technische universität dortmund

LAMARR INSTITUTE FOR MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Observational data in astronomy

DOMINIK ELSÄSSER
TU DORTMUND & LAMARR INSTITUTE FOR ML AND AI
DEPUTY SPOKESPERSON OF THE MAGIC COLLABORATION
DOMINIK.ELSAESSER@TU-DORTMUND.DE

NASA, ESA and the Hubble Heritage Team (STScI/AURA)

NASA

Astronomy and data

- Astronomy is the science of celestial objects and the cosmos
- One of the oldest cultural endeavours of humankind

© Landesamt für Denkmalpflege und Archäologie Sachsen-Anhalt, Juraj Lipták.

Astronomy and data

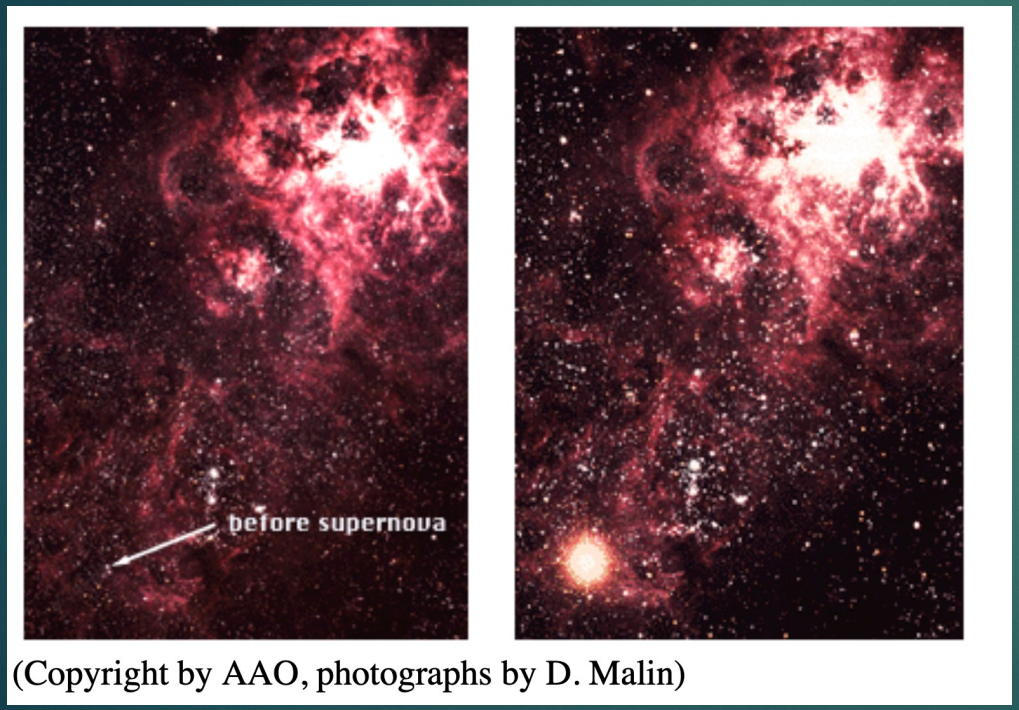
- Astronomy is the science of celestial objects and the cosmos
- One of the oldest cultural endeavours of humankind
- Data: facts or statistical information in the widest sense

Astronomy and data

- Astronomy is the science of celestial objects and the cosmos
- One of the oldest cultural endeavours of humankind
- Data: facts or statistical information in the widest sense

- Astronomers have – over all of human history – collected a treasure of data
- And indeed we even in the 21st century learn from very old „archives“

Progress, one stellar funeral at a time



(Copyright by AAO, photographs by D. Malin)

- Massive stars sometimes do not go gently into the night
- Such events – core-collapse supernovae – are of extreme importance for many fields of astrophysics (e.g. neutrino astronomy)
- Expected to happen in our galaxy about twice per century
- Last certain observation: 1604 (Kepler)

Remnant of Kepler‘s supernova

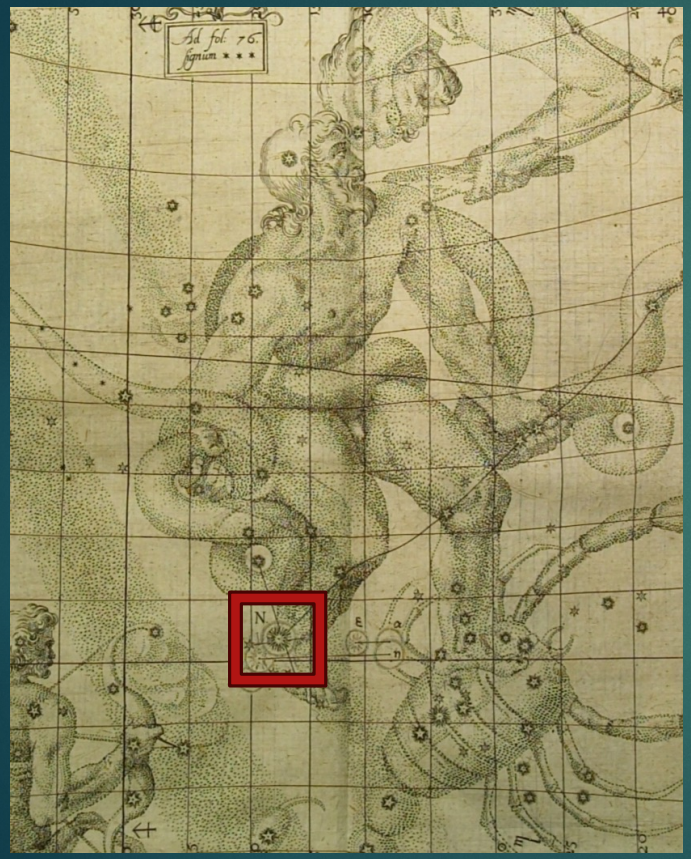
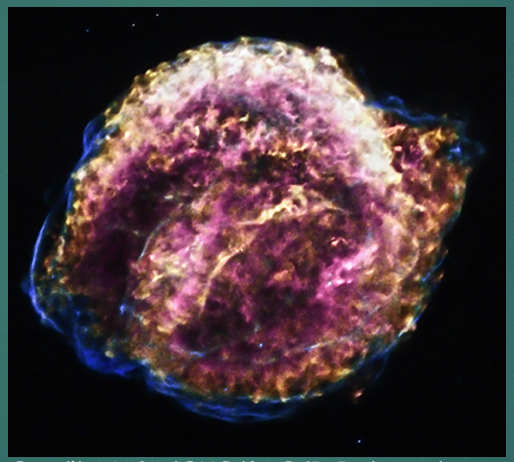


Image credit: Bryce Roberts, from an original print at the Milton S. Eisenhower Library at Johns Hopkins University. SN location marked by DE.



Credit: NASA/CXC/SAO/D.Patnaude

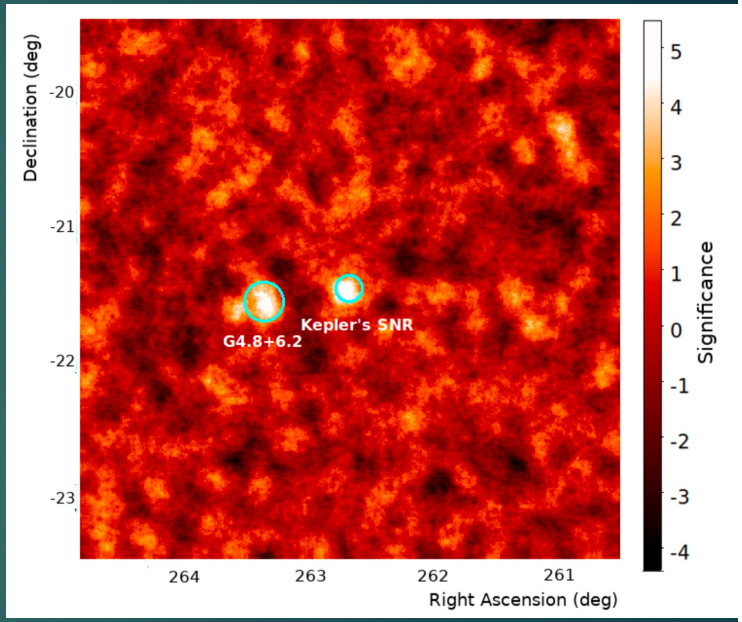


Image credit: H.E.S.S. Collaboration

A potential example dating back to ca. 3600 BC

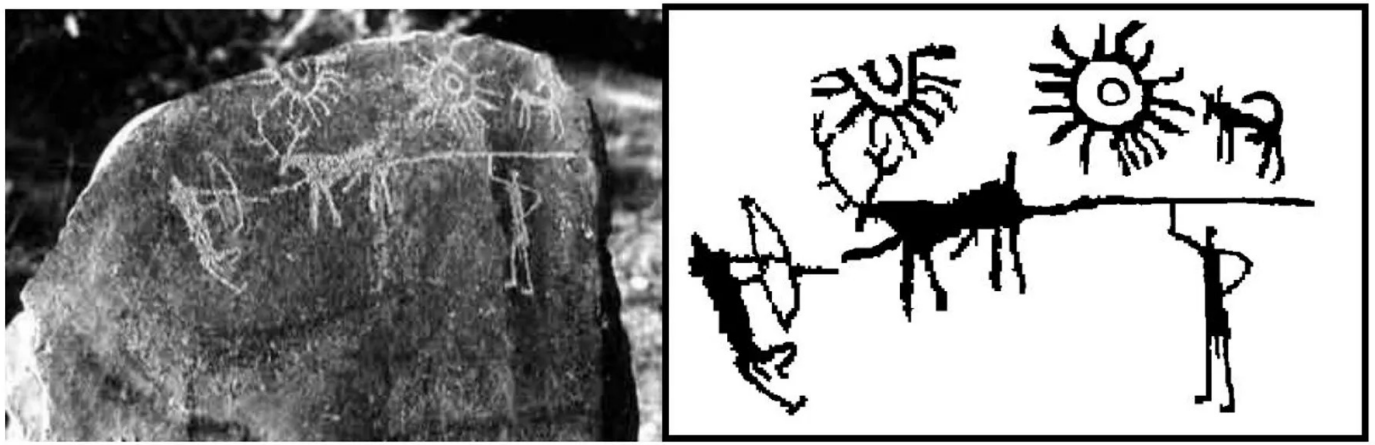


Image courtesy of the Tata of Fundamental Research.Institute

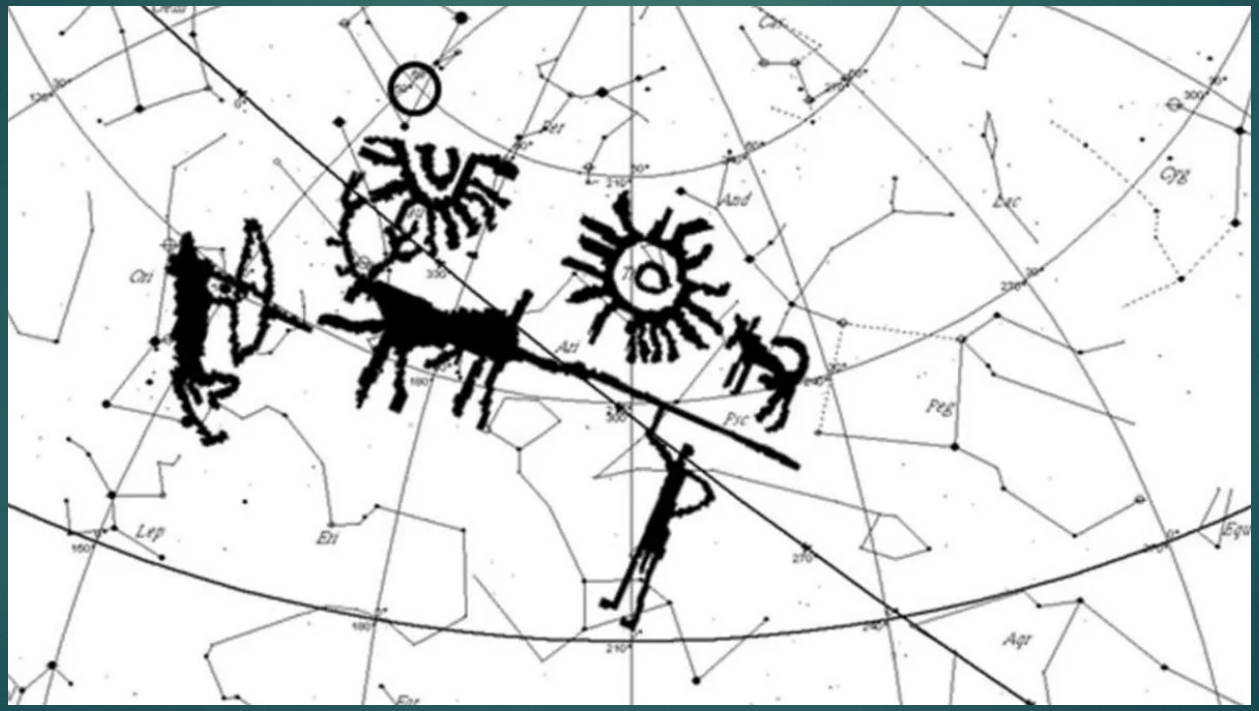
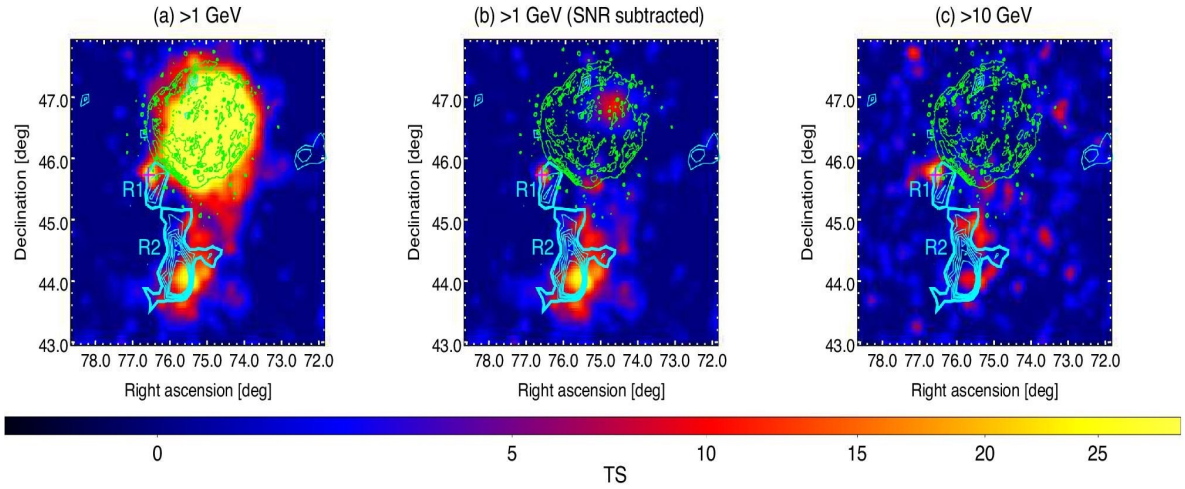


Image courtesy of the Tata of Fundamental Research.Institute

Fig. 1. Gamma-ray TS maps in the vicinity of SNR HB9 observed with Fermi-LAT. All maps are given with square bins of ...



Tomohiko Oka, Wataru Ishizaki

Publ Astron Soc Jpn Nihon Tenmon Gakkai, Volume 74, Issue 3, June 2022, Pages 625–638, https://doi.org/10.1093/pasj/psac024

Astronomy and data

- Astronomy as a science has progressed over the course of history along improvements in instrumentation
- As astronomers, we can (in most cases) not only not do „experiments“ in the sense of „preparing“ our objects
- We sometimes are also reliant on events that – statistically – happen only rarely (...once...much less than once) per lifetime of a given observatory (or observer)
- Some modern telescopes produce data volumes much too large to store the complete stream
- (Almost) uniquely in all of science, we thus have to take decisions on very short timescales, which however will influence the ability of later generations to work with our data (from events which they may themselves never witness)

A personal perspective

- We, with all our modern multi-million (billion) Euro instrumentation and international collaborations, may well be the stone carvers of later days
- (Far-)future generations of scientists may have much more powerful instruments and computing
- Our best theoretical ideas may to them seem utterly trivial, or even naive
- Our observations of rare events and also of seemingly frequent processes (long time series...) however may be of lasting value
- Our task: Keep observational data safe for the long term and ideally in a form that is as independent as possible from theory (e.g. in the form of unfolded energy spectra of gamma-emitting supernova remnants)
- In addition, any reduction of the element of arbitrary human decision taking in the reduction chain is welcome (because the individual reasoning behind that will mostly be lost to future scientists)

Of course this is easy to say…

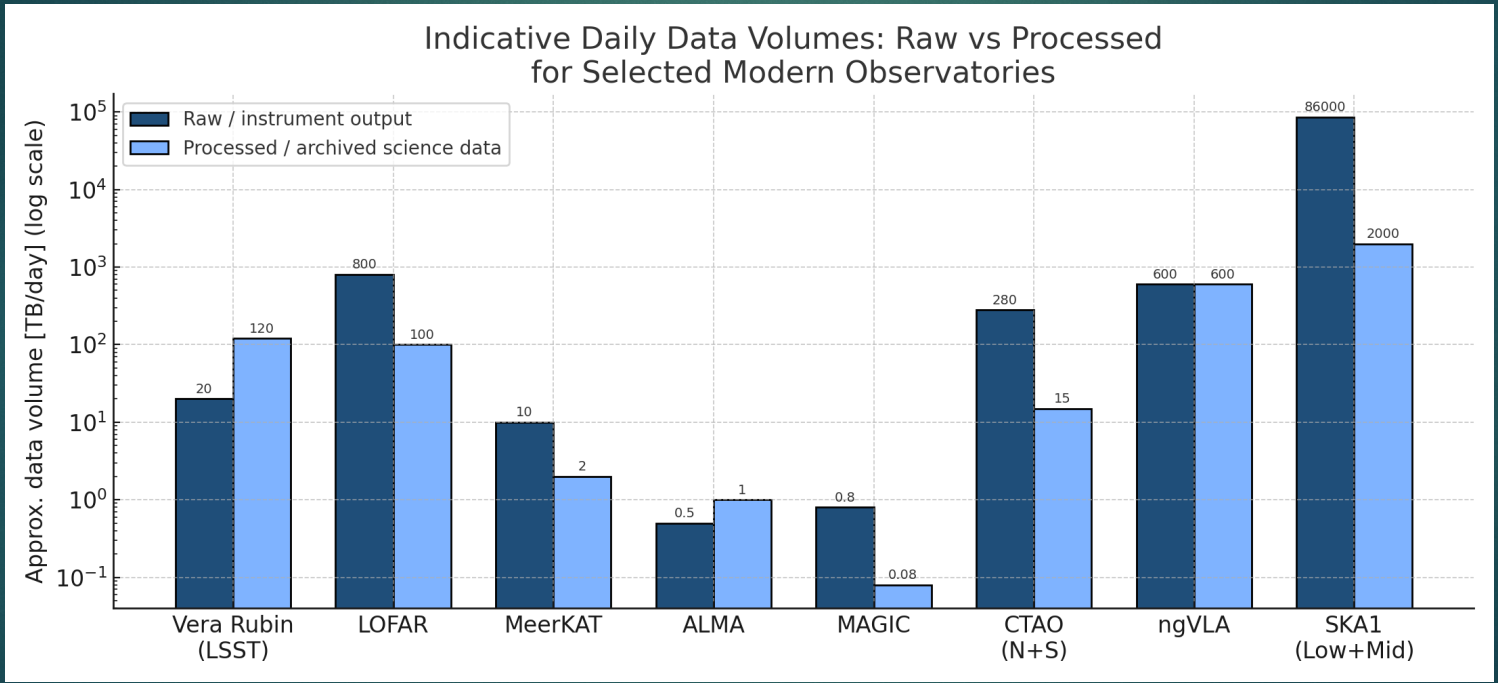
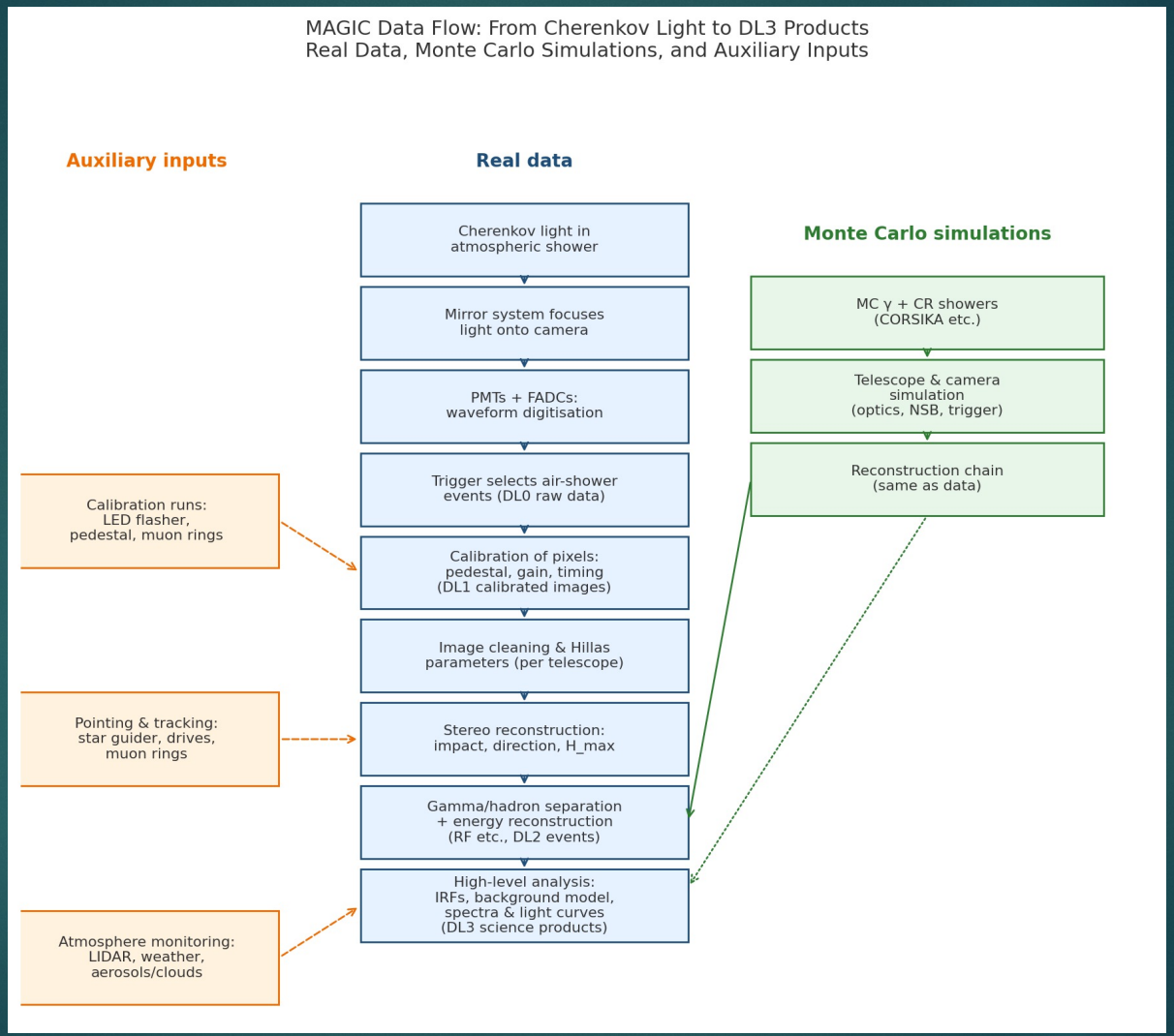
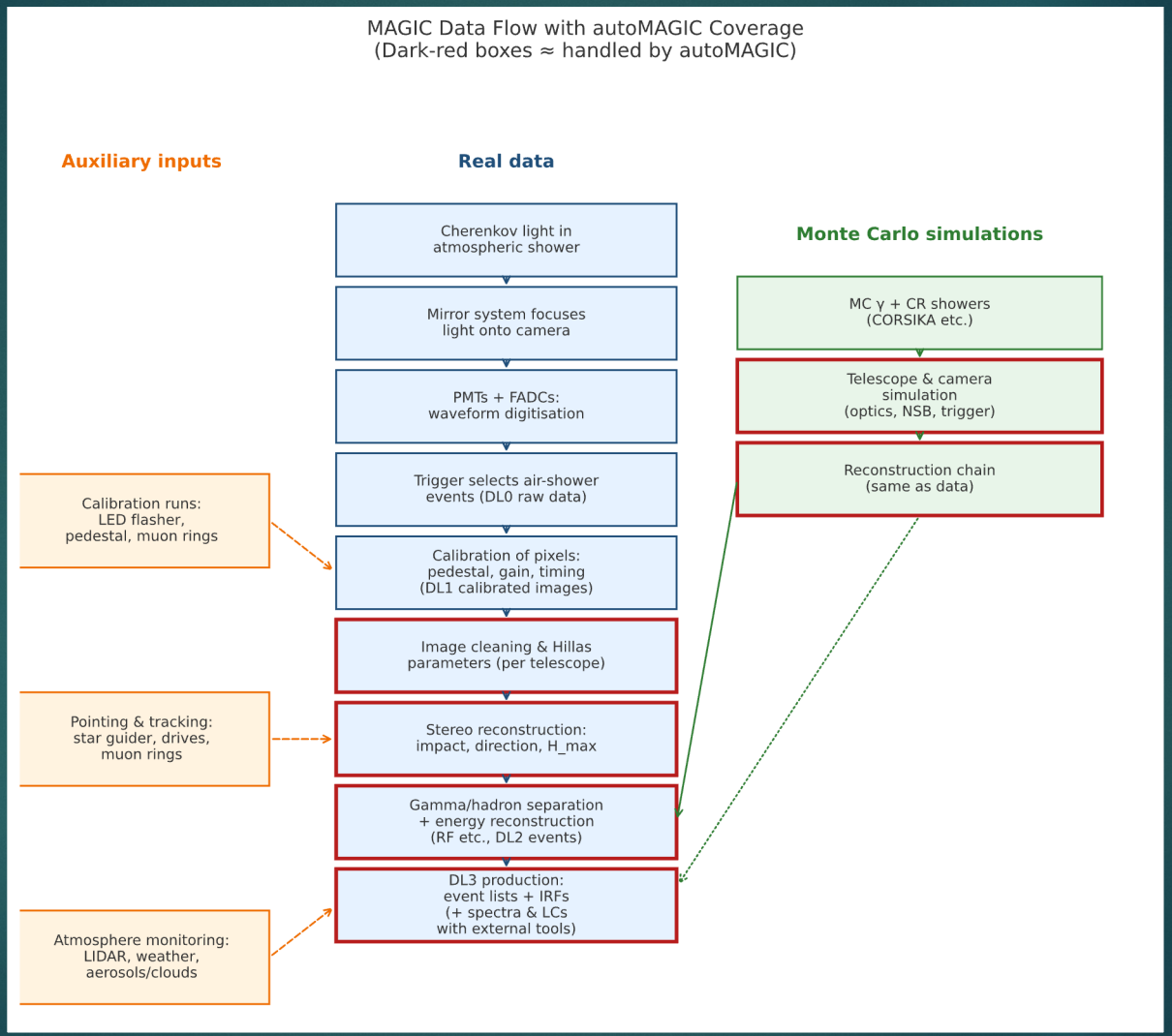


Image Credit: IAC


MAGIC Data Flow: From Cherenkov Light to DL3 Products
Real Data, Monte Carlo Simulations, and Auxiliary Inputs
Auxiliary inputs
Real data
Monte Carlo simulations
Cherenkov light in atmospheric shower
Mirror system focuses light onto camera
PMTs + FADCs: waveform digitisation
Trigger selects air-shower events (DL0 raw data)
Calibration of pixels: pedestal, gain, timing (DL1 calibrated images)
Image cleaning & Hillas parameters (per telescope)
Stereo reconstruction: impact, direction, H_max
Gamma/hadron separation + energy reconstruction (RF etc., DL2 events)
High-level analysis: IRFs, background model, spectra & light curves (DL3 science products)
MC γ + CR showers (CORSIKA etc.)
Telescope & camera simulation (optics, NSB, trigger)
Reconstruction chain (same as data)
Calibration runs: LED flasher, pedestal, muon rings
Pointing & tracking: star guider, drives, muon rings
Atmosphere monitoring: LIDAR, weather, aerosols/clouds



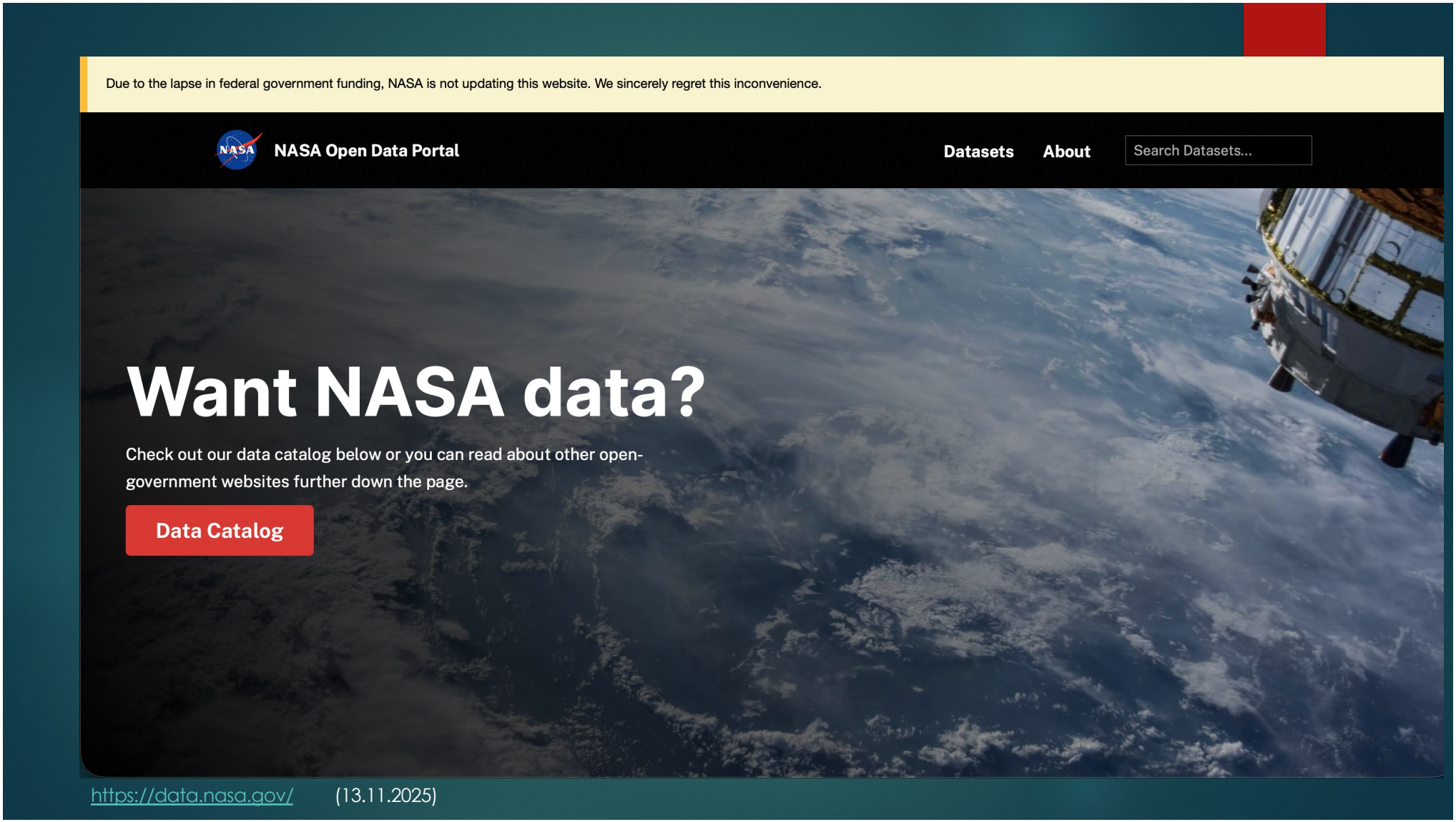
autoMAGIC

L. Linhoff
S. Mender
J.I. Schubert
C. Nigro
S. Fröse
F. Wersig
C. Walther
et al.

Example:
MAGIC Data Management

- Continuous data acquisition (incl. Metadata)
- Full rawn data can be kept for O(yrs)
- Afterwards: tapes „ejected“, pre-processed data left immediately accessible
- Decision: which raw data to keep immediately accessible for longer (dep. on status of analysis & publications → DMPP; can be stringently justified?)
- (1): indefinitely keep well understood (reference) data sets (e.g. Crab Nebula runs)
- (2): indefinitely keep corr. Metadata, simulations, and analysis packages, documentation
- For posterity: scientific publications, data products in public repositories (ideally: publications also open access; cf. DFG / data: 10yrs)
- On-going discussion: Final deposition of remaining data after end of experiment

the MAGIC telescopes

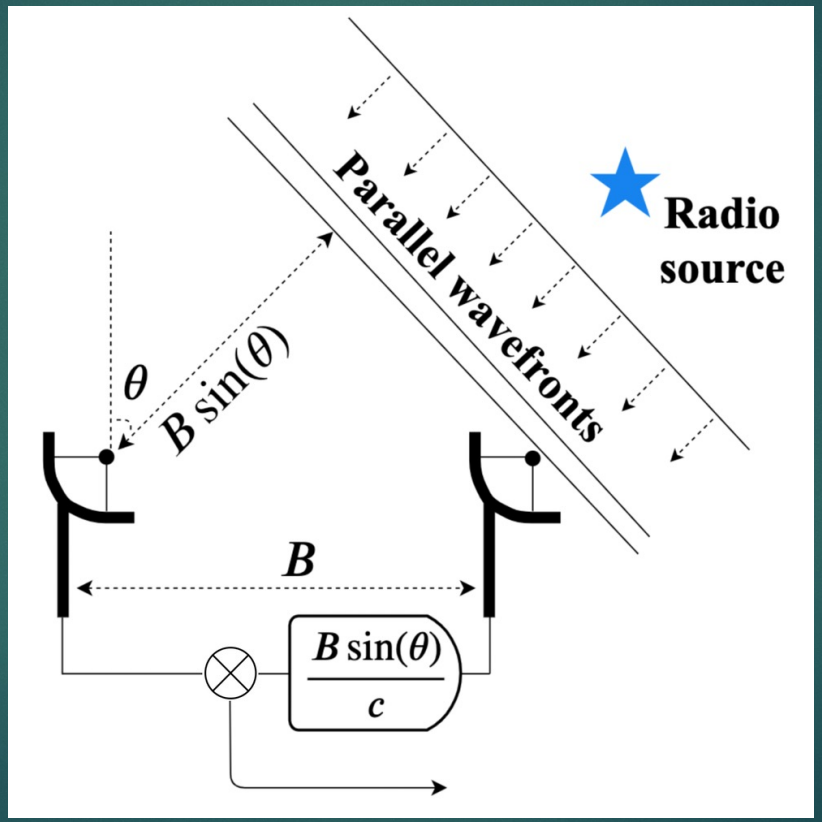


https://data.nasa.gov/ (13.11.2025)

Radio interferometry

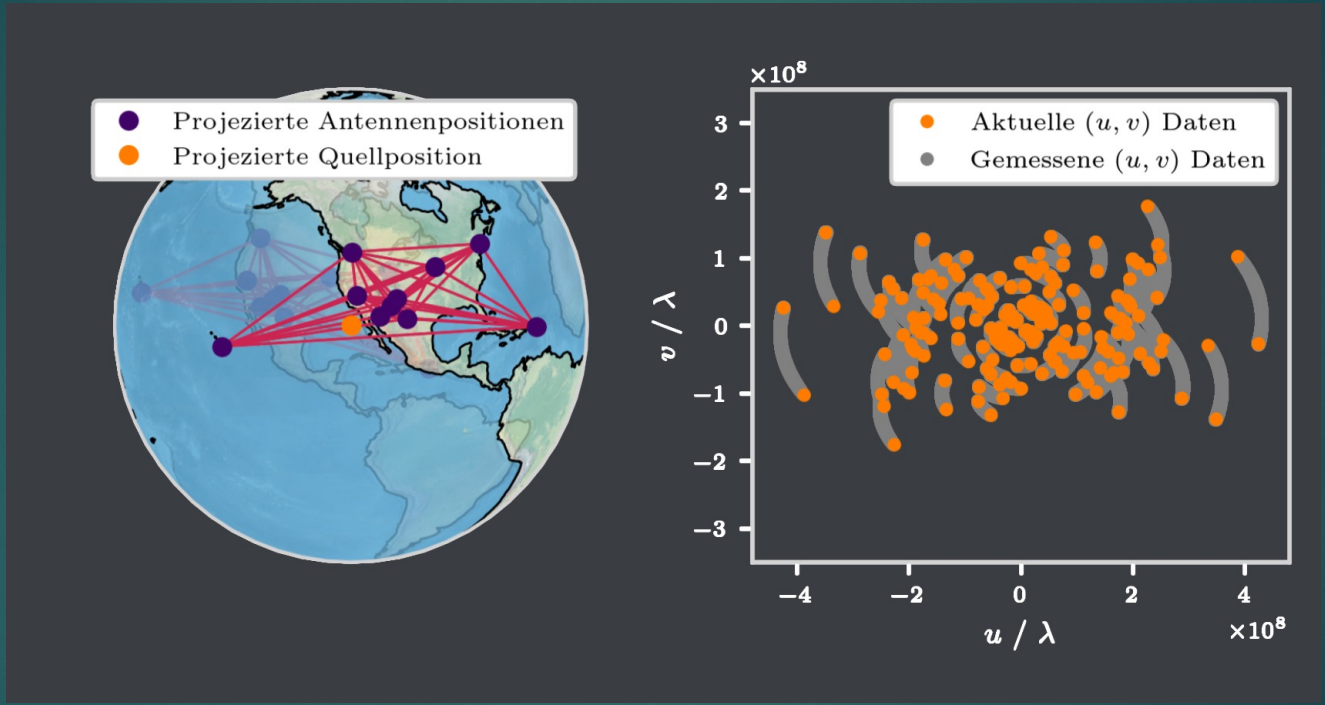
- The long wavelengths in the radio regime blessing and curse all alike
- It is energetically „cheap“ for Nature to produce radio photons
- And radio telescopes need not have reflecting surfaces with sub-micrometer precision, so we can build them very large
- But even a single 100m aperture radio telescope (like at Effelsberg) will in many cases provide a resolution merely comparable to the un-augmented human eye (at a vastly higher sensitivity of course)
- None the less, the highest-resolution „images“ of celestial objects often come from radio telescope systems

Radio interferometry



Credit: SETI Institute / Wael Farah

VLBI



Credit: Kevin Schmitz

Radio interferometry

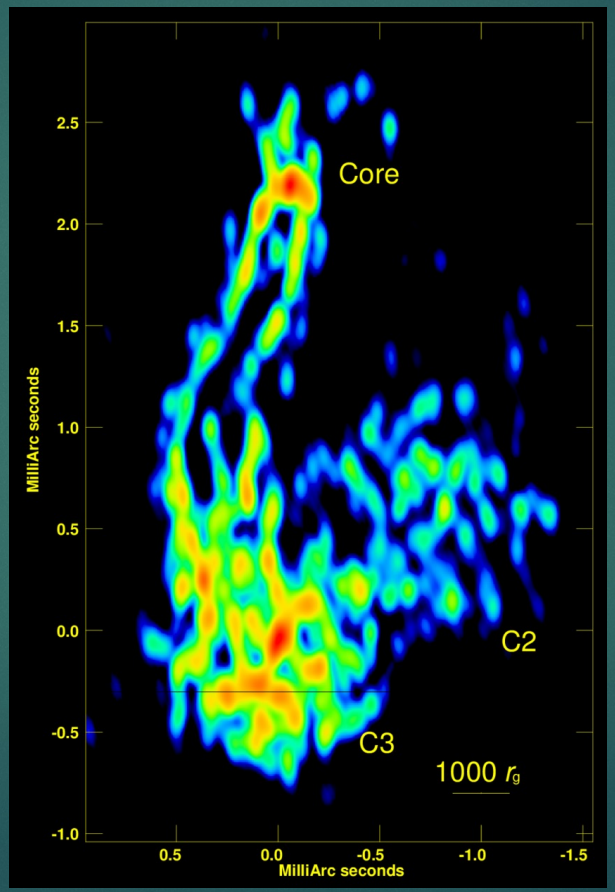
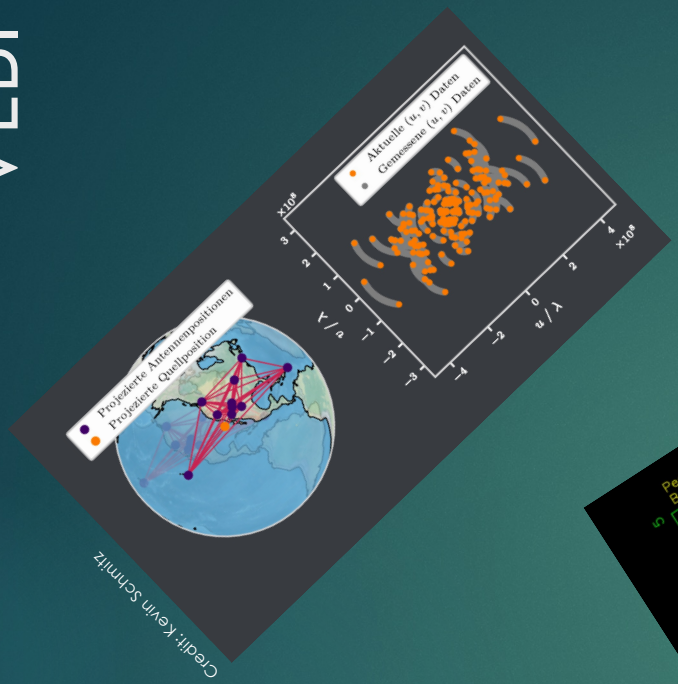
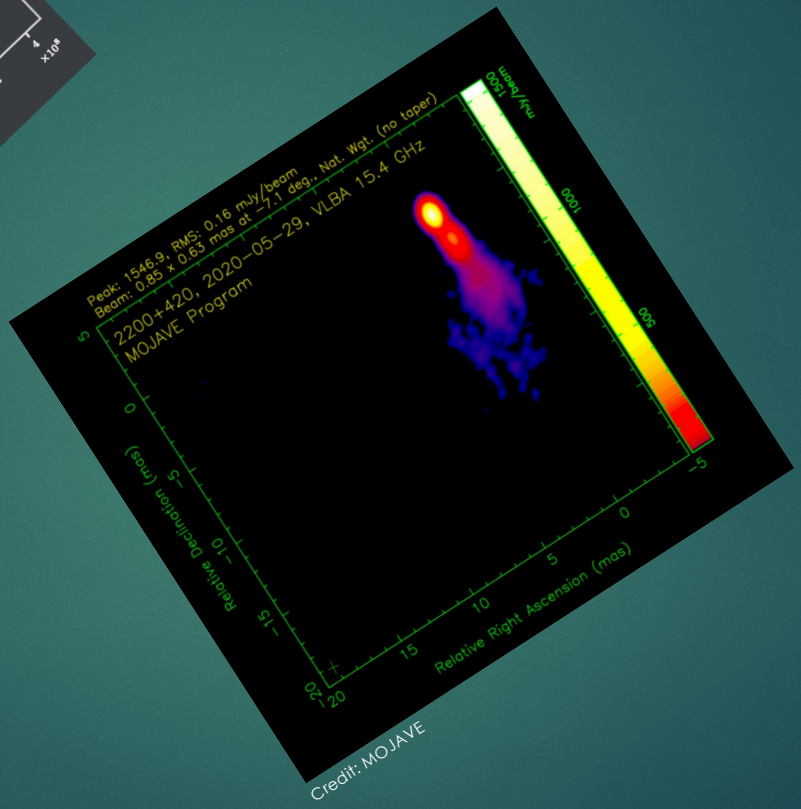


Image From: G. Giovannini et al.
A wide and collimated radio jet in 3C84 on the scale of a few hundred gravitational radii
June 2018Nature Astronomy 2(6)
DOI:10.1038/s41550-018-0431-2

VLBI





Deep Learning based radio interferometric image cleaning and reconstruction: radionets
https://github.com/radionets-project/radionets
K. Schmitz
F. Geyer
S. Fröse
A. Knierim
T. Groß
et al.

Take aways

- Astronomy poses unique challenges; irrevocable decisions need to be taken on short timescales, but will potentially have consequences extending into the far future
- We need to be mindful of that; preservation of data, metadata, simulations and software packages for indefinite timescales; data products best kept (at least) in theory-independent form
- Rapid developments happening with regards to the usage of AI methods in data reduction and analysis → opportunities arising for performance increases and reduction of arbitrary decisions
- One more point for discussion: how to make timescales match in terms of development? Data science and AI moving much more rapidly than design and construction (and funding) of observatories...