Work Package 5: Access and Knowledge Creation

Prof. Steven Tingay
INAF
On behalf of WP5 team

Access and Knowledge Creation

This work package will focus on the interface between a distributed European SKA Data Centre (ESDC) and a distributed body of end users: **this is all about the users and their SKA experience.**

The basic question we want to answer is: What is the right “user interaction model” for such a centre?

- What do scientists want to do with SKA data?
- Where do they want to do it?
- How do they want to do it?

An ESDC should…

- “allow users to dive in and discover new things”;
- “take processing to the data”;
- “be a platform for the community to experiment”;
- “allow users to login to some portal and provide a transparent user experience”.

User interaction with the SKA

Users will interact with the Observatory to propose and design observational programs.

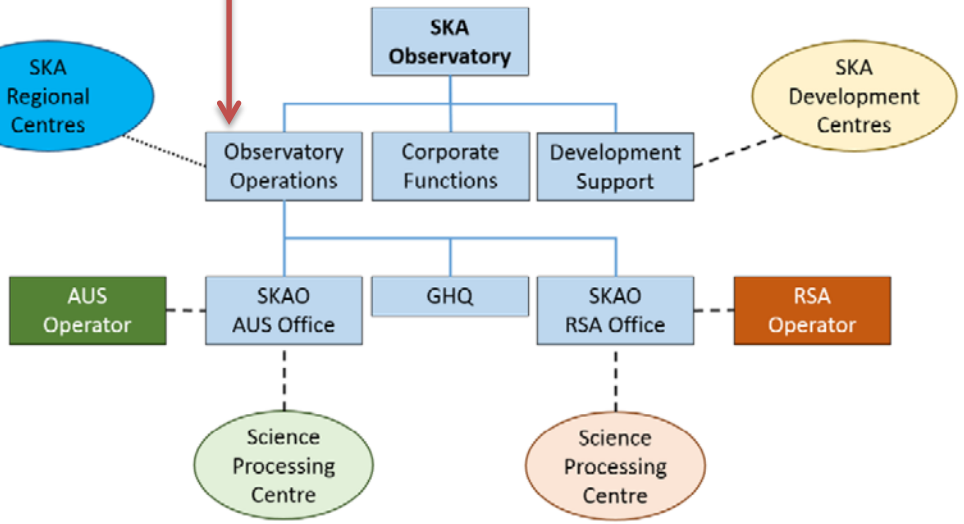


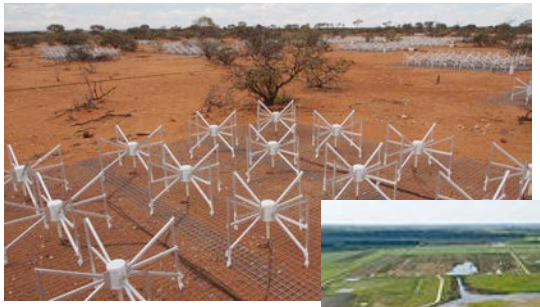
Figure 1: Organisational structure of the SKA Observatory. (Gary Davis, SKAO)

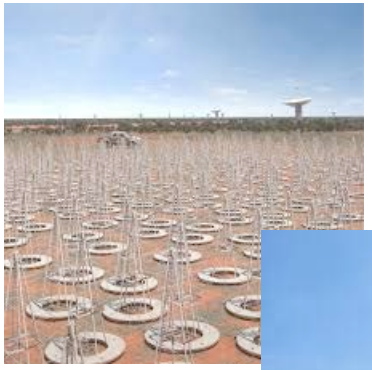
Users will interact with one or more "SKA Regional Centres" to work with SKA data (in general).

A few (not all) fundamental considerations for the ESDC

- What will be the range of science missions supported (presumably we should support SKA-low and SKA-mid, imaging, non-imaging etc) in an ESDC?
- What will be the size and nature of the user community?
- What volumes/types of data will have to be handled?
- What level of compute processing/storage required?
- What will be the division of data processing between the “Tier 0” centres and the ESDC?
- How will SKA Observatory access policies flow into the ESDC? Once users have invested IP in their data in an ESDC, are there any additional policy considerations?
- What is the physical nature of the facility (maturity, location, distributed or single site)?

LOFAR/MWA already >10 PB archives








MeerKAT/ASKAP coming online

WP5: Breakdown and people

- WP5.1: Survey of existing user interaction models for large-scale radio astronomy facilities and integration of WP5 outputs into consolidated ESDC design study (Steven Tingay, INAF-IRA):
  - Undertake a survey of existing relevant user interaction models (both from facility and user points of view);
  - SKA pathfinders/precursors with distributed users and >10 PB data storages (LOFAR/MWA MeerKAT/ASKAP);
  - Existing European distributed user support models (ALMA);
  - Traditional facilities (ATCA/JVLA);
  - What do we want to retain from the traditional facilities?
  - Can we scale the ALMA European Regional Centre model to the SKA/ESDC?
  - If not, where are the gaps and how can we address them?


European ARC nodes
ALMA
EUROPEAN ARC
ALMA Regional Centre
Onsala
Manchester
Leiden
Bonn–Cologne
Ondrejov
ESO Garching
Grenoble
Bologna
Lisbon

WP5: Work breakdown and people

- WP5.2: Recommendations for the design of user interfaces for data discovery, access, and retrieval (Cristina Knapic, INAF-OATs):
  - How will users identify the data they need?
  - How will users access the data?
  - How will users retrieve data? From where, to where?
  - Make concrete recommendations for final report;
- WP5.3: Recommendations for the design of user interfaces for data processing, reprocessing, analysis, and visualisation (Alessandro Costa, INAF,-OACt):
  - How will users access, interact with, and monitor data processing/reprocessing/analysis pipelines?
  - Data processing may need to be distributed across different physical facilities;
  - Data processing may need to be hierarchical and iterative. How will users manage this complexity?
  - Make concrete recommendations for final report;
- WP5.4: Integration with VO interoperability framework (Riccardo Smareglia, INAF-OATs):
  - Ground truth the findings of WP5.1, WP5.2, and WP5.3 as far as possible within one possible solution set;
  - Not intended to be pre-emptive of the results of WP5.2 and WP5.3.

WP5: Work breakdown and people

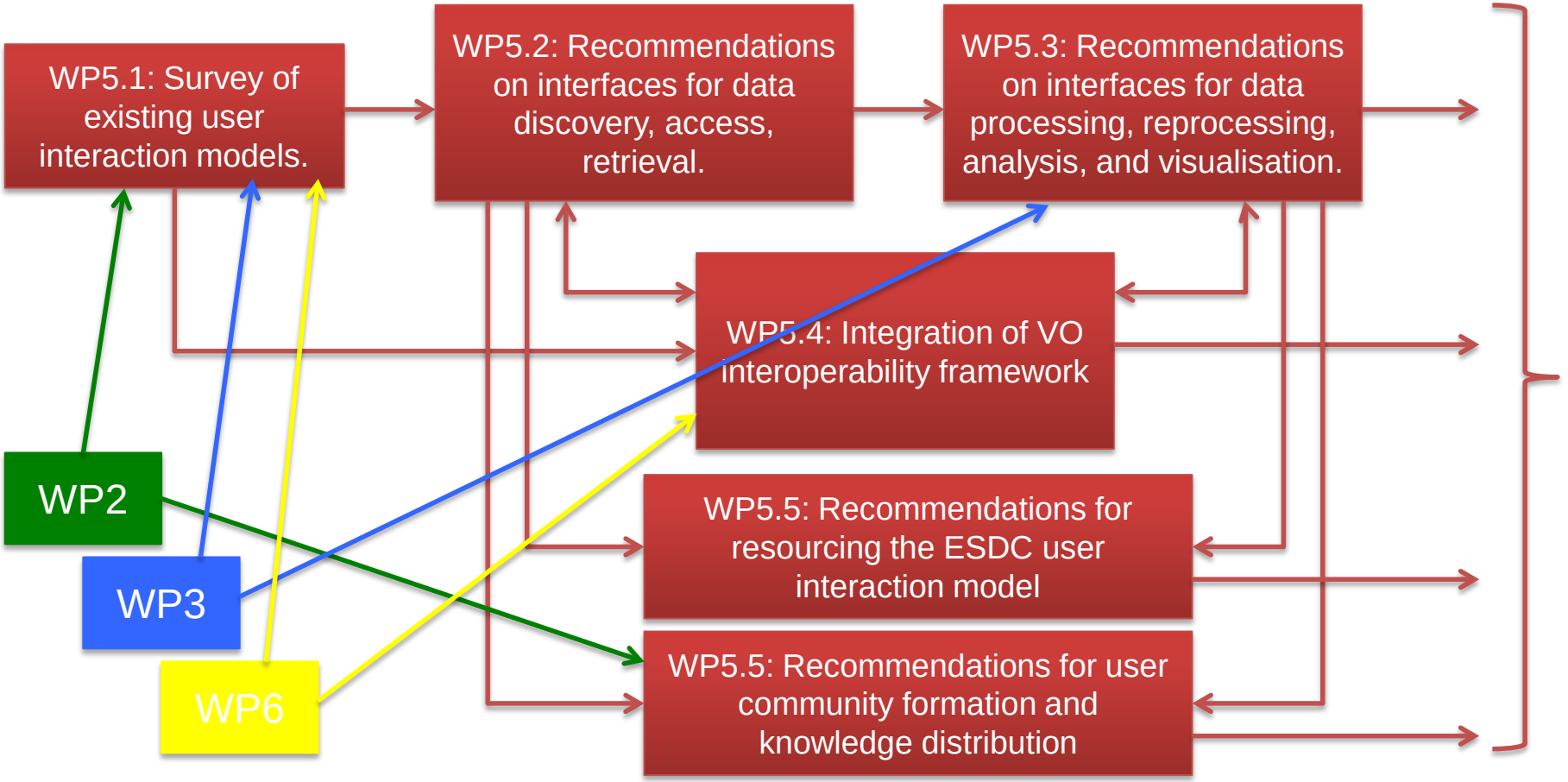
- WP5.5: Recommendations for the resourcing of an ESDC user interaction model (Jan Brand, INAF,-IRA):
  - Based on the conclusions of our work, what resources (human and financial) are likely to be needed to support the ESDC user interaction model?
- WP5.6: Recommendations for a plan of user community formation and knowledge distribution (Marcella Massardi, INAF-IRA):
  - How can we grow the user community for the SKA in Europe?
  - Need to support established radio astronomy users (experts);
  - Growth will come from non-expert users/communities/nations;

WP5: Work breakdown and people

Funded effort – 63 person-months (we will be calling on many others though!!):

- INAF: Steven Tingay, Ugo Becciani, Alessandro Costa, Jan Brand, Marcella Massardi, Cristina Knapic, Riccardo Smareglia, and 3 x new hires (overlap other WPs);
- ASTRON: Yan Grange;
- UCAM: Part of new hire (overlap WP3);
- UMAN: Part of new hire (overlap WP3);
- CNRS: Mark Allen, Françoise Genova, and Bernd Vollmer;
- CSIC: Lourdes Verdes-Montenegro, Susana Sanchez, and Julian Garrido;

Intended workflow



To implement the workflow

- A number of contact points between milestones/deliverables between WP3 and WP5;
- Coordination of a single user/facility survey (WP5.1) that involves (at least) WP2, WP3, WP5, and WP6;
- Effort shared between WP3 and WP5 in the form of new hires across lead institutions;
- All WP5 task leaders in the same organisation.

Summary deliverables

- Survey report and gap analysis: M18
- Design recommendations #1: M24
  (data discovery, access, and retrieval)
- Design recommendations #2: M28
  (data processing, reprocessing, analysis, and visualisation)
- Applicability of VO framework: M28
- User interaction model resourcing: M28
- Growing the ESDC community: M28
- Integration of WP5 outputs into overall design study: M34

Conclusions

- WP5 is an ambitious but important program of work within the AENEAS project;
- High level of dependency across the whole AENEAS project, including WP5 – will need to be carefully managed:
  - Seek to remain flexible.
- High level of dependency on the direction and timescales of wider SKA project – these dependencies will become clear over the course of the AENEAS project:
  - Seek to remain flexible.