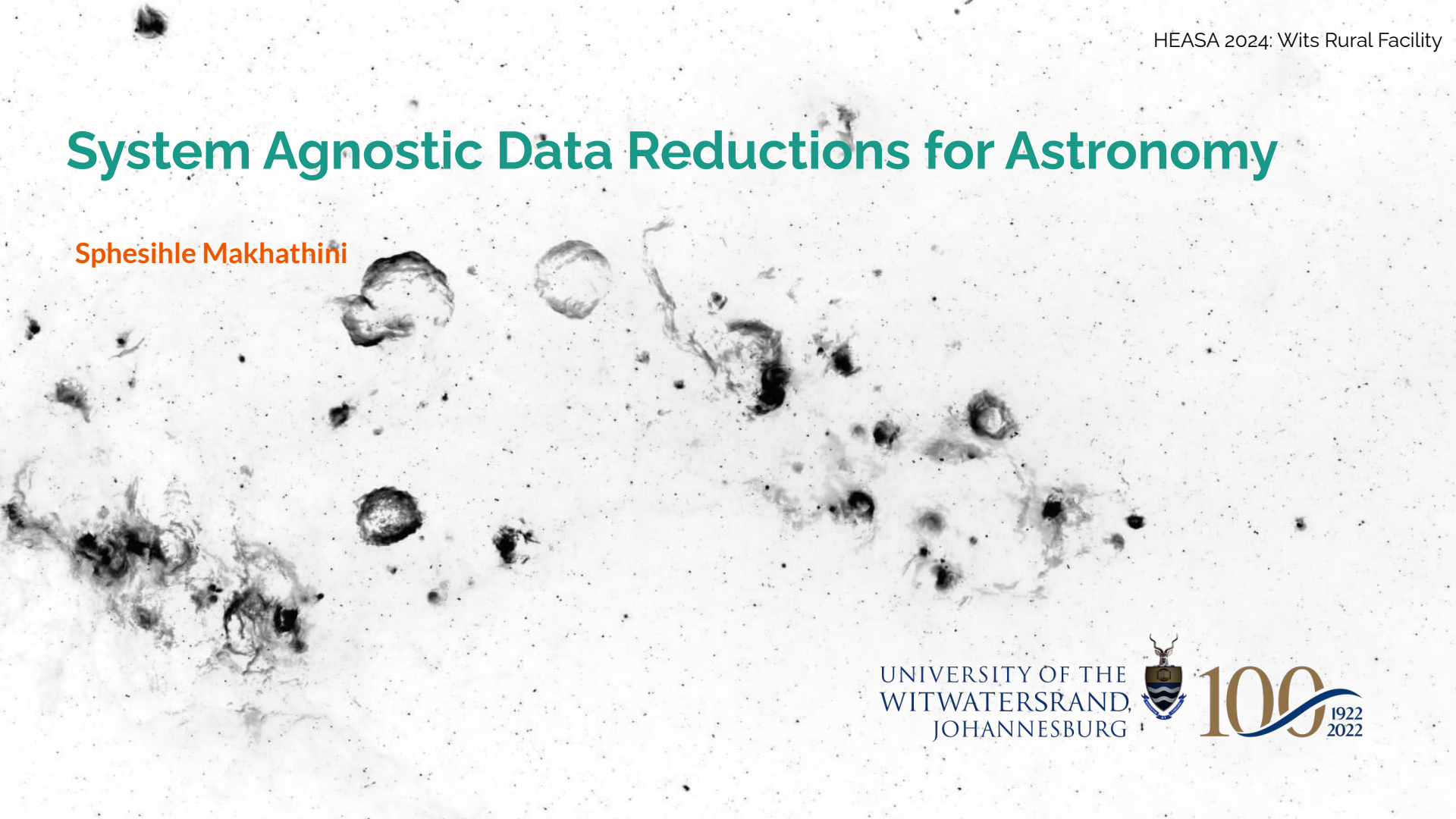
System Agnostic Data Reductions for Astronomy

Sphesihle Makhathini



UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Astronomy Software

01 | Often written by astronomers, not software developers. Unnecessarily complex and opaque

02 | Uses non-standard libraries. Difficult to build and/or install

03 | Difficult to install multiple packages on the same system

04 | No cohesive development
- Libraries
- User interface
- Configuration files

05 | The result is clunky and unreliable software, scripts and pipelines that produce unreliable scientific products. More importantly, it's difficult to check whether some of the results are reproducible.

How can we fix this?

2 - 5

Result from 1

1

Not likely to change.

2-4

Can be fixed for pipelines through **stimela**.

stimela

Python-based (now Yaml) based framework for writing data reduction pipelines.

https://stimela.readthedocs.io/

Nomenclature

Cab

A self contained unit that can perform a **well defined** task on some input and produce some output

Step

An instance of a cab or recipe.

Recipe

A set of steps that perform a **well defined** task.

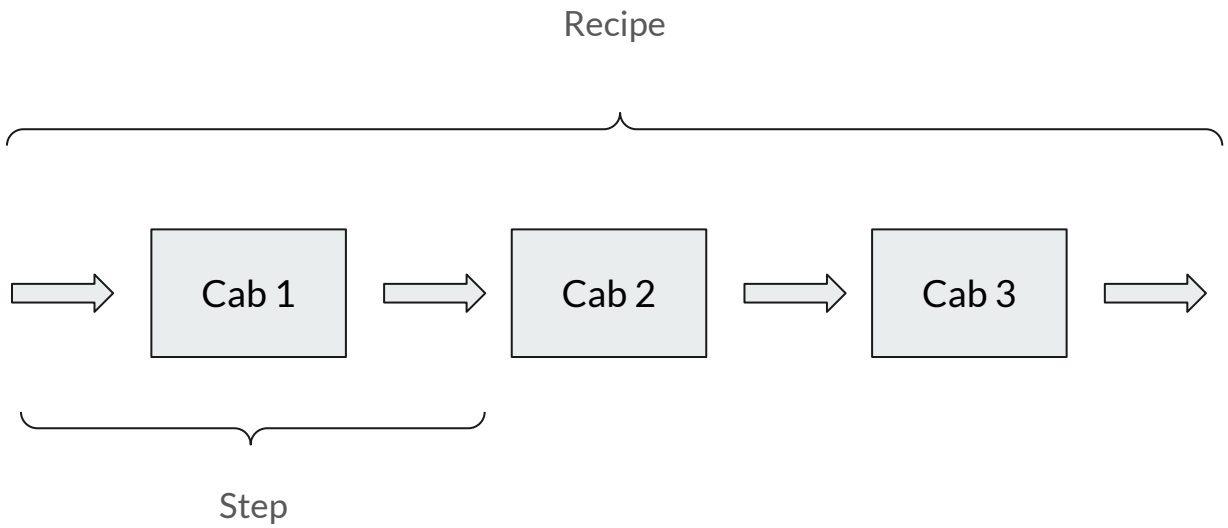
Well defined

Produces a **predictable and reproducible** output given the same set of inputs.

Containerization

OS–level virtualization that allows applications to run in isolated user spaces called **containers (or images)** in any cloud or non-cloud environment, regardless of type or vendor.

Cabs, Steps and Recipes



Anatomy of a Cab

```
4   cabs:
5     fitstool.stack-freq-cube:
6       info: Uses Owlcat fitstool.py (https://github.com/ratt-ru/owlcat) to stack a frequency cube
7       command: fitstool.py
8       image:
9         _use: vars.cult-cargo.images
10        name: python-astro
11      inputs:
12        images:
13          dtype: List[File]
14          required: true
15          policies:
16            positional: true
17            repeat: repeat
18      outputs:
19        cube:
20          dtype: File
21          policies:
22            positional: true
23            format: '--stack={0}:FREQ'
```

Input definition

- Operating System (can be host)
- Libraries
- Application or script

Output definition

Anatomy of a Recipe

```yaml
calibration-recipe:
    info: "a notional recipe for calibration & imaging"

    # this recipe has some input parameters
    inputs:
        ms:
            dtype: MS
            required: true
            info: "measurement set to use"
        image-name:
            dtype: str
            required: true
            info: "base name for output images"
        image-size:
            dtype: int
            default: 4096
            info: "image size, in pixels"
```

```yaml
steps:
    image:
        info: "make initial image and model from DATA column"
        # this is the unerlying tool that the step invokes (defined in my
        cab: imager-tool
        # and these are the parameters of the step...
        params:
            ms: =recipe.ms
            mode: image
            size: =recipe.image-size * 2
            column: DATA
            output.image: '{recipe.image-name}.image-{info.suffix}-{curre
            output.model: '{recipe.image-name}.model-{info.suffix}.fits'
    predict:
        info: "predict model into MODEL_DATA"
        cab: imager-tool
        params:
            ms: =recipe.ms
            mode: predict
            model: =previous.output.model
            column: MODEL_DATA
```

CARACal

Containerized Automated Radio Astronomy Calibration

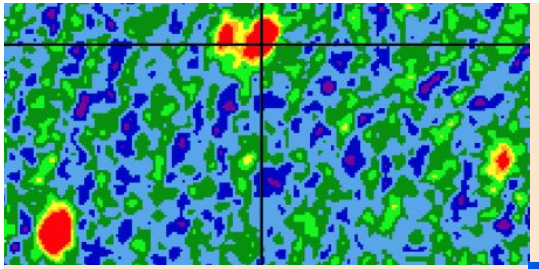
https://caracal.readthedocs.io/

- Radio Interferometry (RI) data reduction pipeline
  - Not just for MeerKAT
- Continuum + line data
- Simple to install, and use
- Configurable
- Stimela-based
  - Portable
  - System agnostic
  - Reproducible results

About CARACal

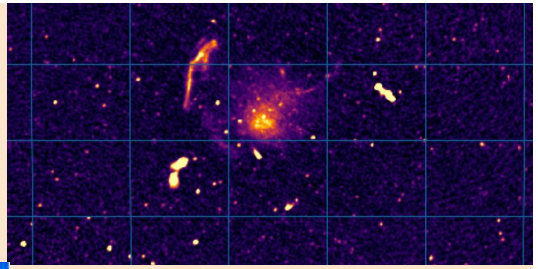
- End-to-End data reduction pipeline
- Well defined set of stimela recipes (called workers)
- Each worker performs a task along the process chain
- Only requires Python and a containerization tool (Singularity, Podman, Docker)
    - No need to install any other software
    - The software are pulled from an online container/image repository and come fully equipped with required libraries and applications
- Allows astronomer to focus on optimizing data reduction settings instead of spending hours setting up the perfect software environment.
- Portable and System Agnostic
    - Share reductions with collaborators via Yaml files. No bash scripts, no Make files
    - Run CARACal with **same configuration file on same data,** and get the **same results** regardless of host system.

Lowers barrier to entry for students



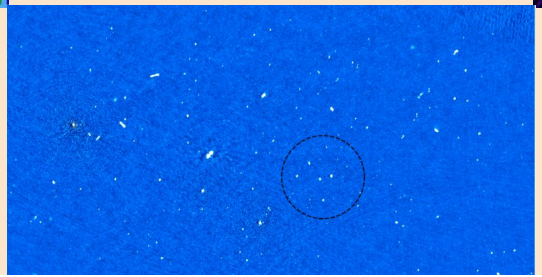
Sinenhlanhla Sikhosana (UKZN, PhD, IP)

Diffuse radio emission in ACTPol clusters (UKZN, PHD,



Julia Healy (Kapteyn/UCT, PhD, IP)

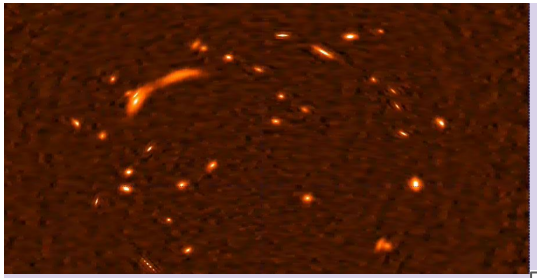
Galaxies in cluster substructures (Norma and Abel2626)



Shilpa Ranchord (UP, MSc, IP)

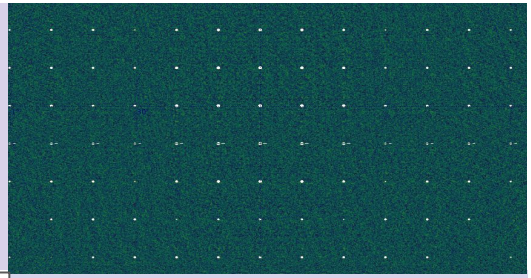
Strong gravitational lensing of HI with galaxy clusters: a deep MeerKAT search

Lowers barrier to entry for students



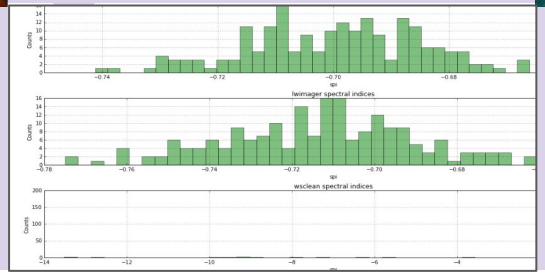
Athanaseus Ramalia (RU MSc, 2019))

Quantifying imaging performance of legacy and novel Imaging software packages



Alessia Verz (INAF, PhD, 2018)

SKA strong gravitational lensing simulations

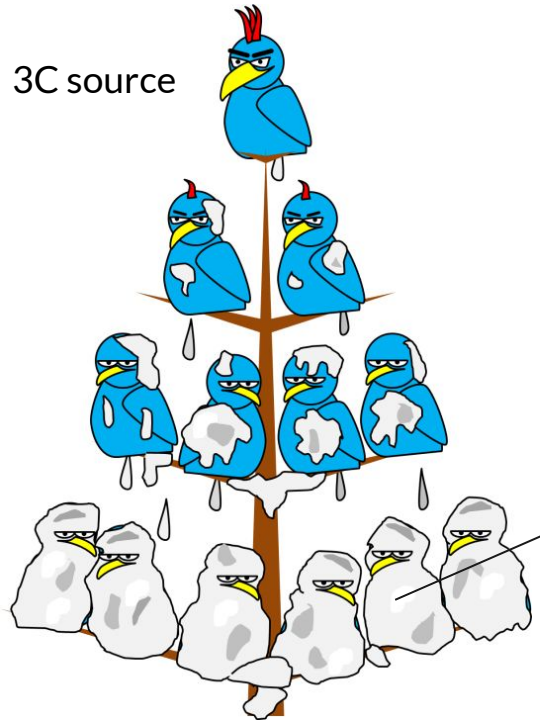


Kelachukwu Iheanetu(RU, PhD, 2020)

Effects of the Primary Beam on image fidelity

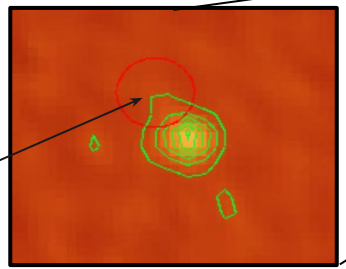
Humbled by 3C283

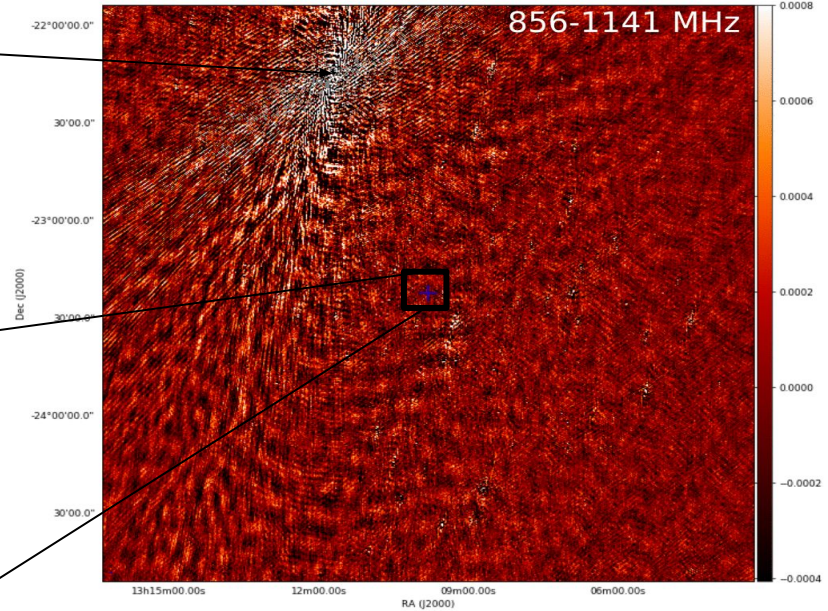
3C source



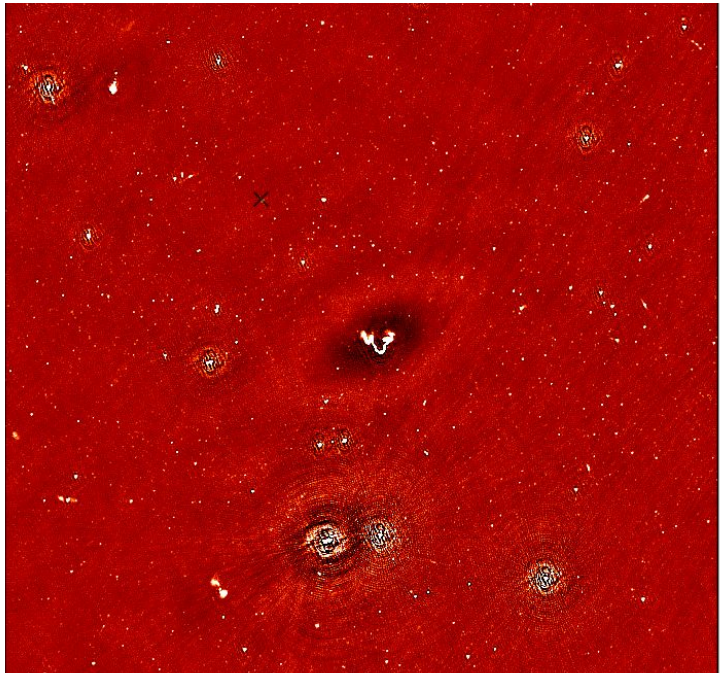
3C283
5.2 Jy @ 1.4 GHz

Very bright source that falls into null of PB as a function of frequency!!

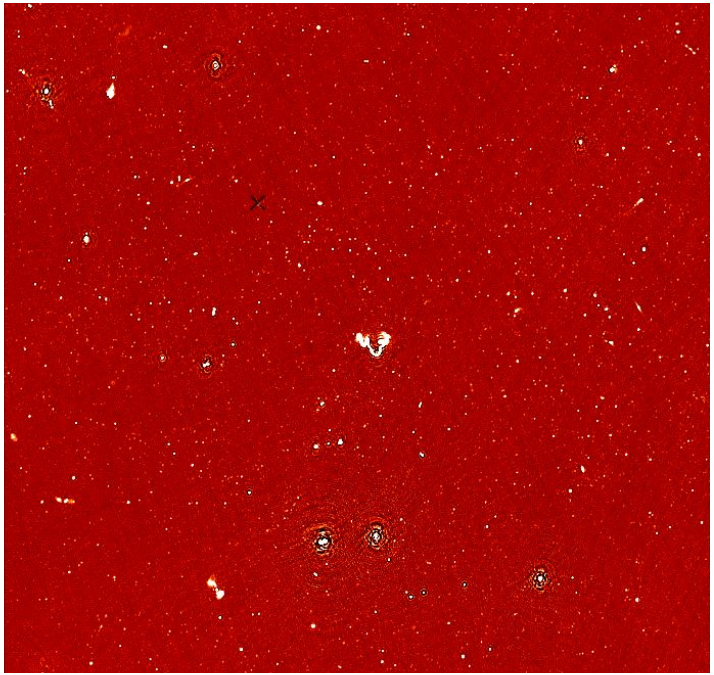




Flexible (write own workers)



Implemented in Cubical (Kenyon et al 2019)



$$\mathbf{V}_{pq} = \mathbf{G}_p \left(\mathbf{dE}_p \mathbf{E}_p \mathbf{X}^{(1)}_{pq} \mathbf{E}^H_q \mathbf{dE}^H_q + \mathbf{X}^{(2)}_{pq} + \cdots \right) \mathbf{G}^H_q$$

GW170817 afterglow

The Astrophysical Journal Letters, 848:L12 (59pp), 2017 October 20 https://doi.org/10.3847/2041-8213/aa91c9

OPEN ACCESS

Multi-messenger Observations of a Binary Neutron Star Merger*

LIGO Scientific Collaboration and Virgo Collaboration, Fermi GBM, INTEGRAL, IceCube Collaboration, AstroSat Cadmium Zinc Telluride Imager Team, IPN Collaboration, The Insight-HXMT Collaboration, ANTARES Collaboration, The Swift Collaboration, AGILE Team, The 1M2H Team, The Dark Energy Camera GW-EM Collaboration and the DES Collaboration, The DLT40 Collaboration, GRAWITA: GRAvitational Wave Inaf TeAm, The Fermi Large Area Telescope Collaboration, ATCA: Australia Telescope Compact Array, ASKAP: Australian SKA Pathfinder, Las Cumbres Observatory Group, OzGrav, DWF (Deeper, Wider, Faster Program), AST3, and CAASTRO Collaborations, The VINROUGE Collaboration, MASTER Collaboration, J-GEM, GROWTH, JAGWAR, Caltech-NRAO, TTU-NRAO, and NuSTAR Collaborations, Pan-STARRS, The MAXI Team, TZAC Consortium, KU Collaboration, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS: Transient Robotic Observatory of the South Collaboration, The BOOTES Collaboration, MWA: Murchison Widefield Array, The CALET Collaboration, IKI-GW Follow-up Collaboration, H.E.S.S. Collaboration, LOFAR Collaboration, LWA: Long Wavelength Array, HAWC Collaboration, The Pierre Auger Collaboration, ALMA Collaboration, Euro VLBI Team, Pi of the Sky Collaboration, The Chandra Team at McGill University, DFN: Desert Fireball Network, ATLAS, High Time Resolution Universe Survey, RIMAS and RATIR, and SKA South Africa/MeerKAT

The Astrophysical Journal Letters, 868:L11 (8pp), 2018 November 20 https://doi.org/10.3847/2041-8213/aaeda7

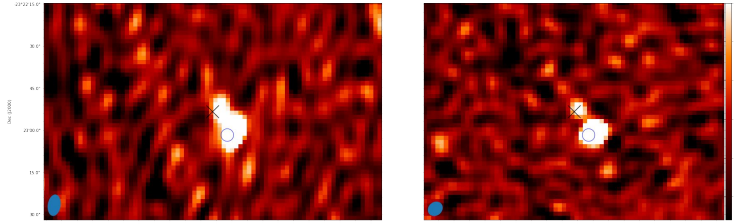
A Strong Jet Signature in the Late-time Light Curve of GW170817

K. P. Mooley[1,2,14], D. A. Frail[1], D. Dobie[3,4], E. Lenc[4], A. Corsi[5], K. De[2], A. J. Nayana[6], S. Makhathini[7,8], I. Heywood[8,9], T. Murphy[3], D. L. Kaplan[10], P. Chandra[6], O. Smirnov[7,8], E. Nakar[11], G. Hallinan[2], F. Camilo[7], R. Fender[9,12], S. Goedhart[7], P. Groot[12,13], M. M. Kasliwal[2], S. R. Kulkarni[2], and P. A. Woudt[12]
[1] National Radio Astronomy Observatory, Socorro, NM 87801, USA; kunal@astro.caltech.edu
[2] Caltech, 1200 E. California Blvd. MC 249-17, Pasadena, CA 91125, USA
[3] Sydney Institute for Astronomy, School of Physics, University of Sydney, Sydney, New South Wales 2006, Australia

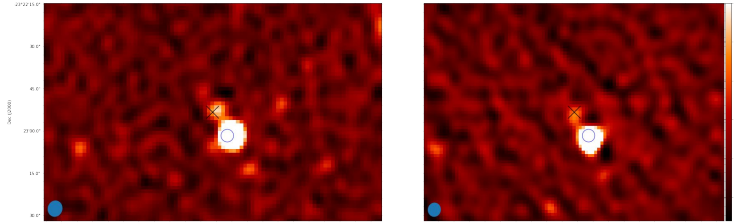
The Panchromatic Afterglow of GW170817: The Full Uniform Data Set, Modeling, Comparison with Previous Results, and Implications

Makhathini, S.; Mooley, K. P.; Brightman, M.; Hotokezaka, K.; Nayana, A. J.; Intema, H. T.; Dobie, D.; Lenc, E.; Perley, D. A.; Fremling, C.; Moldón, J.; Lazzati, D.; Kaplan, D. L.; Balasubramanian, A.; Brown, I. S.; Carbone, D.; Chandra, P.; Corsi, A.; Camilo, F.; Deller, A.; Frail, D. A.; Murphy, T.; Murphy, E. J.; Nakar, E.; Smirnov, O.; Beswick, R. J.; Fender, R.; Hallinan, G.; Heywood, I.; Kasliwal, M.; Lee, B.; Lu, W.; Rana, J.; Perkins, S.; White, S. V.; Józsa, G. I. G.; Hugo, B.; Kamphuis, P.

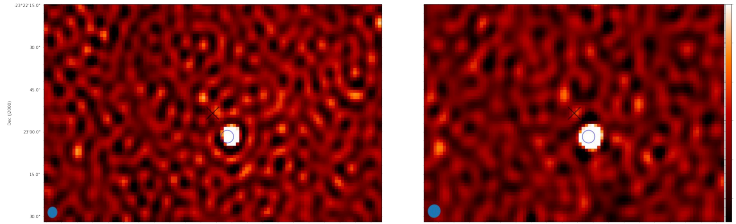
MeerKAT detections



Jan 19, 02 March 2018



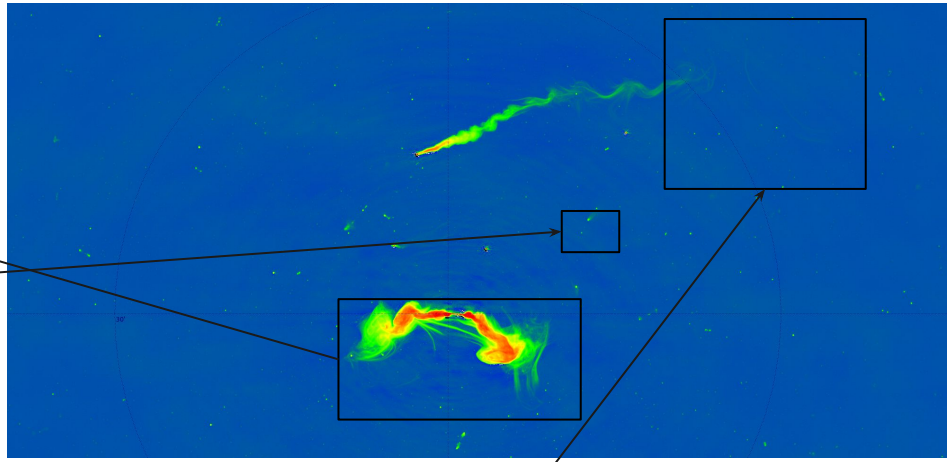
26 April, 05 May 2018



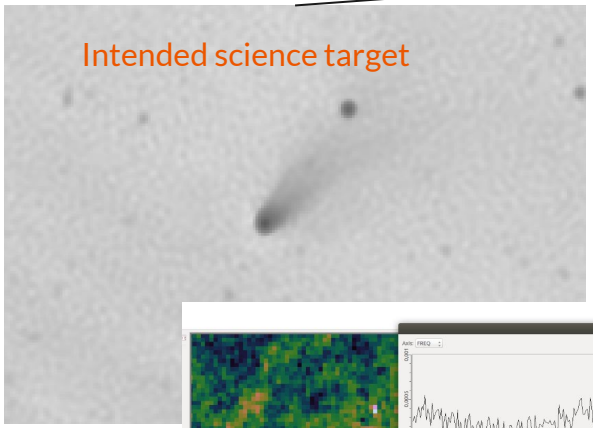
07 July, 01 Sept 2018

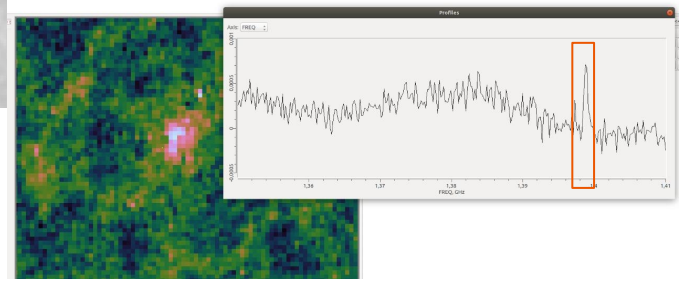
CTN discovery

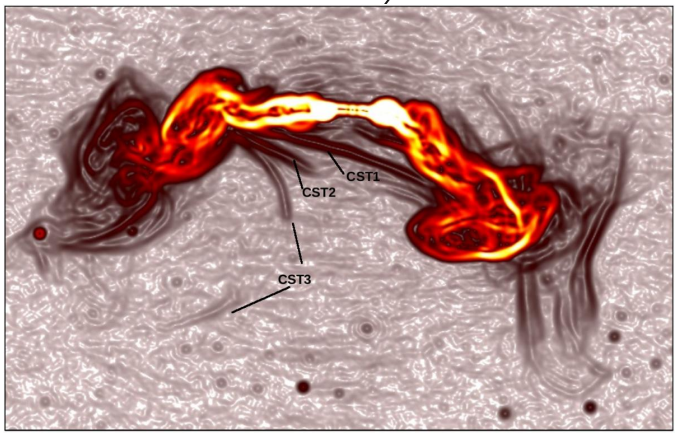
Collimated synchrotron threads linking the radio lobes of ESO 137-006 (Ramatsoku+ 2020)



Intended science target







Thank you.

SARAO/NASA