



Adventures in AI with the Long Wavelength Array

Jayce Dowell
University of New Mexico

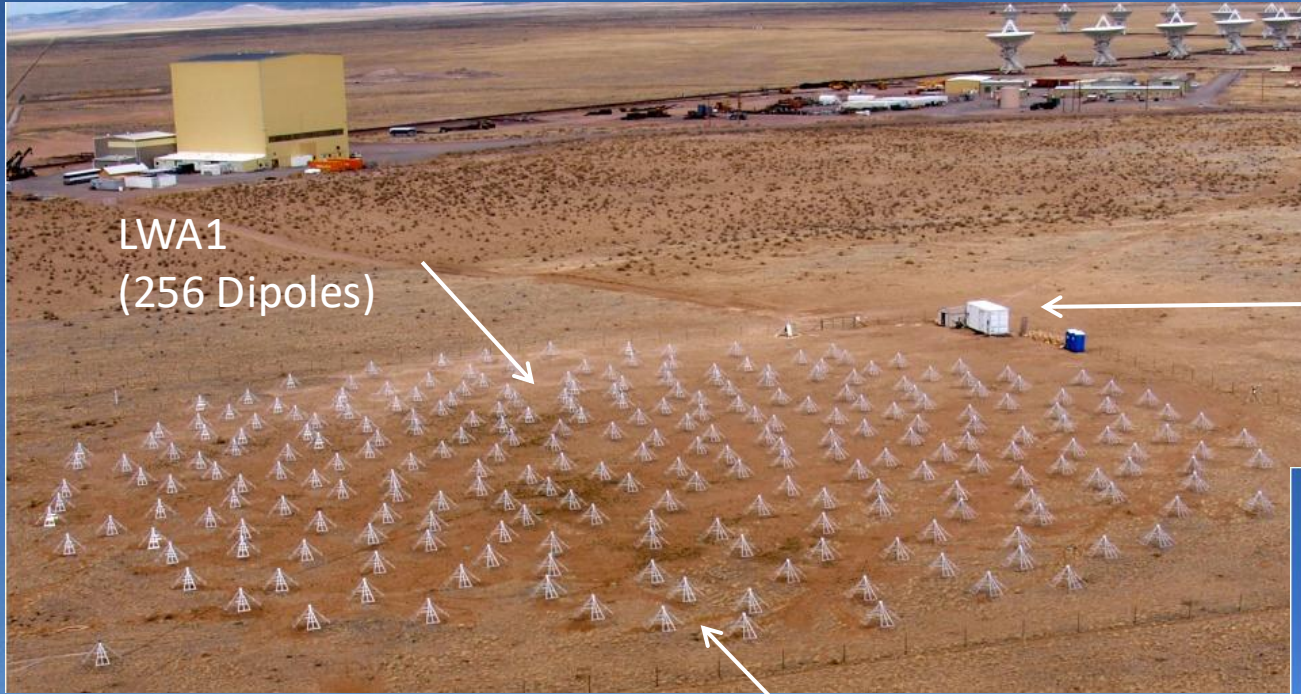
Cosmic Horizons 2
July 16, 2026



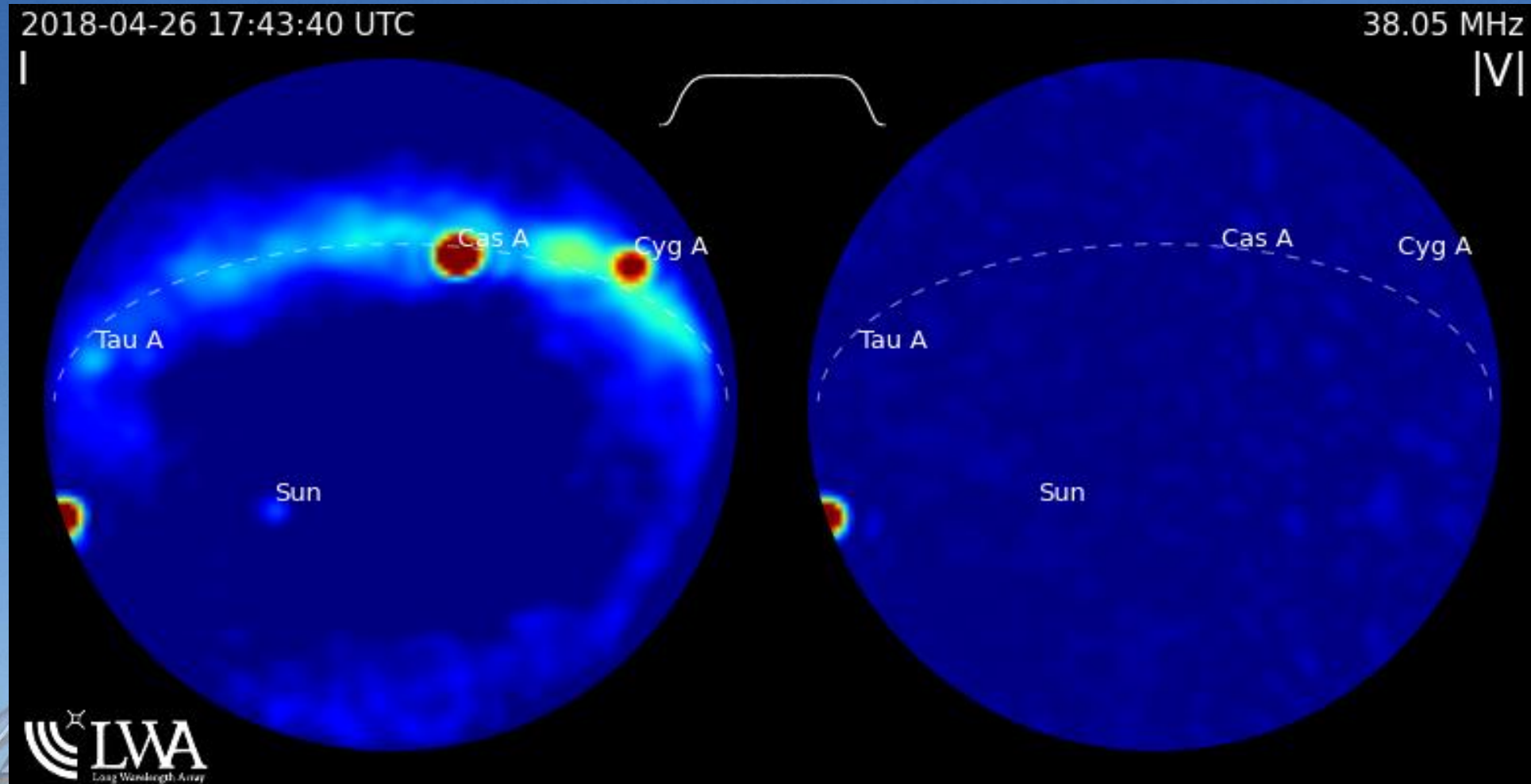


What is the LWA?

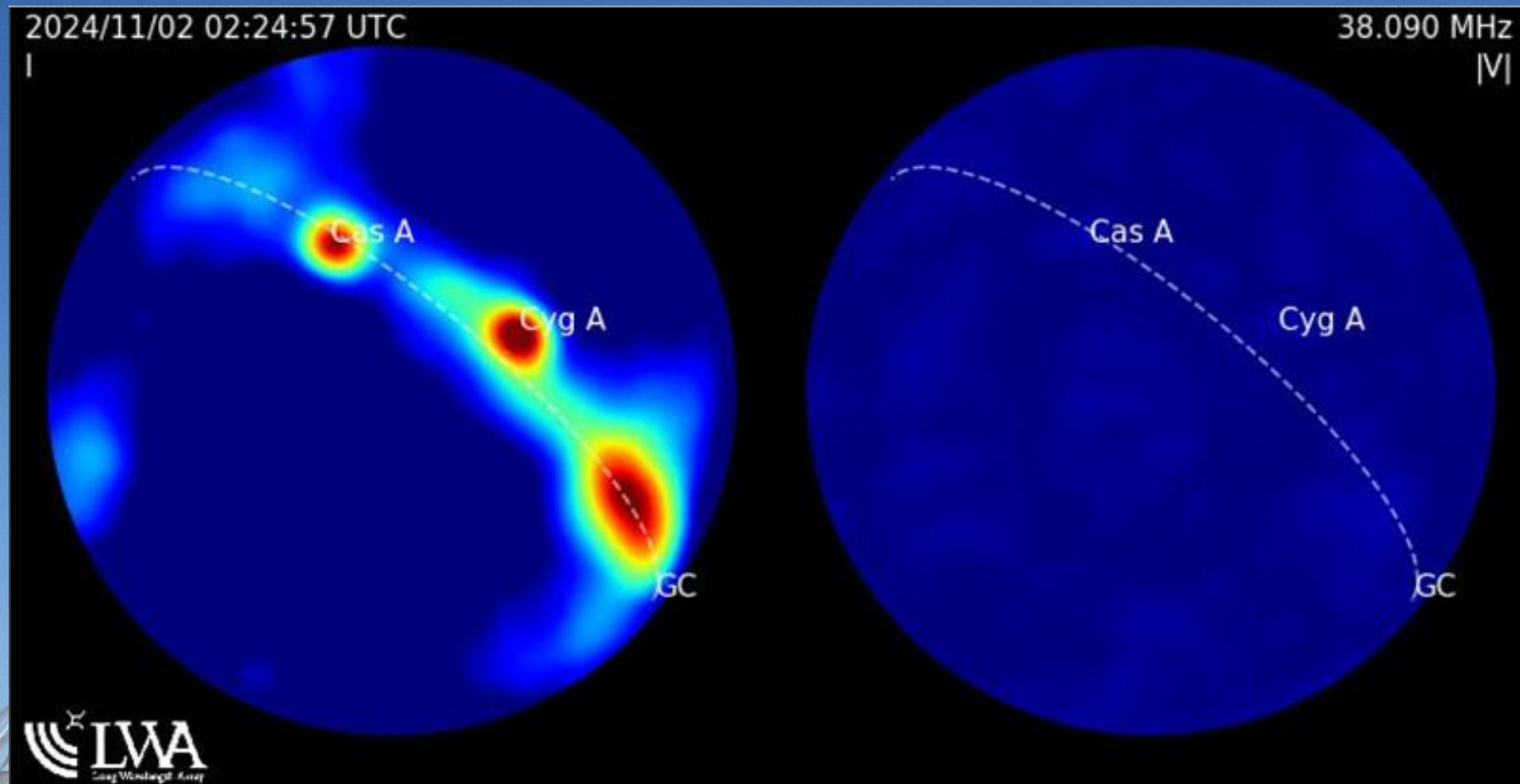
- Radio telescope operating between the ionospheric cutoff and the FM band
 - ~10 to 88 MHz
- Two “flavors” of station:
 - Full station with 256 dual pol. antennas
 - Mini station with 64 dual pol. antennas
 - Both are ~100 m in diameter
- Interferometer of interferometers called the LWA Swarm



Interferometry – Level 1



Interferometry – Level 1



All-Sky Transient Searches

- All-sky imaging is useful for transient searches
 - LWA1 has been making 5 s snapshots of the sky since 2011
 - Started with a single 100 kHz image, recent upgrades have pushed that to ~700 images at a time
- Data quality across time/frequency/station adds complexity
 - Instrumental problems, RFI, Solar/Jovian bursts, lightning, etc.
 - Some are broadband, others are narrow band
- Need a data quality pre-filter to help focus on the best data
 - Ideally a single filter for all frequencies/stations
 - Bonus: monitor station health and overall interference environment over time

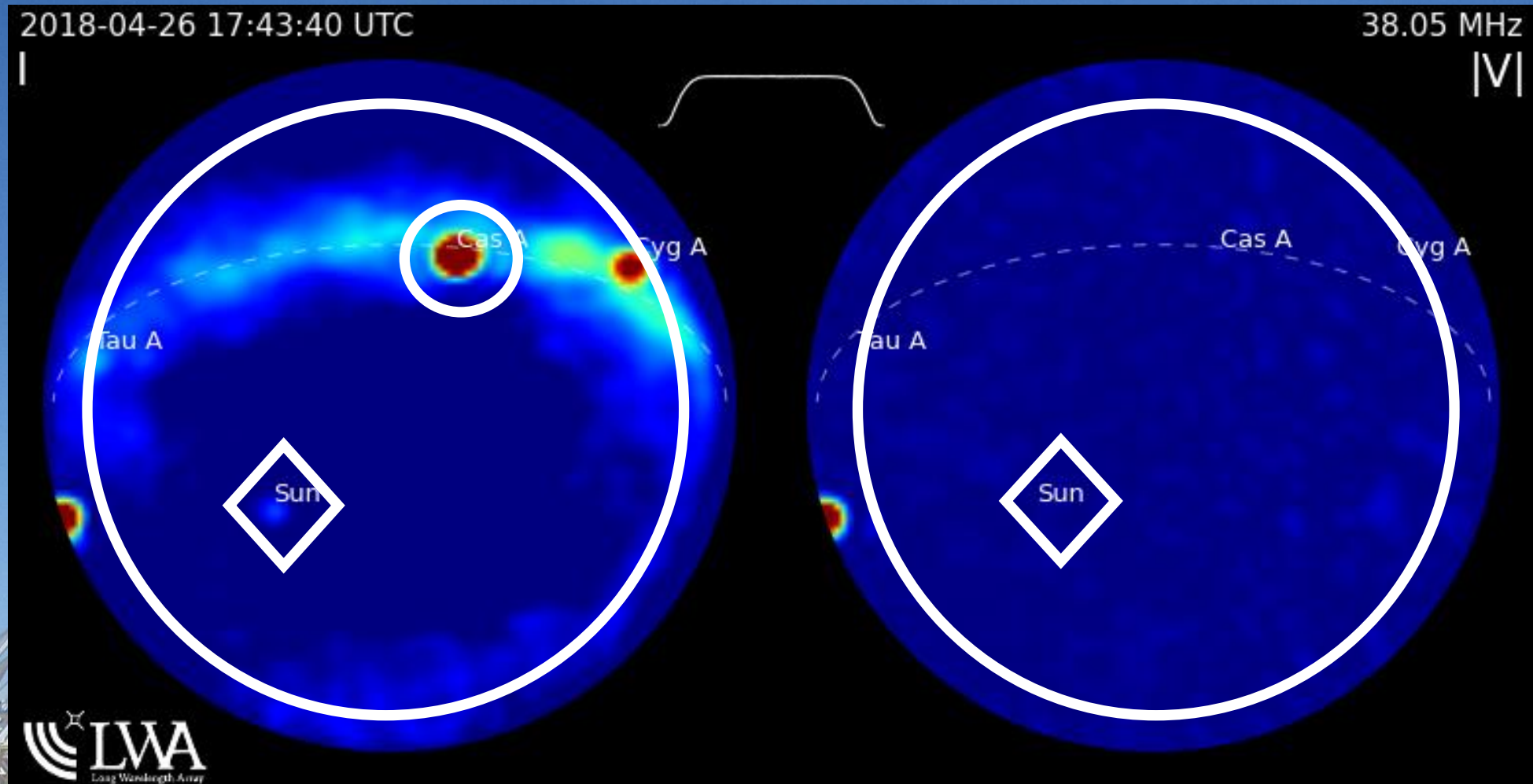


All-Sky Pre-Filter: Oarfish

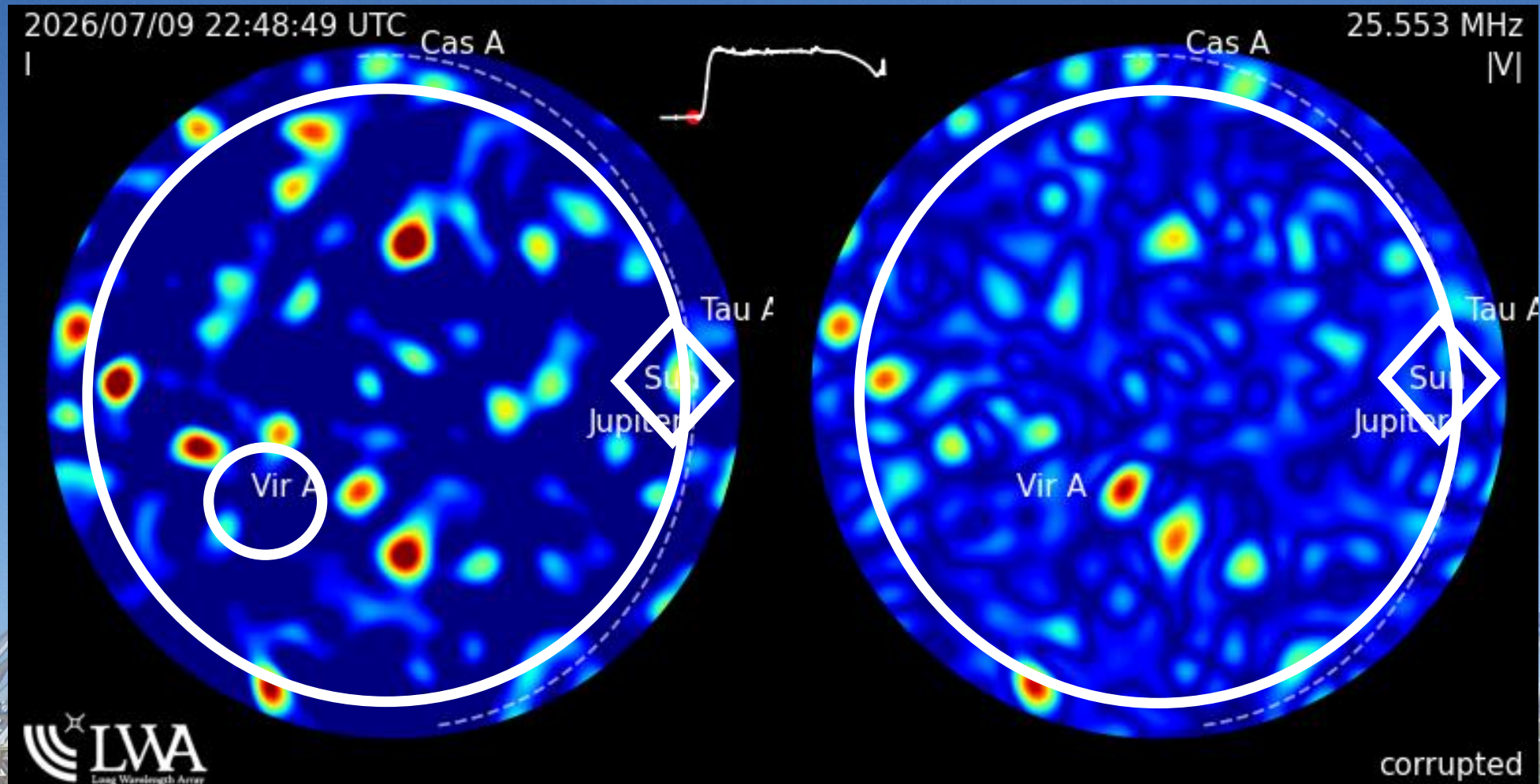
- Oarfish is an image classifier that scores LWA images for quality
 - CNN using a combination of image analysis and extracted features
 - Uses both Stokes I and $|V|$ all-sky images
- Two stages:
 - First classifier determines good/high quality vs bad/low quality
 - Second classifier refines bad/low quality into why the image is bad
 - The two classifiers are also weighted and combined to create a numeric quality score
- Implemented with PyTorch with a 0MQ-based client/server model
 - Client/server to help separate the software for the classifier from whatever the image provider uses



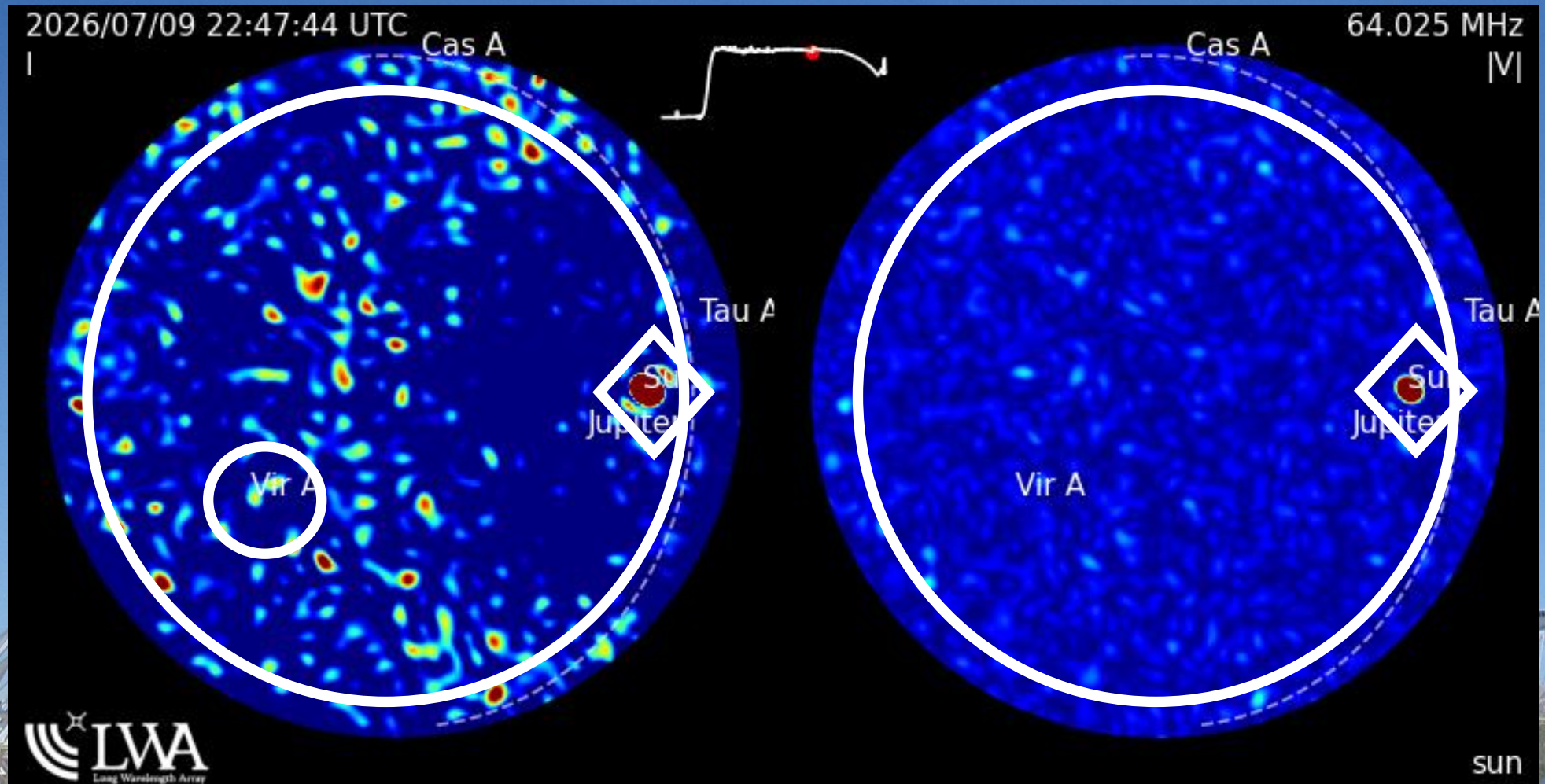
Oarfish: What does it look for?



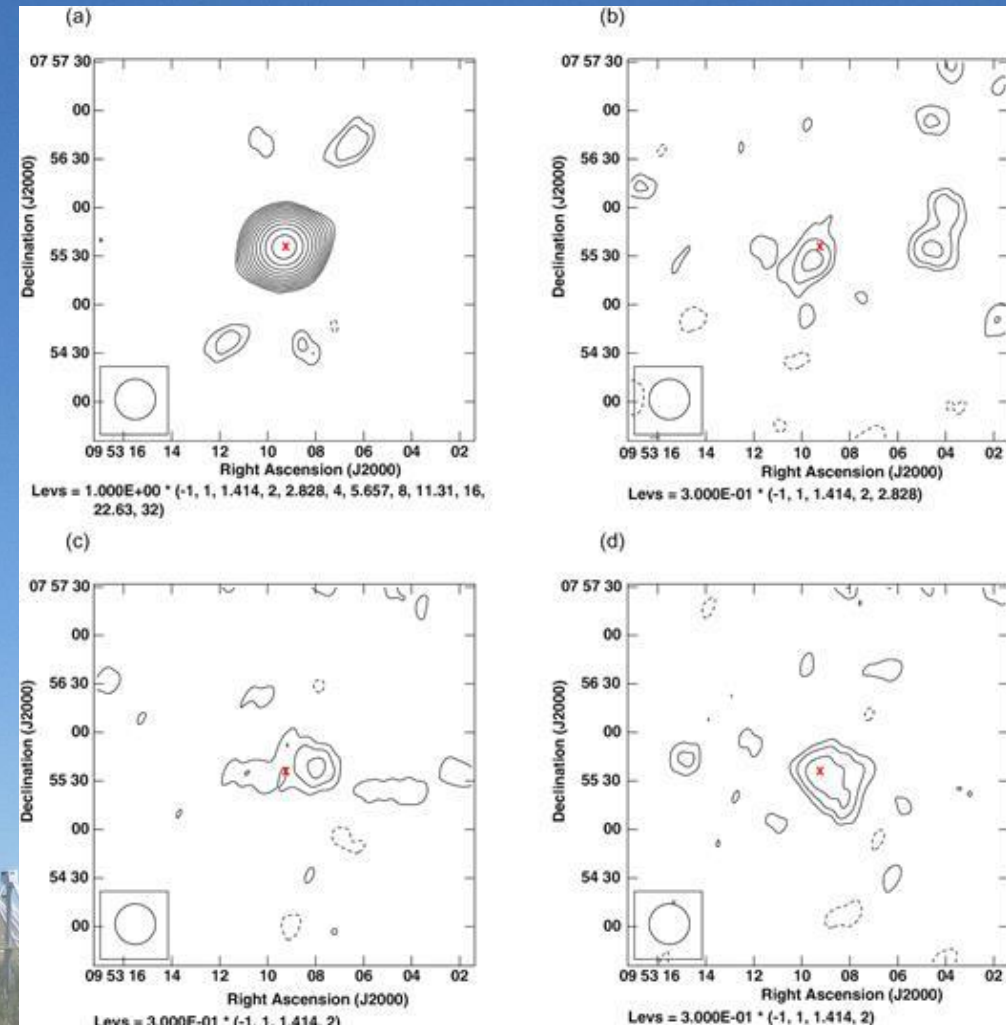
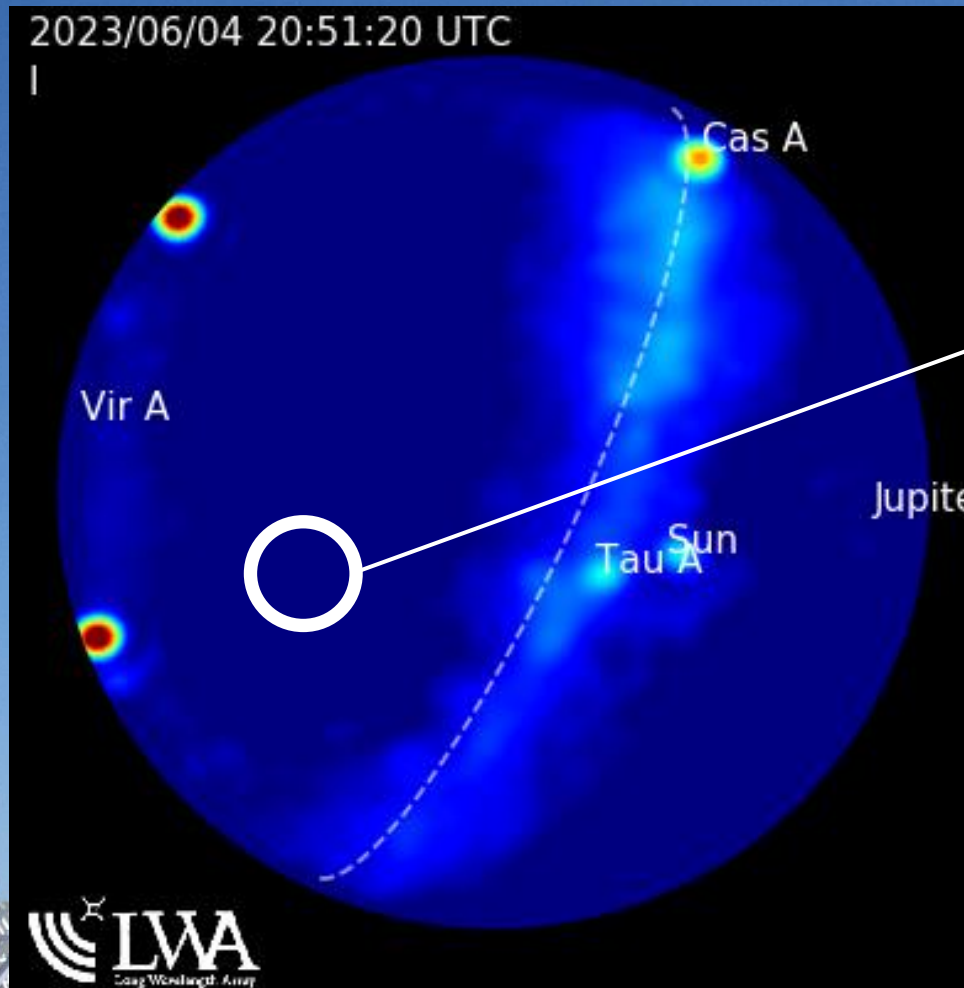
Oarfish: What does it look for?



Oarfish: What does it look for?



Interferometry – Level 2



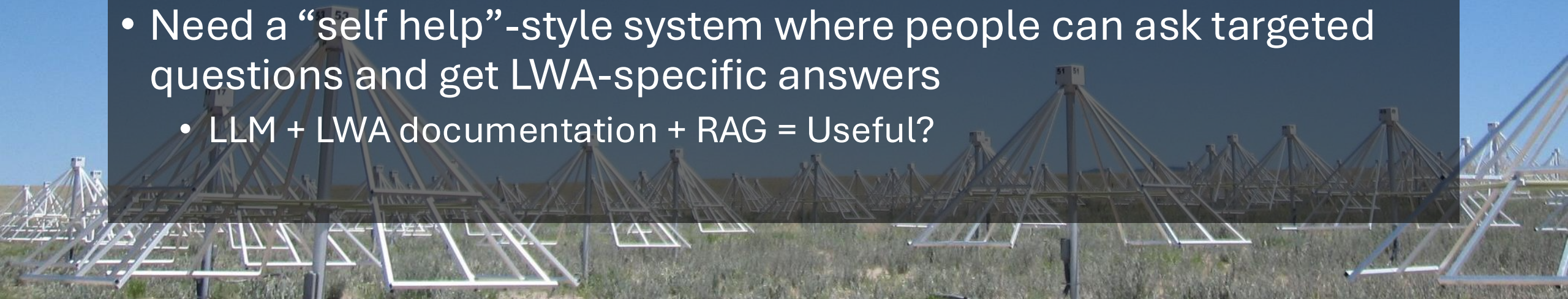
Ruan et al. (2020)

The LWA Swarm



New Stations, New Partners: New Problems

- The LWA Swarm is distributed at many levels, including the construction and maintenance
 - Steep learning curve associated with going from zero to your own LWA station
 - After construction there is monitoring, maintenance, troubleshooting, getting students started, etc.
- Documentation exists but it is often easier to not read it
- Need a “self help”-style system where people can ask targeted questions and get LWA-specific answers
 - LLM + LWA documentation + RAG = Useful?



Alan:

The LWA Observatory Assistant



Alan

Observatory Assistant

● Chat



Ask Alan anything about the LWA

Station status, observation planning, schedules,
documentation, and more.

Type your message...

Default ▾

Send

Alan is an AI assistant for the LWA and may occasionally make mistakes. Verify important information through official documentation or with LWA staff.

[Provide Feedback](#)

Powered by 

Copyright © 2026 The LWA Consortium

Alan: The Nuts and Bolts

- RAG system + LWA and radio astronomy-specific tools
 - "You are Alan, an expert assistant for the Long Wavelength Array (LWA). You support LWA users across their entire workflow: learning about LWA capabilities and its science case, planning and scheduling observations, building and operating stations, troubleshooting operational issues, and accessing and reducing data. You do this by finding relevant documentation and accessing operational systems."
- Knowledgebase consists of a curated set:
 - LWA publications, memos, subsystem documentation, and guides
 - Data reduction scripts
 - A glossary of LWA-specific acronyms or terms
- Self hosted using apache and ollama
 - Currently limited to running with LLMs with ~30 B parameters

Alan:

The LWA Observatory Assistant

● Chat

How is LWA1?

High

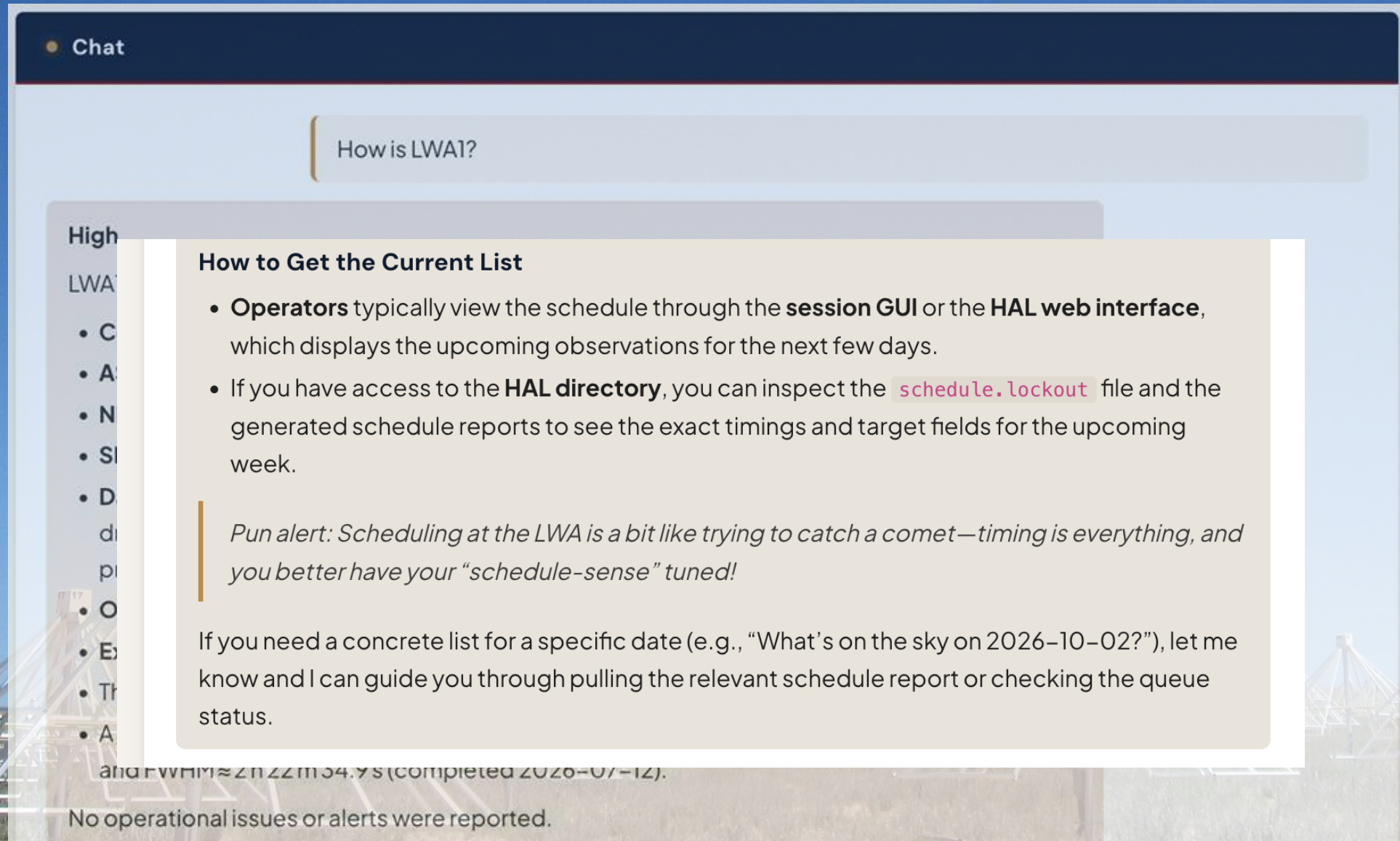
LWA1 is currently **operating normally**.

- **Core subsystems:**
- **ASP** – Status = NORMAL, last updated 2026-07-15 15:50:00 Z.
- **NDP** – Status = NORMAL, last updated 2026-07-15 15:50:01 Z.
- **SHL** – Status = NORMAL (warning conditions cleared), last updated 2026-07-15 15:50:00 Z.
- **Data recorders** (DR1-DR5): all report **NORMAL** and are **Idle**. Each recorder has two attached drives with ample free space (used % ranging from ~2 % to ~20 %). No storage warnings are present.
- **Overall station status: NORMAL.**
- **Extra health information:**
- The most recent health-check completed about **5.2 hours ago** (2026-07-15 10:42 Z).
- A recent stress-test using CygA was performed at 74 GHz with measured SEFD $\approx 79\,077$ Jy and FWHM $\approx 2\text{ h }22\text{ m }34.9\text{ s}$ (completed 2026-07-12).

No operational issues or alerts were reported.

Alan:

The LWA Observatory Assistant



Alan: The Challenges

- Papers and technical documentation are dense
 - Finding the right passages for a question is hard
 - Context injection → Iterative knowledgebase search + acronym resolution
- Many possible tools leads to tool overload
 - Want to cover as many use cases as possible, but tool schemas eat context and not all are relevant to every question
 - All tools at every turn → Down selection based on the current turn
- Models and context sized are limited by the current backend hardware
 - Easy for Alan to lose the plot, particularly on multi-turn chats
 - Current setup → New hardware? Active context management? Compression/context summarization?



Next Steps

- Integrate oarfish into transient search pipelines and look at AI/ML techniques for the searches themselves
 - Automatic event classification on 5D data?
 - “Simple” find events and then extract time-frequency information for each
 - Multi-dimensional YOLO
 - Fusion across/between stations
- Further improve Alan’s performance through improved search and context management.
 - Look into way to also collect/collate/store “institutional knowledge”