Title: StarAI - empowering access to astronomy data handled by the Canadian Astronomy Data Centre and its CANFAR platform with artificial intelligence

Authors: Adrian Damian, Hossen Teimoorinia, Mrunal Mustapure, Serhii Zautkin

Proposed track: Science platforms in the big data era

Abstract:

We present **StarAI**, a prototype system that orchestrates large language models (LLMs) to streamline discovery and access of astronomy data hosted by the Canadian Astronomy Data Centre (CADC) and use of its CANFAR science platform. Natural-language queries are translated into precise ADQL against CADC's TAP services, returning observation records with resolvable links to preview images and **FITS** products. StarAI is built on the **Model Context Protocol (MCP)**: a TypeScript/Express client steers model tool use, while an MCP server exposes Skaha API services such as ADQL-based search, DataLink resolver, target coordinate resolution, CADC table discovery, and retrieval-augmented generation (RAG) search. The RAG service connects to a **Qdrant** vector database of JWST and HST proposal texts and to a CANFAR user knowledge base, enabling plain-language requests (e.g., "Infrared images of M31", "Observations related to … (quasars, Supernovae) etc.") to yield reproducible queries and valid observation records. The web UI (Next.js/React/TypeScript) renders results and packaged products with stable URIs, and the services are containerized for CI/CD deployment; models are accessed via commercial APIs or locally hosted open-source checkpoints. Early user journeys suggest that StarAI can reduce platform onboarding time, lower friction when moving from questions to ADQL, and scale from quick-look exploration to bulk, asynchronous, agentic jobs (work in progress). We outline the roadmap from an alpha testbed toward a CANFAR-integrated beta and invite community feedback and experimentation via the live pre-alpha instance.

Keywords: large-language models, astronomy software, artificial intelligence, science platforms, CADC, CANFAR, CAOM2, TAP, ADQL

Artifacts: Demo - https://canfar.testapp.ca/star-ai