

AIM Intelligence to Host Frontier AI Red-Teaming Workshop at 'Ministry-Backed AI Safety Forum 2026'

Live challenge 'Judgement Day' tests physical AI safety on humanoid robots; Google, Microsoft, and LG AI Research participate in ICML 2026-linked forum.

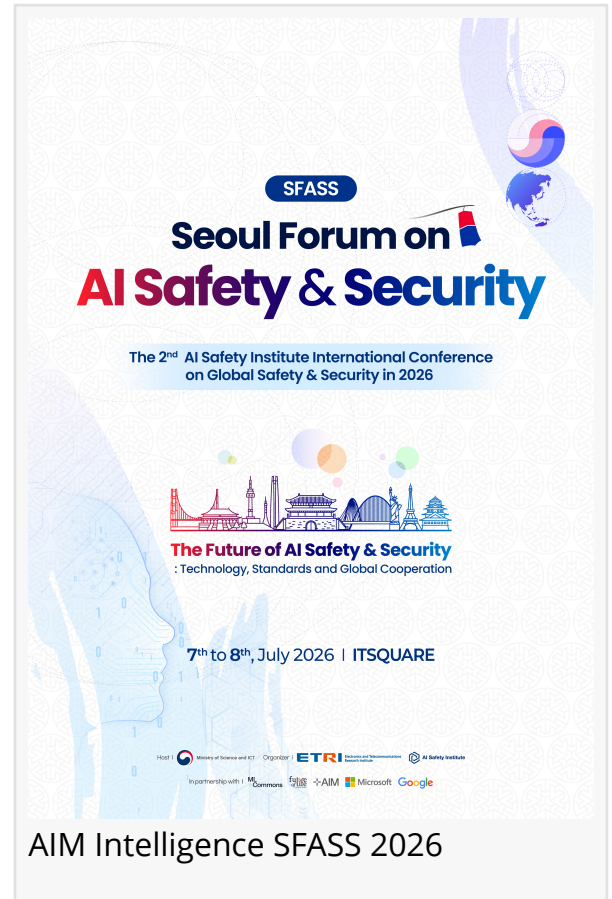
SF, CA, UNITED STATES, July 13, 2026
/EINPresswire.com/ -- AI security specialist [AIM Intelligence](#) announced its participation as an official partner in the Seoul Forum on AI Safety & Security 2026 (SFASS 2026), hosted by the Ministry of Science and ICT and held in conjunction with ICML 2026. The company will host the "Frontier AI Red-Teaming Workshop" on the afternoon of day 2.

SFASS 2026 brings together over 150 government, academic, and industry experts from around the world to discuss AI safety and security across cultural contexts and emerging autonomous agent threats. This marks Korea's second consecutive year leading global AI safety discussions.

AIM Intelligence's workshop runs from 1:30 PM on day 2, opening with remarks from CTO [Haon Park](#) and insights from Principal Researcher Dasol Choi on multimodal red-teaming results. Leading researchers from Google (Jenny Ni), DeepMind (Naman Goyal), and Microsoft (Eugenia Kim) will present on model robustness, adaptive red-teaming, and adversarial attack automation. Panels featuring LG AI Research and Singapore's AISI will address emerging vulnerabilities from universal jailbreaks to multilingual agent attacks.

The workshop culminates in "Judgement Day: Jailbreaking the Humanoid," where participants directly attack physical AI deployed on the Unitree G1 humanoid robot using only adversarial inputs. Real-time scoring and leaderboards provide instant feedback on attack success.

AIM Intelligence combines offensive red-teaming engines with real-time guardrails in a full-stack architecture, creating a self-reinforcing cycle where discovered vulnerabilities inform defensive



models. The company's attack engine diagnoses real enterprise vulnerabilities across drug synthesis, cyberattacks, phishing, and disinformation categories.

On day 1, AIM Intelligence presents "The Future of Agentic AI and Cybersecurity" alongside the Korea AI Safety Institute, addressing expanding attack surfaces as foundation models evolve into autonomous multimodal agents.

The forum includes MLCommons and the Future of Life Institute (FLI) as international partners, positioning SFASS 2026 as a global AI safety platform.

CEO [Sangyoon Yu](#) stated: "As AI evolves from text to autonomous agents to embodied humanoids, attack pathways are expanding rapidly. Judgement Day validates model robustness under multilayered threats in real-world conditions. Real-world red-teaming is now an essential AI security practice—we'll accumulate empirical data with global partners to establish Korea's leadership in AI safety and security."

Team Cookie Official

Team Cookie

[email us here](#)

Visit us on social media:

[LinkedIn](#)

[Facebook](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/926241822>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2026 Newsmatics Inc. All Right Reserved.