

Boosting AI for Science by Highly Available Cloud Services in the Era of Agentic AI

Wednesday, 18 March 2026 15:06 (22 minutes)

The shift from conventional computing, networking, and storage to AI-driven scientific discovery (AI4S) calls for a new generation of intelligent infrastructure. In this era of agentic AI, scalable access to models, tools, data, and agents has become critical—serving as essential utilities powering next-generation research. To meet these demands, the HepAI team has developed Qionwu, a core technology that supports unified integration, dynamic authorization, and high-concurrency API services for large models, scientific tools, and diverse AI agents. The system reliably enables various AI applications—including data agents, physics analysis assistants, and conversational bots.

This foundation paves the way for large-scale autonomous exploration and collaborative research involving thousands of intelligent agents. In this talk, we will share the design and implementation of our highly available cloud service infrastructure based on Qionwu, with a special emphasis on how it powers physics analysis agents such as Dr.Sai. We will demonstrate its capabilities through real use cases in scientific domains and discuss future directions for agentic AI in accelerating scientific breakthroughs.

Primary author: ZHANG (张), Zhengde (Institute of High Energy Physics (IHEP), CAS, China)

Co-authors: QI, Fazhi; GANG, Chen

Presenter: QI, Fazhi

Session Classification: Artificial Intelligence (AI) - I

Track Classification: Track 10: Artificial Intelligence (AI)