

AI Scientist Challenge Rules

[← Back to Call for Challengers](#)

Version v2025.11.11 · Released November 2025

This page summarizes the official AI Scientist Challenge Rules (Version 2025.11.11). All participants must comply with the submission workflow, technical constraints, and evaluation procedures described below.

1. Submission Requirements

1.1 Submission Channel

Participants must provide:

- A `.zip` file containing the complete project, structured according to the required format.
- If deployment requires extra environment variables (e.g., academic API keys), include a complete `.env` file inside the same package.
- Teams competing in multiple tracks must submit a separate questionnaire for each track, with exactly one code package per submission.

The Organizing Committee distributes a private questionnaire link by email before **21:00 on November 11 (UTC+8)**, enabling upload of the `.zip` package.

1.2 Code Structure Requirements

Projects must follow the required structure and be submitted as a `.zip` file through the questionnaire link. An official example project is available at [GitHub](#).

Repository requirements (root directory):

- **Dockerfile** (image build script)
 - Base image **must** be:

```
python:3.12-slim-bookworm
```
 - Additional dependencies are unrestricted (e.g., local vector stores, OpenAI SDK, LangChain, CrewAI, AutoGen).
 - The Docker image must be buildable within mainland China's network environment.
- **docker-compose.yml** for service deployment
 - The service must expose **port 3000** and map it correctly.

ICAIS 2025 Nov 23-25 @Beijing China

- SCI_EMBEDDING_BASE_URL (to be set to <https://dashscope.aliyuncs.com/compatible-mode/v1>)
- SCI_EMBEDDING_API_KEY (filled in by the Organizing Committee)
- SCI_MODEL_API_KEY (filled in by the Organizing Committee)
- SCI_LLM_MODEL ("deepseek-chat")
- SCI_LLM_REASONING_MODEL ("deepseek-reasoner")
- SCI_EMBEDDING_MODEL ("text-embedding-v4")

Participants can use DeepSeek's official services ([deepseek-chat](#) / [deepseek-reasoner](#)) and Alibaba Cloud Bailian's [text-embedding-v4](#) model for local development and testing to keep interfaces consistent with the evaluation environment.

If extra environment variables are required, include a complete `.env` file in the submission package. It must still contain the seven model-related variables above, with values left empty for the organizers. All model information must be read via environment variables.

Daily Deployment Schedule

The system deploys all new submission packages at **11:00 (UTC+8) each day**. Each team should upload no more than one update per day.

- If an updated version is submitted, the new build replaces the previous deployment in the arena. Past evaluation data is retained, and the leaderboard adds a date suffix to distinguish versions.
- If no new submission is uploaded, the current version continues to participate without change.

2. Compliance

2.1 Deployment Environment

- All services run on dedicated Alibaba Cloud ECS instances located in mainland China with **2 cores / 4 GB RAM / 50 GB disk**.
- Each repository is loaded onto a dedicated server for code scanning, image building, and deployment.
- The environment ships with the latest Docker, docker-compose, and the `python:3.12-slim-bookworm` base image.
- Participants must ensure external APIs required by their service are accessible from within mainland China.

2.2 Access to External Services

2.2.1 Access to LLM Services

- The Organizing Committee provides official LLM and embedding models **only during deployment and evaluation**; API keys are not shared during development.
- The following models are used during evaluation:
 - `deepseek-chat`
 - `deepseek-reasoner`
 - `Qwen/Qwen3-Embedding-4B`
- Systems must rely solely on these models during evaluation. Use of any external or third-party model APIs is strictly prohibited.
- Reference documentation:

- [Example submission project](#)

2.2.2 Access to Academic APIs

Only the following academic data APIs may be accessed. Participants must implement throttling and error handling to respect any rate limits.

- [OpenAlex](#)
- [Semantic Scholar](#)
- [Unpaywall](#)
- [Crossref](#)
- [OpenReview](#)
- [arXiv](#)

No other academic APIs may be used without explicit approval. All submitted code is scanned and verified to preserve fairness.

3. API Specifications

Services must return responses in streaming Markdown via Server-Sent Events (SSE). Any other response format will be filtered.

3.1 Literature Review API

- **Endpoint:** POST `http://<agent_service_host>:3000/literature_review`
- **Input:**

```
{
  "query": "Please help me comprehensively sort out the latest advancements in the field of diffusion language models"
}
```

- **Output:** SSE Markdown chunks ending with `data: [DONE]`.
- **Error format:**

```
400 Bad Request
{
  "error": "Bad Request",
  "message": "Query is required"
}
```

- **Timeout:** 15 minutes.

3.2 Paper QA

- **Endpoint:** POST `http://<agent_service_host>:3000/paper_qa`
- **Input:**

```
{
  "query": "Please carefully analyze and explain the reinforcement learning training methods used in this article.",
  "pdf_content": "xxx"
}
```

ICAIS 2025 Nov 23-25 @Beijing China

- **Timeout:** 15 minutes.

3.3 Ideation

- **Endpoint:** POST `http://<agent_service_host>:3000/ideation`
- **Input:**

```
{
  "query": "Please help me come up with an innovative idea for spatiotemporal data prediction using LLM technology."
}
```

- **Output:** SSE Markdown streaming.
- **Timeout:** 10 minutes.

3.4 Paper Review

- **Endpoint:** POST `http://<agent_service_host>:3000/paper_review`
- **Input:**

```
{
  "query": "Please provide a brief review of this paper",
  "pdf_content": "xxx"
}
```

- **PDF content:** Base64-encoded string for a single PDF file.
- **Output:** SSE Markdown streaming with required sections:
 - Summary
 - Strengths
 - Weaknesses / Concerns
 - Questions for Authors
 - Score section with Overall (10), Novelty (10), Technical Quality (10), Clarity (10), Confidence (5)
- **Timeout:** 20 minutes.

4. Real-Time Ranking

Participants can monitor standings at <http://39.97.229.86/leaderboard>, which is updated continuously with the latest arena evaluations.

Final submissions close at **20:00 on November 22 (UTC+8)**. Official results are determined by the leaderboard state at **20:00 on November 23 (UTC+8)**.

ICAIS 2025 Nov 23-25 @Beijing China

Exploring the frontiers of automated scientific discovery	Program
	Committees
	Register

Conference Service Provider

Suzhou Henuo Culture Media Co., Ltd.

© 2025 The 1st International Conference on AI Scientists (ICAIS). All rights reserved.