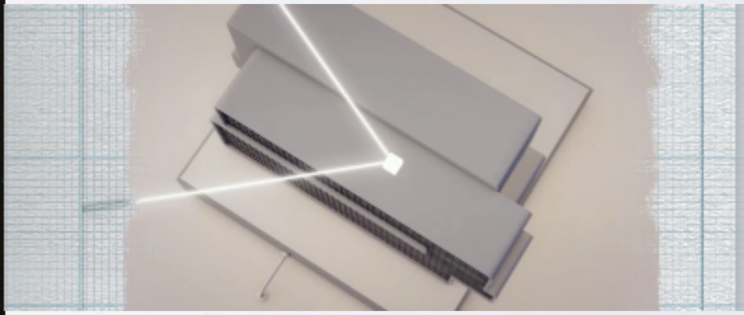
What are the HI community needs for the SKA Regional Centres (SRCs)?

Lourdes Verdes-Montenegro
Instituto de Astrofísica de Andalucía (CSIC)

Co-chair of the HISWG (together with Sarah Blyth, UCT)
Member AENEAS project & of the SRC Coordination Group

Blok et al 2015, Picture courtesy B. Koribalski



Diamond & Bolton, SC17

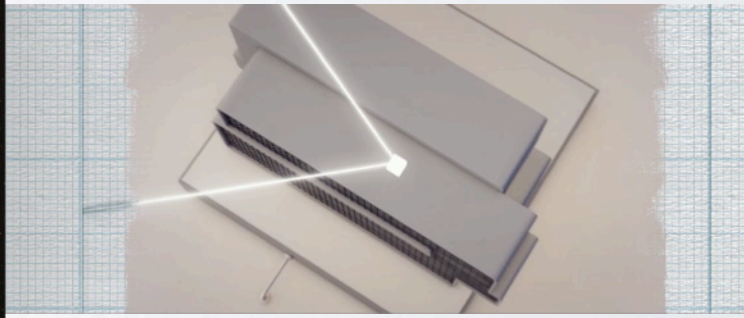
12th September 2018, The HI/Story of the Nearby Universe (Groningen)

What are the HI community needs for the SKA Regional Centres (SRCs)?

Lourdes Verdes-Montenegro
Instituto de Astrofísica de Andalucía (CSIC)

Co-chair of the HISWG (together with Sarah Blyth, UCT)
Member AENEAS project & of the SRC Coordination Group

Blok et al 2015, Picture courtesy B. Koribalski



Diamond & Bolton, SC17

12th September 2018, The HI/Story of the Nearby Universe (Groningen)

OUTLINE

- **SDP, SRCs, AENEAS, SWGs** and all these acronyms
- **Initial set of processing considerations, data products and associated volumes, formats, and tools:** **your input is needed!**
- As written in my abstract

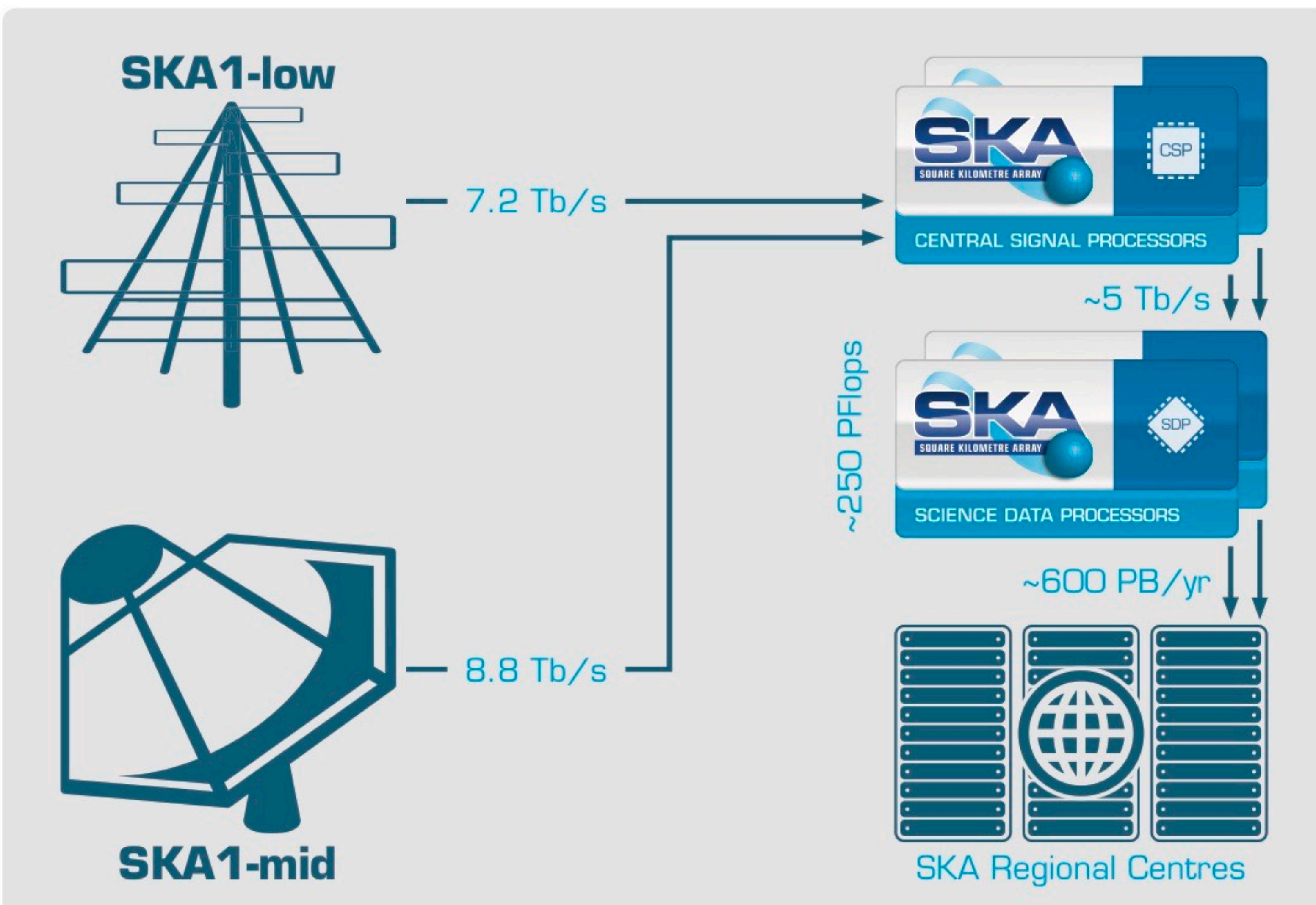
  "The HI/StoryNU workshop seems a right place to present this talk, and hopefully **motivate further discussions** and hence make sure that **the needs of the HI community are well integrated within the SRCs early in the process**."

SKA Regional Centres (SRCs)

The data journey

Basic data products will be produced and stored in Cape Town for SKA1-mid and Perth for SKA1-low. From there, they will be delivered to a global alliance of SKA Regional Centres for further processing and archiving and access by the user community.

New data models will need to be developed as current software packages for radio astronomy data reduction don't have the capabilities to handle the SKA's large bandwidths and Field of View datasets.



The SKA will be:

Distributed

One observatory operating two telescopes on three continents for a global scientific community.

Accessible

Common software and user interface. Pre-programmed algorithms. Training at the SKA Regional Centres.

Open

Open Access to non-proprietary data.

1010
0100
0110

Citizen-ready

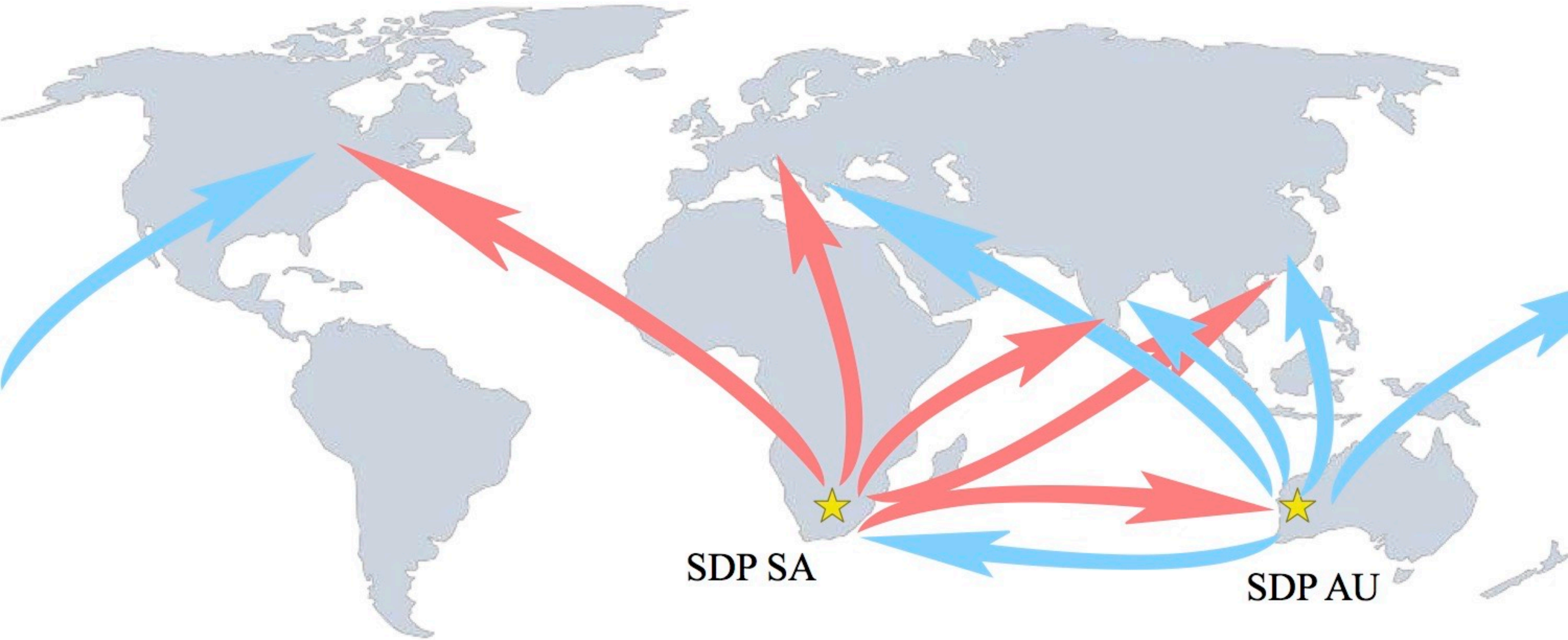
Access to SKA public data for citizen science projects.

@ska_telescope @SKA_telescope Square Kilometre Array YouTube The Square Kilometre Array

1.2 August 2018

SKA Fact sheets. August 2018. skatelescope.org

SKA1 Data flow

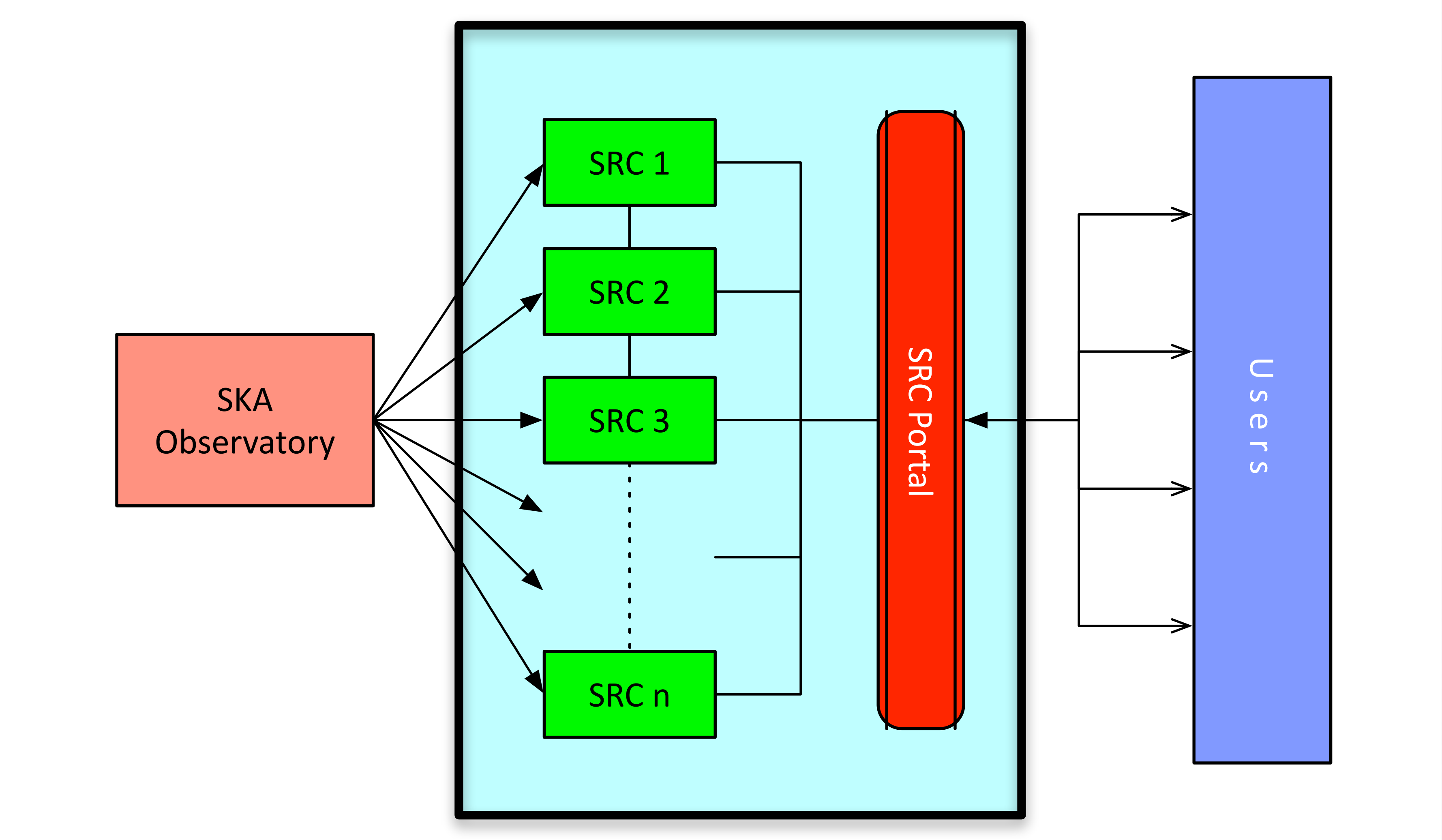


Observatory Data Products flow from the Science Data Processors in Perth and Cape Town to Science Regional Centres around the globe

J. Wagg. PHISCC 2018

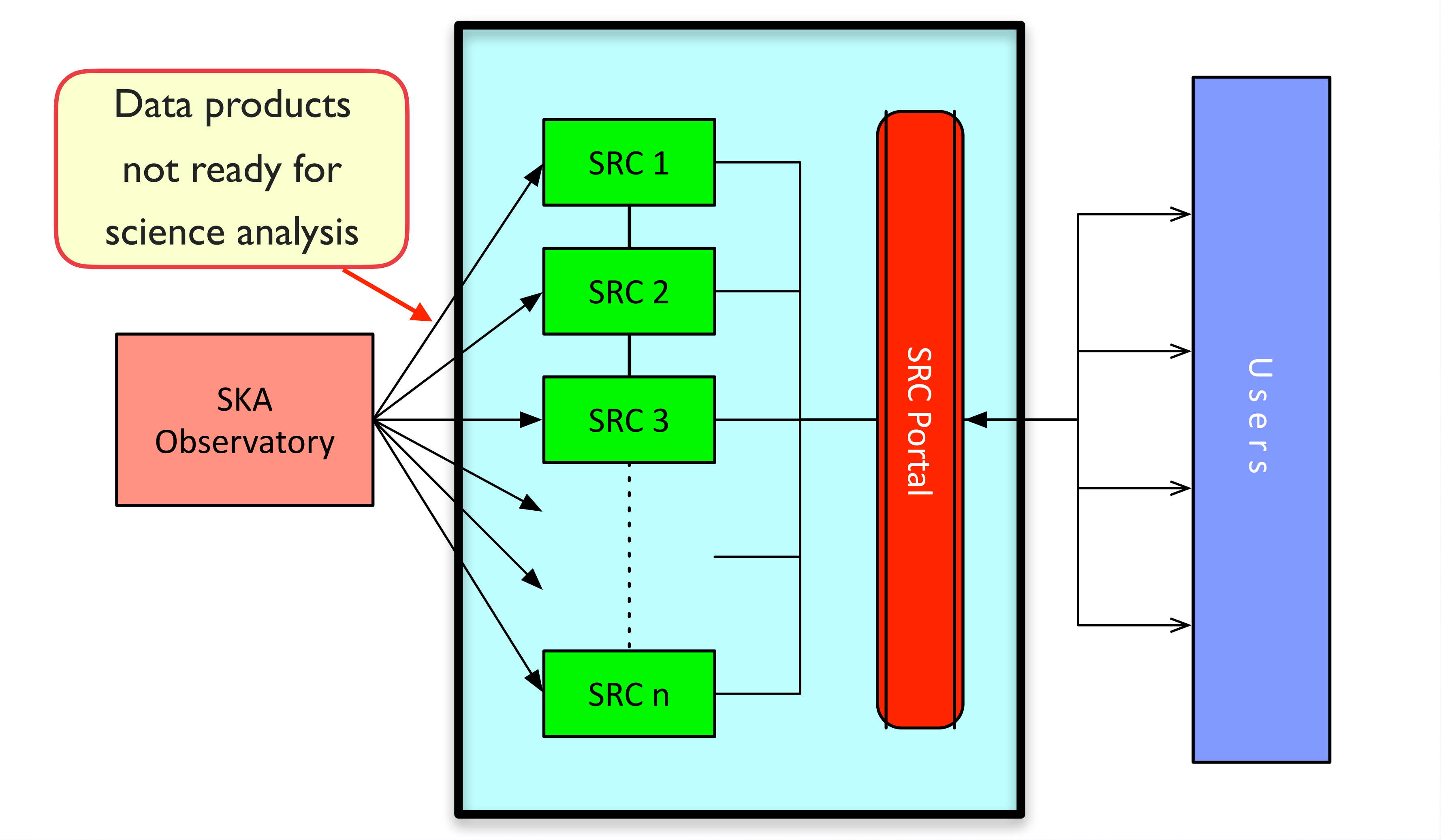
SKA Regional Centres (SRCs)

- SRCs will provide **access** to the SKA community to **data products** they are authorised to, as well as the **tools and processing power** to generate and analyse Advanced Data Products (ADPs)



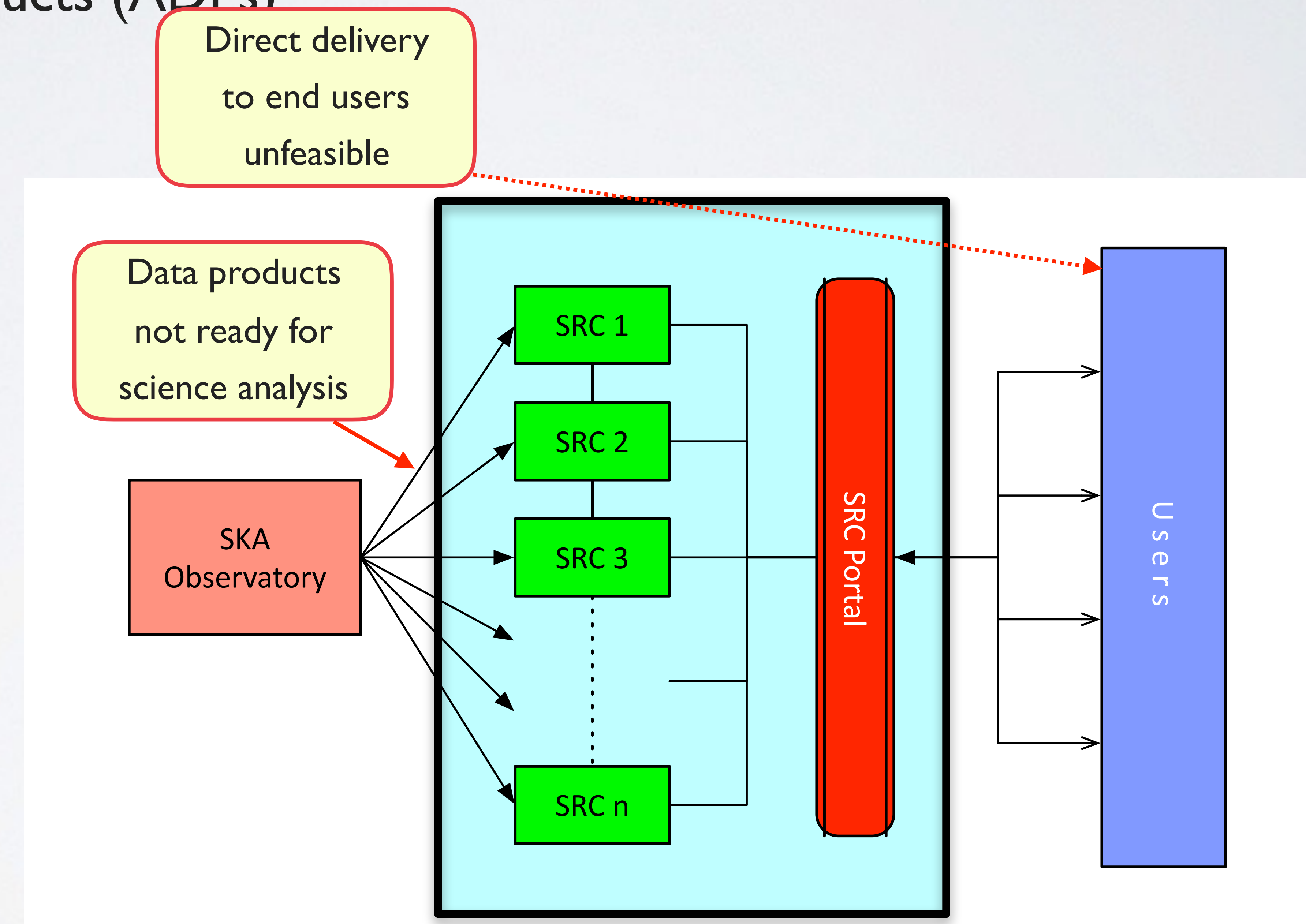
SKA Regional Centres (SRCs)

- SRCs will provide **access** to the SKA community to **data products** they are authorised to, as well as the **tools and processing power** to generate and analyse Advanced Data Products (ADPs)



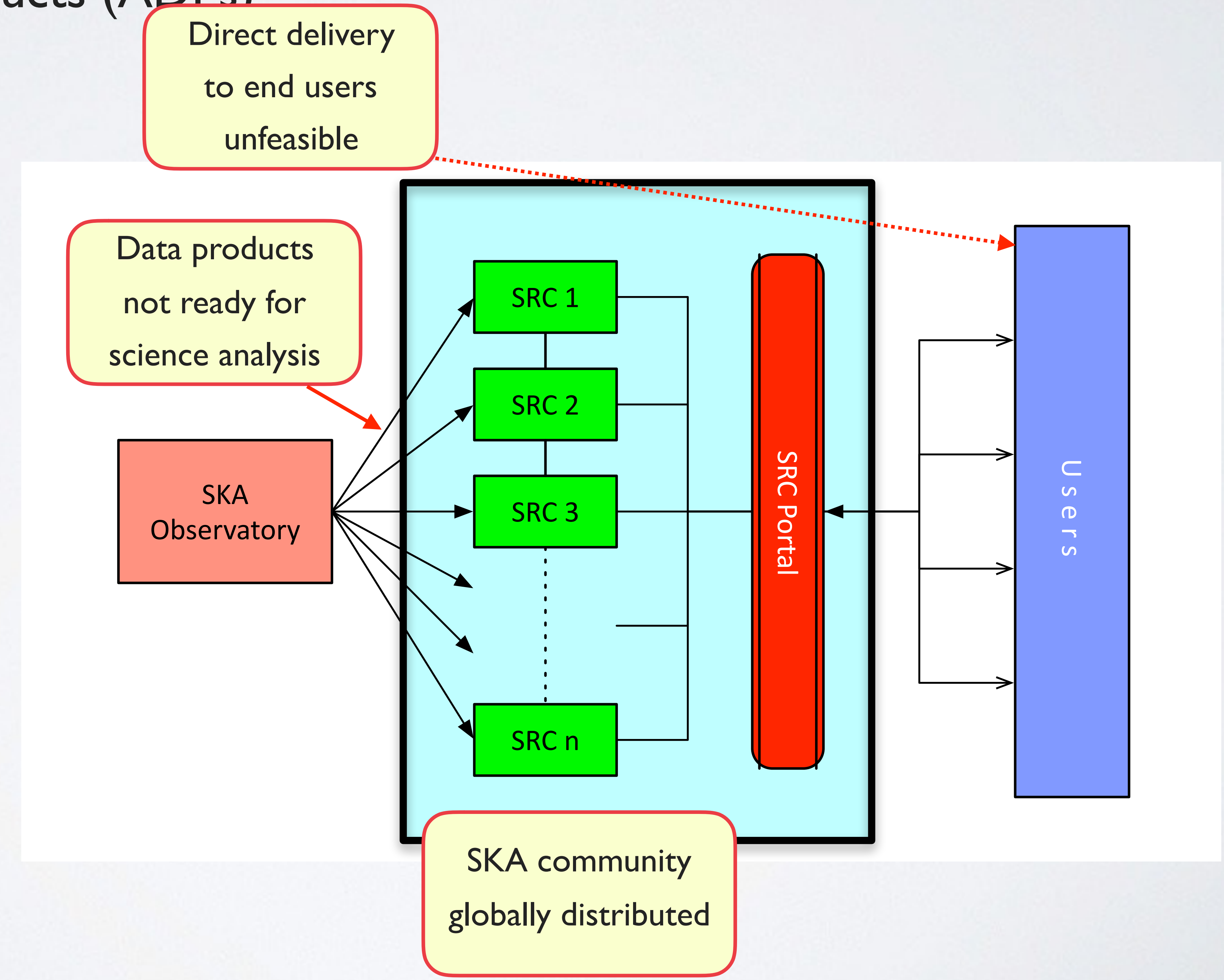
SKA Regional Centres (SRCs)

- SRCs will provide **access** to the SKA community to **data products** they are authorised to, as well as the **tools and processing power** to generate and analyse **Advanced Data Products (ADPs)**



SKA Regional Centres (SRCs)

- SRCs will provide **access** to the SKA community to **data products** they are authorised to, as well as the **tools and processing power** to generate and analyse **Advanced Data Products (ADPs)**



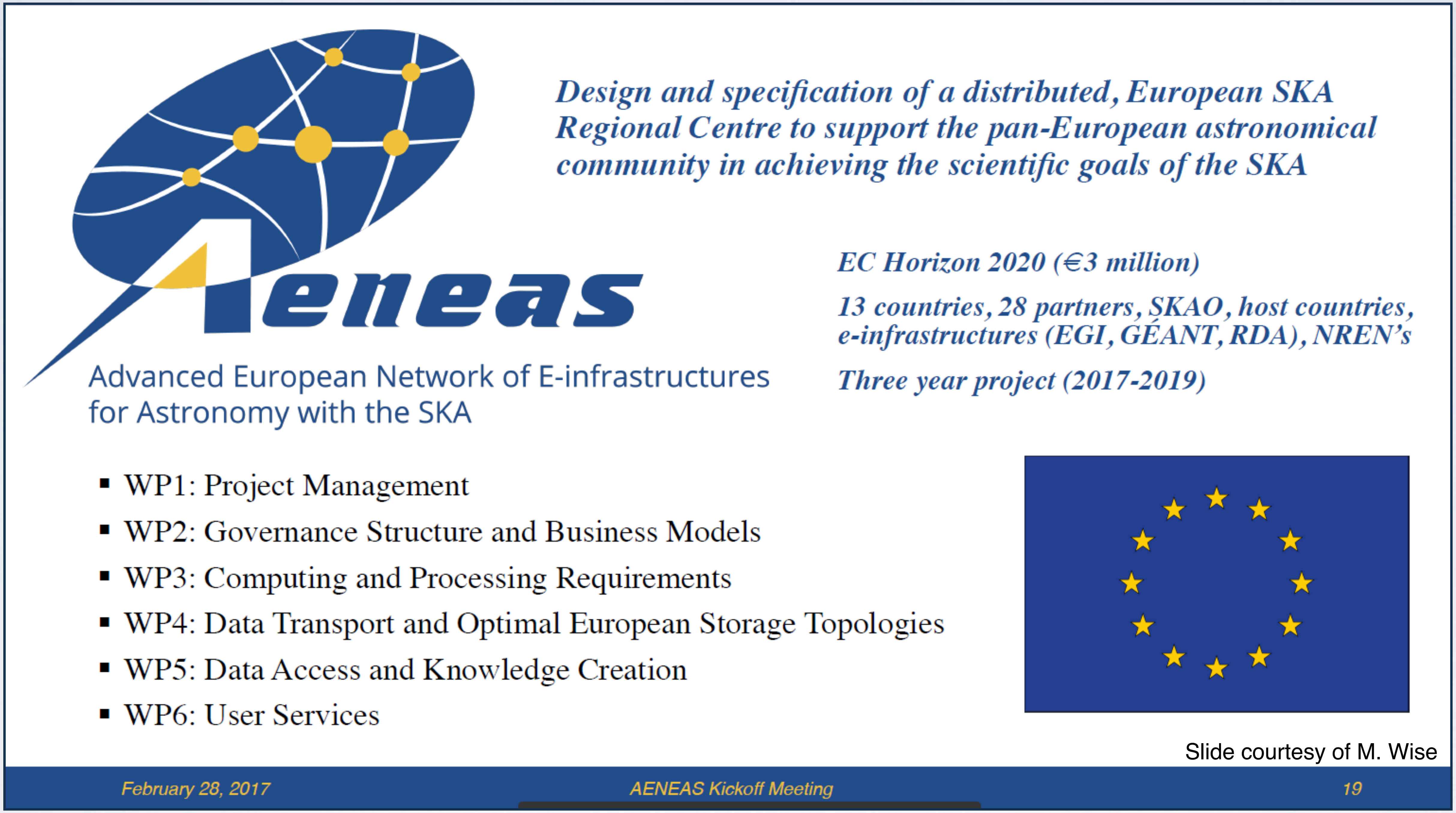
SKA Regional Centres (SRCs)

- SRCs will provide **access** to the SKA community to **data products** they are authorised to, as well as the **tools and processing power** to generate and analyse Advanced Data Products (ADPs)

<u>Essential functions:</u>

- Provision of a **common platform** to enable the access to SDP Data Products and the creation of ADPs of high science value
- **User support, curation and preservation** of SKA and user-generated **data products** and **workflows** leading to advanced science data products
- Provision of **resources for post-processing analysis and data visualization**
    - Provide platform for continued development of software
    - Provide a scientific platform for user to allow innovation in research methodologies
- Application of SKAO data policies and procedures for access to SKA data
- **Sharing data products, processing resources and workflows between SRCs**
- **Supporting VO services and protocols**

SKA Regional Centres (SRCs)



Getting inputs from the SKA scientific community is key (SWGs, FGs)

Processing requirements and technologies?

Interfaces, tools and techniques for analysis?

The Science Working Groups

SKA Science Working Groups & Focus groups

The Science Working Groups (SWGs) and Focus Groups (FGs) are scientific advisory bodies that provide input to the SKA Organisation on issues related to the design, construction, and future operations of the SKA that are likely to affect the Observatory's scientific capability, productivity and user relations. In addition, the FGs have a more specific, technical focus.

If you are interested in participating in any of the groups, please contact the current chairs or corresponding project scientists via the website link below.

- **Cosmology**
- **Cradle of Life**
- **Epoch of Reionization**
- **Extragalactic Continuum (galaxies/AGN, galaxy clusters)**
- **Extragalactic Spectral Line**
- **HI galaxy science**
- **High Energy Cosmic Particles (FG)**
- **Magnetism**
- **Our Galaxy**
- **Pulsars**
- **Solar, Heliospheric & Ionospheric Physics**
- **Transients**
- **VLBI (FG)**

@ska_telescope

@SKA_telescope

Square Kilometre Array

YouTube The Square Kilometre Array

For more, visit

astronomers.skatelescope.org/science-working-groups

SKA Fact sheets. August 2018. skatelescope.org

SKA1 Science Goals. HI Science

Science Goal	SWG	Objective	SWG Rank
1	CD/EoR	Physics of the early universe IGM - I. Imaging	1/3
2	CD/EoR	Physics of the early universe IGM - II. Power spectrum	2/3
3	CD/EoR	Physics of the early universe IGM - III. HI absorption line spectra (21cm forest)	3/3
4	Pulsars	Reveal pulsar population and MSPs for gravity tests and Gravitational Wave detection	1/3
5	Pulsars	High precision timing for testing gravity and GW detection	1/3
6	Pulsars	Characterising the pulsar population	2/3
7	Pulsars	Finding and using (Millisecond) Pulsars in Globular Clusters and External Galaxies	2/3
8	Pulsars	Finding pulsars in the Galactic Centre	2/3
9	Pulsars	Astrometric measurements of pulsars to enable improved tests of GR	2/3
10	Pulsars	Mapping the pulsar beam	3/3
11	Pulsars	Understanding pulsars and their environments through their interactions	3/3
12	Pulsars	Mapping the Galactic Structure	3/3
13	HI	Resolved HI kinematics and morphology of ~10^10 M_sol mass galaxies out to z~0.8	1/5
14	HI	High spatial resolution studies of the ISM in the nearby Universe.	2/5
15	HI	Multi-resolution mapping studies of the ISM in our Galaxy	3/5
16	HI	HI absorption studies out to the highest redshifts.	4/5
17	HI	The gaseous interface and accretion physics between galaxies and the IGM	5/5
24	Cradle of Life	Survey all nearby (~100 pc) stars for radio emission from technological civilizations.	3/5
25	Cradle of Life	The detection of pre-biotic molecules in pre-stellar cores at distance of 100 pc.	4/5
26	Cradle of Life	Mapping of the sub-structure and dynamics of nearby clusters using maser emission.	5/5
27	Magnetism	The resolved all-Sky characterisation of the interstellar and intergalactic magnetic fields	1/5
28	Magnetism	Determine origin, maintenance and amplification of magnetic fields at high redshifts - I.	2/5
29	Magnetism	Detection of polarised emission in Cosmic Web filaments	3/5
30	Magnetism	Determine origin, maintenance and amplification of magnetic fields at high redshifts - II.	4/5
31	Magnetism	Intrinsic properties of polarised sources	5/5
32	Cosmology	Constraints on primordial non-Gaussianity and tests of gravity on super-horizon scales.	1/5
33	Cosmology	Angular correlation functions to probe non-Gaussianity and the matter dipole	2/5
34	Cosmology	Map the dark Universe with a completely new kind of weak lensing survey - in the radio.	3/5
35	Cosmology	Dark energy & GR via power spectrum, BAO, redshift-space distortions and topology.	4/5
36	Cosmology	Test dark energy & general relativity with fore-runner of the 'billion galaxy' survey.	5/5
37	Continuum	Measure the Star formation history of the Universe (SFHU) - I. Non-thermal processes	1/8
38	Continuum	Measure the Star formation history of the Universe (SFHU) - II. Thermal processes	2/8
39	Continuum	Probe the role of black holes in galaxy evolution - I.	3/8
40	Continuum	Probe the role of black holes in galaxy evolution - II.	4/8
41	Continuum	Probe cosmic rays and magnetic fields in ICM and cosmic filaments.	5/8
42	Continuum	Study the detailed astrophysics of star-formation and accretion processes - I.	6/8
43	Continuum	Probing dark matter and the high redshift Universe with strong gravitational lensing.	7/8
44	Continuum	Legacy/Serendipity/Rare.	8/8

Table 1. Collated list of science goals. Within each science area, the entries are ordered in the rank provided by the SWG Chairs. The eight different groups of SWG contributions are listed in the Table in an arbitrary sequence.

SKA1 SCIENCE PRIORITY OUTCOMES

Document number SKA-TEL-SKO-0000122
Context .... SCI-REQ-RE
Revision 1
Author........ R. Braun, T. Bourke, J. Green, J. Wagg
Date 2014-09-25
Document Classification........ FOR PROJECT USE ONLY
Status........ Released

2014/09/25

PREVIOUS RELATED DISCUSSIONS

- 2015 SDP/SWG meeting
- A document was circulated on "Science Data Processor: anticipated data products"
- The HISWG gave comments/requests to the SDP
  - this to be the beginning of further discussions on what could be realistic,
  - Post processing by SWGs will have to be done by computing nodes or centres outside the central processor

Note that:

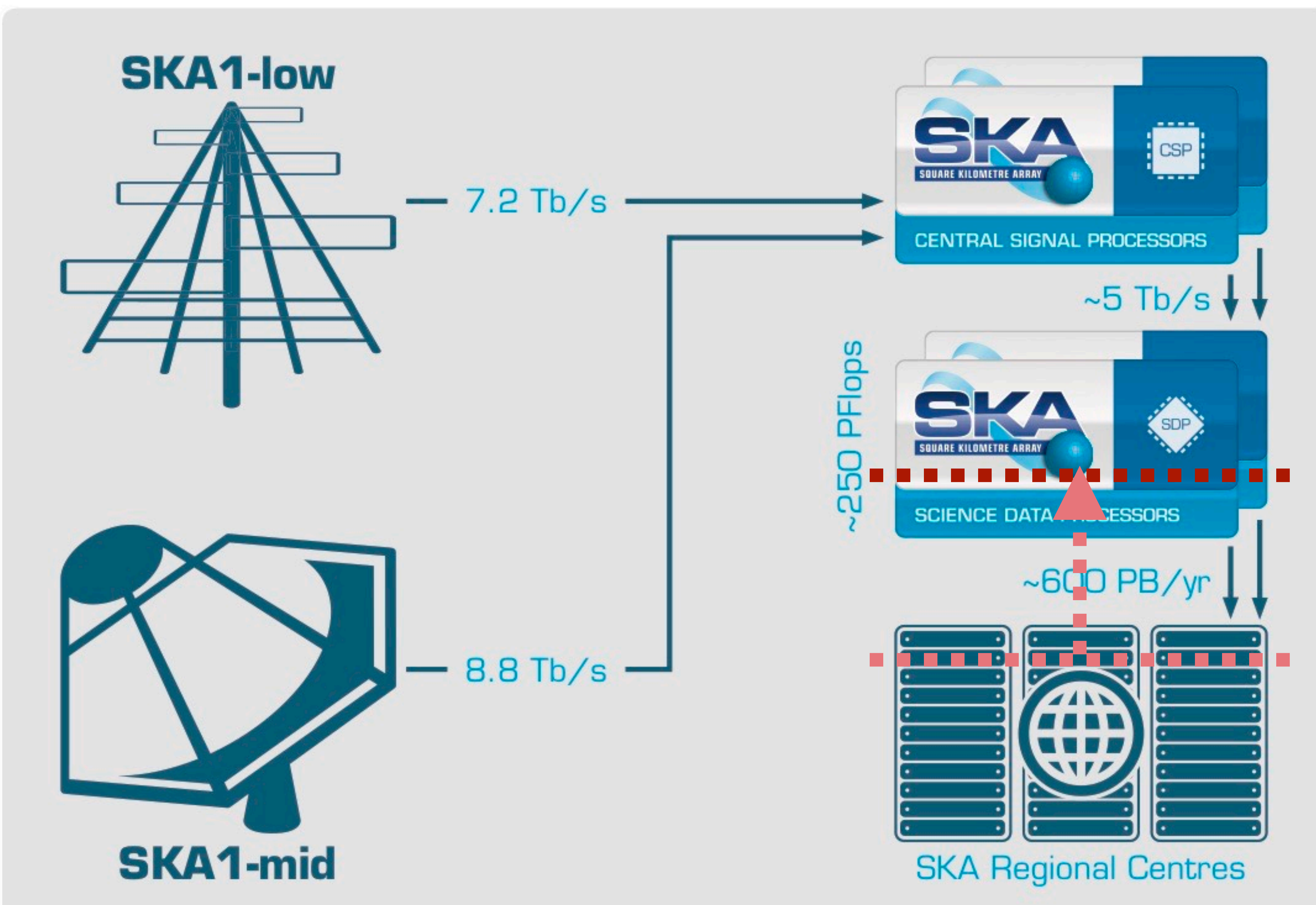
- This meeting was previous to the decision by the Board (April 2016) confirming its preference for a regional network model for provision of science data
- So now **the feedback should have as a starting point the boundary between the SDP and the SRCs**

SKA Regional Centres (SRCs)

The data journey

Basic data products will be produced and stored in Cape Town for SKA1-mid and Perth for SKA1-low. From there, they will be delivered to a global alliance of SKA Regional Centres for further processing and archiving and access by the user community.

New data models will need to be developed as current software packages for radio astronomy data reduction don't have the capabilities to handle the SKA's large bandwidths and Field of View datasets.



The SKA will be:

Distributed

One observatory operating two telescopes on three continents for a global scientific community.

Accessible

Common software and user interface. Pre-programmed algorithms. Training at the SKA Regional Centres.

Open

Open Access to non-proprietary data.

Citizen-ready

Access to SKA public data for citizen science projects.

@ska_telescope @SKA_telescope Square Kilometre Array The Square Kilometre Array

1.2 August 2018

SKA Fact sheets. August 2018. skatelescope.org

• Initial set of requirements extracted from the sample use cases scenarios indicated below (that intended to serve to guide requirements for SKA1 design)

SQUARE KILOMETRE ARRAY

SKA1 SCIENTIFIC USE CASES

Document number SKA-TEL-SKO-0000015
Revision 03
Author .... SKA Science Working Groups (SKAO Contact: Jeff Wagg)
Date 2016-01-29
Document Classification .... FOR PROJECT USE ONLY
Status .... Released

January 2016

**SKA-TEL-SKO-0000015_REV_03_SKA1_SCIENCE_USE_CASES_COMBINED-PART-1-SIGNED.PDF

• Initial set of requirements extracted from the sample use cases scenarios indicated below (that intended to serve to guide requirements for SKA1 design)

SQUARE KILOMETRE ARRAY

SKA1 SCIENTIFIC USE CASES

Document number	SKA-TEL-SKO-0000015
Revision	03
Author	SKA Science Working Groups (SKAO Contact: Jeff Wagg)
Date	2016-01-29
Document Classification	FOR PROJECT USE ONLY
Status	Released

These are not real (Key) Science projects

TABLE OF CONTENTS

**SKA-TEL-SKO-0000015_Rev_03_SKA1_Science_Use_Cases_Combined-part-1-signed.pdf

3.4 SKA1 All-Sky HI Survey

PROJECT DETAILS	
Title	**SKA1 All-Sky HI Survey**
Principal Investigator	**Oort**
Co-Authors	**HI-Galaxy SWG**
Time Request	**10,000 hrs**
FACILITY	Preconditions
SKA1-LOW	
? SKA1-MID	

RECEIVER(S) REQUIRED		Time (hrs)
	SKA1-LOW	
	SKA1-MID Band 1	
?	SKA1-MID Band 2	10,000
	SKA1-MID Band 3	
	SKA1-MID Band 4	
	SKA1-MID Band 5	

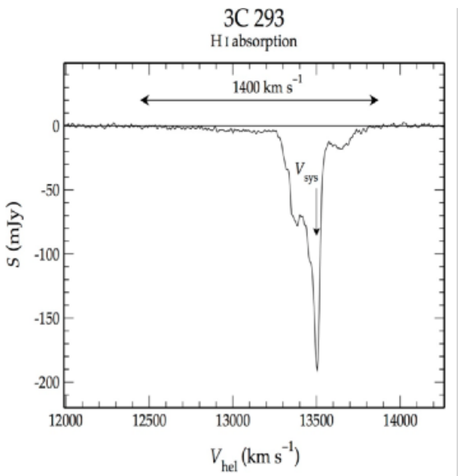
OPERATIONAL MODE (as defined in Concept-of-Operations)		Details
?	Normal	Mosaic observations
	Fixed schedule (give cadence)	
	Time-critical override	
	Custom Experiment	
?	Commensal	Continuum, polarisation, cc
	Collaborative & Coordinated	
	Sub-arrays required	

DATA ANALYSIS	
Procedures required	RFI mitigation, flagging, calibration, continuum subtraction, widefield imaging, mosaicing, source-finding and source parameterisation
Processing considerations	Likely processing issues: • Large data volumes • Widefield images with non-coplanar baselines • Accurate primary beam for mosaicing • Flagging of RFI • Requirement to search cubes multiple times (source detection). • Requirement to stack at a given set of coordinates
Data products	Stokes I data cubes; image cut-outs; spectra; minicubes; catalogues
Pipeline	• Collect visibilities on multiple days • Apply barycentric correction • Apply flagging • Calibrate visibilities • Peel strong continuum sources • Subtract remaining continuum sources using global sky model • Make daily cubes at multiple resolutions for each pointing • Combine cubes (and beams) for individual pointings in the image domain • Residual polynomial continuum subtraction in image domain • Linear mosaicking of multiple fields, followed by cutouts in RA, Dec, Freq • Multiscale deconvolution of strongest sources
Quality assessment plan & cadence	• Inspect RFI occupancy plots • Examine rms and histogram of pixel values in daily cubes. • Compare flux densities of known sources in daily cubes.
Latency (Desired time lag between ... being available in the archive. e.g. This could range from 'a few seconds' for transient detections using the fast imaging pipeline, to 'upon completion of scheduling block and pipeline reduction' (approximately 24 hours), to 'at completion of the full project'.)	On completion of observations and data reduction

This document was prepared when only the SDP "existed"

16	*HI*	HI absorption studies out to the highest redshifts.

3.3 A blind HI 21-cm absorption line survey at 3<z<6
SKA1-LOW, 5000 hrs
50 unique pointings x 100 h
3.7 An all-sky absorption survey at z~1 - 3
Mid- Band 1, 1,000 h, Individual pointings



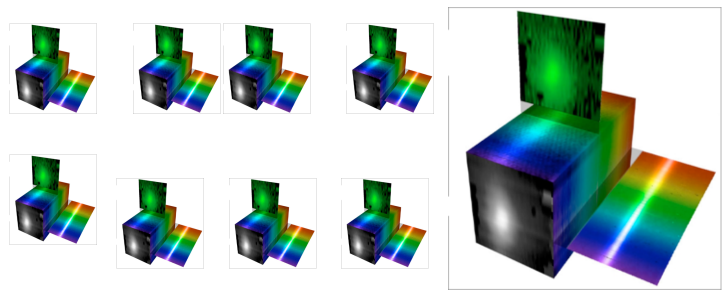
Morganti et al 2015

- RFI mitigation and flagging
- Gain calibration
- Generate continuum visibility dataset
- Generate continuum image
- Apply the self-calibration solutions and direction-dependent corrections to the spectral line visibility dataset
- Subtract continuum from the line dataset
- Doppler correct (CVEL) line dataset
- Generate stokes-I spectral-line cube for full FoV at 4.6 kHz resolution
- Deconvolve channels with line detections, if needed.

———-

Calibrated continuum and spectral line visibilities to be **combined with the data from other observing runs** to generate `final' stokes-I spectral-line cube(s).

3.4 SKA1 All-Sky HI Survey
3.6 Medium-Deep HI Imaging Survey
3.9 Deep HI Imaging Survey
3.10 Medium-Wide HI Imaging Survey
Mid-Band 2, Mid- Band 1 x 310,000 h, 2,000 h, 3000 h, 2000h
Mosaic observations,
20,000 targets, One deep field, Single pointing

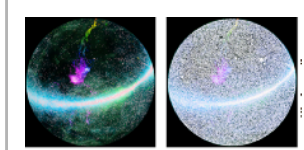


- Collect visibilities on multiple days
- Apply barycentric correction
- Apply flagging
- Calibrate visibilities
- Peel strong continuum sources
- Subtract remaining continuum sources using global sky model / Polynomial continuum subtraction

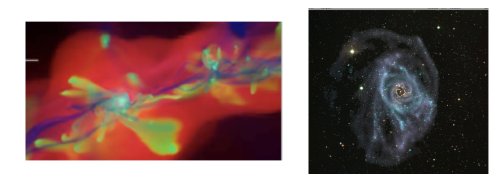
————

- Daily cubes at **multiple resolutions** for each pointing (will SDP produce the required number?)
- **Combine cubes** (and beams) for individual pointings in the image domain
- Residual cont. subtraction in image domain
- **Multiscale deconv.** of strongest sources
- Residual polynomial continuum subtraction in image domain
- **Linear mosaicking** of multiple fields, followed by cutouts in RA, Dec, Freq

3.5 Deep Galactic and Magellanic HI Survey
Mid- Band 2, 4,500 h, Maps through multiple fields of view, 1,200 targets
3.8 Cosmic Web: The extended environment of galaxies and the IGM
Mid- Band 2, 100 h, Individual pointings with multiple objects
3.11 High spatial resolution imaging of the HI in nearby galaxies
Mid- Band 2, 300 h, multiple fields of view



McClure-Griffiths et al 2015



Agertz, Teyssier, Moore 2009

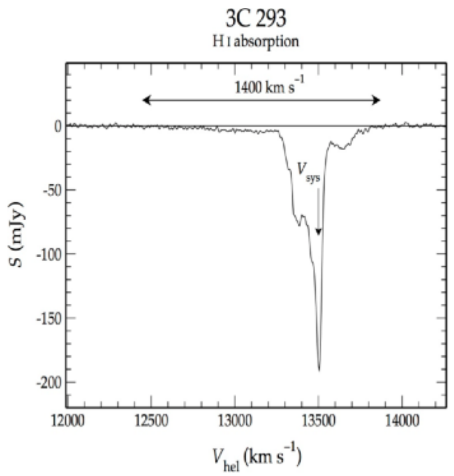
- Calibration, flagging, imaging
- **Combination of spectral line data cubes** from different observing runs
- Gridding the UV data so that new data can be combined in this grid as it is observed.

PIPELINES

Inputs!

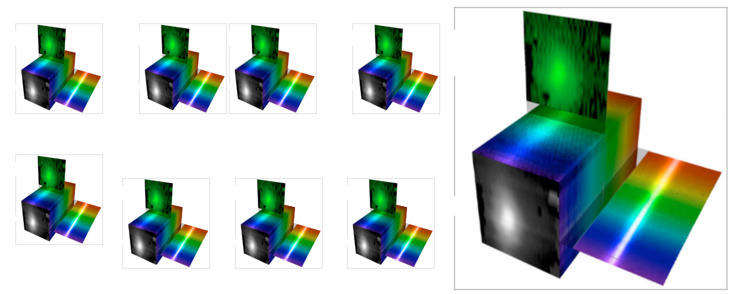
13	*HI*	Resolved HI kinematics and morphology of ~10^10 M_sol mass galaxies out to z~0.8

3.3 A blind HI 21-cm absorption line survey at 3<z<6
SKA1-LOW, 5000 hrs
50 unique pointings x 100 h
3.7 An all-sky absorption survey at z~1 - 3
Mid- Band 1, 1,000 h, Individual pointings

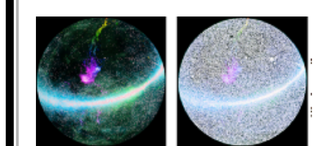


Morganti et al 2015

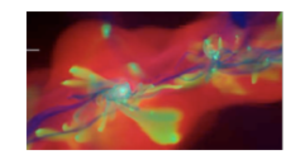
3.4 SKA1 All-Sky HI Survey
3.6 Medium-Deep HI Imaging Survey
3.9 Deep HI Imaging Survey
3.10 Medium-Wide HI Imaging Survey
Mid-Band 2, Mid- Band 1 x 310,000 h, 2,000 h, 3000 h, 2000h
Mosaic observations,
20,000 targets, One deep field, Single pointing



3.5 Deep Galactic and Magellanic HI Survey
Mid- Band 2, 4,500 h, Maps through multiple fields of view, 1,200 targets
3.8 Cosmic Web: The extended environment of galaxies and the IGM
Mid- Band 2, 100 h, Individual pointings with multiple objects
3.11 High spatial resolution imaging of the HI in nearby galaxies
Mid- Band 2, 300 h, multiple fields of view



McClure-Griffiths et al 2015



Agertz, Teyssier, Moore 2009

- RFI mitigation and flagging
- Gain calibration
- Generate continuum visibility dataset
- Generate continuum image
- Apply the self-calibration solutions and direction-dependent corrections to the spectral line visibility dataset
- Subtract continuum from the line dataset
- Doppler correct (CVEL) line dataset
- Generate stokes-I spectral-line cube for full FoV at 4.6 kHz resolution
- Deconvolve channels with line detections, if needed.

———-

Calibrated continuum and spectral line visibilities to be **combined with the data from other observing runs** to generate `final' stokes-I spectral-line cube(s).

- Collect visibilities on multiple days
- Apply barycentric correction
- Apply flagging
- Calibrate visibilities
- Peel strong continuum sources
- Subtract remaining continuum sources using global sky model / Polynomial continuum subtraction

———

- Daily cubes at **multiple resolutions** for each pointing (will SDP produce the required number?)
- **Combine cubes** (and beams) for individual pointings in the image domain
- Residual cont. subtraction in image domain
- **Multiscale deconv.** of strongest sources
- Residual polynomial continuum subtraction in image domain
- **Linear mosaicking** of multiple fields, followed by cutouts in RA, Dec, Freq

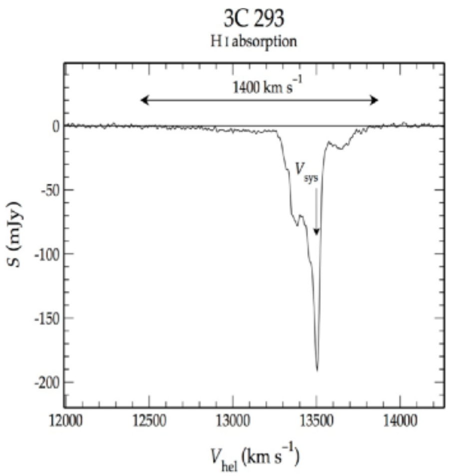
- Calibration, flagging, imaging
- **Combination of spectral line data cubes** from different observing runs
- Gridding the UV data so that new data can be combined in this grid as it is observed.

PIPELINES

Inputs!

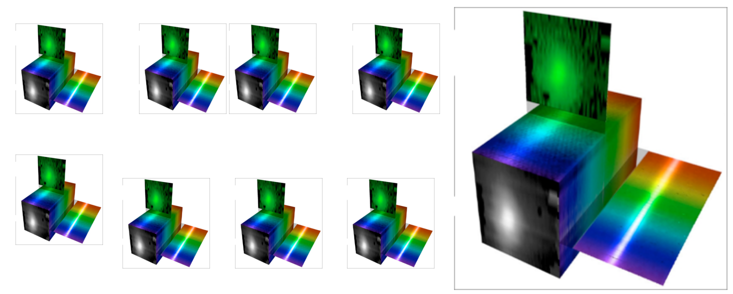
15	*HI*	Multi-resolution mapping studies of the ISM in our Galaxy

3.3 A blind HI 21-cm absorption line survey at 3<z<6
SKA1-LOW, 5000 hrs
50 unique pointings x 100 h
3.7 An all-sky absorption survey at z~1 - 3
Mid- Band 1, 1,000 h, Individual pointings



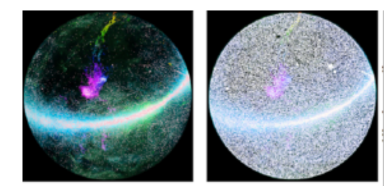
Morganti et al 2015

3.4 SKA1 All-Sky HI Survey
3.6 Medium-Deep HI Imaging Survey
3.9 Deep HI Imaging Survey
3.10 Medium-Wide HI Imaging Survey
Mid-Band 2, Mid- Band 1 x 310,000 h, 2,000 h, 3000 h, 2000h
Mosaic observations,
20,000 targets, One deep field, Single pointing



3.5 Deep Galactic and Magellanic HI Survey
Mid- Band 2, 4,500 h, Maps through multiple fields of view, 1,200 targets

3.8 Cosmic Web: The extended environment of galaxies and the IGM
Mid- Band 2, 100 h, Individual pointings with multiple objects
3.11 High spatial resolution imaging of the HI in nearby galaxies
Mid- Band 2, 300 h, multiple fields of view



McClure-Griffiths et al 2015

- RFI mitigation and flagging
- Gain calibration
- Generate continuum visibility dataset
- Generate continuum image
- Apply the self-calibration solutions and direction-dependent corrections to the spectral line visibility dataset
- Subtract continuum from the line dataset
- Doppler correct (CVEL) line dataset
- Generate stokes-I spectral-line cube for full FoV at 4.6 kHz resolution
- Deconvolve channels with line detections, if needed.

———-

Calibrated continuum and spectral line visibilities to be **combined with the data from other observing runs** to generate `final' stokes-I spectral-line cube(s).

- Collect visibilities on multiple days
- Apply barycentric correction
- Apply flagging
- Calibrate visibilities
- Peel strong continuum sources
- Subtract remaining continuum sources using global sky model / Polynomial continuum subtraction

———

- Daily cubes at **multiple resolutions** for each pointing (will SDP produce the required number?)
- **Combine cubes** (and beams) for individual pointings in the image domain
- Residual cont. subtraction in image domain
- **Multiscale deconv.** of strongest sources
- Residual polynomial continuum subtraction in image domain
- **Linear mosaicking** of multiple fields, followed by cutouts in RA, Dec, Freq

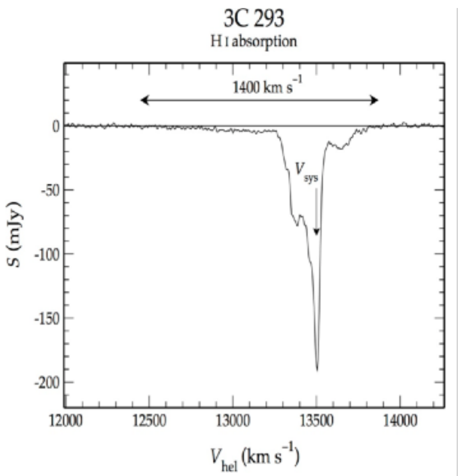
- Calibration, flagging, imaging
- **Combination of spectral line data cubes** from different observing runs
- Gridding the UV data so that new data can be combined in this grid as it is observed.

PIPELINES

Inputs!

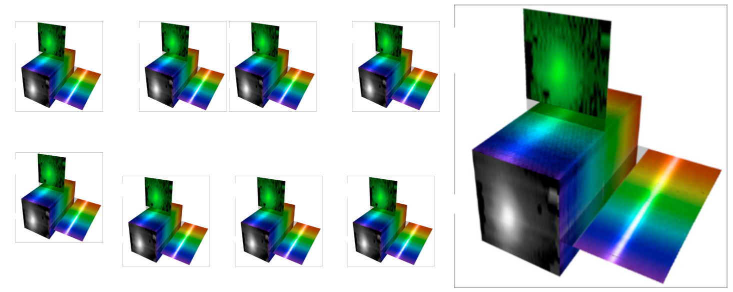
17	*HI*	The gaseous interface and accretion physics between galaxies and the IGM

3.3 A blind HI 21-cm absorption line survey at 3<z<6
SKA1-LOW, 5000 hrs
50 unique pointings x 100 h
3.7 An all-sky absorption survey at z~1 - 3
Mid- Band 1, 1,000 h, Individual pointings

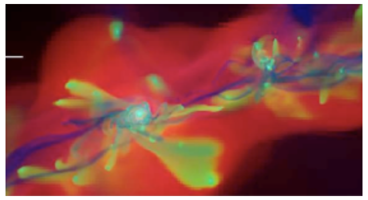


Morganti et al 2015

3.4 SKA1 All-Sky HI Survey
3.6 Medium-Deep HI Imaging Survey
3.9 Deep HI Imaging Survey
3.10 Medium-Wide HI Imaging Survey
Mid-Band 2, Mid- Band 1 x 310,000 h, 2,000 h, 3000 h, 2000h
Mosaic observations,
20,000 targets, One deep field, Single pointing



3.5 Deep Galactic and Magellanic HI Survey
Mid- Band 2, 4,500 h, Maps through multiple fields of view, 1,200 targets
3.8 Cosmic Web: The extended environment of galaxies and the IGM
Mid- Band 2, 100 h, Individual pointings with multiple objects
3.11 High spatial resolution imaging of the HI in nearby galaxies
Mid- Band 2, 300 h, multiple fields of view



Agertz, Teyssier, Moore 2009

- RFI mitigation and flagging
- Gain calibration
- Generate continuum visibility dataset
- Generate continuum image
- Apply the self-calibration solutions and direction-dependent corrections to the spectral line visibility dataset
- Subtract continuum from the line dataset
- Doppler correct (CVEL) line dataset
- Generate stokes-I spectral-line cube for full FoV at 4.6 kHz resolution
- Deconvolve channels with line detections, if needed.

——-

Calibrated continuum and spectral line visibilities to be **combined with the data from other observing runs** to generate `final' stokes-I spectral-line cube(s).

- Collect visibilities on multiple days
- Apply barycentric correction
- Apply flagging
- Calibrate visibilities
- Peel strong continuum sources
- Subtract remaining continuum sources using global sky model / Polynomial continuum subtraction

————

- Daily cubes at **multiple resolutions** for each pointing (will SDP produce the required number?)
- **Combine cubes** (and beams) for individual pointings in the image domain
- Residual cont. subtraction in image domain
- **Multiscale deconv.** of strongest sources
- Residual polynomial continuum subtraction in image domain
- **Linear mosaicking** of multiple fields, followed by cutouts in RA, Dec, Freq

- Calibration, flagging, imaging
- **Combination of spectral line data cubes** from different observing runs
- Gridding the UV data so that new data can be combined in this grid as it is observed.

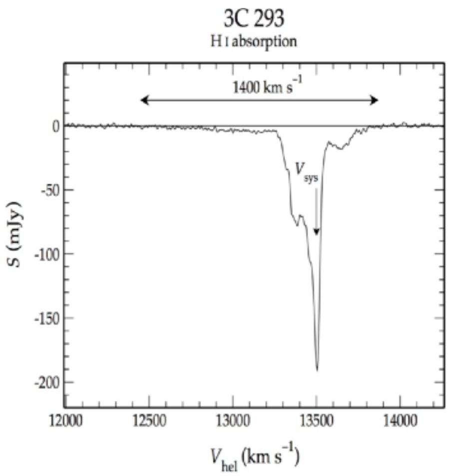
PIPELINES

Inputs!

14	HI	High spatial resolution studies of the ISM in the nearby Universe.

3.3 A blind HI 21-cm absorption line survey at 3<z<6
SKA1-LOW, 5000 hrs
50 unique pointings x 100 h

3.7 An all-sky absorption survey at z~1 - 3
Mid- Band 1, 1,000 h, Individual pointings



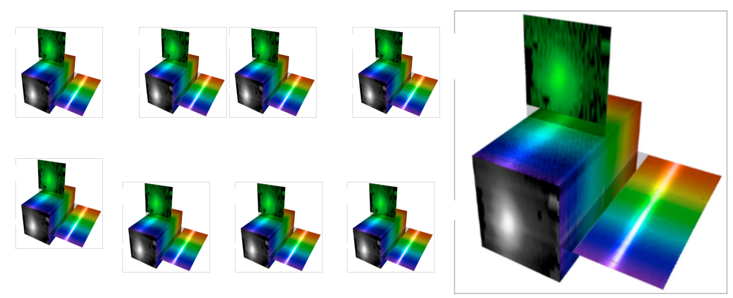
Morganti et al 2015

- RFI mitigation and flagging
- Gain calibration
- Generate continuum visibility dataset
- Generate continuum image
- Apply the self-calibration solutions and direction-dependent corrections to the spectral line visibility dataset
- Subtract continuum from the line dataset
- Doppler correct (CVEL) line dataset
- Generate stokes-I spectral-line cube for full FoV at 4.6 kHz resolution
- Deconvolve channels with line detections, if needed.

———-

Calibrated continuum and spectral line visibilities to be **combined with the data from other observing runs** to generate `final' stokes-I spectral-line cube(s).

3.4 SKA1 All-Sky HI Survey
3.6 Medium-Deep HI Imaging Survey
3.9 Deep HI Imaging Survey
3.10 Medium-Wide HI Imaging Survey
Mid-Band 2, Mid- Band 1 x 310,000 h, 2,000 h, 3000 h, 2000h
Mosaic observations,
20,000 targets, One deep field, Single pointing



- Collect visibilities on multiple days
- Apply barycentric correction
- Apply flagging
- Calibrate visibilities
- Peel strong continuum sources
- Subtract remaining continuum sources using global sky model / Polynomial continuum subtraction

———

- Daily cubes at **multiple resolutions** for each pointing (will SDP produce the required number?)
- **Combine cubes** (and beams) for individual pointings in the image domain
- Residual cont. subtraction in image domain
- **Multiscale deconv.** of strongest sources
- Residual polynomial continuum subtraction in image domain
- **Linear mosaicking** of multiple fields, followed by cutouts in RA, Dec, Freq

3.5 Deep Galactic and Magellanic HI Survey
Mid- Band 2, 4,500 h, Maps through multiple fields of view, 1,200 targets

3.8 Cosmic Web: The extended environment of galaxies and the IGM
Mid- Band 2, 100 h, Individual pointings with multiple objects

3.11 High spatial resolution imaging of the HI in nearby galaxies
Mid- Band 2, 300 h, multiple fields of view

Blok et al 2015, Picture courtesy B. Koribalski

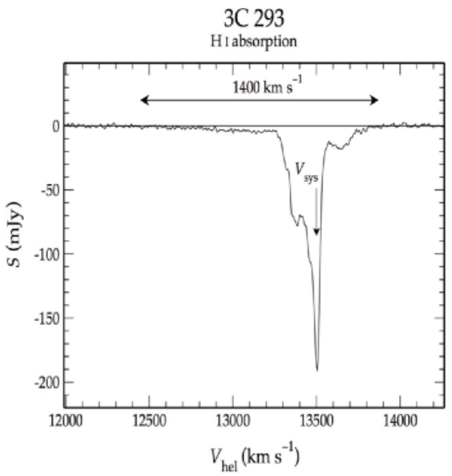
- Calibration, flagging, imaging
- **Combination of spectral line data cubes** from different observing runs
- Gridding the UV data so that new data can be combined in this grid as it is observed.

PIPELINES

Inputs!

14	*HI*	High spatial resolution studies of the ISM in the nearby Universe.

3.3 A blind HI 21-cm absorption line survey at 3<z<6
SKA1-LOW, 5000 hrs
50 unique pointings x 100 h
3.7 An all-sky absorption survey at z~1 - 3
Mid- Band 1, 1,000 h, Individual pointings



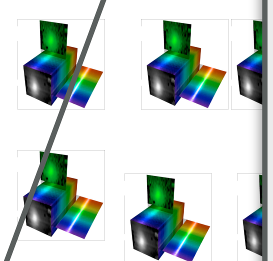
Morganti et al 2015

- RFI mitigation and flagging
- Gain calibration
- Generate continuum visibility dataset
- Generate continuum image
- Apply the self-calibration solutions and direction-dependent corrections to the spectral line visibility dataset
- Subtract continuum from the line dataset
- Doppler correct (CVEL) line dataset
- Generate stokes-I spectral-line cube for full FoV at 4.6 kHz resolution
- Deconvolve channels with line detections, if needed.

———-

Calibrated continuum and spectral line visibilities to be combined with the data from other observing runs to generate `final' stokes-I spectral-line cube(s).

3.4 SKA1 All-Sk
3.6 Medium-D
3.9 Deep HI In
3.10 Medium-W
Mid-Band 2, Mic
3000 h, 2000h
Mosaic observat
20,000 targets, (



- Collect visibil
- Apply baryce
- Apply flaggin
- Calibrate visi
- Peel strong c
- Subtract rem global sky mo subtraction

———

- Daily cubes a pointing (will number?)
- **Combine cu** pointings in t
- Residual cont
- **Multiscale deconv.** of strongest sources
- Residual polynomial continuum subtraction in image domain
- **Linear mosaicking** of multiple fields, followed by cutouts in RA, Dec, Freq

- Collect visibilities on multiple days
- Apply barycentric correction
- Apply flagging
- Calibrate visibilities
- Peel strong continuum sources
- Subtract remaining continuum sources using global sky model / Polynomial continuum subtraction

———

- Daily cubes at **multiple resolutions** for each pointing (will SDP produce the required number?)
- **Combine cubes** (and beams) for individual pointings in the image domain
- Residual cont. subtraction in image domain
- **Multiscale deconv.** of strongest sources
- Residual polynomial continuum subtraction in image domain
- **Linear mosaicking** of multiple fields, followed by cutouts in RA, Dec, Freq

PIPELINES

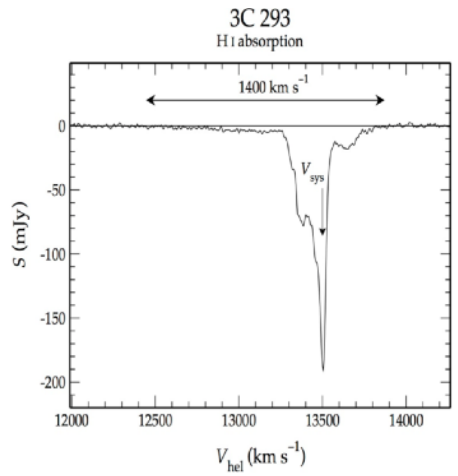
Inputs!

Procedures required

Processing considerations

Data products

3.3 A blind HI 21-cm absorption line survey at 3<z<6
SKA1-LOW, 5000 hrs
50 unique pointings x 100 h
3.7 An all-sky absorption survey at z~1 - 3
Mid- Band 1, 1,000 h, Individual pointings



Morganti et al 2015

Procedures required, not in the SDP

—

Processing considerations

Reprocessing of calibrated visibilities

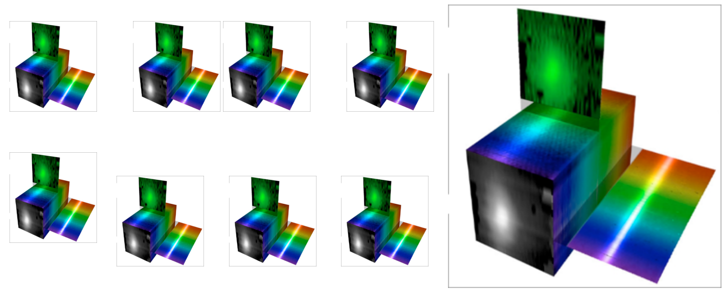
Data products

- Stokes I continuum visibility datasets and images at 225, 275 and 325 MHz
- Stokes I spectral-line cube over 200 – 350 MHz with 4 kHz resolution
- Cubelets and spectra towards all the sources brighter than 10 mJy in the FoV along with the RFI flags applied to the data.

——-

Continuum source and spectral line catalogs
Continuum image with 30% bandwidth centered at 600 MHz. This will be used to check the total flux density of the background source which should further be **logged in a public database for future sky-model reference**

3.4 SKA1 All-Sky HI Survey
3.6 Medium-Deep HI Imaging Survey
3.9 Deep HI Imaging Survey
3.10 Medium-Wide HI Imaging Survey
Mid-Band 2, Mid- Band 1 x 310,000 h, 2,000 h, 3000 h, 2000h
Mosaic observations,
20,000 targets, One deep field, Single pointing



Procedures required, not in the SDP

- source-finding and source parameterisation
- data combination to created integrated deep cube (uv and image domain)

Processing considerations

Large data volumes, search cube multiple times (source detection), stacking

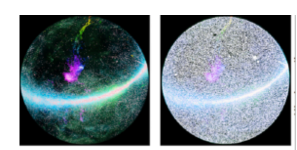
Data products

- Stokes I data cubes
- Calibrated, imaged, continuum subtracted datacubes

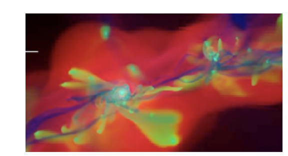
————

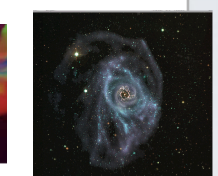
- image cut-outs
- spectra
- minicubes
- catalogues
- moment maps
- **masks used to make moment maps**
- **signal-to-noise maps**

3.5 Deep Galactic and Magellanic HI Survey
Mid- Band 2, 4,500 h, Maps through multiple fields of view, 1,200 targets
3.8 Cosmic Web: The extended environment of galaxies and the IGM
Mid- Band 2, 100 h, Individual pointings with multiple objects
3.11 High spatial resolution imaging of the HI in nearby galaxies
Mid- Band 2, 300 h, multiple fields of view



McClure-Griffiths et al 2015





Agertz, Teyssier, Moore 2009

Procedures required, not in the SDP

—

- Image cubes at multiple resolutions (will SDP produce the required number?).
- Addition of single dish data for imaging.
- joint deconvolution of mosaic;
- Multi-scale deconvolution;
- Addition of single-dish data for imaging

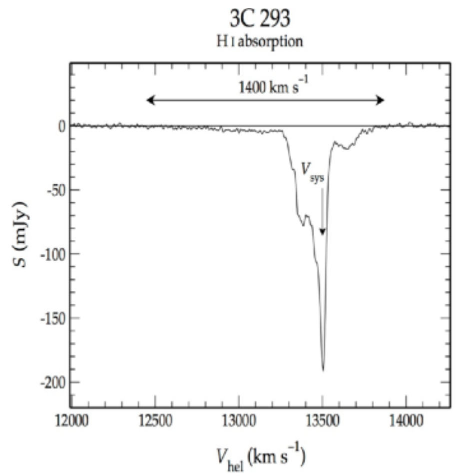
Processing considerations

Large data volumes due to full spectral resolution?

Data products

- Fully calibrated I, Q, U and V cubes at full spectral resolution
- Image cubes, moment maps, images of the PS
- Data cubes, Total HI image, velocity field, velocity dispersion map. At various resol.
- **masks used to make moment maps**
- **signal-to-noise maps**

3.3 A blind HI 21-cm absorption line survey at 3<z<6
SKA1-LOW, 5000 hrs
50 unique pointings x 100 h
3.7 An all-sky absorption survey at z~1 - 3
Mid- Band 1, 1,000 h, Individual pointings



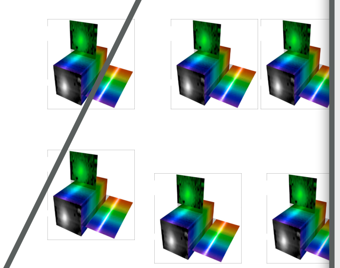
Morganti et al 2015

Procedures required, not in the SDP
—
Processing considerations
Reprocessing of calibrated visibilities
Data products
- Stokes I continuum visibility datasets and images at 225, 275 and 325 MHz
- Stokes I spectral-line cube over 200 – 350 MHz with 4 kHz resolution
- Cubelets and spectra towards all the sources brighter than 10 mJy in the FoV along with the RFI flags applied to the data.

———-

Continuum source and spectral line catalogs
Continuum image with 30% bandwidth centered at 600 MHz. This will be used to check the total flux density of the background source which should further be **logged in a public database for future sky-model reference**

3.4 SKA1 All-Sky HI Survey
3.6 Medium-Deep
3.9 Deep HI Imagi
3.10 Medium-Wid
Mid-Band 2, Mid- Ba
3000 h, 2000h
Mosaic observations
20,000 targets, One



Procedures require
- source-finding an
- data combination deep cube (uv ar
Processing conside
Large data volumes (source detection),

Data products
- Stokes I data cub
- Calibrated, image datacubes

———

- image cut-outs
- spectra
- minicubes
- catalogues
- moment maps
- **masks used to m**
- **signal-to-noise maps**

3.5 Deep Galactic and Magellanic HI

signal-to-noise maps

Procedures required, not in the SDP
- source-finding and source parameterisation
- data combination to created integrated deep cube (uv and image domain)

Processing considerations
Large data volumes, search cube multiple times (source detection), stacking

Data products
- Stokes I data cubes
- Calibrated, imaged, continuum subtracted datacubes

———

- image cut-outs
- spectra
- minicubes
- catalogues
- moment maps
- **masks used to make moment maps**
- **signal-to-noise maps**

Inputs!

MAIN CHARACTERISTICS OF THE DATA PRODUCTS

- **Volumes:**
  - SKA1MID, Band 2: Discovery cube size 2.6Pbytes. HI science a fraction down to 1/10 of the max. Extracted data products at least 10 times smaller (moment maps, pos.vel cuts, spectra)
  - continuum data products or spectral postage stamp cubes would be orders of magnitude smaller than the discovery cube
  - Data storage and compute access for reducing and calibrating data can be an issue (feedback from PHISCC 2018)
- **Formats:**
  - ongoing work by ICRAR-IT on the jp2 and jpx formats could feed perfectly into the requirements of efficiently extracting sub-cubes of various resolutions from a huge master cube.
  - …
- **Metadata:**
  - E.g. for spectra, description of the parameters that went into making them, like the size of the extraction region.

Inputs!

POTENTIAL SOFTWARE TO INTEGRATE IN THE SRCs SCIENCE PLATFORM

- Developed by precursors/pathfinders, KSPs, etc:
  - **Analysis**, e.g. SoFiA, TiRiFiC, GalAPAGOS, GIPSY/GuiPSY, 2DBAT, FAT, MAGMO, Barolo, GBKFit, CASA, etc (TBD)
  - **Visualization**: e.g SlicerAstro, VISIONS, X3D

Inputs!

- **Interaction/connection with the VO** (e.g. in order to request complementary data or input data for modeling;)

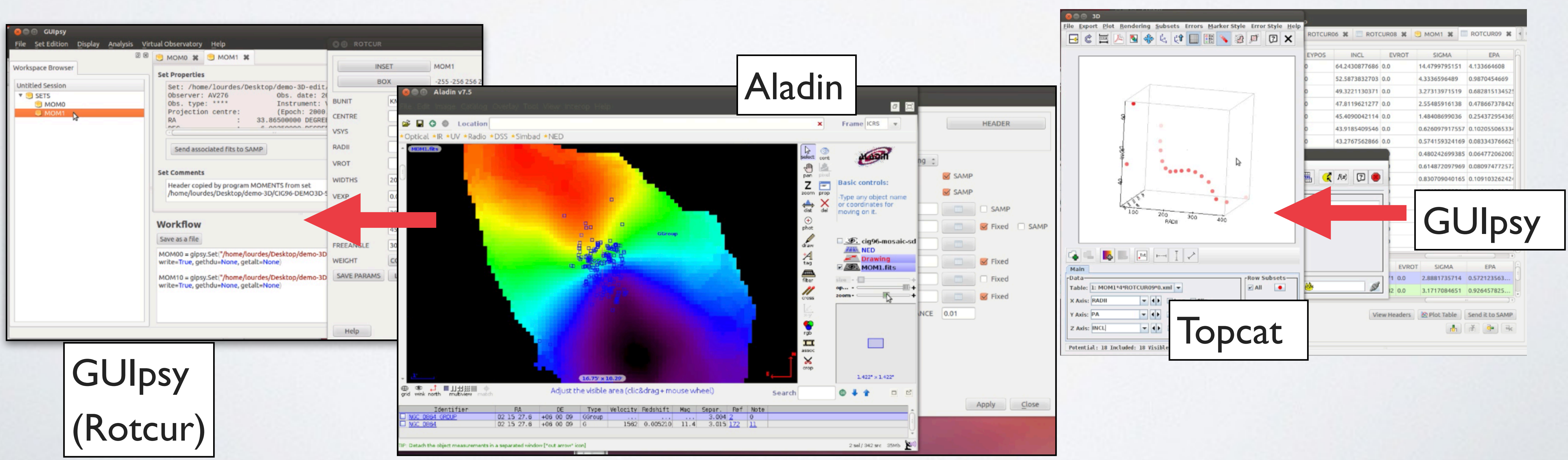
POTENTIAL SOFTWARE TO INTEGRATE IN THE SRCS SCIENCE PLATFORM

- Developed by precursors/pathfinders, KSPs, etc:
  - **Analysis**, e.g. SoFiA, TiRiFiC, GalAPAGOS, GIPSY/GuiPSY, 2DBAT, FAT, MAGMO,Barolo, GBKFit, CASA, etc (TBD)
  - **Visualization**: e.g SlicerAstro, VISIONS, X3D

Inputs!

- **Interaction/connection with the VO** (e.g. in order to request complementary data or input data for modeling;)

GUIpsy a VO compliant tool for the kinematical modelling of HI datacubes
Sánchez-Exposito, S.; Ruiz, J.E.; Vogelaar, M.G.R; Terlouw, J.P.; Verdes-Montenegro, L.; Santander-Vela, J.D.; **van der Hulst, J.M**; Garrido, J.
3GC4: HI Fidelity, 28th October, 2016. Port Alfred, South Africa



Not HI specific!

OPEN SCIENCE AT THE SRCs

- SKA data exploitation through a platform that **facilitates:**
  - **collaboration among international teams** in order to extract the maximum scientific knowledge
  - **data sharing + re-use & re-purposing of the analysis tools**
  - the accuracy and the **reproducibility of our scientific methods**

Not HI specific

Open Science at the SRCs

- SKA data exploitation through a platform that **facilitates:**
  - **collaboration among international teams** in order to extract the maximum scientific knowledge
  - **data sharing + re-use & re-purposing of the analysis tools**
  - the accuracy and the **reproducibility of our scientific methods**

SQUARE KILOMETRE ARRAY

PUBLIC

SKA REGIONAL CENTRE REQUIREMENTS

Document Number	SKA-TEL-SKO-0000735
Document Type	RSP
Revision	01
Author	R. C. Bolton and the SRCCG
Date	2017-09-27
Document Classification	UNRESTRICTED
Status	Released

Ellaborated by the SRCCG

Approved by Gary Davis

Not HI specific

Open Science at the SRCs

- SKA data exploitation through a platform that **facilitates:**
  - **collaboration among international teams** in order to extract the maximum scientific knowledge
  - **data sharing + re-use & re-purposing of the analysis tools**
  - the accuracy and the **reproducibility of our scientific methods**


SQUARE KILOMETRE ARRAY

SKA REGIONAL CENTRE REQUIREMENTS

Document Number.......SKA-TEL-SKO-0000735
Document Type.......RSP
Revision.......01
Author.......R. C. Bolton and the SRCCG
Date.......2017-09-27
Document Classification.......UNRESTRICTED
Status.......Released

- (REQ) **Open Access:** enabling users to provide **public links to SKA science data products in their research publications**.
- (REQ) **Reproducibility:** saving the complete **workflow and provenance associated with any ADP***, in such a way that they can be queried, viewed and the associated **workflows can be re-used to create new ADPs.**
- (Goal) **Advanced data product re-generation:** Preserving the software environment associated with the provenance and workflow of an ADP that is required to re-execute the workflow

*ADP = Advanced Data Product

Inputs!

OTHERS

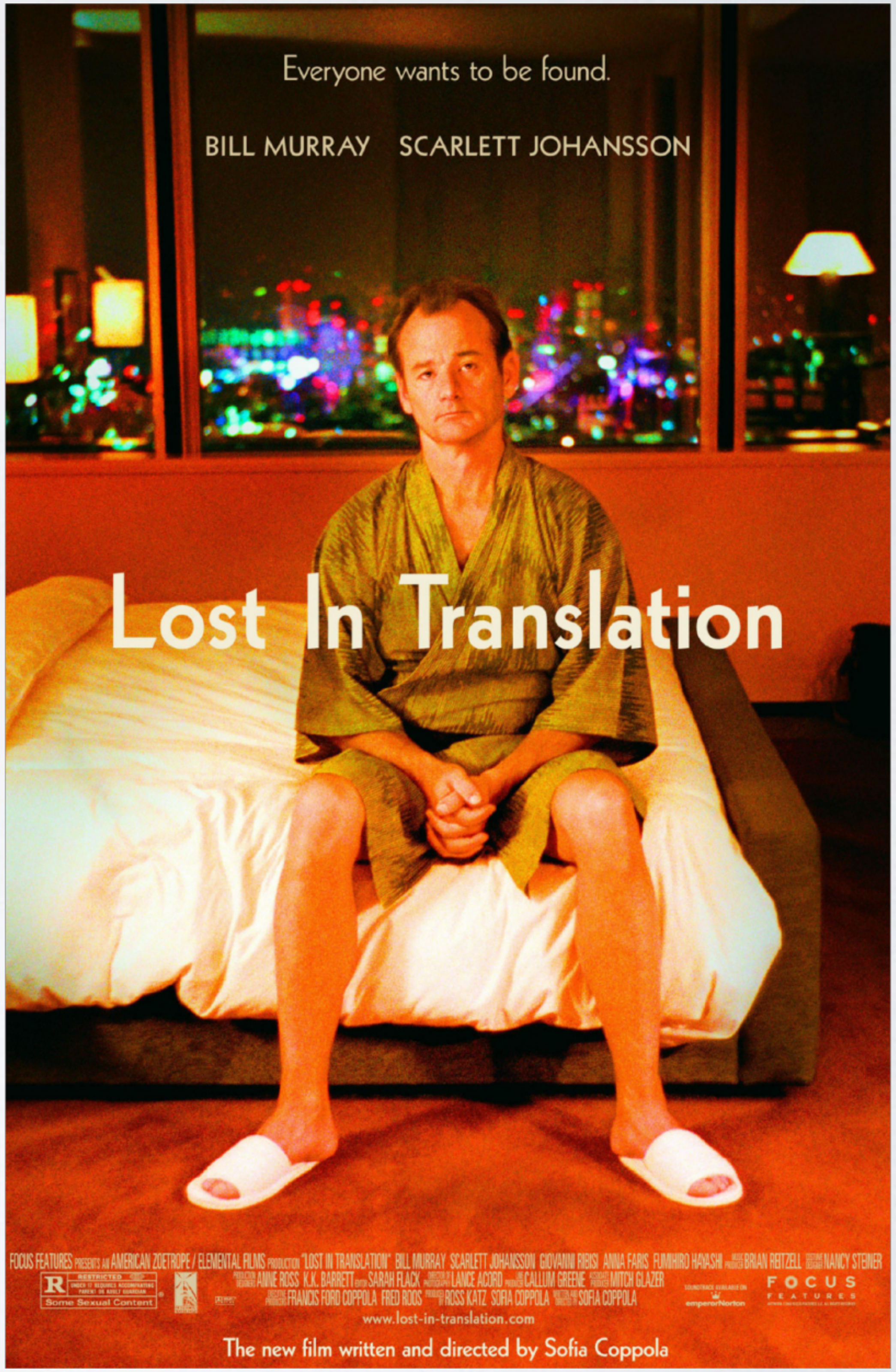
- **Other items/feedback:**
  - Interested in **reuse of SDP pipelines in the SRCs** (e.g imaging with different parameters)
  - Commensality?
  - Will need flexibility to adapt the processing strategy once the precursors/pathfinders are underway
  - Get details on the computing systems already envisaged at this stage
  - Best QA metrics to be using? (Feedback from PHISCC 2018)

(MY) CONCLUSIONS

In order to allow mutual feedback between astronomers/SWGs and SRCs designers there is something we should avoid:

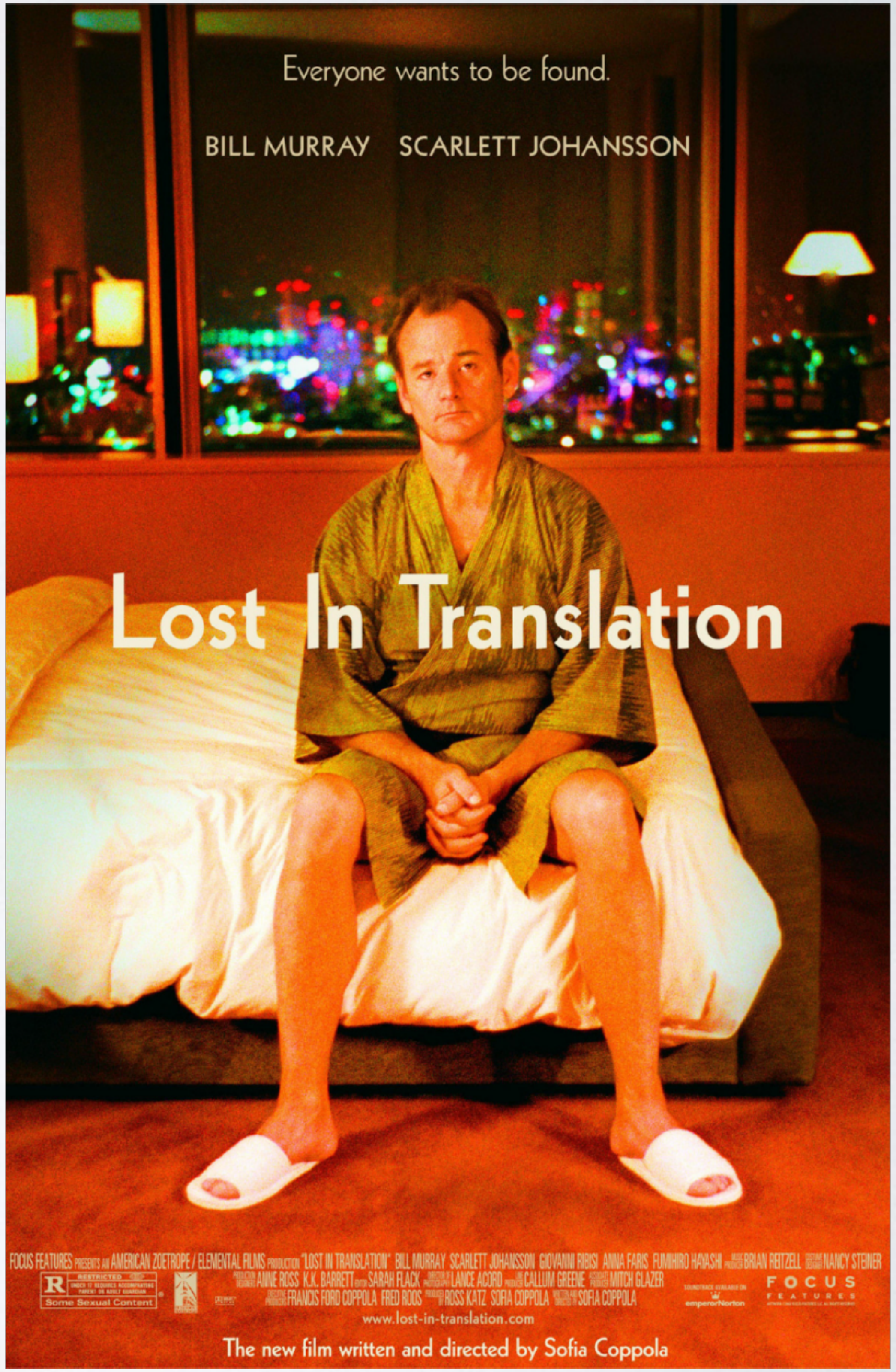
(MY) CONCLUSIONS

In order to allow mutual feedback between astronomers/SWGs and SRCs designers there is something we should avoid:



(MY) CONCLUSIONS

In order to allow mutual feedback between astronomers/SWGs and SRCs designers there is something we should avoid:



Specific session in the next SKA Science meeting?

"New Science enabled by New Technologies in the SKA Era"

SKA HQs @ Jodrell Bank, April 8th - 12th, 2019

- Key Science Projects
- Science Data Processor
- SKA Regional Centres Coordination Group
- Initiatives to prepare for the SRCs
- SKA data challenges
- How do PI projects fit in

WHAT'S NEXT?

Please go back to the slides for further details

Any feedback will be welcome!

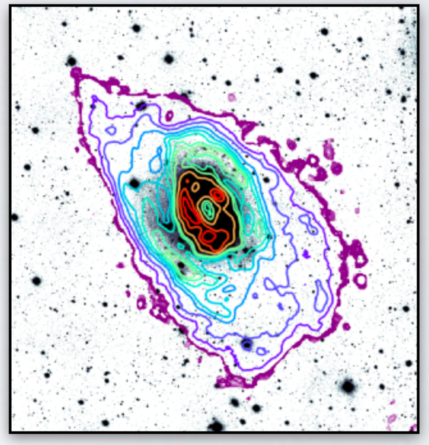
lourdes@iaa.es

SKA
SQUARE KILOMETRE ARRAY
Spain

PostDoctoral Position
IAA-CSIC, Granada, Spain

Lourdes Verdes-Montenegro
lourdes@iaa.es

Origin of asymmetries in isolated galaxies: study of their outskirts with deep optical images and HI interferometric data



- Job conditions
  - To start as soon as possible
  - Duration: till 31/12/2019, with good chances to extend it further.
- Candidates
  - Expertise in reducing deep optical images and/or HI interferometric data is required
- Work environment
  - AMIGA group (Analysis of the interstellar Medium of Isolated GAlaxies, amiga.iaa.es)
  - AMIGA PI: coordinator of the Spanish participation in the SKA & co-Chair of the SKA HI SWG
  - The IAA-CSIC has recently obtained the Center of Excellence Severo Ochoa distinction and AMIGA team leads the development of a prototype of an SKA Regional Centre fully engaged with Open Science