Particles, Universe, NuClei and Hadrons for the NFDI

A consortium in the NFDI.

PUNCH 4NFDI

LOFAR, MEERKAT & SKA OPEN DATA ASPECTS

Michael Kramer & Dominik J. Schwarz

Max-Planck-Institut für Radioastronomie

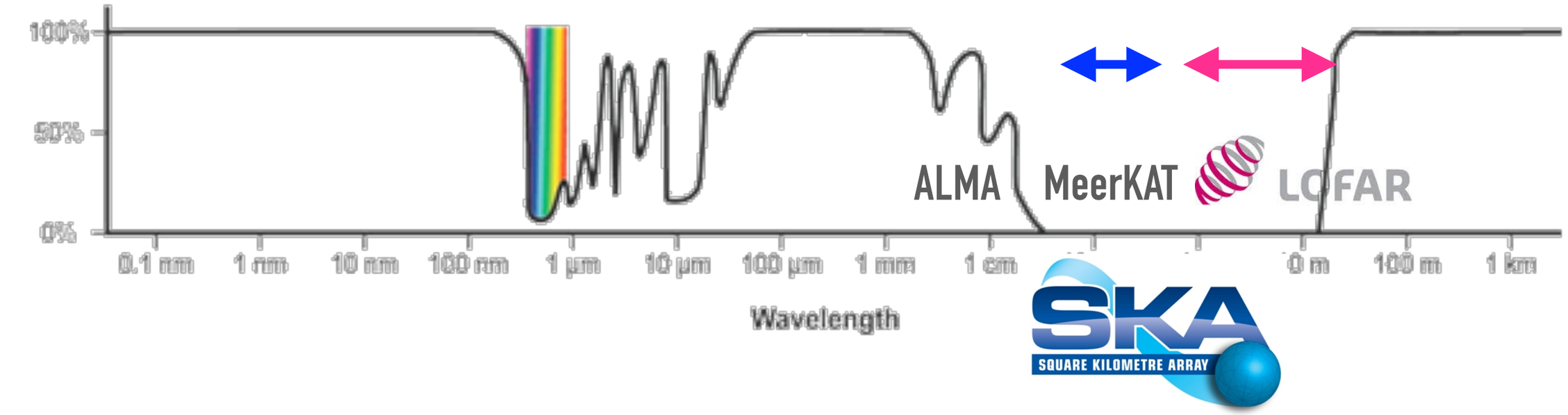
UNIVERSITÄT BIELEFELD

PUNCH4NFDI Open Data Workshop, Feb 2021

RADIO ASTRONOMY

- Study the Universe by means of electromagnetic radiation at frequencies between 10 MHz and 1 THz
- Broad range of science questions:
  from **first stars** and **black holes** to **cradle of life**,
  from **exoplanets** to **largest cosmic structures**,
  from **fast radio bursts** to **gravitational waves**,
  from the **laws of gravity** to **cosmic magnetism**, and **more**

transparency of Earth's atmosphere/ionosphere



LOW FREQUENCY ARRAY

LOFAR

➤ The ***largest radio telescope*** of the world

International LOFAR Telescope ILT: NL, DE, PL, FR, IR, IT, LAT, SW, UK

WHAT IS LOFAR?

LOFAR Superterp, NL

WHAT IS LOFAR?

International LOFAR Station DE609, Norderstedt

UNIVERSITÄT BIELEFELD

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

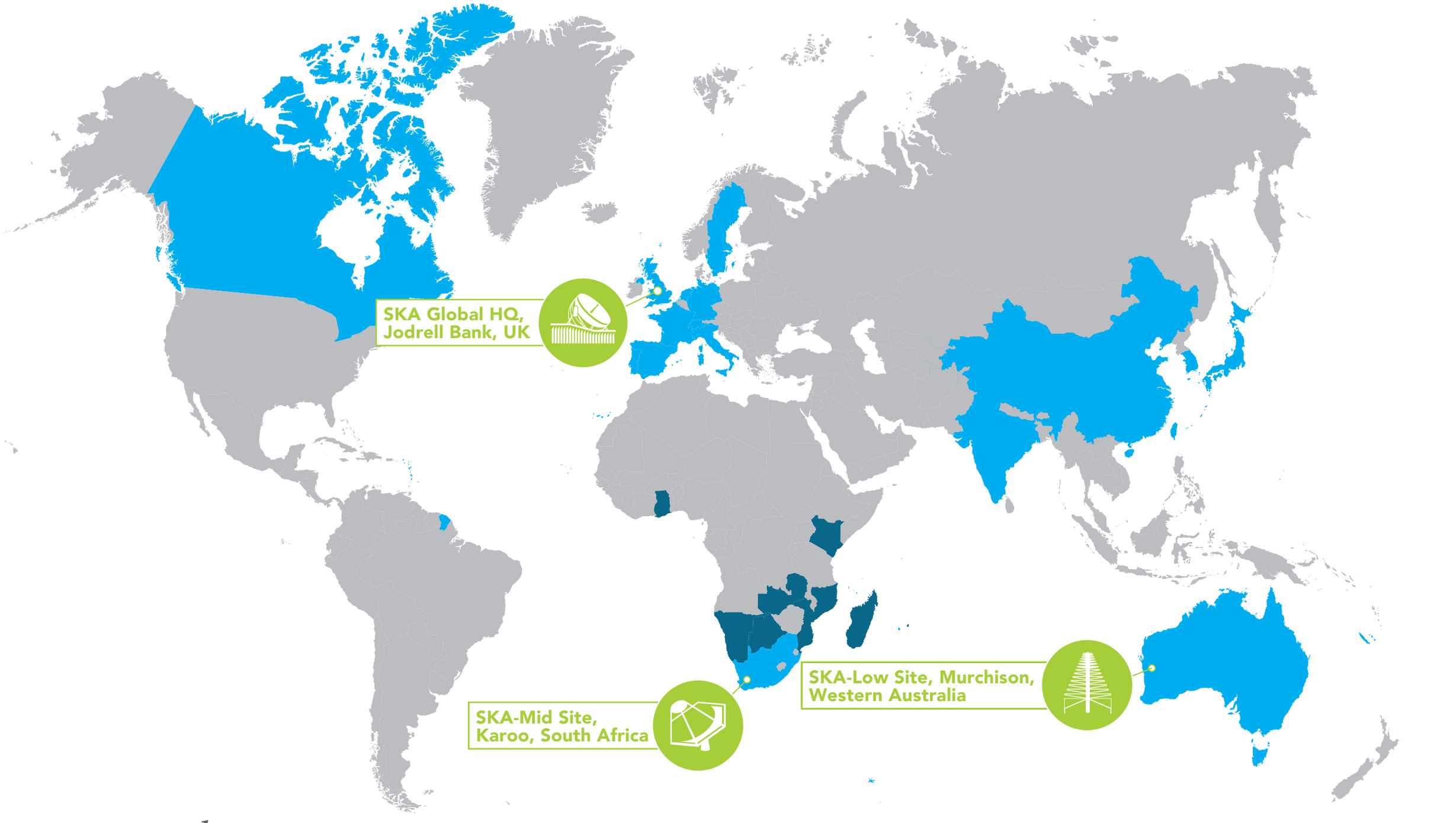
MEERKAT = MEER KAROO ARRAY TELESCOPE

National Research Foundation

SARAO
South African Radio Astronomy Observatory

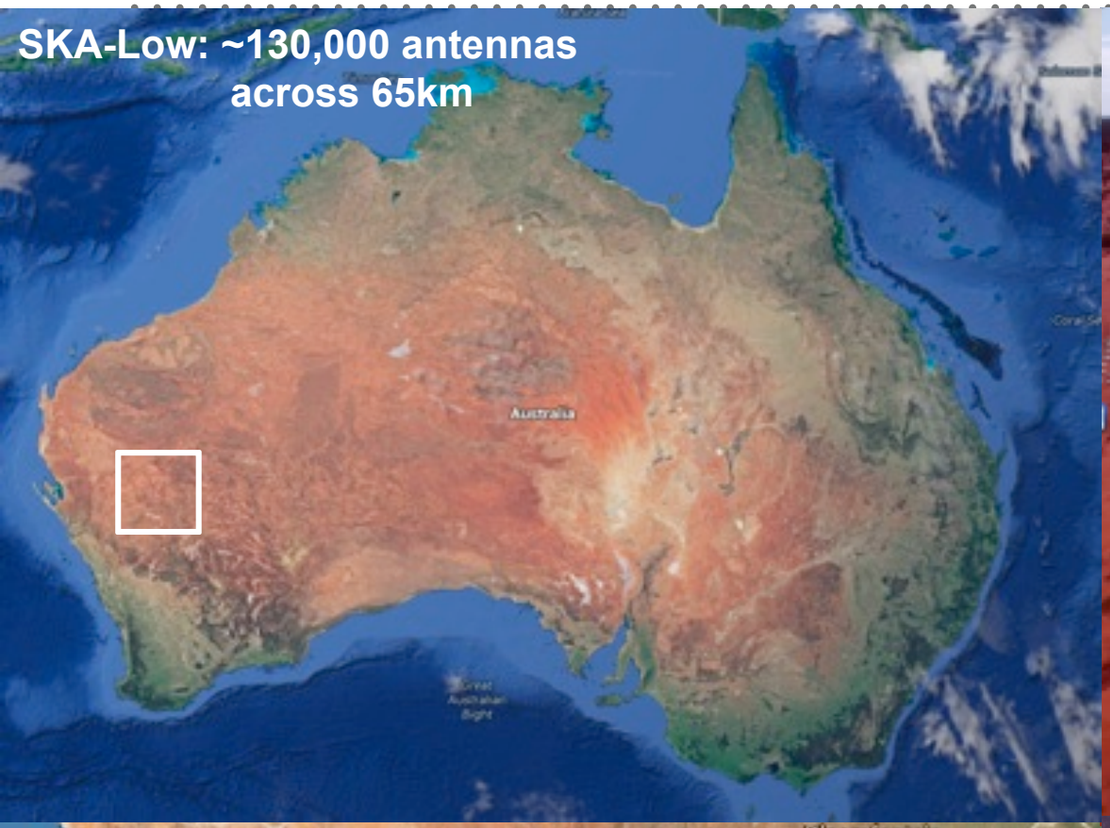
MAX-PLANCK-GESELLSCHAFT

SQUARE KILOMETRE ARRAY (SKA)



SKA Observatory est. 2021

SQUARE KILOMETRE ARRAY (SKA)



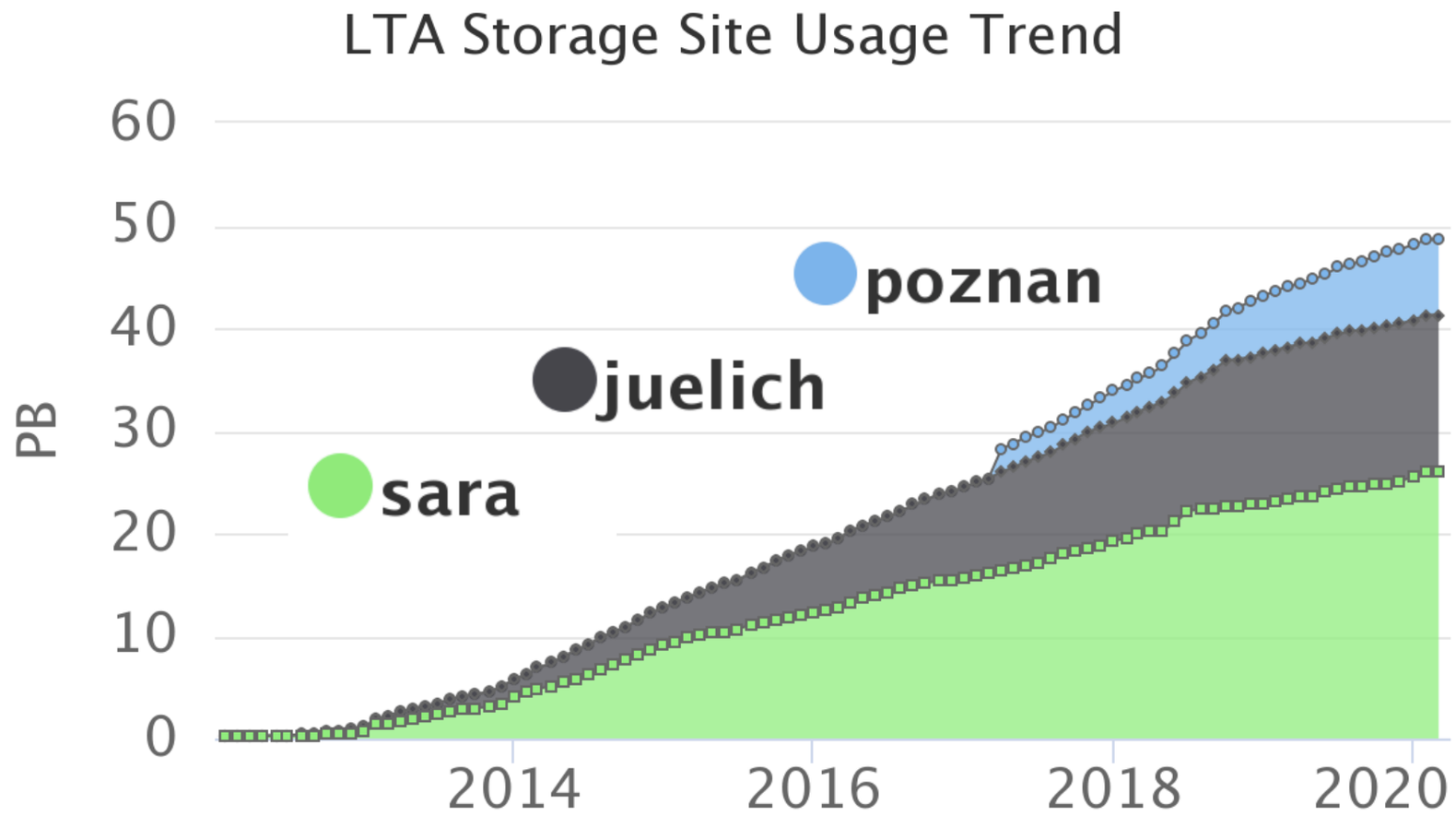
LOFAR RESOURCE ALLOCATION

- LOFAR operates in cycles (1 semester), currently cycle 15
- **Cycle Proposal** and **Long Term Proposal** (4 cycles) calls
- **Proposals are open world wide**
- National consortia have **reserved share of access**
- Proposals include requests for **observing time**, **computing time**, **data volume,** and **user support**
- **Programme committee** (PC) ranks proposals and allocates resources together with observatory
- **ILT board** endorses PC decisions

LOFAR DATA POLICY

- Categories of data: **ILT data** and **station owner data**
- **All ILT data become public**, all ILT observing and standard pipeline processing **data remain ILT property**
- Station owner data (10% of observing time) remain at discretion of station owners
- **Derived data are property of scientists**,
  e.g. LOFAR survey catalogues are released by Key Science Projects
- **Proprietary data period** for proposer is **one year**
- The PC can change data rights and scope for each ILT project
- **All ILT data are stored in the LOFAR Long Term Archive**

LONG TERM ARCHIVE (LTA)



- Total data volume > 50 PB
- Growth 7 PB/a
- LTA mostly tapes
- Requires HPC facilities at LTA site (PUNCH4NFDI TA2 – 4)

LONG TERM ARCHIVE (LTA)

lta.lofar.eu

Lofar Long Term Archive

LOFAR **Long Term Archive**

LOGIN

HOME SEARCH DATA BROWSE PROJECTS HELP

Search

- Basic search
- Advanced search
- Browse projects
  - All projects
  - Commissioning
  - Cycle 0
  - Cycle 1
  - Cycle 10
  - Cycle 11
  - Cycle 12
  - Cycle 13
  - Cycle 14
  - Cycle 15
  - Cycle 2
  - Cycle 3
  - Cycle 4
  - Cycle 5

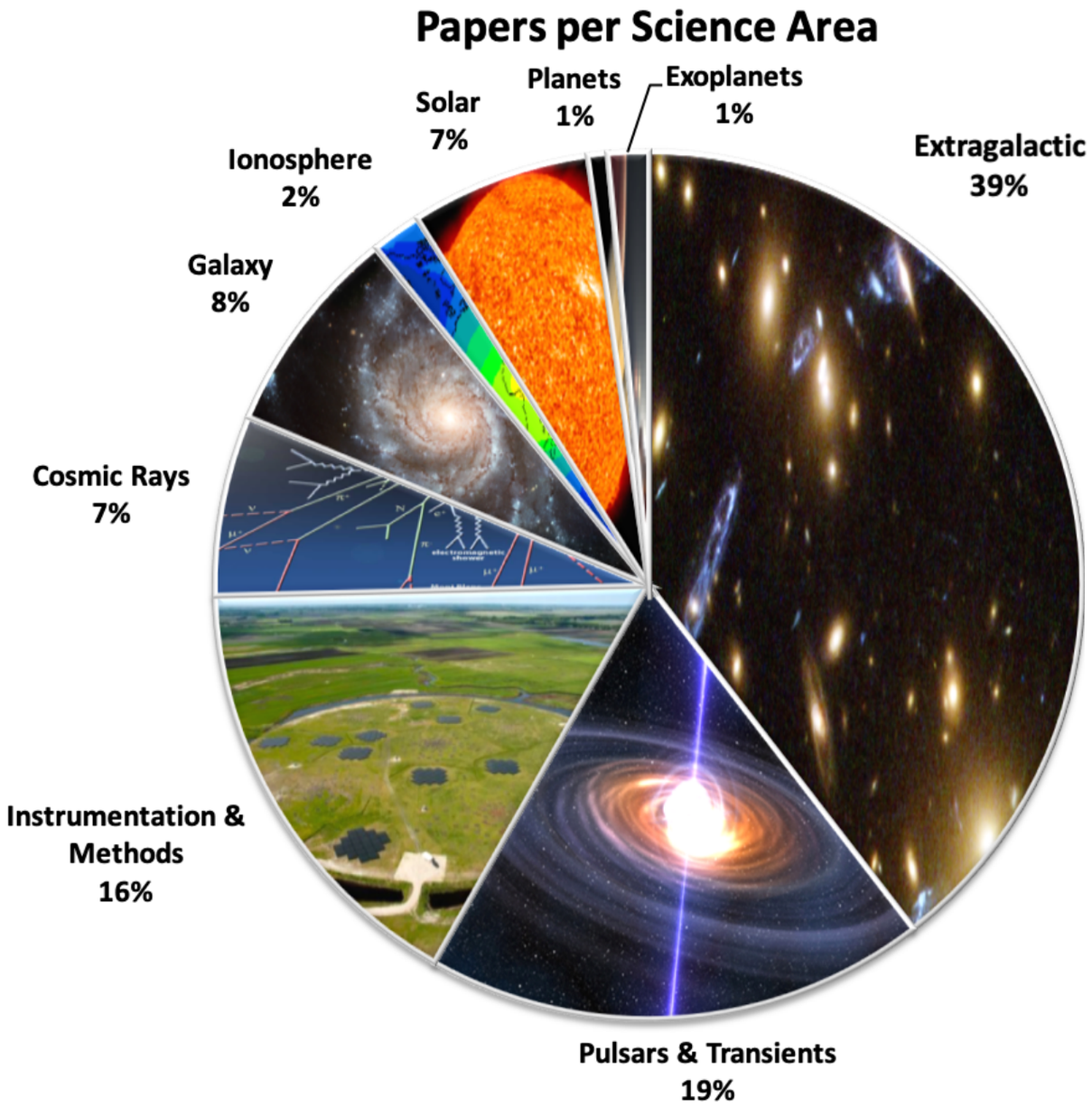
Cycle 11 projects Number of projects: 18

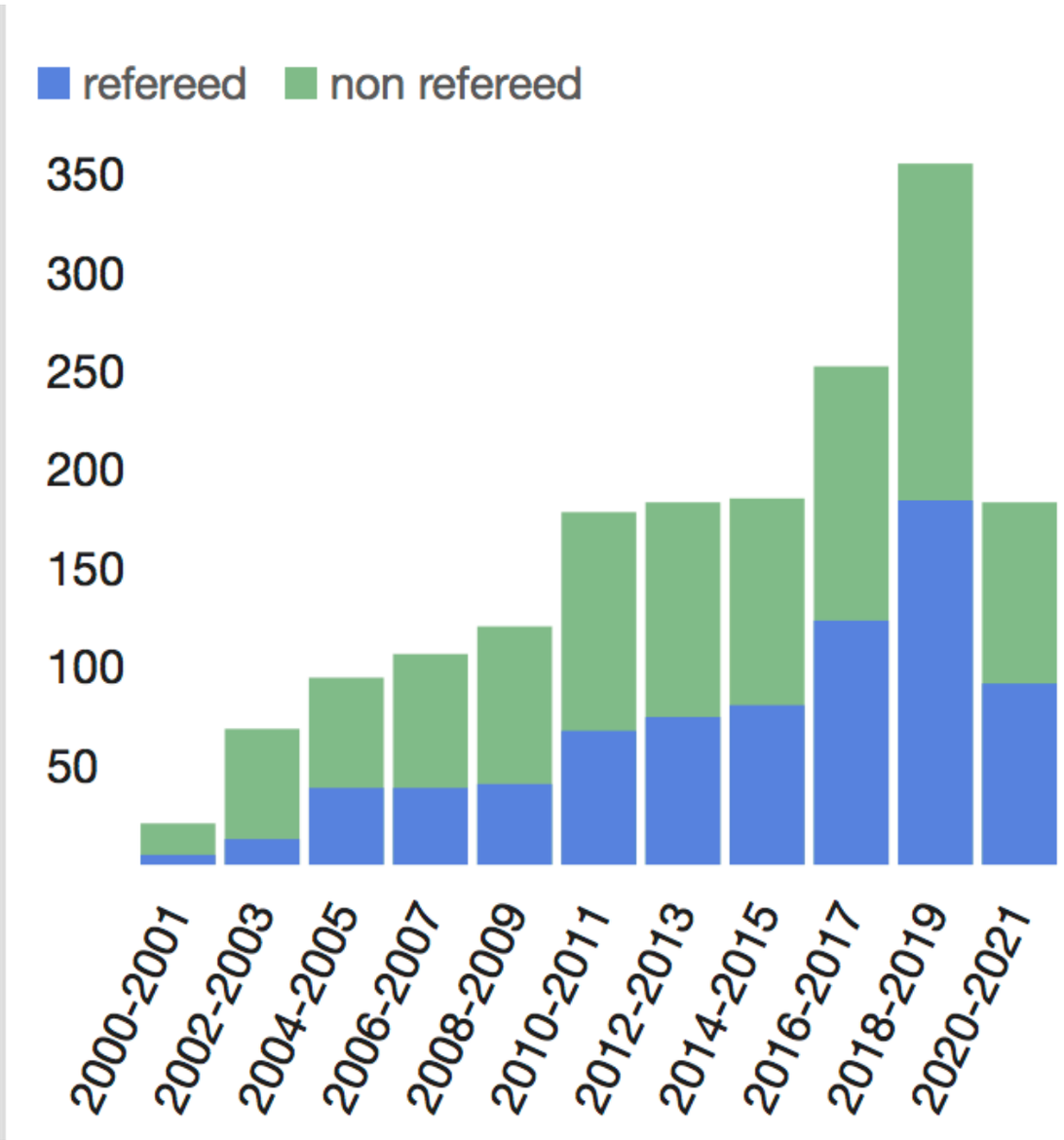
In the table below a project can selected to restrict all data searches to that project only. Use the 'search' button to select the project and go to the search page, use the 'show data' button to select the project and to show all data in it. Alternatively click on the project name to view the project details.

The first column shows a when you are a member of the project or a for public projects.

M	Project	Description	Release date	Actions
	DDT11_001	Characterization and localization of a new low-frequency repeating Fast Radio Burst source	2018-12-05	set search show data
	LC11_001	Interferometric imaging and tied-array beam observations of the SUN and Coronal Transients with the aid of pulsars.	2020-08-22	set search show data
	LC11_002	Rapid follow-up of Gamma-Ray Bursts	2020-08-06	set search show data
	LC11_004	Non-thermal shocks from the jets of massive YSOs	2020-01-08	set search show data
	LC11_005	LOFAR observation of the radio lighthouse: CU Vir	2020-04-11	set search show data
	LC11_007	Continuing 3d-VLBI of scattering-induced echoes in B1508+55	2020-05-18	set search show data
	LC11_009	Searching for Ultra Steep Spectrum Radio Halos in low-mass clusters of galaxies	2020-05-30	set search show data
	LC11_010	The LOFAR eRosita survey of the Virgo cluster	2020-04-11	set search show data
	LC11_011	Spectral long-baseline modeling of 3C 196	2020-01-29	set search show data
	LC11_013	A LOFAR HBA survey of SINGS galaxies	2020-03-22	set search show data
	LC11_014	Revealing the origin of the diffuse radio emission in 2A0335+096	2020-04-11	set search show data
	LC11_015	Towards a First Sample of Revived Fossil Plasma Sources in Galaxy Clusters	2020-10-18	set search show data

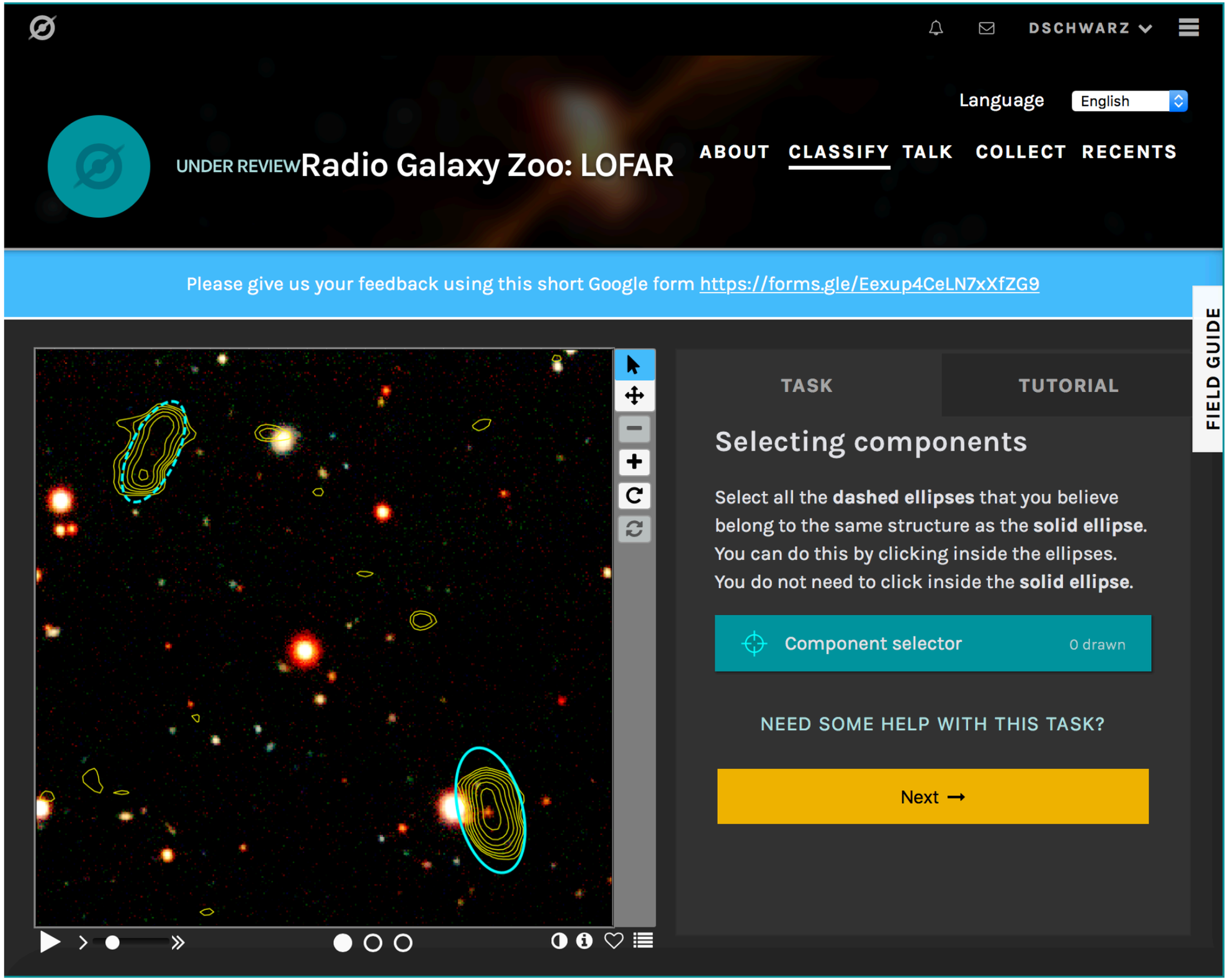
LOFAR — PUBLICATIONS





1800 publications from or on LOFAR (ADS, Feb 2021)

CITIZEN SCIENCE: RADIO GALAXY ZOO



MEERKAT

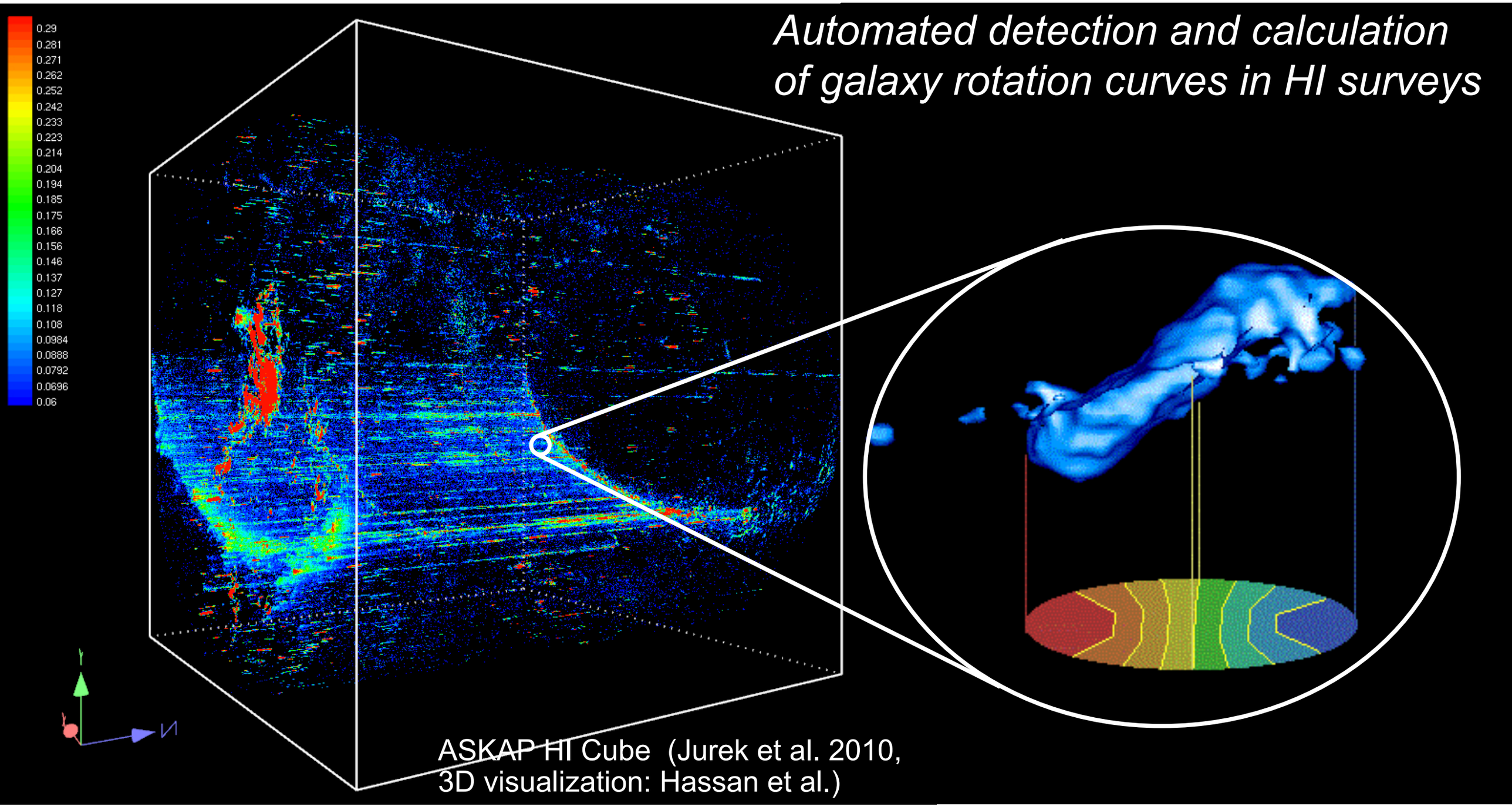
- **Large Science Projects** (co-)led by SA PIs,
  German participation based on individual or institutional contribution in many of them
- Neither access nor data are open
- MeerKAT offers **small fraction of Open Time**:
  open to world wide community (so far only two calls)
- No archive like LOFAR LTA
- Data are moved to compute centre in Cape Town, from where the Projects have to retrieve it within a certain time frame
- Huge data volumes, if low level data are needed,
  shipping discs is often still the best option

SQUARE KILOMETRE ARRAY

- Main data challenges
  - **Data irreversibility** (online reduction in near-realtime) [PUNCH4NFDI TA5]
    - raw data: ~1 Petabit per sec
      - archive: ~ 300 Petabyte per yr
    - **Data monsters:** up to ~1 Petabyte per single 3D image
- SKA Regional Centre Steering Committee
- Open Data in the classical sense will be impossible
- Open Data must be restricted to higher level data products

SKA „DATA MONSTERS“

Automated detection and calculation of galaxy rotation curves in HI surveys



ASKAP HI Cube (Jurek et al. 2010,
3D visualization: Hassan et al.)

3D images: up to 1 Petabyte per cube

- *Outstanding resolution: 36.600 x 36.600 x 250.000 voxel*

CONCLUSIONS

- Open Data is not the world wide standard in radio astronomy
- Openness can be in conflict with resource allocation, granting rules or missions of organisations
- Commitment is typically rewarded by privileged access to data
- Various levels of openness are realised
- Open Data does not automatically imply FAIR Data
- Access and reproducibility are challenging due to enormous data rates and volumes (PUNCH4NFDI TA2-4)
- Problem of data irreversibility (PUNCH4NFDI TA5)