2026 Workshop: Database to Discovery in CEDAR

Long title
From Database to Discovery in CEDAR Research: Interoperability, Metadata, and AI Enabled Workflows Across Geospace Repositories
Conveners
Enrique Rojas
Katherine Cariglia
erojasv@mit.edu
Description

CEDAR science increasingly depends on combining data and metadata across multiple repositories, instruments, and communities, but researchers and database teams still face recurring barriers in discovery, access, interpretation, interoperability, and reproducible analysis. This session will bring together database administrators, cyberinfrastructure developers, and research users from Madrigal, SuperDARN, SuperMAG, OMNIWeb/CDAWeb, GIRO, AMISR, and related resources to discuss practical advances in geospace data infrastructure. Topics will include interoperable APIs and metadata (for example HAPI and SPASE aligned workflows), AI assisted data navigation and interpretation, community repositories of analysis tools and notebooks, scientific visualization, and strategies to assess how database infrastructures improve research productivity and usability for the CEDAR community.

Agenda

10:00-10:05 Introduction
10:05-10:15 Pablo Reyes The AMISR database and REproducible Software Environment
10:15-10:25 Alex Chartier Ensuring data access, interoperability and reproducibility for dynamic NSF research facilities
10:25-10:35 Giorgio Picanco GIRO/DIDBase: Current Capabilities, Challenges, and Future Directions for Data Products and Ionospheric Services
10:35-10:45 James Lewis From Metadata to Meaning: Lessons from Building Cross-Mission Analysis Tools
10:45-10:55 Buffer

10:55-11:05 Jesper Gjerloev The physics view: Extracting relevant information from large repositories to understand processes.
11:05-11:15 Katherine Cariglia CEDAR Madrigal Database Updates 2026
11:15-11:25 Jennifer Knuth Data Access in Heliophysics: HAPI and LaTiS from a User's Perspective
11:25-11:35 Enrique Rojas From Data Discovery to Idea Discovery: Papers, Workflows, and Agents as a New Layer for Madrigal
11:35-12:00 Open discussion

Justification

This session is motivated by a shared CEDAR need: frontier atmosphere-ionosphere-magnetosphere studies now require multi-source data synthesis, but the effort required to discover, access, align, interpret, and reuse data remains high and uneven across repositories. The proposed session is designed to be useful for both database teams and research scientists. For database administrators, it provides a venue to compare implementation strategies for metadata, APIs, provenance, discoverability, and user support. For researchers, it focuses on practical workflows that reduce time to science, improve reproducibility, and make cross-database studies easier to execute. We specifically want to invite perspectives from various database teams so the discussion is centered on common design patterns, pain points, and opportunities for coordinated progress across CEDAR and neighboring heliophysics data ecosystems. A useful outcome would be a short community roadmap identifying near-term pilot collaborations, shared priorities for metadata/access improvements, and candidate evaluation metrics for user-facing tools.

Related to CEDAR Science Thrusts:
Encourage and undertake a systems perspective of geospace
Develop observational and instrumentation strategies for geospace system studies
Fuse the knowledge base across disciplines in the geosciences
Manage, mine, and manipulate geoscience/geospace data and models
Workshop format
Short Presentations
Panel Discussion
Keywords
data infrastructure, interoperability, metadata and APIs, AI
View PDF