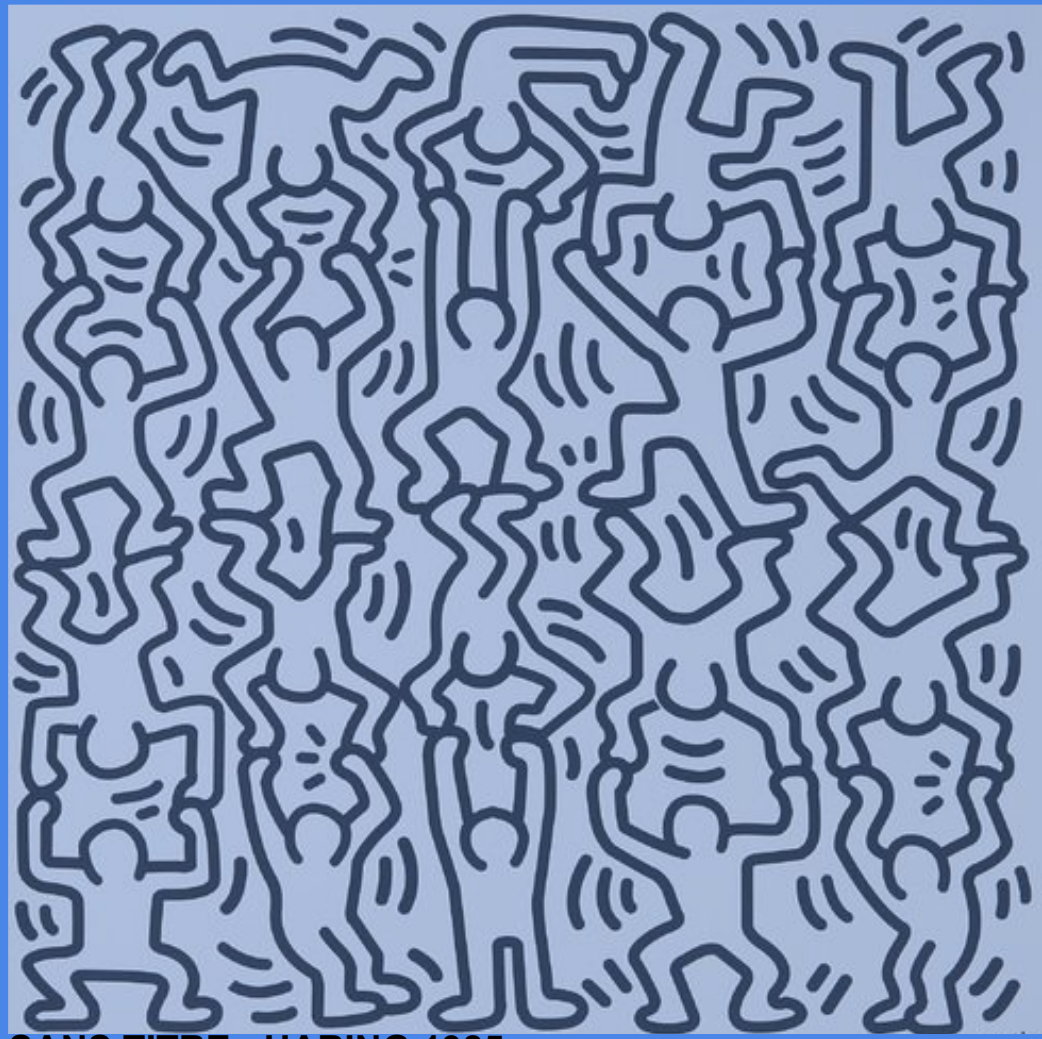


SANS TITRE , HARING 1985

WP5 - Access and Knowledge Creation

Marcella Massardi

(INAF- IRA / Italian node of the European ARC)

Advanced European Network of E-infrastructures for Astronomy with the SKA

While developing the design of the ESDC, the AENEAS team is ***analyzing existing products, technologies, services, best practices and standards and gather the user requirements.***
This will enhance the possibilities for the European scientists to exploit the SKA data, maximizing the impact of the telescope on the scientific production, together with fuelling innovation in networking, computing and storage technologies.

WP5 is focused on the ***interface between a distributed ESDC and a distributed body of end users*** *whose goal is the exploitation of SKA data for knowledge creation.*
WP5 is therefore studying the design of "user interaction models" that could be implemented for the ESDC.

Task 5.1 Survey of existing user interaction models for large-scale radio astronomy facilities and **integration of WP5 outputs into consolidated ESDC design study** (responsible M. Massardi)

Task 5.2 Recommendations for the design of user interfaces for data discovery, access, and retrieval (responsible V. Galluzzi)

Task 5.3 Recommendations for the design of user interfaces for data processing, reprocessing, analysis, and visualization (responsible A. Costa)

Task 5.4 Integration with VO Interoperability Framework (responsible R. Smareglia)

Task 5.5 Recommendations for the resourcing of an ESDC user interaction model (responsible J. Brand)

Task 5.6 Recommendations for a plan of user community formation and knowledge distribution (responsible M. Massardi)

THANKS TO ALL THE WP5 MEMBERS !!!!

M0 M12 M18 M24 M34

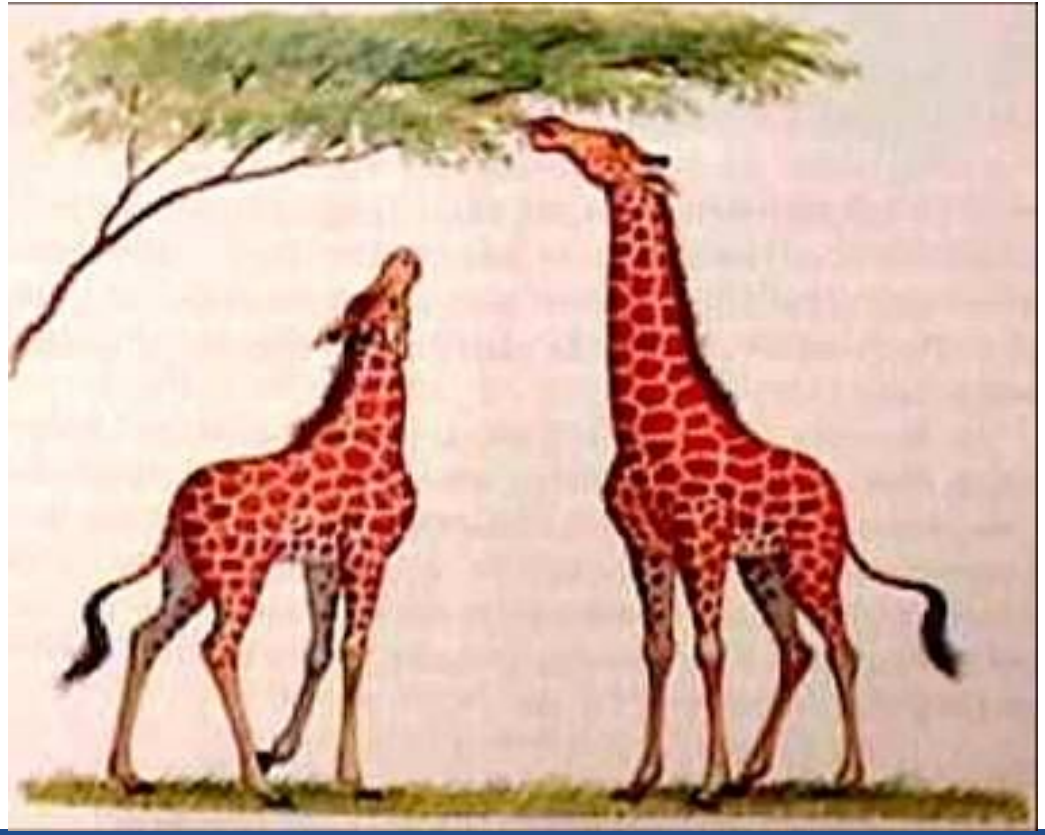
Learning from the past...

1) System needs
(goals towards user&tel)
2) User definition
(community/mentality)
3) Services provided
(duties/activities/policy/limitations)
4) Accessibility
(human interaction/interfaces)
5) Resources
(personnel/tools/infrastructures)

The ESRC is the interface,
access the archive,
offers the computation platform.
It must be trustworthy and resilient.

...suggesting new ideas

SKA IS AN EVOLUTIONARY JUMP!



Only a restricted portion of users have the conditions/knowledge/skills to access the resources.

This limits the telescope outcome.

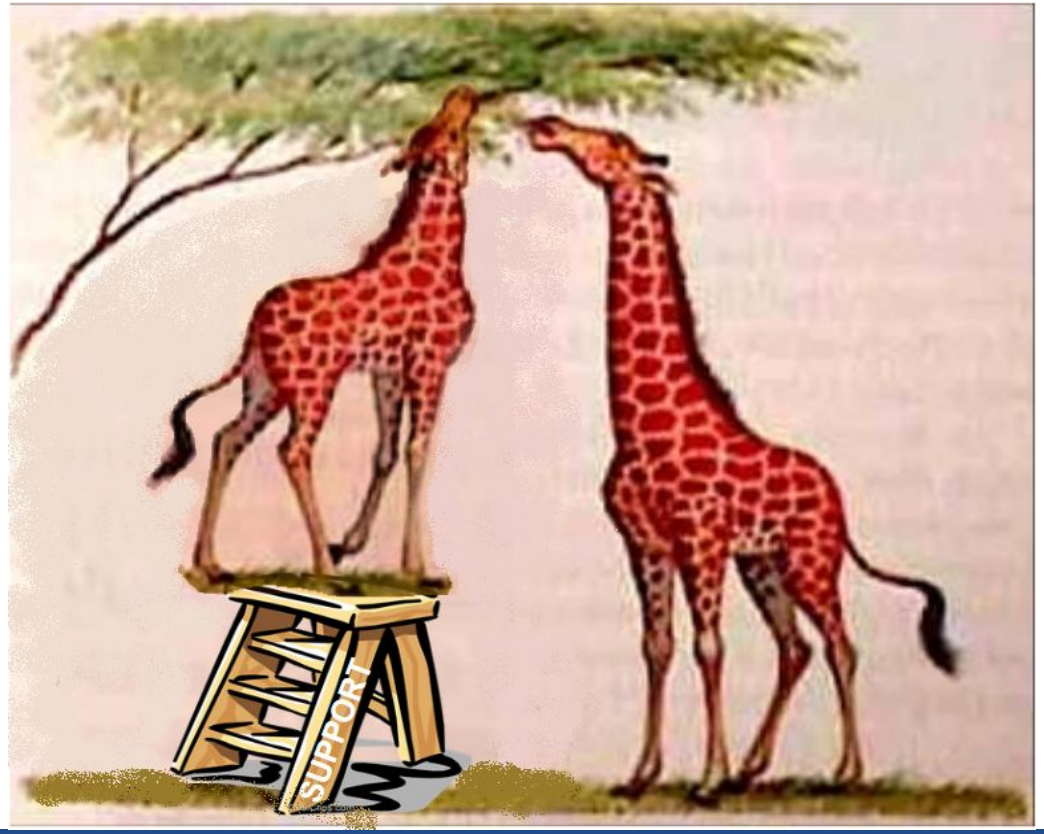
Community growing is slow.

Knowledge sharing is limited to those that already have the skills.

THE ORIGIN OF SPECIES - DARWIN 1859

SKA IS AN EVOLUTIONARY JUMP!

Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016



SUPPORT and USER INTERFACES

allow the access to the resources to a LARGER COMMUNITY

PROS TO THE TELESCOPE:

larger user base

= larger production

= bigger success

= faster development

PROS TO THE USERS:

larger user base

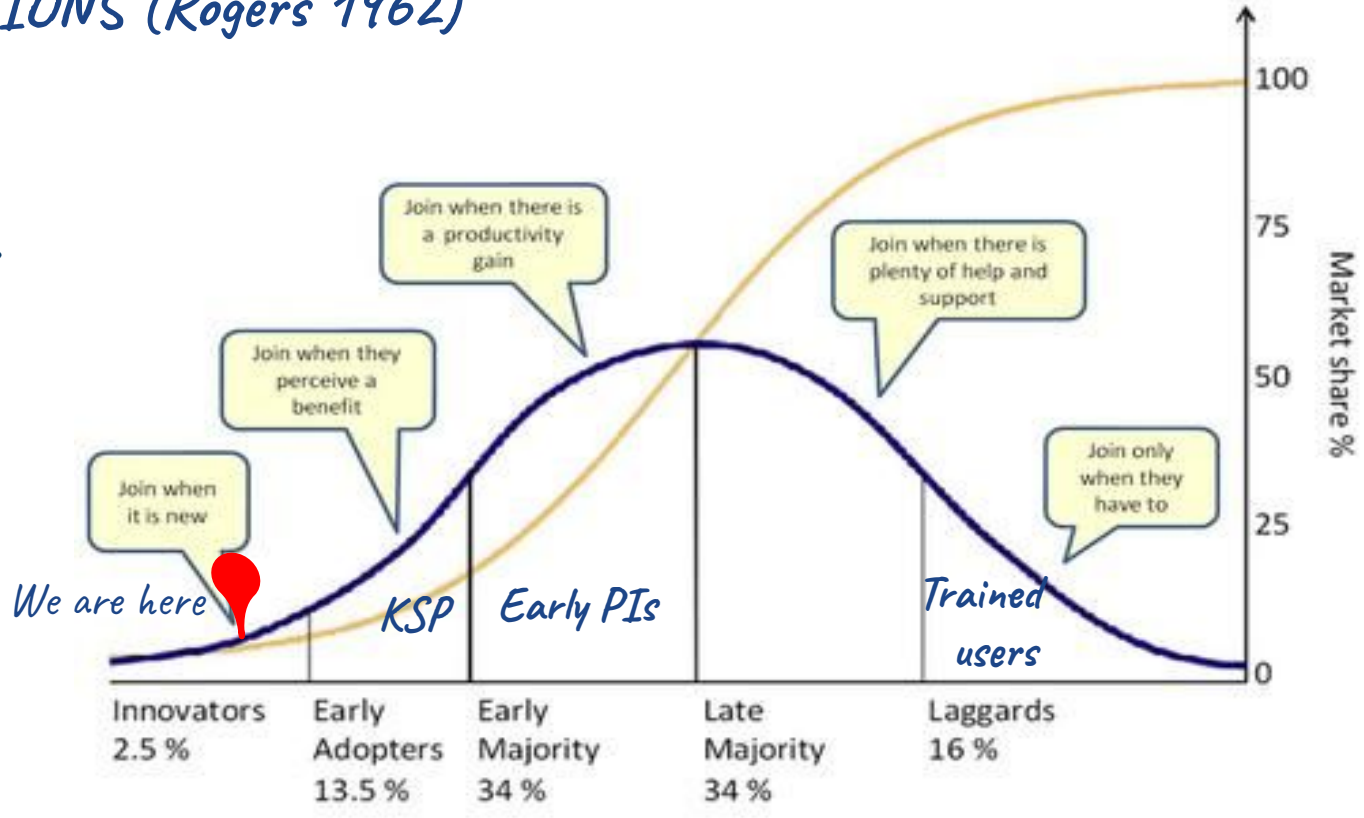
= better networking

= easy persistence of knowledge in time and space

= faster results

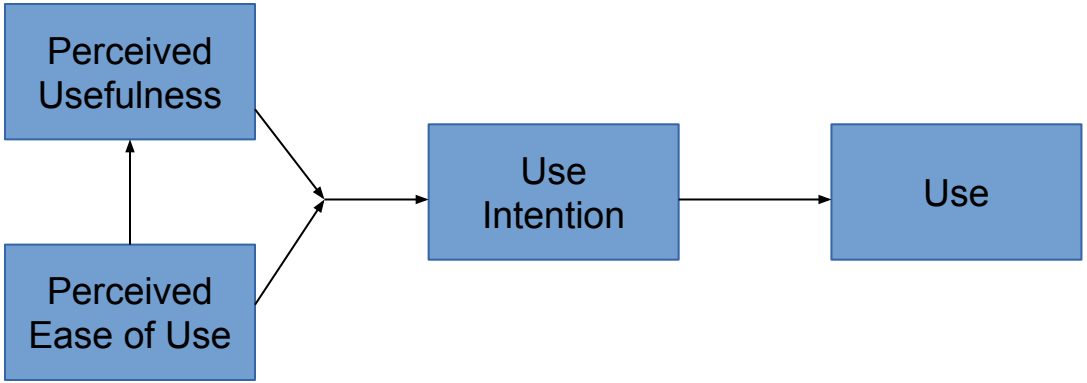
THEORY OF DIFFUSION OF INNOVATIONS (Rogers 1962)

We cannot modify the curve, but shorten the time of the different phases



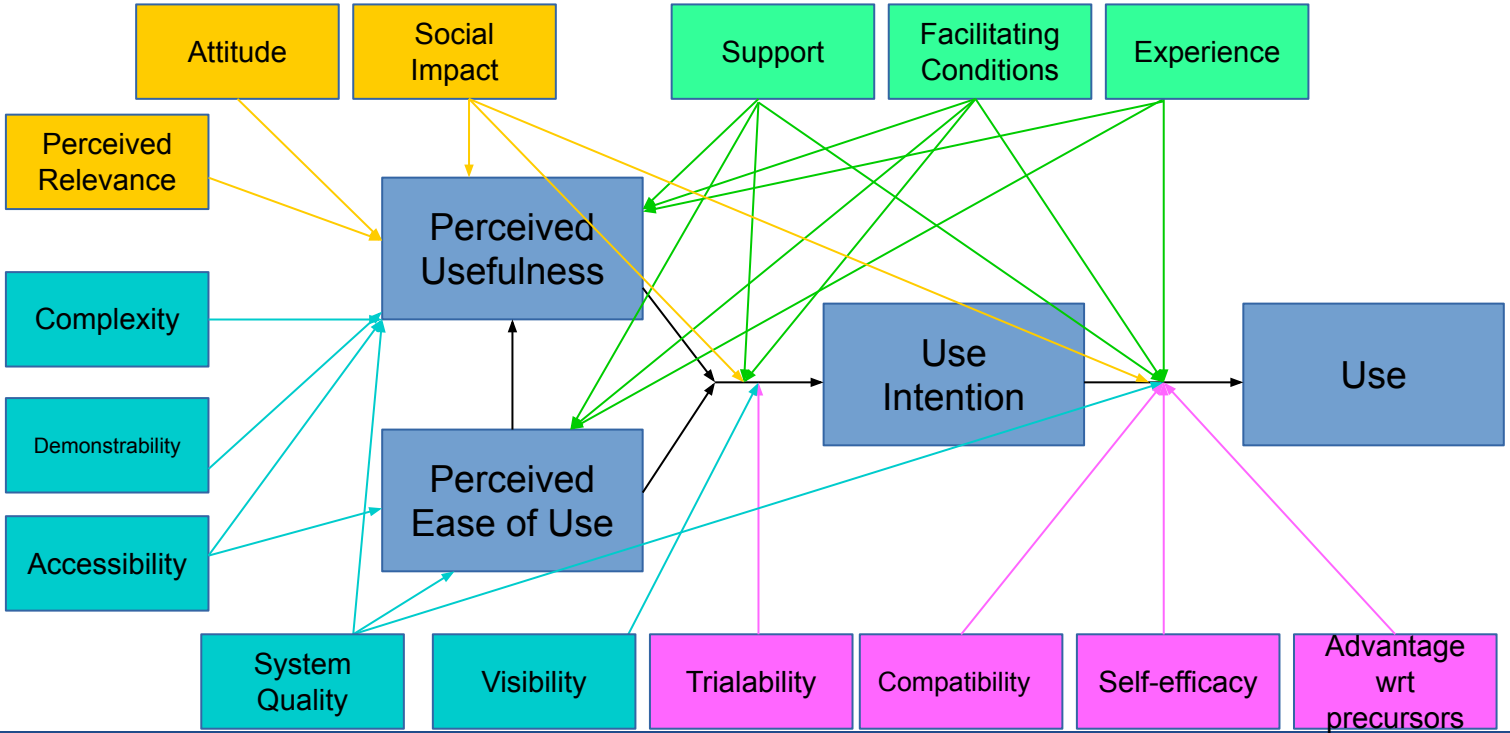
TECHNOLOGY ACCEPTANCE MODELS (Davis 1989 et al.)

Parameters that affect the acceptance of a new technology



TECHNOLOGY ACCEPTANCE MODELS (Davis 1989 et al.)

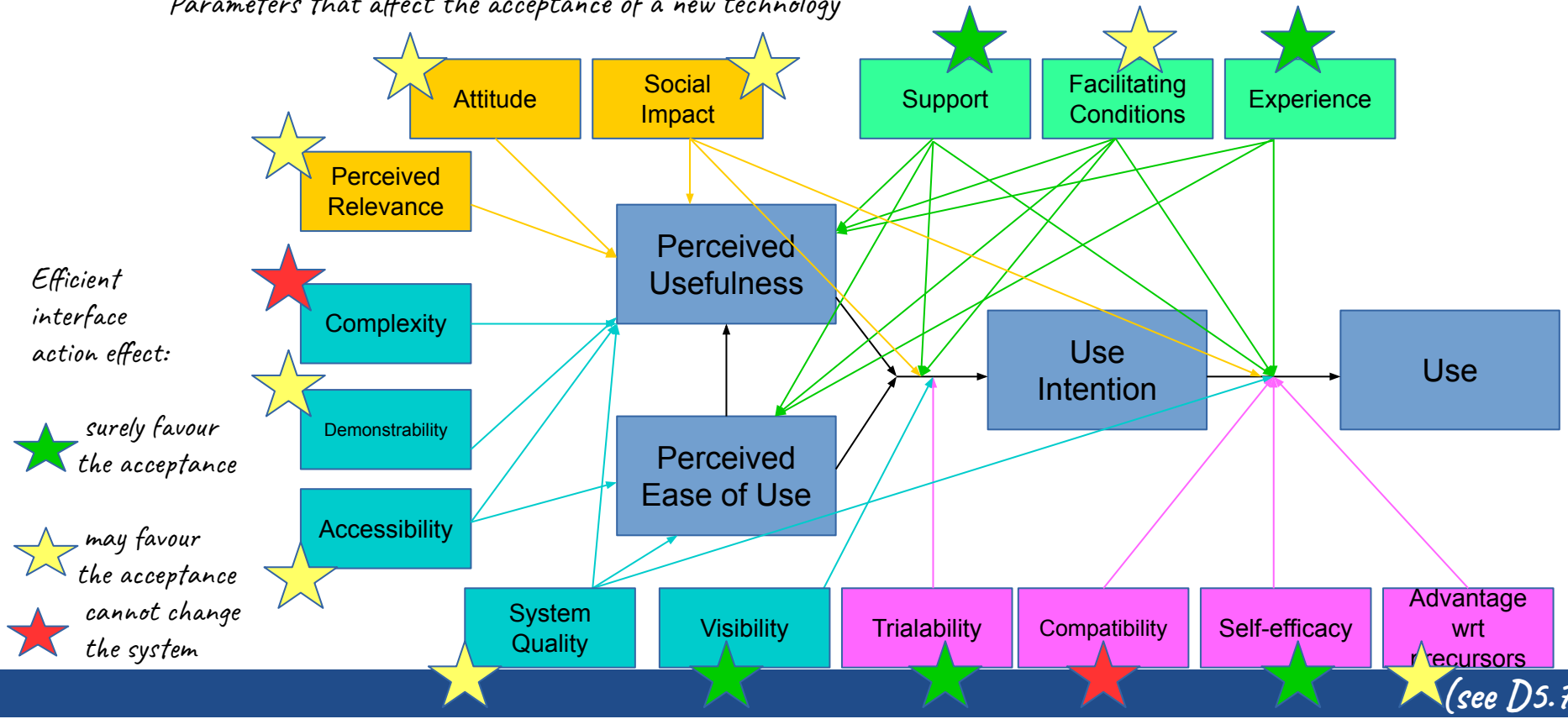
Parameters that affect the acceptance of a new technology



(see D5.7)

TECHNOLOGY ACCEPTANCE MODELS (Davis 1989 et al.)

Parameters that affect the acceptance of a new technology



(see D5.7)

INGREDIENTS FOR A TELESCOPE INTERACTION MODEL

Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016

- Information
- Proposal
- Project tracking
- Archive
  - Data
  - Metadata
- Data Processing Software
- Data Handling Hardware
- One point gateway access
  (with federated access)

VEGETABLES IN A BOWL - ARCIMBOLDO 1590

INGREDIENTS FOR A TELESCOPE INTERACTION MODEL

Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016

- Information
- Proposal
- Project tracking
- Archive
  - Data
  - Metadata
- Data Processing Software
- Data Handling Hardware
- One point gateway access (with federated access)

Expert people with broad variety of capabilities and access to resources

VEGETABLES IN A BOWL - ARCIMBOLDO 1590

Interface = Sharing Knowledge WITH/AMONG Users

- ★ **Observers** are the proposers of new observations, single or coordinated in small groups or large collaborations. This category will evolve with the telescope development.

- ★ **Archive Users/Data Scientists** exploit the archive product content for their own purposes. This class might partially overlap with the previous, and is potentially the largest, depending on the accessibility of the archive.

- ★ **Software Engineers** includes developers of software and tools. Might belong also to the above classes, but are mostly interested in technical details.

The ALMA Regional Centre Experience

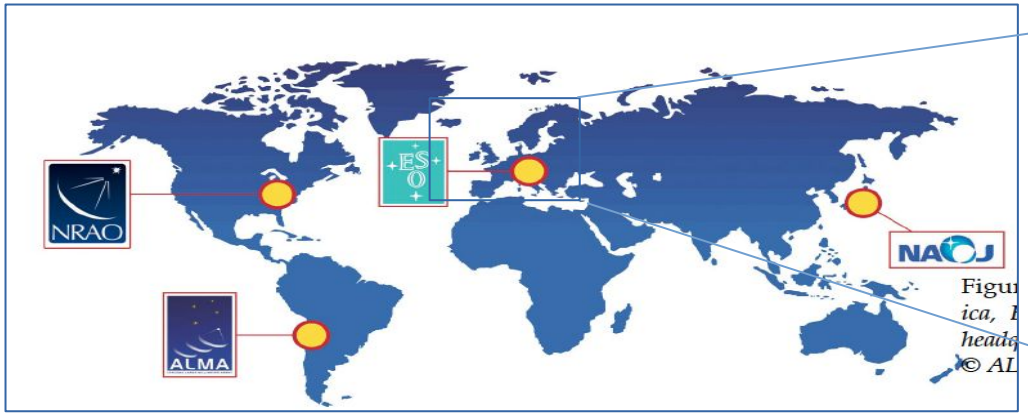
The Atacama Large Telescope Array (ALMA) is:

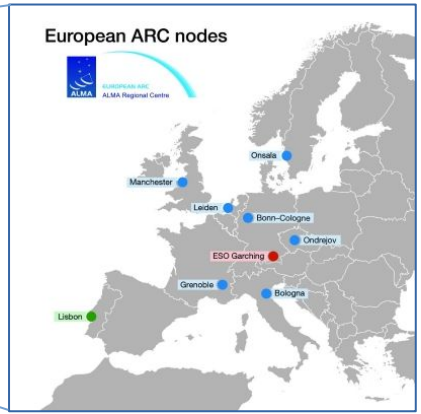
- a telescope matching the capabilities of 2 arrays, single dishes and phased for VLBI operating at ~5000m on the Chilean Andes
- the first case of calibrated and raw data and images in a radio-astronomical archive
- the first example of global collaboration and support network for a ground based telescope

The ALMA Regional Centre Experience

The ALMA Regional Centre in its first 8yrs of activity has been:

- a new approach/interface towards a new telescope
- a reference point for users of any level of career and expertise
- a support for improving the observations and the archive usage while the telescope was evolving
- a distributor of knowledge, exploiting the capillarity of the network



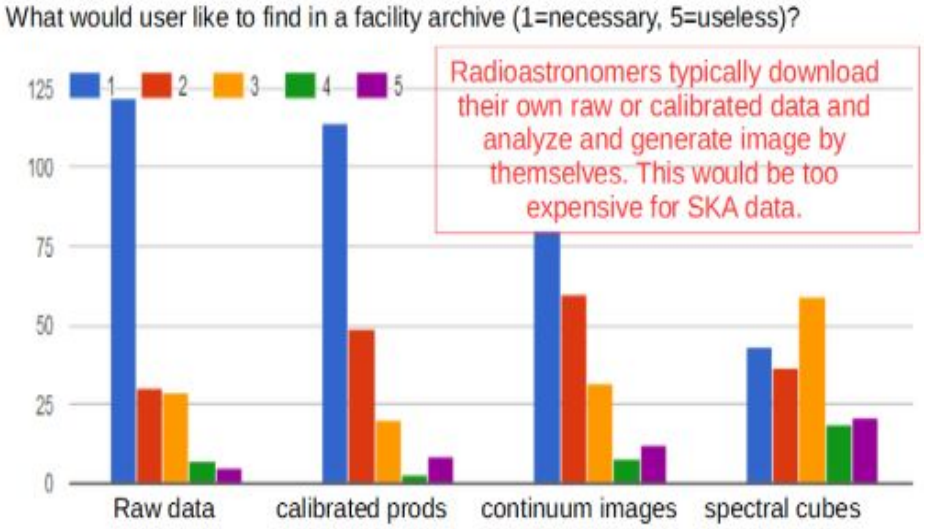


Gaps to mind towards the SKA: the reasons for an SRC

Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016

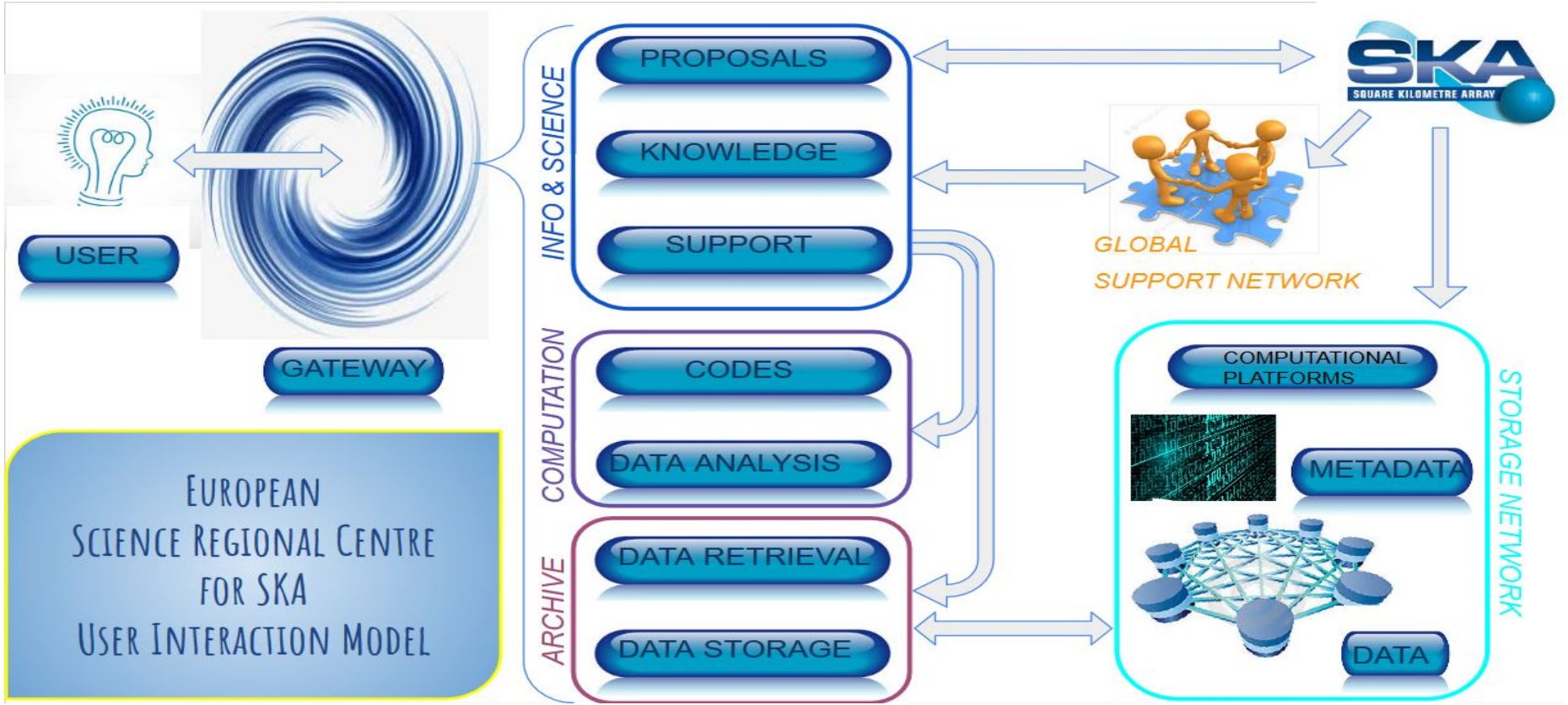
USER NEEDS AND HABITS
SYSTEM REQUIREMENTS

- DATA SIZE → STORAGE
  → TRANSFER
  → COMPUTATION
  → TIME
- VARIOUS DATA AND PRODUCT FORMAT
  (e.g. observatory vs advanced products)



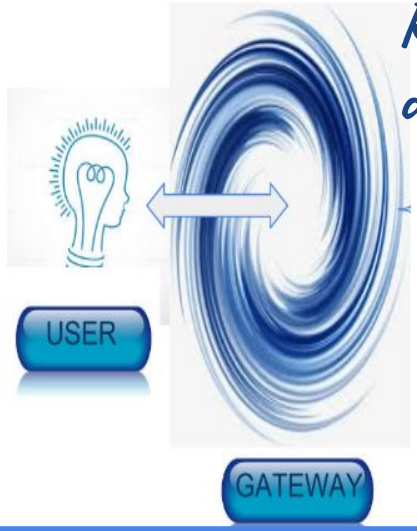
A Regional Centre IS the interface, accesses the archive,
offers the computational platforms, grows the community.
Must be trustworthy, up-to-date, resilient.
An efficient Regional Centre IS A RESOURCE FOR THE USERS

(see D5.1 and D5.2)

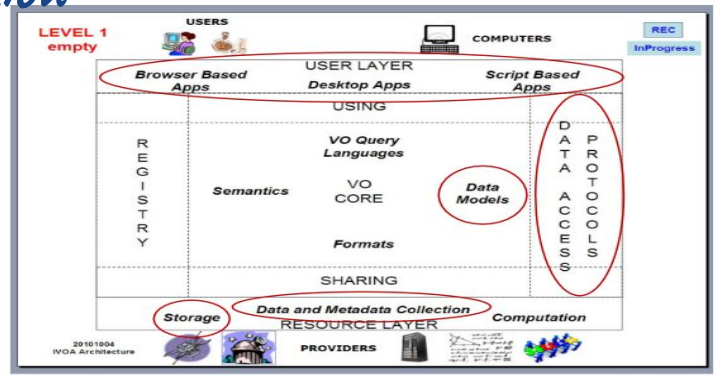

USER
GATEWAY
INFO & SCIENCE
PROPOSALS
KNOWLEDGE
SUPPORT
COMPUTATION
CODES
DATA ANALYSIS
ARCHIVE
DATA RETRIEVAL
DATA STORAGE
SKA
SQUARE KILOMETRE ARRAY
GLOBAL
SUPPORT NETWORK
COMPUTATIONAL PLATFORMS
METADATA
DATA
STORAGE NETWORK
EUROPEAN
SCIENCE REGIONAL CENTRE
FOR SKA
USER INTERACTION MODEL

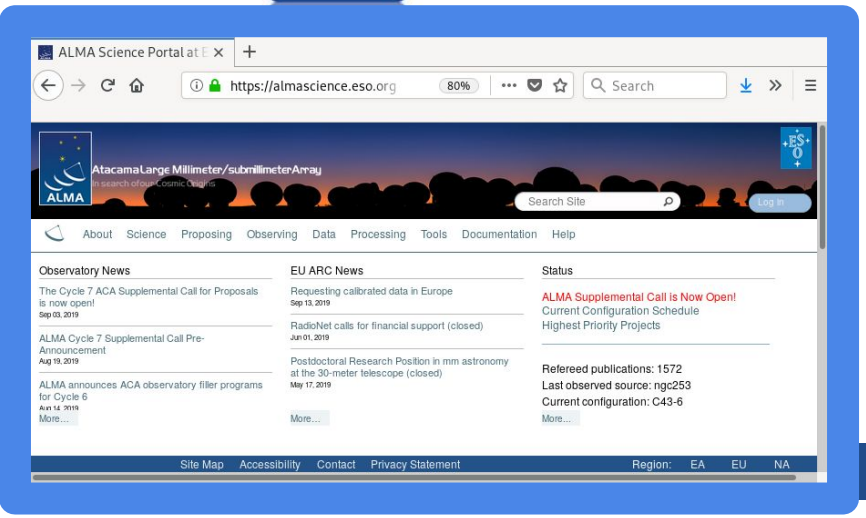
Recommendations on data analysis and visualization

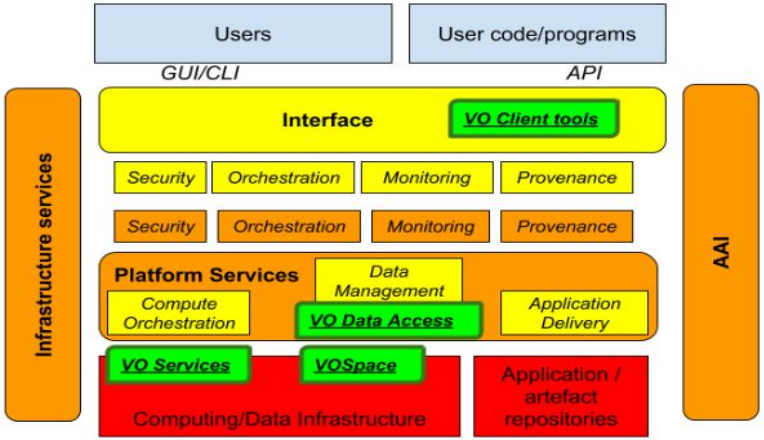
Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016



The GATEWAY is a one-stop shop for the user with easy access to all the resources







(see the talk by Costa and D5.3)

Recommendations on data archive and retrieval

The **ARCHIVE INTERFACE** should

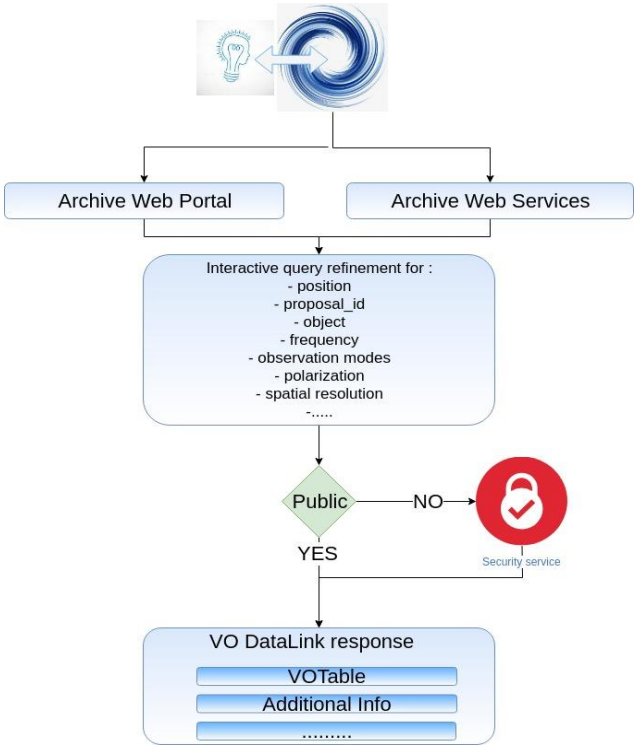
- be single, uniform, intuitive and multi-mask
- be able to handle a list of search keywords, and scriptable
- allow the extraction of portions of datasets
- VO compliant

METADATA should be

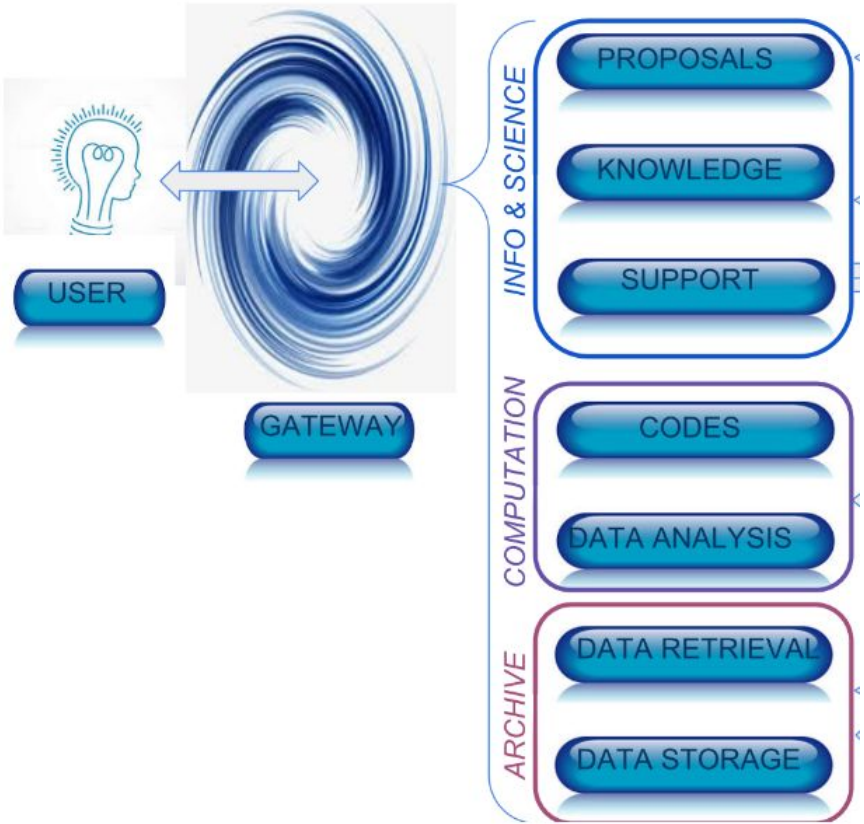
- properly defined across the whole product workflow
- well documented for any user to access them

A **GOOD DATA FORMAT** should:

- be explicitly versioned (open, well documented, maintained);
- be self-describing;
- be instrument-independent;
- support parallelization;
- support streaming features with multiple data representation;
- be fine for storage and processing and allow portioning;
- be standardized.



(see the talk by Galluzzi, D5.3, D5.5 and Archive survey and Metadata reports)



COMMON USER'S REQUESTS:

- Submitting proposals
- Accessing info about telescope
  Observations and
  Procedures
- Interacting with human support

- Direct access to archive (with limitations)
- Direct access to computational Platforms
  (with limitations)
- Request of advanced products
  - exploiting existing code
  - using user-provided code

Repositories for data & codes

Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016

Given the SKA size, platforms and archives will be continentally (and hopefully world-wide) distributed

SKA ARCHIVE NETWORK: Hosts the telescope products
World-wide distributed

ADVANCED PRODUCTS REPOSITORY:
Hosts the products generated by SRC with distributed codes
Content is quality assured

CODES REPOSITORY: Hosts the pipelines and the SW
Used/ingested by the SRC
Content is quality assured

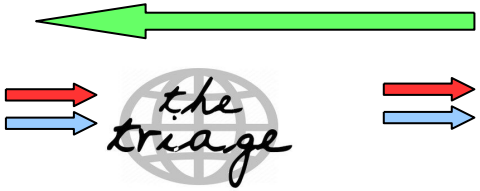
APELLES PAINTING CAMPASPE - VAN HAECHT 1630

(See talk by Garrido)

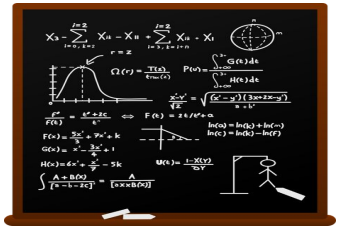
REQUEST FOR ADVANCED PRODUCTS

SRC resources should be assigned through dedicated calls

GENERAL REQUEST

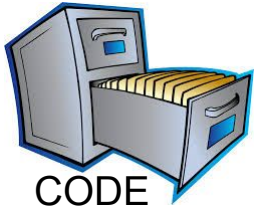


RESPONSIBLE



COMPUTATION

ADVANCED PRODUCT REPOSITORY



CODE REPOSITORY

ARCHIVE

CODES
QUESTION
RESPONSE
DATA

STAFFING

Archive manager: looks after content of archive; storage, accessibility of the data
Database administrator: takes care of software, hardware, user accounts
System administrator: takes care of computer cluster, storage system, security; access to computing platforms
Archive operator: handles data requests, extracts data from archive; contact with PI

Software analysts: to install and manage data reduction/analysis programs; maintaining web interface and Helpdesk.

Support astronomers: Assistance with extraction of scientific information. Competence in wide range of research interests required.

Administrative support

Minimum requirement: 11 FTE (to be filled by 15 people)

4 FTE in archive management; 2 software analysts; 4 astronomers; 1 administrative support
2 persons to share 1 FTE for support astronomer (to facilitate pursuing scientific interests and a career in research)

The timeline towards the SRC

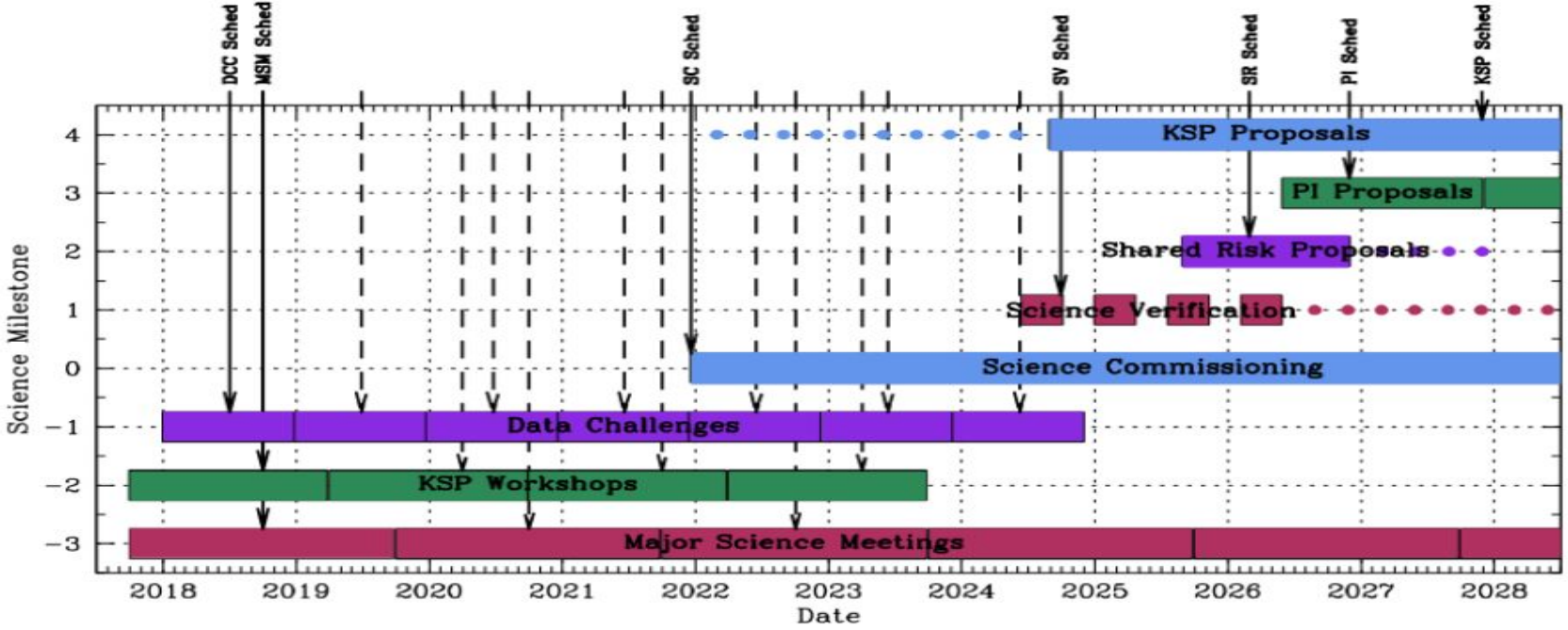
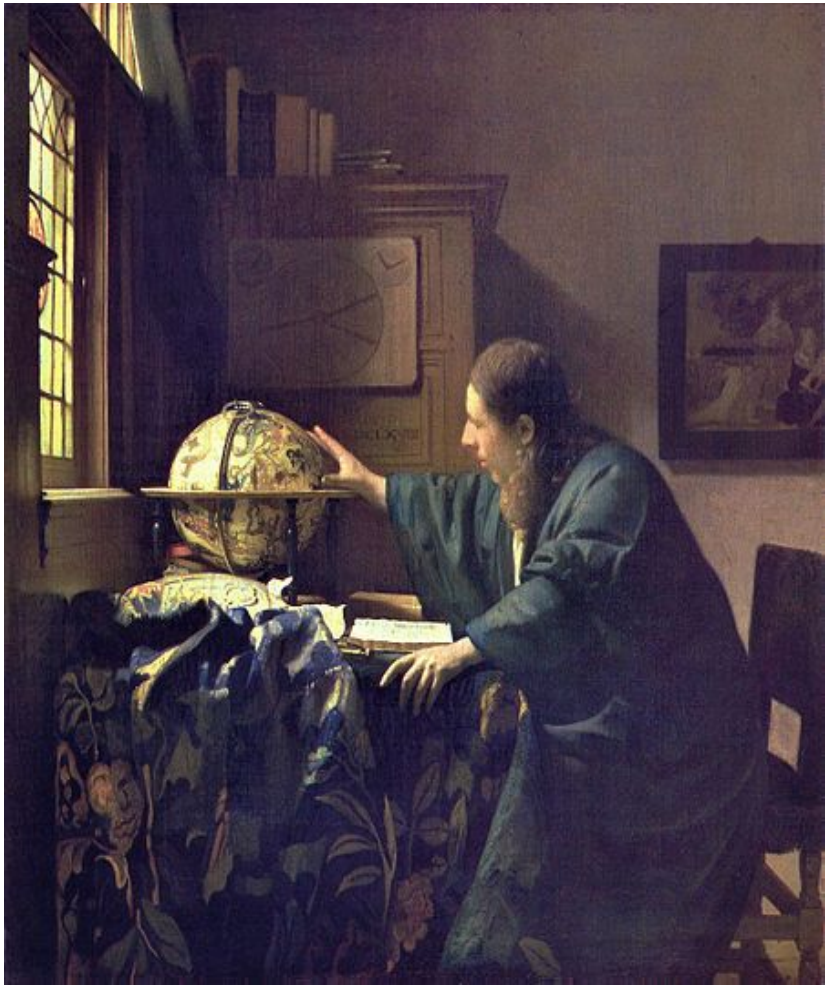


Fig. Timeline of Major SKA1 Science Milestones for an Assumed Availability Date of AA2 and AA4 (Science Planning Document (Rev 02, 15/11/17))



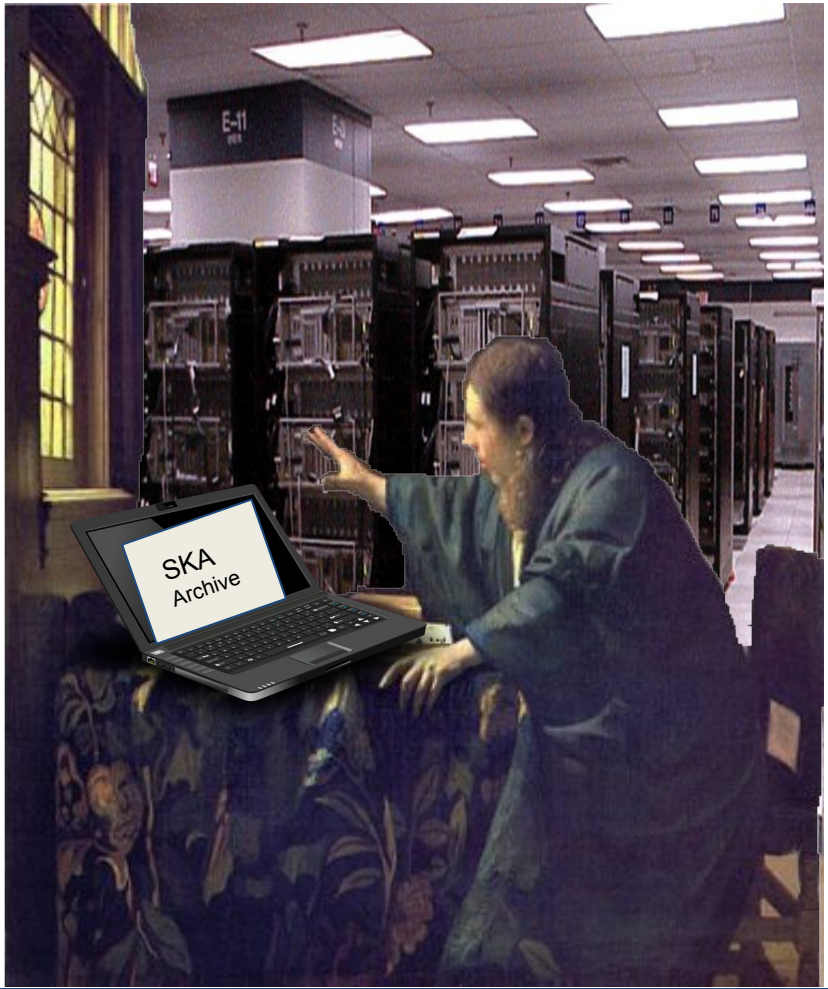
THE ASTRONOMER - VERMEER 1688

A NEW APPROACH IS NECESSARY:

- large collaborations (see the KSP)
- open access and knowledge sharing
- merging between data science and astronomy
  - common goals
  - metalanguage
  - broader perspectives

-> New professional profiles are necessary to staff the Regional Centre and to maximize the impact of new generation of telescopes

-> The community needs to be formed and informed taking advantage of the telescope development



A NEW APPROACH IS NECESSARY:

- large collaborations (see the KSP)
- open access and knowledge sharing
- merging between data science and astronomy
  - common goals
  - metalanguage
  - broader perspectives

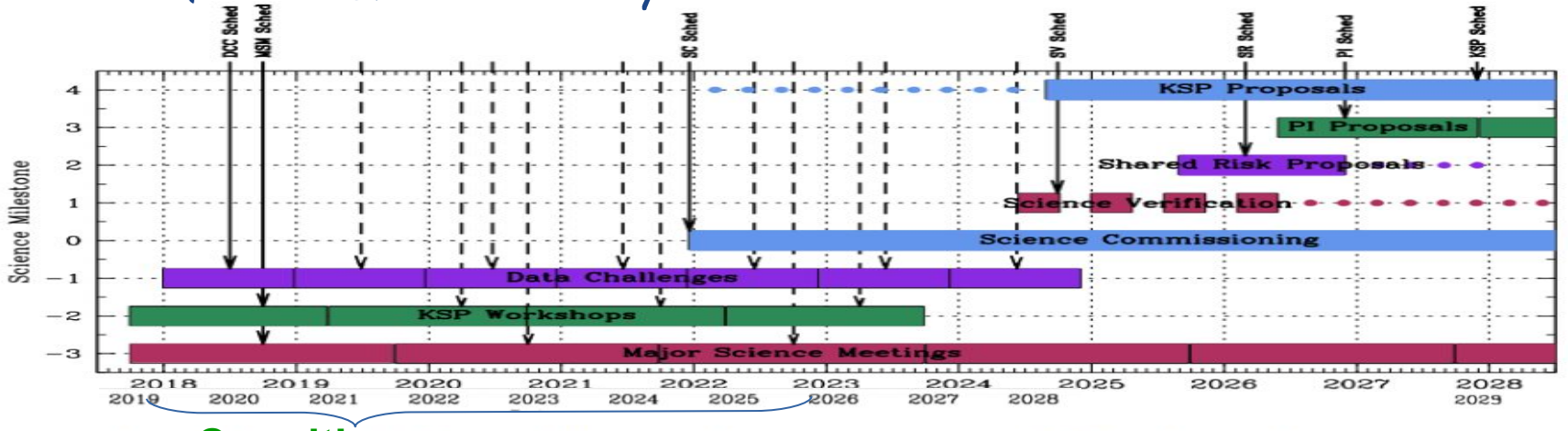
-> New professional profiles are necessary to staff the Regional Centre and to maximize the impact of new generation of telescopes

-> The community needs to be formed and informed taking advantage of the telescope development

(see Burkutean's talk)

A training plan for the SKA community

Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016



Training phases:

Cognitive

-acquiring the general overview and basic knowledge on the system capabilities (i.e. data structure and handling)

-building awareness on the gaps and issues

Practice

- dirty hands-on simulations, data from precursors and lately on SKA data
- test tools/interfaces and workflows

Active

- proposing as PIs
- mining the archive
- maintaining and distributing the knowledge

(see D5.7 and SKA Data Challenge Workshop- Bologna 30 sep-2 oct 2019)

The SKA Data Challenge Workshop

Bologna 30 sep - 2 oct

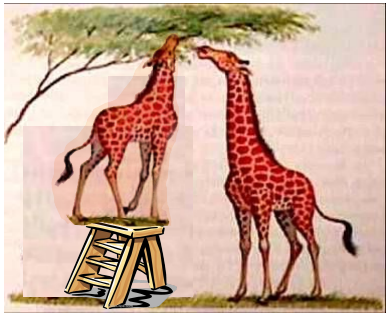
Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016

Organized in collaboration with the Italian SKA Board - INAF UTG II Radioastronomy

Fostering the Astronomical Community in Europe towards the SKA

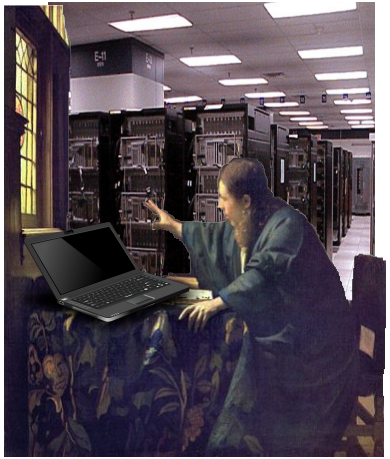
THE VILLAGE SCHOOL - THE SCHOOL EXAM - ANKER ~1860

Conclusions



A Regional Centre IS the interface, accesses the archive, offers the computational platforms, grows the community. Must be trustworthy, up-to-date, resilient.

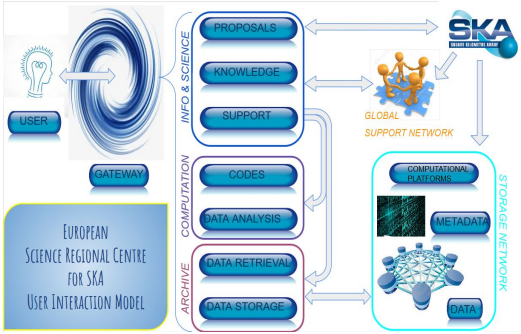
Support and efficient interfaces allow a broader community to access the resources offered by SKA



Recommendations for a user support interface:

1) **System needs** (goals towards user&tel)
2) **User definition** (community/mentality)
3) **Services provided** (duties/activities/policy/limitations)
4) **Accessibility** (human interaction/interfaces)
5) **Resources** (personnel/tools/infrastructures)

Time is a resource that we could exploit testing on SKA-like data, building pipelines and tools, forming a new user approach and professional figures



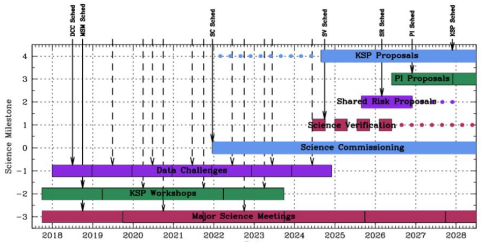


Fig. Timeline of Major SKA1 Science Milestones for an Assumed Availability Date of AA2 and AA4 (Science Planning Document (Rev 02, 15/11/17))