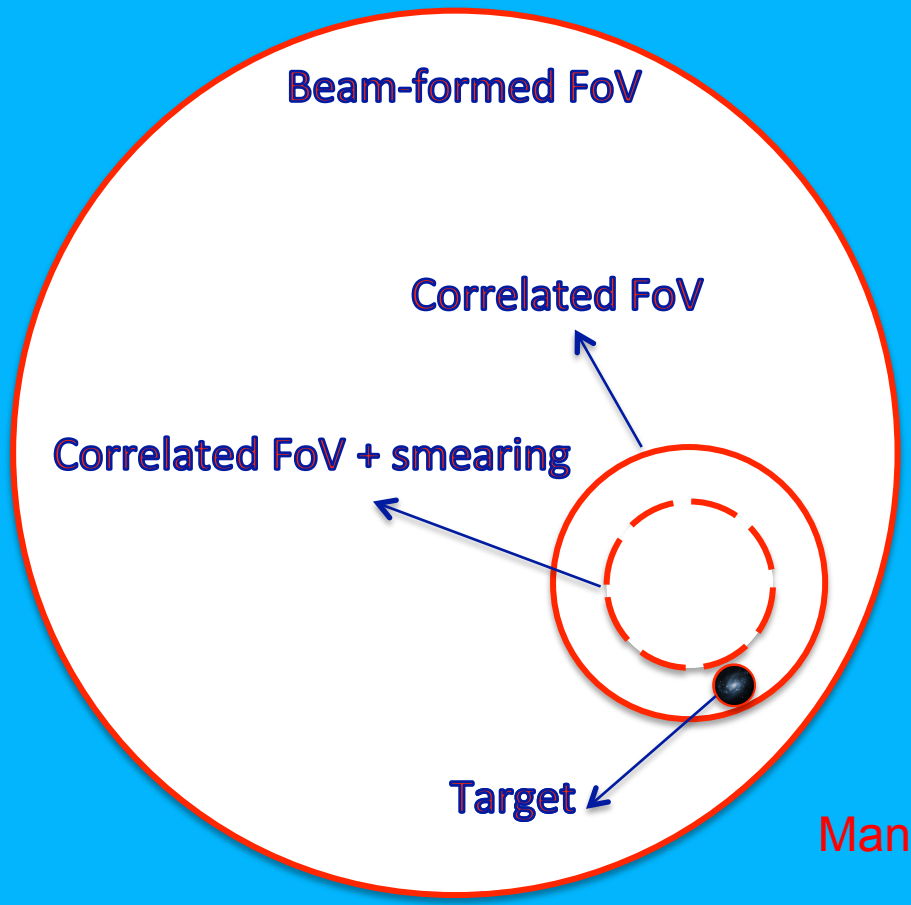
- computationally expensive, need data ➔ not practical

Approximated reconstruction:

- only metadata needed ➔ enough for data selection

The ADP & new tooling

Characterization tools needed when dealing with level 0/1 data products from wide-field interferometers ➔ accounting for beam-shapes of different modes for searches and selection but also for processing / analysis



Observed / processed data characterization in terms of:

- uv-coverage
- beam pattern (FoV & sensitivity)

Accurate reconstruction:

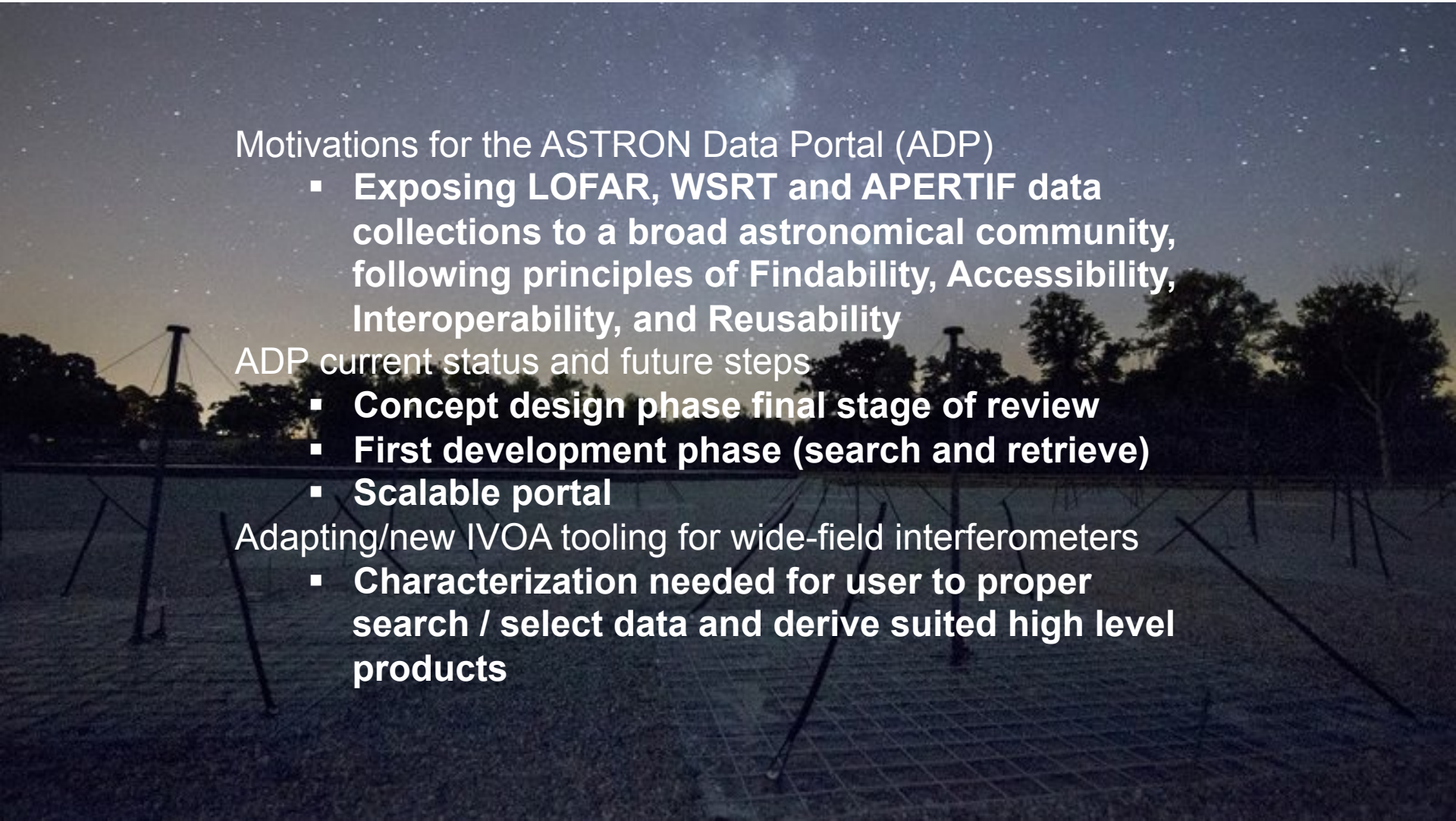
- computationally expensive, need data ➔ not practical

Approximated reconstruction:

- only metadata needed ➔ enough for data selection

ObsCore modeling . .
Managing VO standards regions for visualization . .

Summary



Motivations for the ASTRON Data Portal (ADP)

- **Exposing LOFAR, WSRT and APERTIF data collections to a broad astronomical community, following principles of Findability, Accessibility, Interoperability, and Reusability**

ADP current status and future steps

- **Concept design phase final stage of review**
- **First development phase (search and retrieve)**
- **Scalable portal**

Adapting/new IVOA tooling for wide-field interferometers

- **Characterization needed for user to proper search / select data and derive suited high level products**