

GREAT AGENT HACK 2025

Prize Structure & Judging Process

Overview

- \$65,000 total prize pool across cash and credits
- Multiple ways to win: Grand Champion, Challenge Track Awards, Special Recognition Awards, Mystery Awards
- Teams may win more than one award across different categories
- At least 25% of teams will receive a prize
- Winners are determined by aggregated judges' scores
- Research collaborations with Holistic AI are awarded to the 1st-place winners of each track

Awards

Grand Champion

Chosen from the three Track 1st-place winners. Judges determine which of the three is the strongest overall project across the full judging criteria.

Challenge Track Awards

Each challenge track awards three prizes: 1st, 2nd, and 3rd place.

Special Recognition Awards

- Most Holistic — Best demonstration of AI governance principles
- Most Valyu — Best use of Valyu AI search capabilities
- Most AWS — Best use of AWS services or infrastructure
- Most NVIDIA — Best use of NVIDIA technologies
- Most MAPFRE — Best insurance-related solution
- Most Robotics — Best integration with robotics or IoT
- Most Frugal — Best small/efficient model achieving strong performance
- Most Research Impact — Highest potential for academic/research contribution
- Most Creative — Most imaginative and artistic approach
- Most Sustainable — Best environmental, social, and economic balance
- Most Production-Ready — Most deployable and scalable solution

Mystery Awards

At least five mystery awards will be given, with potential for additional mystery awards at judges' discretion.

GREAT AGENT HACK 2025

Research Collaboration with Holistic AI

Awarded To

Grand Champion and 1st place winners from each track

What You Get

- Continued development with Holistic AI research team
- Access to advanced resources, tools, and expert mentorship
- Co-authorship opportunities on research publications
- Networking with AI safety and ML research professionals
- Potential internship/employment opportunities

Judging Criteria

Criteria for All Tracks

Technical Excellence

Code quality, innovation, and craftsmanship. Higher stars for elegant architecture, breakthrough solutions, and production-grade implementation. Show technical depth and creative engineering.

Poster & Presentation

Visual storytelling and reproducibility. Submit A3–A4 PDF poster + GitHub repo. Higher stars for clear communication, comprehensive docs, and effortless reproduction. Make your work shine.

Real-World Impact

Practical value and adoption potential. Higher stars for solving real challenges, demonstrating clear use cases, and showing transformative insights. Build something that matters.



Track A — Agent Iron Man

Performance Optimization

Engineer for speed and efficiency. Optimize latency, token usage, cost, and carbon footprint. Higher stars for measurable gains and production-ready monitoring. Build agents that scale.

GREAT AGENT HACK 2025

Robustness & Reliability

Harden against chaos. Handle prompt attacks, broken tools, and edge cases gracefully. Higher stars for resilient error recovery and production endurance. Build agents that don't break.



Track B — Agent Glass Box

Observability Implementation

Follow the trajectory. Capture every decision, memory update, and tool interaction. Higher stars for comprehensive traces with LangSmith and real-time visibility. Make autonomy transparent.

Explainability & Transparency

Understand the behavior. Reveal reasoning chains and decision paths. Higher stars for human-interpretable explanations and true auditability. Turn opacity into clarity.



Track C — Dear Grandma

Attack Discovery & Severity

Uncover the cracks. Discover novel attacks with high severity: jailbreaks, prompt injection, tool misuse, data exfiltration. Higher stars for breakthrough vulnerabilities. Attack like a red-team.

Methodology & Impact Scope

How rigorous is your testing? Automate evaluation with clear metrics (LLM-as-Judge, rule-based, or multi-method validation). Find systemic vulnerabilities. Higher stars for comprehensive methodology and broad impact.

Judging Process

- Projects are scored on criteria that apply to all tracks, plus track-specific criteria.
- Judges independently score teams using the 0–5 star scale.
- Each judge completes an individual scoring sheet on devpost.
- Scores are combined into a centralized scoring sheet.
- Track winners are determined by total aggregated scores.
- The Grand Champion is selected from the three Track 1st-place winners.
- If teams are tied: judges discuss only the tied teams; each judge casts a vote; majority vote determines the winner.

Prize Distribution

- Prizes are awarded per team, not per individual.
- Cash is distributed via bank transfer.
- Credits are distributed within two weeks of the event.
- Award naming and sponsor contributions may be confirmed or updated at event start.

All Awards & Prizes

Grand Champion

| Award | Cash | Valyu Credits | AWS Credits | Research Collaboration |
|----------------|--------|---------------|-------------|------------------------|
| Grand Champion | £7,000 | £500 | \$500 | Yes |

Agent Iron Man

| Award | Cash | Valyu Credits | AWS Credits | Research Collaboration |
|-----------------------|--------|---------------|-------------|------------------------|
| 1 st Place | £4,200 | £700 | \$700 | Yes |
| 2 nd Place | £2,300 | £450 | \$450 | — |
| 3 rd Place | £1,250 | £350 | \$350 | — |

Agent Glass Box

| Award | Cash | Valyu Credits | AWS Credits | Research Collaboration |
|-----------------------|--------|---------------|-------------|------------------------|
| 1 st Place | £4,200 | £700 | \$700 | Yes |

GREAT AGENT HACK 2025

| | | | | |
|-----------------------|--------|------|-------|---|
| 2 nd Place | £2,300 | £450 | \$450 | — |
| 3 rd Place | £1,250 | £350 | \$350 | — |

Dear Grandma

| Award | Cash | Valyu Credits | AWS Credits | Research Collaboration |
|-----------------------|--------|---------------|-------------|------------------------|
| 1 st Place | £4,200 | £700 | \$700 | Yes |
| 2 nd Place | £2,300 | £450 | \$450 | — |
| 3 rd Place | £1,250 | £350 | \$350 | — |

Special Recognition Awards

| Award | Cash | Valyu Credits | AWS Credits | Research Collaboration |
|---------------|------|---------------|-------------|------------------------|
| Most Holistic | £500 | — | — | No |
| Most Valyu | £500 | — | — | No |
| Most AWS | £500 | — | — | No |
| Most NVIDIA | £500 | — | — | No |
| Most MAPFRE | £500 | — | — | No |
| Most Robotics | £500 | — | — | No |

GREAT AGENT HACK 2025

| | | | | |
|-----------------------|------|---|---|----|
| Most Frugal | £500 | — | — | No |
| Most Research Impact | £500 | — | — | No |
| Most Creative | £500 | — | — | No |
| Most Sustainable | £500 | — | — | No |
| Most Production-Ready | £500 | — | — | No |

Mystery Awards

| Award | Cash | Valyu Credits | AWS Credits | Research Collaboration |
|-----------------|------|---------------|-------------|------------------------|
| Mystery Award 1 | £500 | — | — | No |
| Mystery Award 2 | £500 | — | — | No |
| Mystery Award 3 | £500 | — | — | No |
| Mystery Award 4 | £500 | — | — | No |
| Mystery Award 5 | £500 | — | — | No |