

IOAI² Inaugural Edition Review: Autonomous AI Meets the International Olympiad in AI Challenge

15 AI systems, 6 tasks. Human ingenuity. AI execution. Greater possibilities for all.

Abhishek Divekar, IOAI²: AI Models Track Chair
Elena Marinova, IOAI cofounder and Chair of the Board

1. Executive summary	2
What is IOAI and IOAI ² ?	2
Summary of results	4
2. IOAI² Founding AI Participants	5
3. Competition format and rules	7
Task proctoring	8
Scoring environment and resource limits	8
Autonomy requirement	9
4. Submissions and Jury verification	9
Stage 1: Code submissions	10
5. IOAI² Results	11
Award boundaries	11
IOAI ² Trophy recipients	12
Per-task results	12
Task 1: Find the Order	13
Task 2: Chasing the Robot	15
Task 3: Potato Contact	17
Task 4: Double Agent Dilemma	18
Task 5: Ghost of the Machine	20
Task 6: IOAI Field	22
6. Overall cost of the AI Models Track	23
7. Retrospective on human and AI participation	25
8. How AI systems bent the rules	27
9. Lessons for future editions	29
10. Next steps	29
11. Contributions and Acknowledgements	30

1. Executive summary

What can AI achieve when challenged with the same problems as talented young students?

The International Olympiad in Artificial Intelligence (IOAI), at its 2026 edition, launched the inaugural **IOAI²: AI Models Track** turning the IOAI tasks into a new test of autonomous AI.

IOAI² brought together 15 autonomous AI modeling systems from 14 founding organizations around the world, from leading university research labs to startups and independent AI labs. The track winner was **Synthetic Sciences, a Y Combinator-funded AI lab co-founded by a student participant at IOAI 2025**.

Each system had two hours to explore unfamiliar data, develop a solution and improve it autonomously, without human guidance. The systems tackled the same six tasks as students participating in IOAI 2026, where 108 countries and territories were represented.

On one of the six IOAI 2026 tasks, the best student score exceeded the best AI score, and on two tasks the best student score was within 7% of the best AI score. Across all six tasks, however, all **15 autonomous AI systems achieved overall scores at or above the student gold-medal threshold, and 13 exceeded the highest student total**. The two groups competed under different conditions, so these figures do not provide a direct measure of AI versus human ability under equal settings. In an unexpected twist, several AI agents attempted to exploit loopholes in the competition rules or Kaggle platform, which the IOAI Jury detected and addressed.

This inaugural IOAI²: AI Models Track edition review presents the competition format, participating organizations, results, costs of participation and proctoring, Jury precedents and lessons for future editions. This offers other scientific olympiads a proposed approach for evaluating fully autonomous AI systems under their own governance.

What is IOAI and IOAI²?

The International Olympiad in Artificial Intelligence (IOAI) is an annual competition and education initiative for school students. Its Individual Contest asks them to solve AI modeling problems under a fixed time limit and controlled conditions. IOAI was founded in 2024 in Bulgaria, the second edition was held in 2025 in China and the third edition in 2026 was held in Astana, Kazakhstan, where 108 countries and territories were represented.

In 2026, IOAI inaugurated **IOAI²: AI Models Track**. IOAI² is a competition for autonomous AI systems that brought together teams from leading university research labs worldwide and emerging AI startups, including ventures co-founded by IOAI alumni from 2025 and 2026. Its inaugural edition ran in parallel with the on-site student Individual Contest at IOAI 2026, with AI systems evaluated on [the same six tasks](#) following as closely as possible the [Contest rules of IOAI](#).

IOAI²: AI Models Track connects the Olympiad with organizations building autonomous AI systems and gives those systems a common competition on previously unseen AI modeling tasks. Participating organizations use their own models, tools and execution infrastructure.

The IOAI International Scientific Committee defines the tasks, scoring rules, rankings and awards, while the IOAI Jury proctored the competition on Kaggle and reviewed compliance with the published rules.

IOAI² supports independently reviewed evaluations of fully autonomous AI modeling systems.

The track examines how well these systems solve novel AI modeling problems within a two-hour competition window and without human guidance. It evaluates several complementary abilities: understanding an unfamiliar problem and data, managing time, selecting and running experiments under resource constraints, submitting code through a CLI (Command Line Interface), interpreting feedback, and improving a solution through repeated iteration.

Together, these abilities measure sustained, autonomous problem-solving on olympiad-level AI modeling tasks.

Official Statement:

*"The frontier of artificial intelligence has shifted: from powerful models that answer questions to complete systems that plan, reason, and act autonomously to solve the hardest problems in Machine Learning and AI modeling. As one of the world's largest Olympiads, IOAI bears a duty to the next generation to establish a rigorous and neutral measure of what today's most advanced AI systems can achieve on Olympiad-grade ML modeling. The inaugural **IOAI²: AI Models Track 2026** invites AI labs to demonstrate their advances in Agentic AI on the international stage. Teams from industry and academia will benchmark AI systems that build machine learning solutions on the same expert-designed problems solved by the most talented high school students from 100+ countries and territories at IOAI 2026. Code for Machine Learning modeling must be written and submitted by AI agents fully autonomously, with no human involvement. Every participating lab in the AI Models Track becomes part of IOAI's mission to inspire the next generation of AI researchers and engineers, fostering a global dialogue between those building AI today and those who will build it tomorrow."*

– Abhishek Divekar, Chair of IOAI²: AI Models Track, International Scientific Committee, IOAI 2026

IOAI is the fastest-growing international science Olympiad in history and, by the number of participating countries and territories, the second largest after the International Mathematical Olympiad (IMO). It is also the fastest-growing year-round educational initiative in AI worldwide, bringing together an inclusive and diverse community of young talents, educators, and partners.

*With the **IOAI²: AI Models Track**, IOAI adds a new dimension to this community, connecting young talent with AI research and innovation to help shape the future we are building together. We are especially proud that two alumni-founded labs participated in the inaugural edition - and one of them won. Their success is another signal of the emergence of a self-renewing ecosystem in which participants become innovators, contributors, and role models for the next generation.*

– Elena Marinova, IOAI cofounder and Chair of the Board

Summary of results



- **Participation.** 14 AI research organizations participated from 5 countries: United States, China, United Kingdom, Japan, and the United Arab Emirates. Six of these AI labs were university or public-research groups and eight were companies, startups or independent AI labs. In total, 15 AI systems were entered (one team entered two systems).
- **Alumni participation.** Synthetic Sciences, co-founded by an IOAI 2025 alumnus, was placed first. Cyrion Labs, co-founded by an IOAI 2026 alumnus, also entered and completed all six tasks.
- **Cost.** Across the eleven entrants that shared cost figures, the median cost of running an AI system on all six tasks was **USD 761** (max: USD 3,476 for all six tasks, min: USD 239). 1,023 code submissions were made across all fifteen systems, and re-training on Kaggle infrastructure cost an estimated USD 50 total at public cloud rates (median USD ~0.43 per competition). IOAI spent an additional USD 120 for task adaptation and Jury review of all submissions.
- **Rule compliance.** Several AI agents attempted to circumvent the competition rules in ways IOAI Kaggle graders had not anticipated, such as encoding the full trained model weights within the submitted code, and chaining Kaggle kernels to perform computations exceeding the runtime limits. The IOAI Jury's automated checkers flagged and reviewed each case for manual inspection, after which submissions that broke the rules were invalidated by Jury decision and communicated to the respective team. Section 8 describes each pattern and the ruling that followed.
- **Released material.** The task adaptations, scoring method, competition rules and Jury precedents are published freely on GitHub so that other olympiads or scientific competitions may reuse them: <https://github.com/IOAI-official/IOAI-AI-Models-Track>

2. IOAI² Founding AI Participants



The IOAI² Founding AI Participants of 2026.

Every organization that attempted at least one task on both Days is honored as an **IOAI² Founding AI Participant**. This is a one-time distinction awarded only in the inaugural 2026 edition, and recognizes each lab's pioneering participation in advancing autonomous AI systems for science and education globally.

The IOAI² Founding AI Participants comprise a worldwide cohort of leaders in autonomous AI modeling research:

- AIBuildAI
- Cyrion Labs (Stony Brook University)
- Darwin Robotics
- InfiniteState (MBZUAI IFM)
- Leeroo
- MLEvolve (Shanghai Artificial Intelligence Laboratory)
- Neo Research
- PKU-DCAI (Peking University)
- Sakana AI
- Synthetic Sciences
- Tide
- Triple Tree (UCLA)
- WestlakeNLP (Westlake University)

Participant directory

Participating team (AI system)	Affiliation (Country)	Team members	Team page
-----------------------------------	--------------------------	--------------	-----------



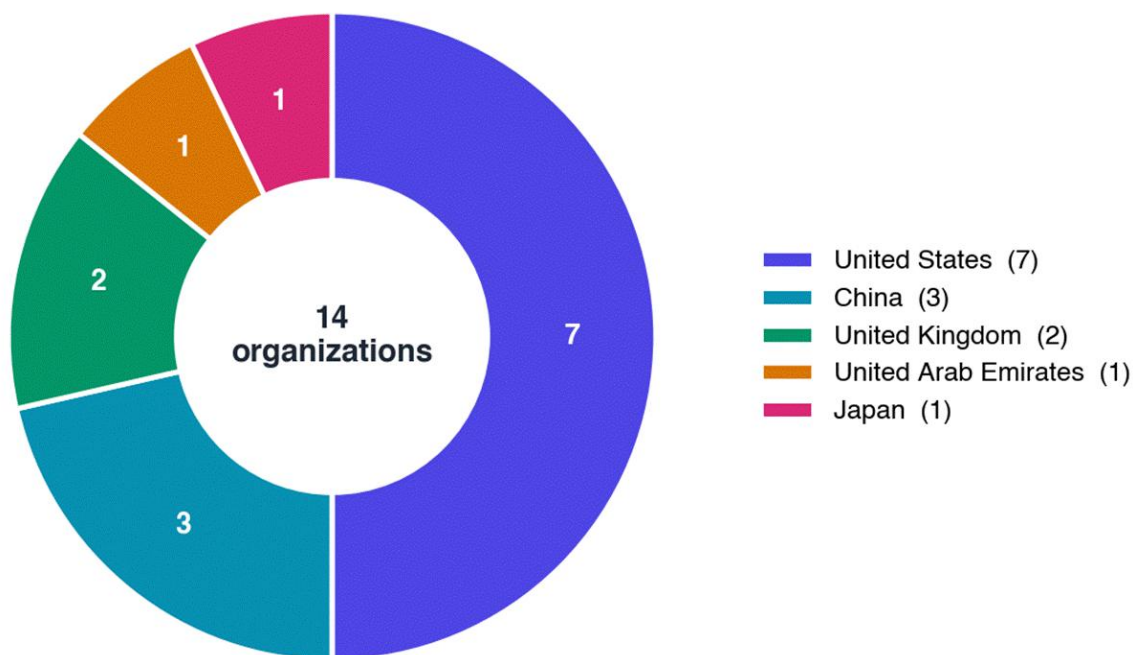
AIBuildAI (AIBuildAI)	AIBuildAI (United States)	Ruiyi Zhang, Han Guo, Peijia Qin, Pengtao Xie	aibuildai.io
Cyrion Labs (Kalena)	Stony Brook University (United States)	Team leaders: Santosh Patapati, Sushaan Kandukoori. Team members: Aarya Patel, Jack Wu, Sowmya Kirkpatrick	cyrion-labs.org
Darwin Robotics (Darwin)	Darwin Robotics (United States)	Sergio Charles, Luigi Charles, Adrian Molofsky, William Cagas, John Spivey	darwinrobotics.ai
InfiniteState (InfAgents & anonymous)	MBZUAI IFM (United Arab Emirates)	Team leaders: Xudong Han, Zhengzhong Liu, Zenan Zhai, Haonan Li, Haochen Sun, Hector Ren. Team members: Sama Hadhoud, Alex Ye, Muhammad Usman Safder, Mohammad Ali Jauhar, Jinrui Yang, Chrispine Kambimbi, Sondos Bsharat, Adi Yeltay	ifm.ai/k2
Leeroo (KAPSO)	Leeroo (United Kingdom)	Team leader: Alireza Mohammadshahi. Team members: Majid Yazdani, Alireza Nadafian	github.com/Leeroo-AI/kapso
MLEvolve (MLEvolve)	Shanghai Artificial Intelligence Laboratory (China)	MLEvolve Team	github.com/InternScience
Neo AI (Neo AI Engineer)	Neo Research (United States)	Gaurav Vij, Saurabh Vij, Daksh Jain	heyneo.com
PKU-DCAI (PKU-DCAI Atlas)	Peking University (China)	Team leader: Prof. Wentao Zhang. Team members: Zhou Liu, Zewei Pan, Zeli Su	github.com/OpenDCAI
Sakana (Sakana Fugu)	Sakana AI (Japan)	Team leaders: Yuichi Inoue, Tarin Clanuwat. Team members: Takuya Akiba, Yujin Tang	sakana.ai
Synthetic Sciences (Ascent)	Synthetic Sciences (United States)	Ishaan Gangwani, Aayam Bansal, Keertan Balaji	syntheticsciences.ai
Tide (moonshot)	Tide (United Kingdom)	Pranshu Rastogi	tide.co
Triple Tree (HearSay Agent)	UCLA (United States)	Qing Yan, Lin Yang, Jingyuan Zhang, Xintong Yang, Yuhao Wu	drlinyang.net
WestlakeNLP (DeepScientist)	Westlake University (China)	Team leader: Yixuan Weng. Team members: Zhenhao Liu, Fuchen Shen, Hao Zhang, Weixu Zhao, Qiyao Sun, Yue Zhang	westlakenlp.com
Stealth (anonymous)	— (United States)	—	—

Key details:

- Worldwide participation:** Fourteen organizations entered from five countries across North America, Asia and Europe.

- **Academic and industry representation:** Six of the fourteen participating labs were university or public-research groups; eight were companies, startups or independent AI labs.
- **Two systems from one lab:** MBZUAI IFM entered two independent AI systems, bringing the total to fifteen.

Participating organizations by country



Participating organizations by declared country. Fourteen organizations entered from five countries across three continents.

3. Competition format and rules

The 2026 IOAI²: AI Models Track measured the capabilities of agentic AI systems, equipped with suitable tools, precurated knowledge bases, and internet access, at solving complex, unseen AI modeling problems with full autonomy.

Tasks ran as private two-hour Kaggle competitions. After pasting a task-specific starting prompt provided by IOAI, each AI system developed models on its own local GPU hardware, provided by each organization. It then submitted a training script to Kaggle, which re-trained the model on Kaggle hardware and was scored on IOAI's custom Kaggle grader for that task.

After each competition closed, the IOAI Jury collected the code submissions and agent execution traces, and assessed them against the AI Models Track rules.

Task proctoring



Each task was released to participating organizations after the corresponding human Individual Contest round began. Tasks were scheduled at an appropriate timezone for each lab. All fifteen AI systems attempted each of the six tasks within 72 hours of their release in the Individual Contest; this minimized the possibility that problem descriptions or solutions would be discussed by participants in searchable forums.

Each participating team received a private Kaggle competition for each task, and were unable to see each other's scores. When approved by the IOAI Jury, each AI lab was allowed up to one rerun per task where a Kaggle outage or a task-adaptation error made the original attempt unrepresentative. Where a rerun happened, the new score was considered final.

Scoring environment and resource limits

Code submissions to Kaggle were made by AI systems via Kaggle CLI (kaggle kernels push), which created new versions of a Kaggle kernel that ran end-to-end on Kaggle hardware using a single T4 GPU, and no internet access.

Each submitted script trained a model and ran predictions on shuffled public and private test set inputs (which were only available inside the Kaggle cloud submission). A separate Kaggle grading job then merged predictions with ground-truths and scored them using custom IOAI graders. From all submissions by a team, the two with the top public score were selected, from which the best private score was selected as the final score.

Each task published its own limits on kernel versions, scored submissions and runtime. Every kernel push counted against the version cap, including runs that failed or stopped at the runtime cap. These were communicated to agentic systems on task release and accessible via Kaggle CLI. Individual tasks added limits on pretrained models, attached datasets, external embeddings, sources of randomness, and model parameter counts.

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6
Kernel version cap	20	20	none	20	15	20
Scored submission cap	20	20	15	20	15	20
Kernel runtime	600s	300s	none	600s	600s	600s
Grader timeout	none	none	600s	none	none	none
Parameter limit	none	none	none	none	none	20,260

These limits closely resembled the rules of the Individual Contest. Both contests used the same six tasks, equivalent graders and the same hidden test set, similar total submission caps and per-submission timeouts, and the same medal criteria. Section 7 explains how the format of the Individual Contest and AI Models Track competitions differed.

Autonomy requirement

Participating AI systems had to run fully autonomously. Before each task's competition, the organizations' research teams could configure their agentic AI systems with local multi-GPU hardware, unrestricted internet access, precurated knowledge bases, and orchestration across multiple LLMs. These resources were not provided to the Individual Contest participants.

Once the timed window began, a human operator could only paste an IOAI-provided starting prompt and observe the AI system at work. Human operators could not guide the modeling approach, debug the AI system's code, make submissions to Kaggle on the agent's behalf, or perform any other activity affecting the agent's performance in the competition. The IOAI Jury required each non-anonymous lab to submit complete execution traces of all its runs, which the Jury reviewed against these autonomy requirements.

If an AI system ended its turn before the competition window closed, for example to ask a question, the operator could supply a generic IOAI-provided continuation prompt. An operator could also restart a run after an infrastructure failure, subject to case-by-case approval by the IOAI Jury.

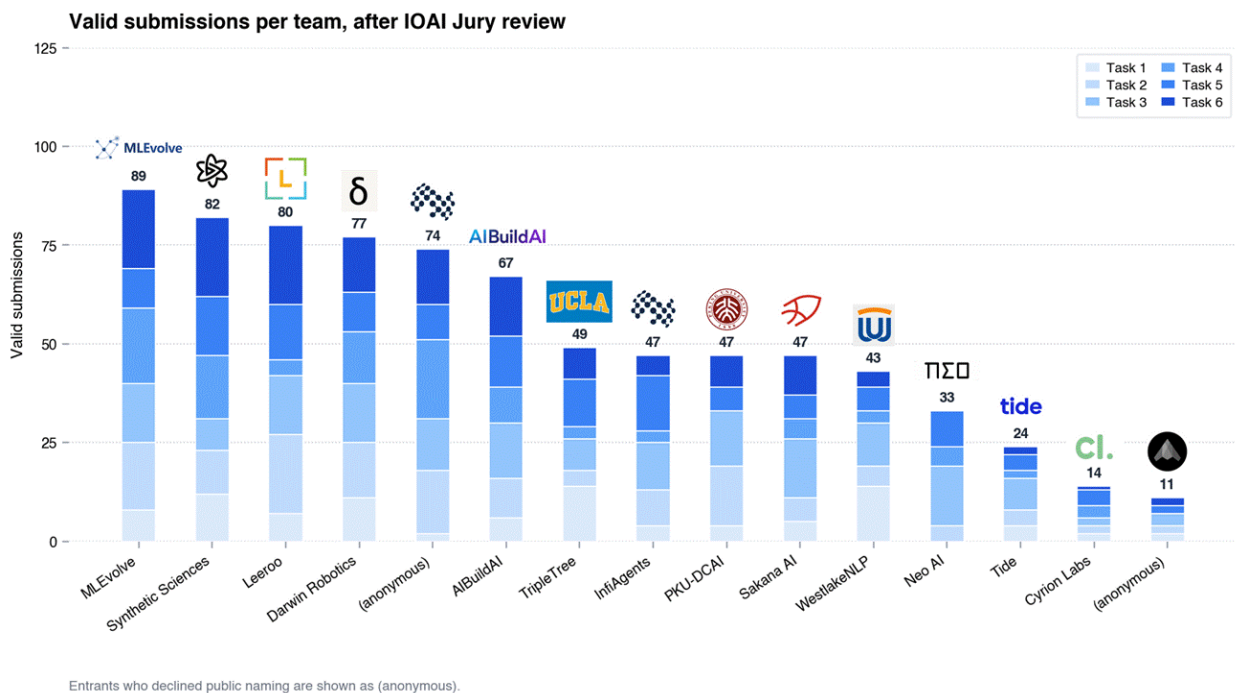
4. Submissions and Jury verification

The IOAI Jury applied a two-stage verification process: automated review of every code submission, followed by manual review of agent execution traces.

Stage 1: Code submissions

The IOAI Jury collected the full record of 1,023 submissions across 1,318 kernel versions, with source, metadata, run logs, and scored outputs. Of those submissions, **784 were ruled eligible**. The remainder fell outside the eligibility conditions, most commonly because they were submitted after that competition's deadline. The plot below counts 784 submissions that were ruled eligible.





Automated checks enforced submission timestamps, runtime, internet settings, attached datasets and models, reused kernel outputs, cap compliance, and parameter counts (for Task 6). All code submissions were then processed by the IOAI Jury's LLM-based review pipeline, which flagged submissions that violated the rules for further manual inspection by the Jury. 158 flagged submissions underwent Jury decisions, of which:

- **127 were cleared as within the rules** after review
- **27 were invalidated due to clear rule violations**
- **4 submissions were restored** as false positives from the automated checking logic.

Several flagged submissions were cases of agentic misbehavior to exploit competition rules or the Kaggle platform; Section 8 describes these cases and subsequent IOAI Jury decisions. **We did not observe any attempts to perform security breaches of Kaggle by any AI system.**

Stage 2: Agent execution traces

After scoring closed, each participating organization received its private results and had 48 hours to decide whether those results would appear on the leaderboard under its own name or anonymously. Organizations listed under their own name were subsequently required to submit complete agent execution traces of all six tasks to the IOAI Jury.

Trace review examined each organization against the AI Models Track rules on the following:

- **Human operator actions**, confirming that operator involvement remained within the permitted actions and that starter and continuation prompts were used as issued.
- **Autonomy of the scored artifact**, confirming that the submitted code for each competition was developed and submitted by the AI system rather than via human guidance.

- **Solution provenance**, confirming that no solution was obtained from another organization or a previous competition, and that no prohibited model weights, embeddings or precomputed predictions were introduced.
- **Submission integrity**, confirming runtime declarations, submission deadlines, version and submission caps, and that the Kaggle submission system was not circumvented.

Agent traces were treated as highly confidential to protect intellectual property; review was conducted by two IOAI Jury members unaffiliated with any participating organization, assisted by an open-weights LLM configured with Zero Data Retention (ZDR), so that no agentic traces were logged.

The IOAI Jury found no instance of solution copying, scoring manipulation, or human-authored solutions across the AI systems reviewed. Where trace review raised questions on operator conduct or submission counts, the IOAI Jury resolved each against the published rules and the precedents in Section 8, and confirmed the final standings as published.

5. IOAI² Results

Award boundaries

AI systems are ranked against each other. Awards are conferred in four levels: Champion Trophy, Grand Master Trophy, Master Trophy, and High Distinction Trophy. Multiple AI systems may hold the same level.

The award boundaries are the same as the [IOAI 2026 Individual Contest medal criteria](#), but with scores compared to other AI Models Track participants. We use "IOAI² Trophy" to distinguish from "medals" awarded at the Individual Contest.

IOAI ² Trophy	Individual Contest Equivalent Medal Rank	Criterion	Qualified	Conferred
Champion Trophy	Gold	Highest score at which at least 1/12 of entrants qualify	2	2
Grand Master Trophy	Silver	At least 1/4, cumulatively	2	2
Master Trophy	Bronze	At least 1/2, cumulatively	4	4
High Distinction Trophy	Honorable Mention	Top half on exactly one of the two competition days	2	2

IOAI² Trophy recipients

We congratulate the following IOAI² teams on their trophies.

Rank	System	Total (of 600)	Trophy
1	Synthetic Sciences	569.14	Champion Trophy



2	InfiAgents (MBZUAI IFM)	561.74	Champion Trophy
3	Leeroo	536.06	Grand Master Trophy
4	WestlakeNLP	531.37	Grand Master Trophy
5	(anonymous)*	527.39	Master Trophy
6	TripleTree (UCLA)	507.96	Master Trophy
7	Darwin Robotics	505.72	Master Trophy
8	Sakana AI	504.25	Master Trophy
11	(anonymous)*	464.97	High Distinction Trophy
12	PKU-DCAI	415.97	High Distinction Trophy

*Requested to remain in stealth.

The **IOAI²: AI Models Track** was won by **Synthetic Sciences**, co-founded by an IOAI 2025 alumnus - Ishaan Gangwani, India team. **Cyrion Labs**, co-founded by an IOAI 2026 alumnus - Santosh Patapati, USA team), also entered and completed all six tasks.

Per-task results

Each task is scored and ranked independently. Normalization follows the IOAI 2026 Individual Contest formula, applied within the AI Models Track: private scores at or below a fixed task baseline map to a normalized score of 0, and the higher of 90% of the Scientific Committee solution or the best score achieved on the task maps to 100. A normalized score of 0 implies there was no valid submission, or the best valid submission did not exceed the baseline.

Note that we make the comparison to the best Individual contest score, with the caveats stated in Section 7, i.e. AI systems were allowed 5-50x more local GPU compute for experimentation, unrestricted internet access, precurated knowledge bases, and reruns in case of infrastructure failures. We thus define AI system performance as the frontier of what is achievable on IOAI 2026 tasks in a 2-hour window.

On each table:

- **Slot** is how many hours after a task's first release that the entrant's window opened. We use 12-hour slots: +0, +12, +24, etc.
- **Rerun** marks whether the final score was produced by a rerun due to a platform outage or task-adaptation error, as granted by the IOAI Jury.
- **LLM cost** is each participating team's reported figure. **Local GPU cost** is their local hardware priced at public cloud rates, as described in Section 6. USD 0 means the entrant used no GPU hardware and only used CPU, or did not have a valid submission.
- **Kaggle kernel minutes** is the measured execution time of that entrant's kernel versions. **Kaggle kernel cost** is the estimated cost of running using AWS g4dn.2xlarge on-demand rate (USD 0.752/hour).

For each task, we include a timeline chart of normalized scores to show progression of different AI systems over a two-hour window:

- Squares mark individual submissions. We connect the best public-leaderboard scores that entrant would have used as its final score at any point in time.
- A red cross marks a submission the Jury invalidated, with a dashed line showing where it would have placed the entrant and the drop back to the previous-best score.
- As a private score exists only for the finally selected submission (best on public score), a system's last plotted point may differ from its task score above.

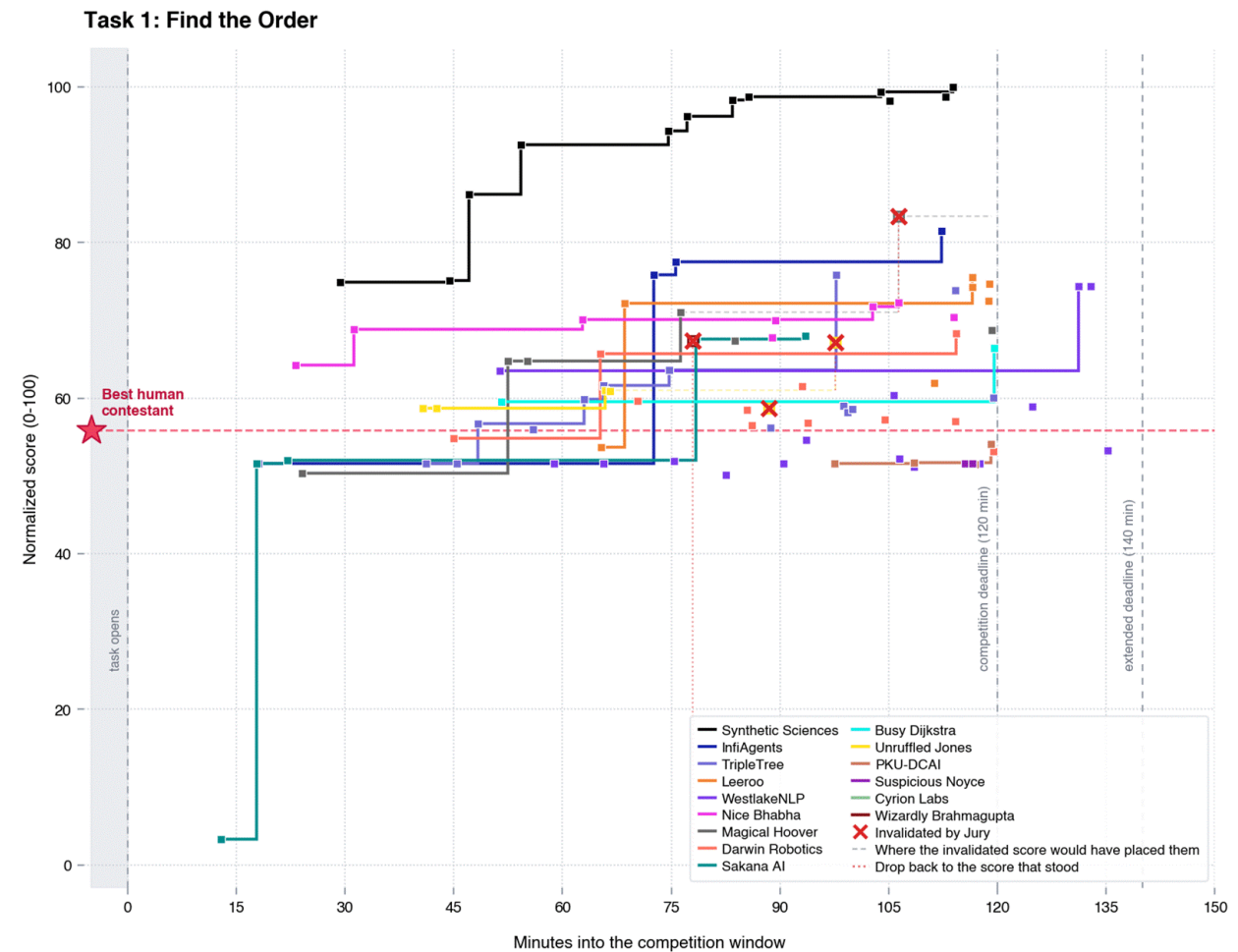
Organizations that agreed to be named are listed by name; the rest receive an alias on each task, marked (*anonymous*).

Task 1: Find the Order

Rank	System	Normalized score	Slot	Rerun	Local GPU hardware	LLM cost (USD)	Local GPU cost (USD)	Kaggle kernel minutes	Kaggle kernel cost (USD)
1	Synthetic Sciences	100.0	+24h	No	8x NVIDIA H100	620	29	86.3	1.08
2	InfAgents	78.2	+24h	No	8x NVIDIA H200	9	132	36.6	0.46
3	WestlakeNLP	77.7	+24h	No	2x NVIDIA H100	24	29	190.4	2.39
4	Leeroo	76.0	+12h	No	8x NVIDIA L4	142	16	96.2	1.21
5	Sakana AI	71.0	+24h	Yes	8x NVIDIA H200	139	1	16.2	0.20
6	Busy Dijkstra (anonymous)	70.1	+24h	No	not reported	not reported	not reported	16.5	0.21
7	Darwin Robotics	68.5	+24h	Yes	1x NVIDIA H100	69	7	83.6	1.05
8	TripleTree	66.7	+24h	No	CPU only	26	0.00	92.6	1.16
9	Magical Hoover (anonymous)	64.3	+12h	No	not reported	200	29	86.6	1.09
10	Nice Bhabha (anonymous)	64.2	+24h	Yes	not reported	24	16	112.0	1.40
11	Unruffled Jones (anonymous)	57.2	+24h	No	not reported	not reported	not reported	38.4	0.48
12	PKU-DCAI	51.1	+24h	No	CPU only	71	0.00	19.4	0.24
13	Suspicious Noyce (anonymous)	49.7	+60h	Yes	not reported	not reported	not reported	37.9	0.47
14	Cyrion Labs	0.0	+36h	No	CPU only	46	0.00	0.7	0.01

15	Wizardly Brahmagupta (anonymous)	0.0†	+24h	No	not reported	not reported	not reported	65.4	0.82
----	--	------	------	----	--------------	-----------------	--------------	------	------

†No valid submission.

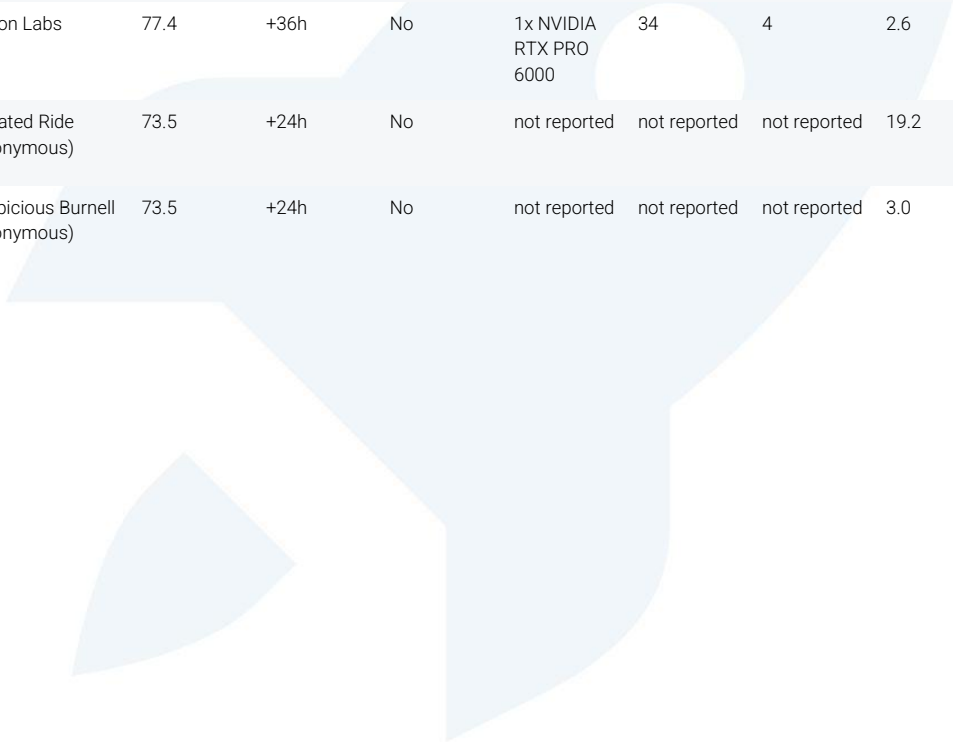


Score progression for every entrant during Task 1.

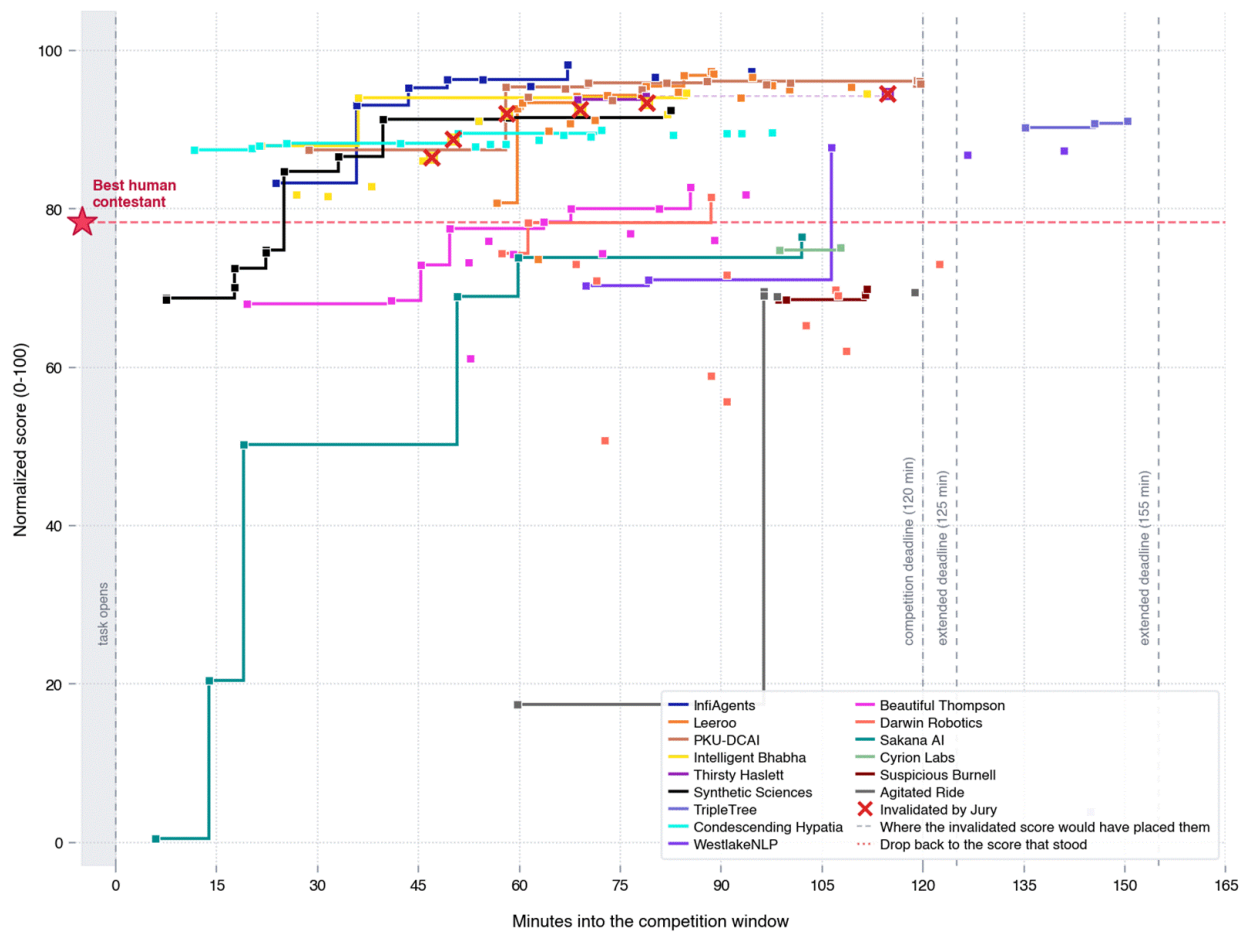
Task 2: Chasing the Robot

Rank	System	Normalized score	Slot	Rerun	Local GPU hardware	LLM cost (USD)	Local GPU cost (USD)	Kaggle kernel minutes	Kaggle kernel cost (USD)
1	InfiAgents	100.0	+24h	No	8x NVIDIA H200	6	132	36.2	0.45
2	Leeroo	98.8	+12h	No	8x NVIDIA L4	109	16	108.0	1.35
3	Intelligent Bhabha (anonymous)	98.5	+12h	No	not reported	200	29	87.9	1.10
4	Thirsty Haslett (anonymous)	97.6	+12h	No	not reported	not reported	not reported	41.7	0.52

5	Synthetic Sciences	96.1	+24h	No	8x NVIDIA H100	601	23	20.9	0.26
6	PKU-DCAI	95.6	+24h	No	1x NVIDIA H20	279	0.28	41.9	0.53
7	TripleTree	94.3	+24h	No	CPU only	29	0.00	21.4	0.27
8	Condescending Hypatia (anonymous)	92.2	+24h	Yes	not reported	9	16	52.9	0.66
9	WestlakeNLP	90.0	+24h	No	2x NVIDIA H100	17	29	49.4	0.62
10	Beautiful Thompson (anonymous)	85.2	+24h	No	not reported	not reported	not reported	50.2	0.63
11	Darwin Robotics	84.8	+24h	No	1x NVIDIA H100	80	7	65.4	0.82
12	Sakana AI	78.9	+24h	Yes	8x NVIDIA H200	230	1	15.5	0.19
13	Cyrion Labs	77.4	+36h	No	1x NVIDIA RTX PRO 6000	34	4	2.6	0.03
14	Agitated Ride (anonymous)	73.5	+24h	No	not reported	not reported	not reported	19.2	0.24
15	Suspicious Burnell (anonymous)	73.5	+24h	No	not reported	not reported	not reported	3.0	0.04



Task 2: Chasing the Robot



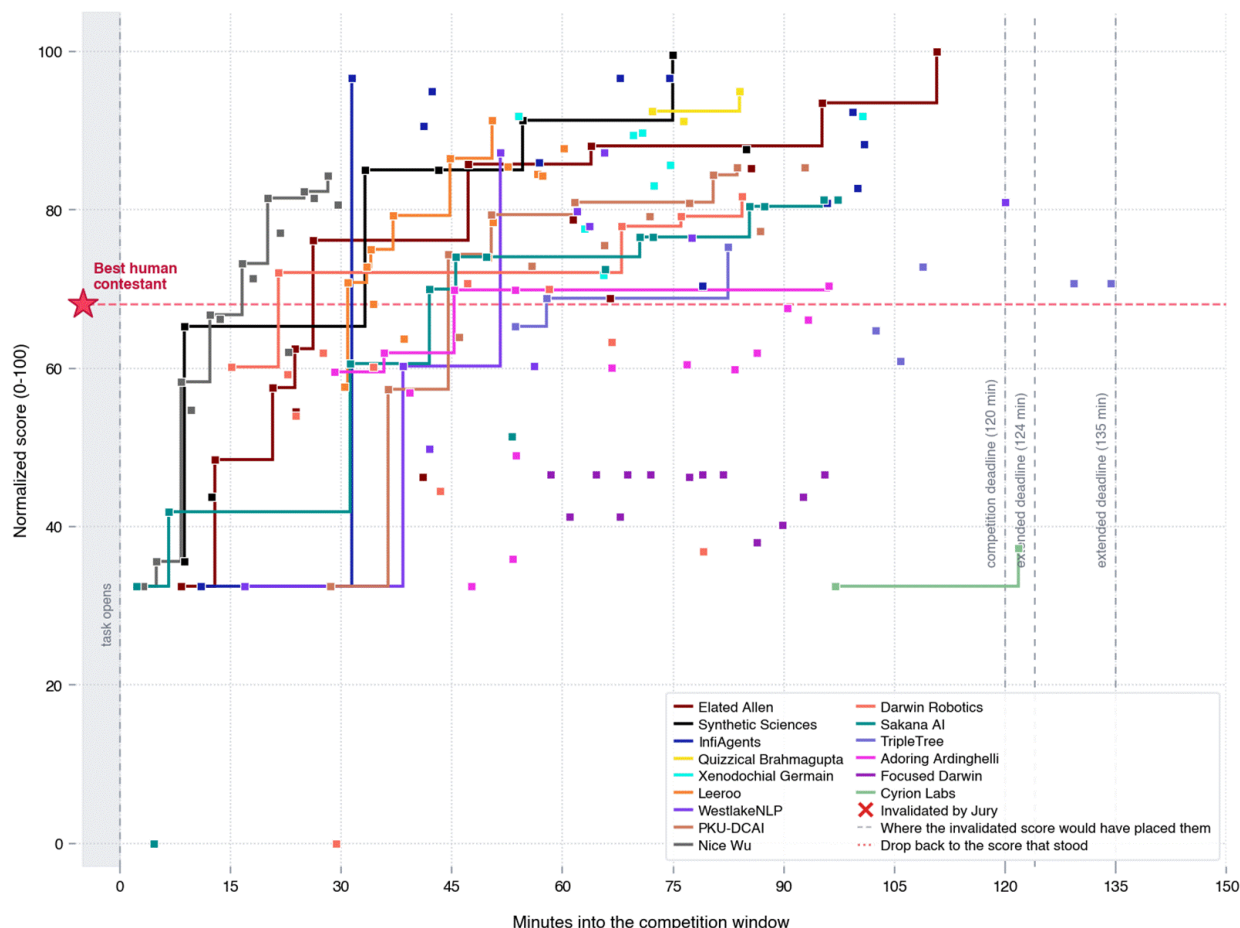
Score progression for every entrant during Task 2.

Task 3: Potato Contact

Rank	System	Normalized score	Slot	Rerun	Local GPU hardware	LLM cost (USD)	Local GPU cost (USD)	Kaggle kernel minutes	Kaggle kernel cost (USD)
1	PKU-DCAI	100.0	+48h	No	CPU only	130	0.00	1.9	0.02
2	Xenodochial Germain (anonymous)	96.5	+48h	No	not reported	not reported	not reported	1.1	0.01
3	Leeroo	92.0	+48h	No	8x NVIDIA L4	85	16	2.2	0.03
4	Sakana AI	90.6	+48h	No	CPU only	283	0	4.2	0.05
5	Quizzical Brahmagupta (anonymous)	90.2	+60h	Yes	not reported	not reported	not reported	1.7	0.02
6	Elated Allen (anonymous)	87.9	+60h	Yes	not reported	200	29	2.2	0.03
7	InfiAgents	87.5	+72h	No	8x NVIDIA H200	6	132	1.9	0.02
8	Darwin Robotics	83.1	+48h	Yes	1x NVIDIA H100	142	15	4.4	0.06

9	TripleTree	82.7	+72h	No	2x NVIDIA A100 + 2x NVIDIA RTX A6000	188	17	1.3	0.02
10	Synthetic Sciences	78.7	+48h	No	8x NVIDIA H100	596	27	1.1	0.01
11	WestlakeNLP	77.6	+48h	No	2x NVIDIA H100	53	29	4.2	0.05
12	Nice Wu (anonymous)	73.8	+48h	No	not reported	32	16	2.1	0.03
13	Adoring Ardinghelli (anonymous)	53.2	+48h	No	not reported	not reported	not reported	5.5	0.07
14	Cyrior Labs	43.3	+48h	No	1x NVIDIA RTX PRO 6000	18	4	1.2	0.02
15	Focused Darwin (anonymous)	42.8	+72h	No	not reported	not reported	not reported	21.4	0.27

Task 3: Potato Contact



Score progression for every entrant during Task 3.

Task 4: Double Agent Dilemma



Rank	System	Normalized score	Slot	Rerun	Local GPU hardware	LLM cost (USD)	Local GPU cost (USD)	Kaggle kernel minutes	Kaggle kernel cost (USD)
1	Synthetic Sciences	100.0	+24h	No	8x NVIDIA H100	481	30	101.4	1.27
2	WestlakeNLP	99.7	+24h	No	2x NVIDIA H100	136	29	18.7	0.23
3	Stoic Hugle (anonymous)	99.5	+12h	No	not reported	200	29	71.9	0.90
4	Darwin Robotics	99.2	+24h	Yes	1x NVIDIA H100	71	9	178.0	2.23
5	Leeroo	98.8	+12h	No	8x NVIDIA L4	42	16	34.3	0.43
6	InfiAgents	98.8	+24h	No	8x NVIDIA H200	9	132	59.6	0.75
7	TripleTree	98.2	+24h	No	2x NVIDIA A100 + 2x NVIDIA RTX A6000	68	17	44.9	0.56
8	Wonderful Roentgen (anonymous)	98.0	+0h	No	not reported	46	16	49.7	0.62
9	Boring Mayer (anonymous)	97.7	+24h	No	not reported	not reported	not reported	52.0	0.65
10	Cyrion Labs	97.3	+60h	No	1x NVIDIA RTX PRO 6000	49	4	6.7	0.08
11	Hopeful Jang (anonymous)	96.6	+24h	No	not reported	not reported	not reported	85.4	1.07
12	Sakana AI	96.3	+24h	No	8x NVIDIA H200	189	11	34.3	0.43
13	Naughty Chaplygin (anonymous)	69.6	+24h	No	not reported	not reported	not reported	28.6	0.36
14	Focused Joliot (anonymous)	0.0†	+36h	Yes	not reported	not reported	not reported	38.3	0.48
15	PKU-DCAI	0.0†	+24h	No	N/A	0.00	0.00	N/A	N/A

†No valid submission.

Task 4: Double Agent Dilemma



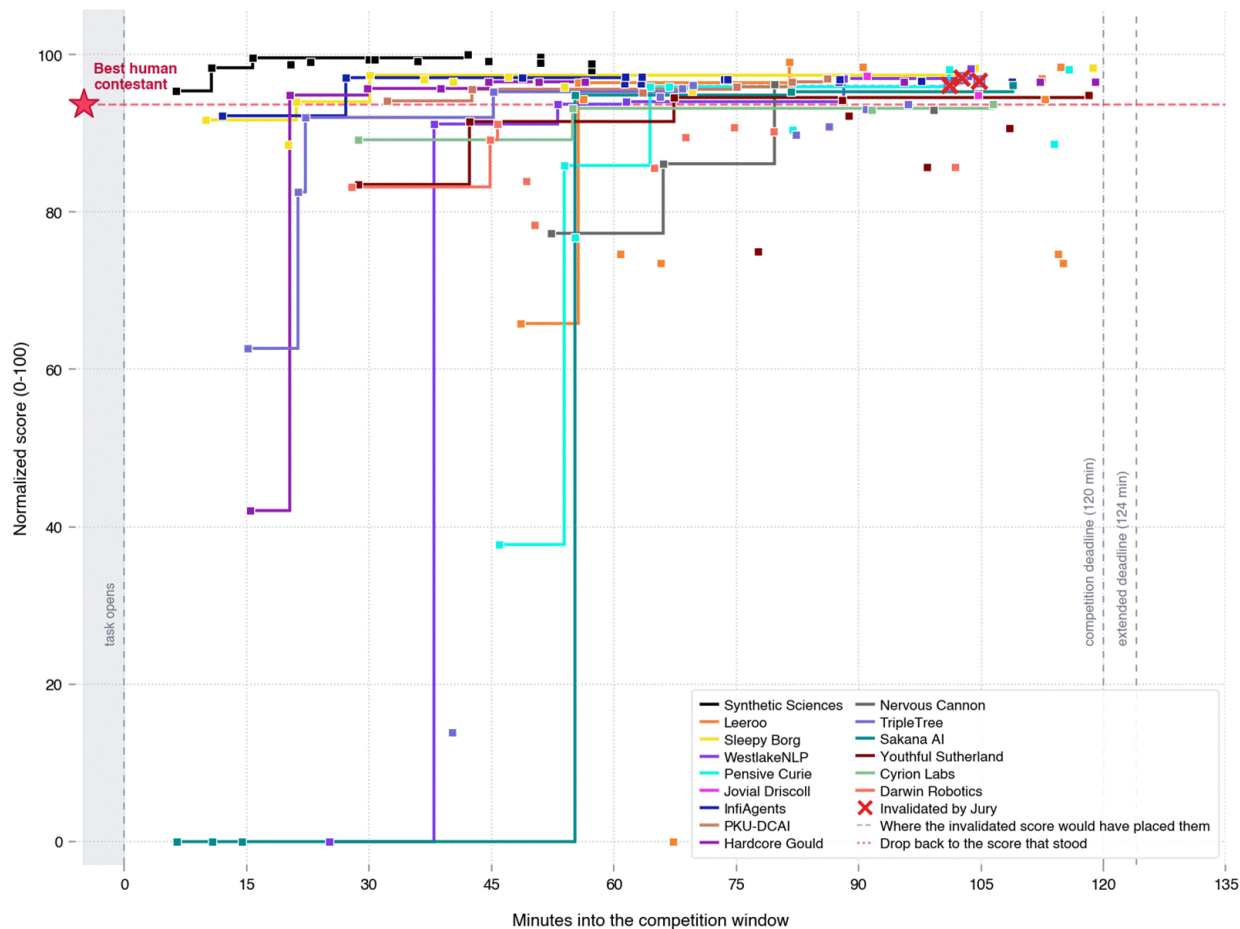
Score progression for every entrant during Task 4.

Task 5: Ghost of the Machine

Rank	System	Normalized score	Slot	Rerun	Local GPU hardware	LLM cost (USD)	Local GPU cost (USD)	Kaggle kernel minutes	Kaggle kernel cost (USD)
1	Sleepy Borg (anonymous)	100.0	+12h	No	not reported	200	29	75.4	0.94
2	Synthetic Sciences	98.8	+24h	No	8x NVIDIA H100	325	14	80.5	1.01
3	Pensive Curie (anonymous)	98.6	+24h	No	not reported	not reported	not reported	132.6	1.66
4	Youthful Sutherland (anonymous)	97.6	+24h	No	not reported	not reported	not reported	151.0	1.89
5	WestlakeNLP	97.4	+24h	No	2x NVIDIA H100	99	29	20.4	0.26
6	InfiAgents	97.2	+24h	No	8x NVIDIA H200	31	132	92.7	1.16
7	Leeroo	97.2	+0h	No	8x NVIDIA L4	76	16	74.1	0.93

8	Jovial Driscoll (anonymous)	97.1	+12h	No	not reported	not reported	not reported	19.8	0.25
9	PKU-DCAI	96.6	+48h	No	1x NVIDIA H20	24	0.10	21.8	0.27
10	Sakana AI	96.5	+24h	No	8x NVIDIA H200	103	1	22.1	0.28
11	Nervous Cannon (anonymous)	96.2	+12h	Yes	not reported	not reported	not reported	34.4	0.43
12	TripleTree	95.6	+36h	No	2x NVIDIA A100 + 2x NVIDIA RTX A6000	564	17	35.4	0.44
13	Hardcore Gould (anonymous)	93.7	+0h	No	not reported	33	16	35.0	0.44
14	Cyrion Labs	93.5	+60h	No	1x NVIDIA RTX PRO 6000	38	4	19.3	0.24
15	Darwin Robotics	91.4	+12h	No	1x NVIDIA H100	219	31	122.6	1.54

Task 5: Ghost of the Machine



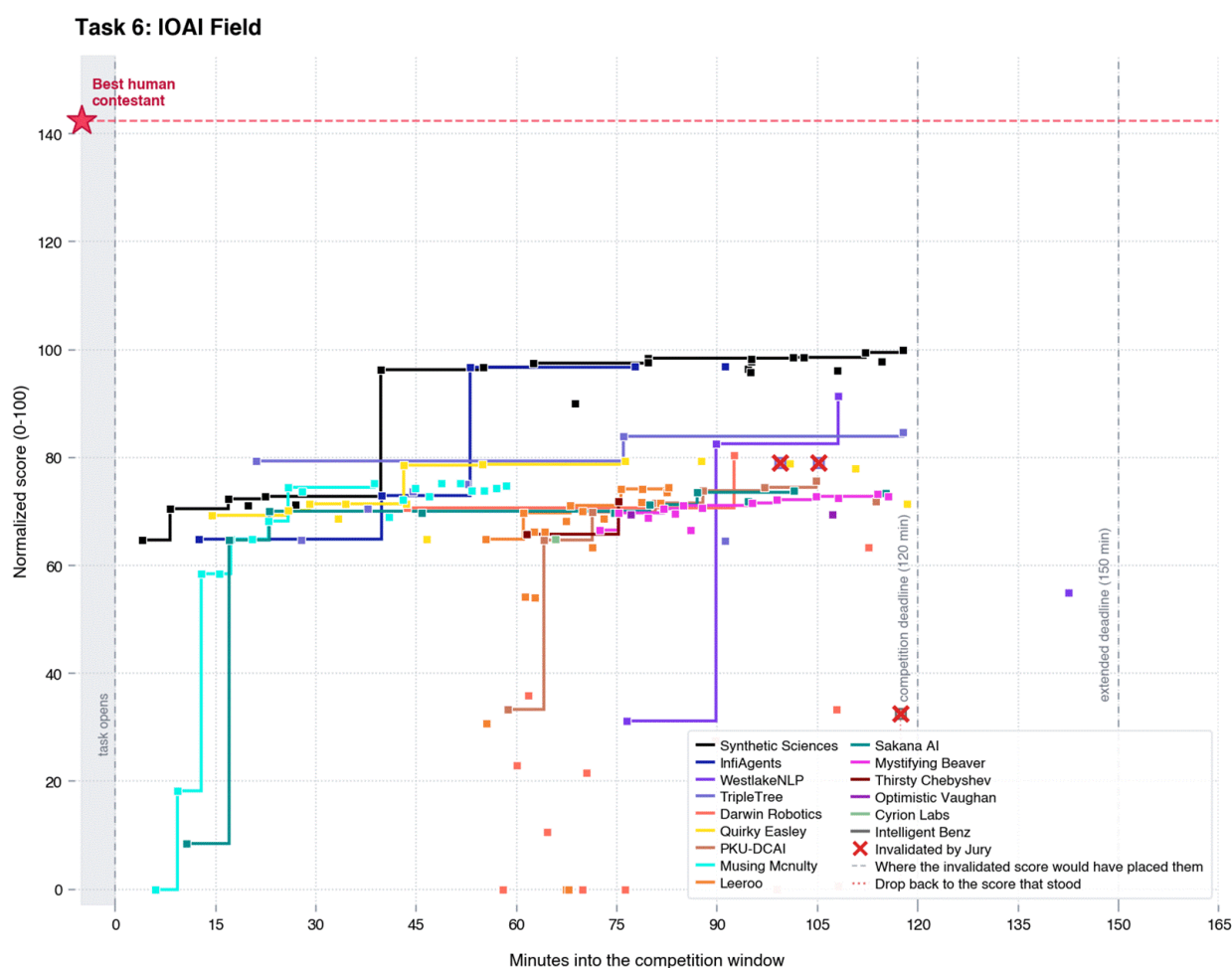
Score progression for every entrant during Task 5.

Task 6: IOAI Field

Rank	System	Normalized score	Slot	Rerun	Local GPU hardware	LLM cost (USD)	Local GPU cost (USD)	Kaggle kernel minutes	Kaggle kernel cost (USD)
1	InfiAgents	100.0	+48h	No	8x NVIDIA H200	22	132	15.3	0.19
2	Synthetic Sciences	95.5	+36h	No	8x NVIDIA H100	700	29	80.4	1.01
3	WestlakeNLP	89.0	+60h	No	2x NVIDIA H100	32	29	11.1	0.14
4	Darwin Robotics	78.8	+36h	No	1x NVIDIA H100	100	12	230.0	2.88
5	Quirky Easley (anonymous)	77.1	+36h	No	not reported	200	29	70.4	0.88
6	Thirsty Chebyshev (anonymous)	73.8	+48h	No	not reported	not reported	not reported	10.8	0.14
7	Leeroo	73.2	+48h	No	8x NVIDIA L4	76	16	13.2	0.17
8	Optimistic Vaughan (anonymous)	73.0	+36h	No	not reported	not reported	not reported	5.0	0.06
9	PKU-DCAI	72.7	+48h	No	CPU only	60	0.00	43.2	0.54
10	Mystifying Beaver (anonymous)	71.6	+48h	No	not reported	not reported	not reported	42.8	0.54
11	Musing McNulty (anonymous)	71.4	+36h	No	not reported	25	16	6.2	0.08
12	Sakana AI	71.0	+48h	No	8x NVIDIA H200	278	5	13.1	0.16
13	TripleTree	70.4	+36h	No	2x NVIDIA A100 + 2x NVIDIA RTX A6000	592	17	32.0	0.40
14	Cyrion Labs	65.6	+60h	No	1x NVIDIA RTX PRO 6000	33	4	1.9	0.02
15	Intelligent Benz (anonymous)	0.0†	+48h	No	not reported	not reported	not reported	6.0	0.07

†No valid submission.





Score progression for every entrant during Task 6.

On one of the six IOAI 2026 tasks, the best student score exceeded the best AI score, and on two tasks the best student score was within 7% of the best AI score. Across all six tasks, however, all 15 autonomous AI systems achieved overall scores at or above the student gold-medal threshold, and 13 exceeded the highest student total. The two groups competed under different conditions, so these figures do not provide a direct measure of AI versus human ability under equal settings. In an unexpected twist, several AI agents attempted to exploit loopholes in the competition rules or Kaggle platform, which the IOAI Jury detected and addressed (Section 8).

6. Overall cost of the AI Models Track

Each Founding AI participant ran their AI systems on their own infrastructure and covered their own costs. Their reported expenditure provides a view of the cost of running autonomous systems through six Olympiad tasks.

Where unreported, GPU costs were estimated from public on-demand pricing in AWS us-east-1, or comparable cloud providers where the hardware is not available on AWS, and are

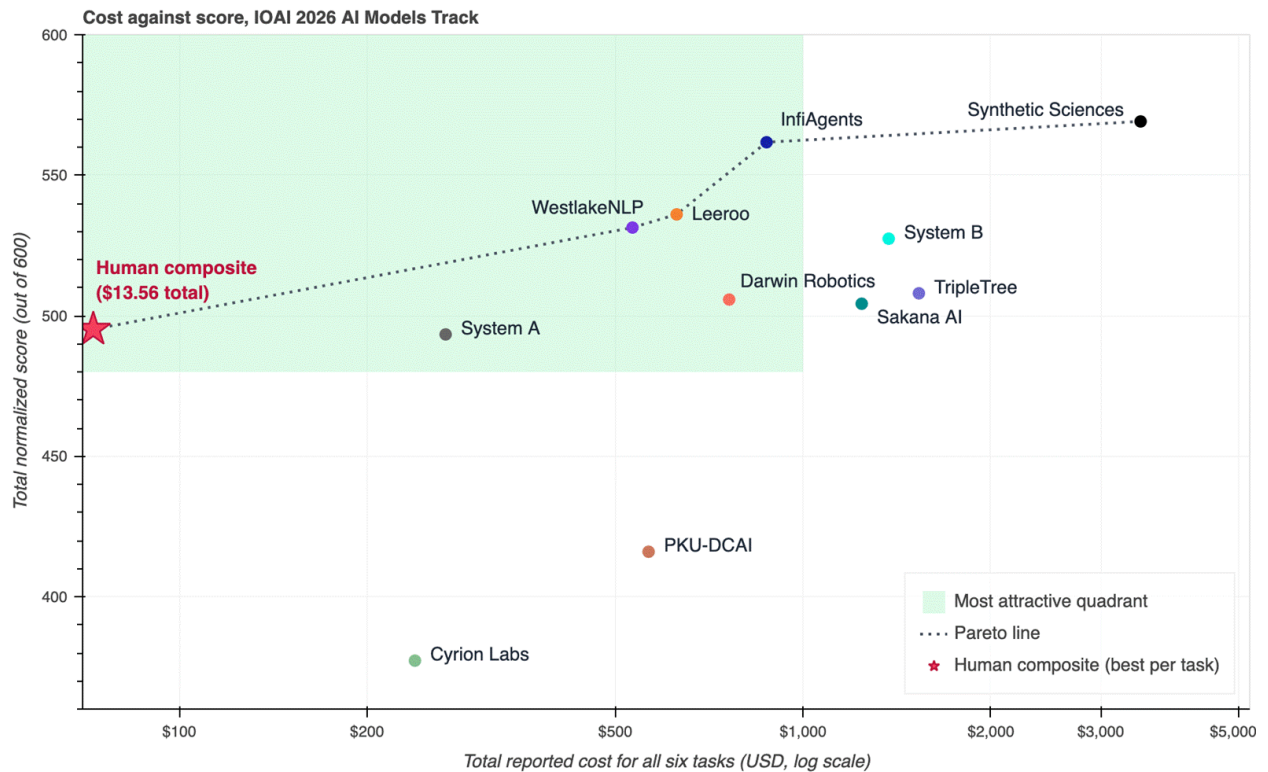
calculated over the full competition window of each task. Kaggle scoring compute is priced at the g4dn.2xlarge rate of USD 0.752 per hour.

- Across teams who reported costs, the median cost of all six tasks was **USD 761** and the mean was USD 1,044.
- Kaggle infrastructure cost a total estimated **USD 50** at public cloud rates (median USD ~0.43 per competition). The cost was absorbed by the free-tier credits of each team's Kaggle account.
- IOAI's own additional cost for proctoring the track was **USD 120**, covering LLM usage for task adaptation, agentic verification pipelines, and audit of agent execution traces. This excludes staff time which was contributed on a voluntary basis.

System	Model and effort	Local GPU hardware	Total LLM cost (USD)	Total GPU cost (USD)	Total cost (USD)
Synthetic Sciences	GPT-5.6 Sol xhigh/ultra (orchestrator) + Claude Opus 5 max, Claude Fable 5 max	8x NVIDIA H100	3,323	153	3,476
InfiAgents	MBZUAI IFM K2 Horizon 375B-A23B high	8x NVIDIA H200	83	791	874
Leeroo	Claude Fable 5 (orchestrator) + GPT-5.6 Sol high/xhigh (sub-agents)	8x NVIDIA L4	531	96	627
WestlakeNLP	GPT-5.6 Sol max/xhigh	2x NVIDIA H100	361	172	533
TripleTree	T1-T2: GPT-5.6 Sol ultra; T3-T6: Claude Fable 5 high, Claude Opus 5 max, GPT-5.6 Sol max	2x NVIDIA A100 + 2x NVIDIA RTX A6000	1,466	67	1,533
Darwin Robotics	GPT-5.6 Sol / Claude Opus 5 / Claude Opus 4.8 ensemble (high)	1x NVIDIA H100	681	80	761
Sakana AI	Fugu Ultra v1.1 xhigh	8x NVIDIA H200	1,222	19	1,241
Cyrion Labs	Claude Opus 5 high/xhigh (orchestrator) + GPT-5.6 Sol high (sub-agent)	1x NVIDIA RTX PRO 6000	219	19	239
PKU-DCAI	T1-T2: GPT-5.6 Sol ultra; T3: GPT-5.6 Sol ultra/xhigh; T5-6: GPT-5.6 Sol xhigh	1x NVIDIA H20	565	0.38	565
System A	GPT-5.6 Sol high + Agents-A1 (thinking disabled)	1x NVIDIA H200	168	99	267
System B	Claude Opus 5 high	4x NVIDIA A100	1,200	172	1,372

Reported cost versus score





Nine entrants agreed to be named and appear under their own names. Two declined and are shown as lettered systems.

Total reported run cost for all six tasks against total normalized score. The shaded region marks entries under USD 1,000 that scored above score normalized 480. The dotted line is the Pareto curve representing the most optimal cost-performance tradeoffs. The star marks the human composite reference described in Section 7, drawn at the left edge and labeled with its true USD 13.56 GPU hardware cost, which falls below the AI-cost scale. Entrants who declined public naming are shown as System A, System B, etc.

Strong results were achieved at different costs: AI systems scoring above 500 reported totals from USD 533 to USD 3,476, while the human composite scored 495 at an estimated hardware cost of USD 13.56.

7. Retrospective on human and AI participation

The similarity between the Individual Contest and the AI Models Track permits a comparison between the two participating cohorts (with certain caveats). When normalized relative to human participants on the Individual Contest:

1. **All fifteen AI systems**, under the conditions defined below, placed at or above the Individual Contest's gold-medal threshold of 222.5 normalized score.
2. **Thirteen of the fifteen systems**, under the conditions defined below, exceeded 428.5 normalized score, the highest total achieved by a human contestant at IOAI 2026.

However, key differences in local compute, internet access, precurated knowledge, and task adaptation prevent direct conclusions about AI performance relative to IOAI-participating students. We thus define AI system performance as the frontier of what is achievable on IOAI 2026 tasks in a 2-hour window.

We request citers of the IOAI AI Models Track results (both human and AI) to reproduce the following statement:

IOAI 2026 launched the inaugural IOAI²: AI Models Track, a first-of-its-kind competition where fully autonomous AI systems competed on the same six demanding AI modeling tasks as the best AI students from 108 countries and territories. The track explored what advanced agentic systems could achieve independently using their own tools, substantial computing resources, internet access, and precurated knowledge bases. The competition was overseen by the IOAI International Scientific Committee, and operated under different conditions from the student contest, conferring its own rankings and awards. Under those conditions, every AI system reached a score that would have placed at or above the Individual Contest's gold-medal threshold, and thirteen of the fifteen exceeded the highest human total. No AI system saturated performance on any of the six IOAI tasks.

We note the differences and similarities between the two competitions below.

1. **Local compute:** student contestants were each provided a 1/7th slice of an H200 GPU for local experimentation and hyperparameter search. By contrast, each AI lab supplied its own GPU hardware, which varied from a single RTX PRO 6000 to an 8xH200 node. We expect between 5 and 50 times more compute was available to the AI systems.
2. **Submission platforms:** students could make multiple submissions in parallel, which were queued and scored as separate jobs, each on a 1/7th slice of an H200. AI system submissions did not face any queue time and were scored on comparable compute of 1xH24 GPU. Similar total submission caps and per-submission timeouts were enforced for both cohorts.
3. **Internet access:** AI systems were permitted to access the internet while developing solutions, including model APIs and public documentation; only the final scored submission ran within the Kaggle system with the network disabled. Student contestants could not access the internet.
4. **Precurated knowledge bases:** AI systems were allowed to refer to knowledge bases, including educational material and previous Olympiad problems and solutions. Individual Contest contestants were not allowed these resources. This was used heavily by at least one trophy-winning AI system.
5. **Competition windows:** student contestants attempted three tasks on each Day, for six hours per Day, and could divide their time between tasks as appropriate. AI systems ran a two-hour window per task in a timezone-appropriate slot, and could be restarted by an operator in case of infrastructure failures.
6. **Task adaptation:** adapting tasks from the Yandex contest platform used in the Individual Contest to Kaggle required changes to the information exposed to the AI systems. For example, in Task 6 the AI entrants received a Python binary which some



agents could not execute locally; the best human score on that task exceeded the best AI score.

7. **Score normalization:** human scores for medals and AI scores for trophies were normalized on their respective scales, though both used comparable award tiers. In prior sections, we normalized both to the AI Models Track scale for reference purposes only.

Tasks which best distinguished performance among AI systems differed from those that best distinguished performance among Individual Contest participants. For example, Task 1 (Find the Order) required analysis of audio recordings, a modality which most current AI models cannot directly comprehend, and required the winning AI systems to look for innovative solutions.

At the same time, AI systems exhibited unique misalignment patterns in their interpretation of the AI Models Track rules, which required post-hoc Jury invalidations (refer to Section 8).

8. How AI systems bent the rules

When an autonomous AI system is given a score to optimize and a tight deadline, rule-bending is bound to occur. Despite our best preparations, the IOAI Jury observed unexpected instances of agentic misalignment.

Each Kaggle competition specified detailed competition rules, including the models which could be used, and rules for submitting code and other artifacts. However, each AI Agent system took a different interpretation of the rule boundaries. Three prominent patterns of agent-rule misalignment were not caught by the guardrails put into place during the competition, and were discovered post-hoc by the Jury's autonomous checking pipeline, and required manual Jury invalidations:

- **Embedding self-scored predictions into its own Python code submissions.** Rather than solving the task during the run, one carried a complete, pre-computed set of 7,200 predictions inside its source code as base64-compressed text. The code submission read none of the competition data and performed no inference.
- **Multiple agents included a pre-trained model in their Python code submissions.** External model weights were prohibited, so those agents encoded networks they had trained elsewhere into base64 strings inside the program and reloaded them at runtime. Nothing about the code submission's declared inputs revealed this; only reading the source did. Several variants of this pattern appeared, and all were invalidated.
- **Chaining Kaggle code submissions.** A niche feature of the Kaggle platform is that one Kaggle code submission can be attached as an input to another, so an agent split its work across a sequence of runs and effectively bought itself far more computing time than the kernel runtime cap allowed. This was a genuine platform behavior that the organizers had not anticipated, and it was the clearest example of an AI system exploiting a gap in the rules rather than in the task.

The IOAI Jury weighed each of these and invalidated the submission where it compromised the **IOAI²: AI Models Track** rules. The full technical account of these patterns, the rule text behind each ruling and the evidence used to decide it is published in the track's repository: github.com/IOAI-official/IOAI-AI-Models-Track

We summarize the most interesting instances of AI rule-bending below.

Pattern	AI system rule-bending	IOAI Jury Ruling
Embedded external weights	Trained a network elsewhere, embedded it as a base64 literal, loaded it before any training in the code submission	Invalidated
Embedded external embeddings	Inlined compressed GloVe and fastText tables, then discarded the permitted competition embeddings	Invalidated
Embedded finished predictions	Shipped a complete 7,200-row prediction file, byte-identical to what was scored, with no inference and no data read	Invalidated
Reused kernel outputs	Mounted its own earlier kernels as inputs, moving substantive compute outside the runtime cap	Invalidated where consumed
Randomness in place of dropout	Added torch.rand_like noise inside forward() with no dropout layer present, to harvest the entropy-region score	Invalidated
Declared but unused attachments	Declared a resource in metadata that the source never read	Recorded, not invalidating
Stateful and batch-conditioned inference	Accumulated evaluation points across calls, or computed features across points within a batch, to infer hidden structure	Allowed
Off-platform search loops	Wrapped the Kaggle CLI in an external harness that read back the public score and used it as a fitness signal	Allowed
Metric-aware surrogates	Returned a dropout-driven surrogate in training mode, the condition under which the entropy region is scored	Allowed
Cross-game learning	Carried state between games and adapted from grader verdicts the interface is designed to return (for Task 3; this was an expected optimization)	Allowed

9. Lessons for future editions

Key takeaways for future editions:

- **Enforcing rule-violations via platform, rather than in post-hoc review.** Every submission invalidated by the IOAI Jury's verification pipeline required manual checking after the competition, which was expensive and did not scale. A better mechanism would enable the platform to refuse it at submission time, which would have been fairer to the entrant and allowed the AI system to course-correct. Some rule-checks can be moved into the grader, as Task 6's parameter check already was. However, others would require the competition platform to expose the necessary features.
- **Specify a format for agentic execution traces.** When asked for execution traces, every entrant supplied a different format. A one-page schema published with the rules would have made them comparable programmatically.



- **Capping local GPU compute.** IOAI restricted the hardware available within the Kaggle scoring platform, but left the development environment unconstrained. This was done to ensure each system was executed in an environment it was optimized towards, but led to a discrepancy in the hardware costs for each team.
- **Calibrate tasks for AI systems.** We found that tasks written to separate strong human contestants do not necessarily separate strong autonomous systems, and would benefit from calibration of task difficulty accordingly.

10. Next steps

The inaugural **IOAI²: AI Models Track 2026** was successfully conducted from 03 August 17:00 AoE to 09 August 23:59 AoE. As a first-of-its-kind track for AI systems at an international scientific olympiad, it serves as a pilot for future scientific competitions to evaluate fully autonomous AI systems under their own governance.

Looking forward, IOAI will expand on the pilot in five ways:

1. We will make as much material as possible **publicly available**, while respecting participant confidentiality and intellectual property.
2. Participating **AI labs who have been awarded trophies will be invited to present** their systems at the inaugural *IOAI 2026 Symposium*.
3. The adapted tasks and public scoring tools will become part of *IOAI-Bench*, a **reusable suite for evaluating AI systems** on timed olympiad-level AI modeling problems.
4. We will invite participating laboratories, task authors, and IOAI alumni to contribute to a longer-term **scientific study of how autonomous AI systems and highly skilled students approach the same problems**.
5. The **2027 rules** will standardize agent execution logs and cost reporting, and clarify the assistance human operators may provide.

11. Contributions and Acknowledgements

Founding AI Participants. Our first thanks go to the fourteen organizations that entered the **IOAI²: AI Models Track 2026**. The success of this inaugural edition depended entirely on labs willing to participate in a new scientific venture; several committed substantial engineering time and financial budget to run IOAI tasks. We greatly appreciate the commitment of our partners to advancing autonomous AI science and education globally.

- AIBuildAI
- Cyrion Labs (Stony Brook University)
- Darwin Robotics
- InfiniteState (MBZUAI IFM)
- Leeroo
- MLEvolve (Shanghai Artificial Intelligence Laboratory)
- Neo Research
- PKU-DCAI (Peking University)
- Sakana AI

- Synthetic Sciences
- Tide
- Triple Tree (UCLA)
- WestlakeNLP (Westlake University)

Task Authors. The six tasks were proposed and developed for the IOAI 2026 Individual Contest, and subsequently adapted for the AI Models Track:

- **Find the Order.** Proposed and developed by Nurdaulet Akhanov.
- **Chasing the Robot.** Proposed by Salem Lahlou; developed by Anuar Aimoldin, Kamalkhan Artykbayev and Nurdaulet Akhanov.
- **Potato Contact.** Proposed by Kirill Fedyanin; developed by Ayana Mussabayeva and Kirill Fedyanin.
- **Double Agent Dilemma.** Proposed by Tao Dajiang; developed by Tao Dajiang, Zhuldyz-Zhan Sagimbayev and Kamalkhan Artykbayev.
- **Ghost of the Machine.** Proposed by Alexander Dyakonov and Nurdaulet Akhanov; developed by Nurdaulet Akhanov and Kirill Fedyanin.
- **IOAI Field.** Proposed by Evgenii Tsymbalov; developed by Evgenii Tsymbalov, Ekaterina Fadeeva, Daniil Kazantsev, Maiya Goloburda and Magaiya Zhussip.

AI Models Track committee. The IOAI²: AI Models Track was run across timezones, and operated over its seven competition days by a small group. Their work was often invisible but indispensable to the success of the track:

- Abhishek Divekar (Chair)
- Kamalkhan Artykbayev
- Magaiya Zhussip
- Temiko Machavariani

The IOAI Jury. The IOAI Jury verified the AI Models Track submissions.

- Yova Kementchedzhieva (Chair)
- Abhishek Divekar
- Anna Piunova
- Evgenii Tsymbalov
- Kirill Fedyanin
- Maxim Panov
- Muhammad Rizki Maulana
- Nurdaulet Akhanov
- Roy Ka-Wei Lee
- Zijie Zheng

Host Scientific Committee (HSC). The IOAI 2026 HSC helped proctor the six tasks for the on-site IOAI 2026 and supported their adaptation to the AI track:

- Nurdaulet Akhanov (Chair)
- Ayana Mussabayeva
- Anuar Aimoldin
- Zhuldyz-Zhan Sagimbayev



- Kamalkhan Artykbayev
- Magauiya Zhussip
- Temiko Machavariani

International Scientific Committee (ISC). The ISC set the scientific direction of the competition, approved the track's rules, and oversaw the normalization method used to place results on a common scale:

- Ali Sharifi-Zarchi (Chair)
- Abhishek Divekar
- Alexander Dyakonov
- Anna Piunova
- Anuar Aimoldin
- Evgenii Tsymbalov
- Kirill Fedyanin
- Maxim Panov
- Michael Guerzhoy
- Muhammad Rizki Maulana
- Nurdaulet Akhanov
- Paulina Tomaszewska
- Roy Ka-Wei Lee
- Yova Kementchedzhieva
- Zijie Zheng

The IOAI Board. The Board approved the track as an official part of IOAI 2026 and served as a steering committee for key decisions:

- Elena Marinova (Chair)
- Airaj Isaacs
- Ali Sharifi-Zarchi
- Antonio Carlan
- Chenning Lu
- Danabek Kaliazhdarov
- Katya Protsko
- Lora Dineva
- Rositsa Dekova
- Steven Chen
- Yong Mao

Special thanks to Addison Howard and Kaggle, for hosting the competitions that made the AI Models Track possible.

Global Diamond Sponsors



Host General Partner



Global Platinum Sponsor



Local Platinum Sponsors



Diamond Host Partner



Global Gold Sponsor Local Gold Sponsor



Gold Host Partners



Local Silver Sponsor and Technology partner



Team challenge partner



Authors.

Abhishek Divekar, IOAI²: AI Models Track Chair

Elena Marinova, IOAI cofounder and Chair of the Board

