Knowledge Discovery Interest Group

Opening Session — IVOA InterOp Nov 2026

Chair: Yihan Tao (NAOC & China-VO)

Vice-Chair: André Schaaff (CDS)

IVOA INTEROP · PERTH, AUSTRALIA · 6 NOVEMBER 2026

Knowledge Discovery IG Topics

InterOp Nov 2026

Knowledge Discovery turns astronomical data into new understanding, spanning data access, visualisation, statistics, machine learning and workflow orchestration.

KDIG tracks how AI, machine learning, data mining and all the knowledge discovery methods are taken up in astronomy, and reflects state of the art to the VO community.

Multimodal foundation models

Models (e.g. AION-1) pre-trained across images, spectra and catalogues serve many tasks; transfer across surveys and instruments remains challenging.

From analysis to observation

AI-driven Telescopes (e.g. StarWhisper Telescope) are exploring how AI can help plan and manage observations, and adapt the schedule in real time.

AI as part of the research workflow

AI systems (e.g. AI Cosmologist) automates data processing and analysis; astronomers provide oversight and validate results, with reliability and sustainable use central.

Looking ahead

A Focus Meeting on 'AI Frontiers for Astronomy', will be held at the XXXIII IAU General Assembly (Rome, August 2027). We invite the community to discuss with us and shape the possible IVOA contribution to the meeting.

Details: aifrontiers.cbpf.br

KDIG Session — This Interop

14:00–15:30 (Perth local time)

Speaker	Title	Time	Abstract	Material
Yihan Tao [NAOC] (remote)	Introduction	2 min		
Bruce Berriman and Rosemary moseley [IPAC]	Harness Engineering & TAP at KOA	15+3 min	This presentation describes the harness architecture of KOkuA, an LLM agent that when complete will provide a natural language assistant for the Keck Observatory Archive. It created and runs queries in ADQL, and responds to users' requests for a simple interface to create complex, custom queries. The goals are accuracy, affordability and resilience against AI vendor (daily!) changes.	
Kuang Mao [Zhejiang Lab] (remote)	A Multimodal Foundation Model for Astronomical Object Search and Scientific QA	10+3 min	We present a multimodal foundation model for astronomy that integrates multi-modal observations into a shared semantic space and couples it with a large language model. By computing offline embeddings as standard data products, the model enables cross-modal search (via natural language or reference objects) and scientific QA directly against observations. This one-step joint retrieval and reasoning approach eliminates reliance on hand-tuned, multi-step pipelines, lowering the barrier to data-intensive astronomy and enabling open-ended discovery.	
Dave morris [University of Manchester] (remote)	Attribution for AI	10+3 min	Developing a set of agent rules to add attribution to content generated using AI	
Dave morris [University of Manchester] (remote)	Alternatives for AI	10+3 min	A report on using DeepSeek open weight models, DeepSeek Harness, and alternative service providers for code generation	
Florian Cabot [EPFL] (remote)	The VIRUP project and AI for Astrophysical Visualization	10+3 min	VIRUP is a real-time astrophysical visualization software designed to handle datasets of various types and scales, ranging from planets and stars to galaxies and large-scale cosmological simulations. In this talk, we will provide an overview of VIRUP and its main capabilities, with a particular focus on recently introduced AI-powered interaction tools, including a chatbot, a coding assistant, and a Model Context Protocol (MCP) server.	
André Schaaff [CDS]	Fight the Bots or Help the Agents ?	7+3 min	A few words around MCP servers, agents, examples in other domains, before the discussion	
André Schaaff, Yihan Tao & all participants	Discussion	8 min		