Alliance for Innovative Regulation

BATTLEFRONT: FRAUD IN THE AGE OF AI

Event Recap Report • June 23–24, 2025, Washington, D.C.

INTRODUCTION

The "Battlefront: Fraud in the Age of AI" event, hosted by the Alliance for Innovative Regulation (AIR) with support from the American FinTech Council (AFC) and Amazon Web Services (AWS), took place in Washington, D.C. on June 23-24, 2025. It convened over 100 in-person and virtual experts from banking, fintech, crypto, cybersecurity, law enforcement, academia, technology, and policy development to confront the growing threat of AI-enabled fraud. Through an immersive AI-powered simulation, panel discussion, and live demonstrations, participants explored cutting-edge fraud attacks and innovative defenses designed by multi-sector teams. This event highlighted urgent challenges created by emerging technologies such as generative and agentic/agentive AI, and the critical need for cross-sector collaboration to tackle fraud on a global scale. This report summarizes key activities, insights, participant feedback, and recommendations to inform future initiatives to combat AI-driven fraud.

> **"Generative AI has achieved mass adoption and now plays a significant role in creating very credible scams and a tsunami of fake narratives. AI has enabled crime as a service for providing automated money laundering, fake identity documents to bypass online KYC, create credible scams and program undocumented families of malicious software."**
> *Vincent Danjean, Head of the Cyberspace and New Technologies Laboratory at INTERPOL*

EVENT OVERVIEW

Background

As AI evolves from being just a tool to an autonomous agent in both fraud and defense, the battleground is rapidly changing. Advancements in AI are enabling increasingly sophisticated fraud schemes, including deepfake scams, synthetic identities, and AI-powered social engineering in both traditional finance and the crypto space. To address these evolving risks, AIR hosted "Battlefront: Fraud in the Age of AI" as an immersive experience applying design principles and empathy mapping to build out a simulation that set fraudsters against defenders to foster collaboration and expert dialogue.

Purpose & Goals

The event was designed to simulate AI-driven fraud scenarios and defense strategies in order to:

- Foster cross-sector knowledge sharing, innovation and collaboration.
- Build partnerships to strengthen fraud prevention and response.
- Empower regulators, policymakers and practitioners with the knowledge and strategies needed to combat fraud effectively.
- Enhance consumer protection by exploring the development of tools that could help protect individuals against fraud.
- Stay abreast of emerging technologies by sharing insights and showcasing tech solutions to help regulators and industry understand and leverage these advances in combatting fraud.

EVENT FORMAT & ACTIVITIES

Simulation

Ten teams representing fraudsters and defenders developed AI-powered attacks and defense mechanisms using generative, agentic and agentive AI tools. Participants engaged in a series of interactive sessions where they assumed roles as either fraudsters or defenders to design or dismantle AI-powered fraud. Using AI, individuals worked in cross-functional teams to design sophisticated fraud schemes and develop advanced defense strategies, encountering fast moving developments and spy in the camp scenarios.

The event featured real-world use cases, including:

- Deepfake-enabled scams in traditional finance, cryptocurrency, and public payment systems.
- GenAI-powered social engineering, ransomware, and Trojan attacks.
- AI-generated synthetic identities targeting public sector programs.
- Consumer-facing GenAI bots for decentralized defense.
- Agentive/agentic AI in post-fraud response, asset tracing and recovery, and redress.

Showcase

The event culminated in a live, interactive face-off where teams deployed their AI strategies in real-time, demonstrating emerging fraud threats and countermeasures. The teams showcased various AI tools used to create fraudulent attacks including personalized multilanguage communications, websites, corporate documents, smart contracts, crypto whitepapers, celebrity endorsement videos, Proof of Attendance Protocols (POAPs), Application Programming Interfaces (APIs) and apps. Defenders presented AI-enabled solutions such as AI chatbots and agents to identify emotional manipulation, build biometric safeguards and digital watermarks, embed invisible signatures in voices or other physical and behavioral characteristics to help detect deepfakes, enable post-fraud recovery activities, conduct analysis to flag fake content and red flags, and support one-click reporting of fraud to law enforcement authorities.

The table below provides a summary of all solutions showcased during Battlefront:

Use Case	😈 Fraudsters	🛡️ Defenders
Use Case 1: Deepfake - The use of deepfake-enabled fraud and scams in traditional finance and crypto/DeFi markets.	**DAO-or-DAI:** Showcases a sophisticated AI-agent driven crypto pump-and-dump scam that lures victims via multilingual outreach, drains wallets through smart contracts and malicious POAPs, and fabricates legitimacy using fake decentralized autonomous organizations (DAO) infrastructure and celebrity promotions.	**TrapWeaver:** Comprehensive, multilingual scam defense system that seamlessly integrates across platforms to provide real-time, multi-layered protection against phishing and impersonation attacks on all devices.
Use Case 2: Social Engineering - The evolution of social engineering schemes, ransomware, and Trojan attacks through Generative AI.	**Team Trust Fall**: A teenage scammer uses AI tools to socially engineer and defraud a solo crypto investor by impersonating a trusted influencer and directing her to a fake platform, demonstrating how easily next-gen fraud exploits emotional and technical vulnerabilities.	**Evolving Vigilant AI (Eva):** An AI-powered fraud defense system that disarms next-generation scams by combining emotional intelligence with technical detection, offering real-time protection through software development kits (SDKs), browser plugins, and standalone apps for banks, fintechs, and vulnerable communities.
Use Case 3: Synthetic IDs & Public Sector Fraud - The threat of AI-generated synthetic identities and documents in large-scale public sector fraud.	**Team Mission Impossible:** Simulates a large-scale synthetic identity attack exploiting emergency financial relief systems using AI-generated companies, voices, and documents to overwhelm IRS analysts and siphon millions undetected.	**Government ID Guardians:** Proposes a four-layer AI defense system to combat IRS tax credit fraud by detecting synthetic identities and prioritizing high-risk claims for human review, enabling overwhelmed analysts to focus on the most serious threats.
Use Case 4: GenAI Defense - Development of consumer-facing GenAI bots for decentralized fraud defense.	**Fraud 2.0:** Depicts a multi-stage AI-enabled scam targeting a music artist through fake legal threats and IP claims, ultimately tricking her into revealing sensitive data and financial credentials under the guise of resolving copyright infringement.	**DataSentinel.AI:** A personalized GenAI-powered defense bot that proactively protects individuals from voice cloning, data theft, and financial fraud through multi-layered monitoring, alerts, and optional countermeasures.
Use Case 5: Redress - Agentic/Agentive AI-driven victim assisted scams using social engineering and defenses to aid in prevention and recovery.	**IdentifAI Labs:** A rogue AI identity provider uses agentic AI to exploit onboarding systems, siphoning personal and financial data at scale to steal and launder funds through synthetic accounts and shell companies.	**Team Guardian Angel:** An agentic AI system designed to proactively prevent and reactively address fraud through intelligent, adaptive intervention strategies.

Participants

Battlefront had over 100 participants and observers, both in person and online, representing financial institutions, decentralized finance (DeFi), crypto exchanges, RegTechs, regulators, law enforcement agencies, academia, behavioral science, technology, AI, and cybersecurity firms.

Virtual Observer Experience

Remote participants observed the event via livestream, gaining strategic insights into AI's impact on fraud risk and prevention, while also interacting with onsite attendees by submitting questions during the showcase. They also had the opportunity to vote for the winning team in each use case.

OUTCOMES & IMPACT

Key Insights

Generative and agentic AI is fundamentally transforming financial fraud:

- The increased personalization and sophistication of generative and agentic AI's persuasive capabilities pose both risks to individuals and opportunities for consumers to protect themselves against fraud.
- AI advances not only accelerate the speed and scale of fraud attacks but also significantly reduce the costs of carrying out iterative or multi-pronged schemes using credible fraudulent identities, businesses, and culturally tailored multi-lingual communications and websites created in minutes.
- Deepfake-enabled scams and social engineering are rapidly evolving across both traditional finance and emerging cryptocurrency ecosystems.
- GenAI bots can strengthen decentralized defense but also introduce new ethical and regulatory challenges.
- AI-generated synthetic identities threaten public-sector benefit programs, requiring enhanced detection and verification tools.

> **"Battlefront provided a unique hands-on experience that deepened my understanding of AI fraud threats."**
> *Financial Institution Participant*

- Consumers need education on authenticating their own AI agents to prevent fraud.
- As AI-driven anti-fraud solutions are developed, they must account for data security and incorporate personal data-protection safeguards.
- Multi-layered defense systems are essential to address AI-generated fraud, whether deployed on local devices or as browser extensions.

Innovative Defense Strategies

The event demonstrated the potential of AI-driven collaboration between the private and public sectors to develop innovative counter-fraud strategies such as:

- AI-powered, emotion-aware agents to detect fraud.
- Honeypot-style systems to proactively track and identify fraudsters as part of defense strategies.
- Large language models and translation tools that enable culturally aware, cross-language communication and promote easier AI - human interactions.
- AI-enabled, real-time validation of transactions through device-level integration.
- AI bots and agents to detect the use of GenAI in generating 200+ fake businesses, automated form fillers, and call centers designed to overwhelm systems.
- Integration of AI agents at the operating-system (OS) level to deliver specialized risk alerts.

Winning Teams from the Battlefront Showcase

During the interactive showcase, teams demonstrated innovative AI-powered strategies designed to both carry out advanced fraud attacks and defend against them. Each team developed and presented thought provoking schemes and solutions. The winning teams, which included fraudsters or defenders, stood out for their creativity, speed, effectiveness and innovative use of AI tools in generating or combatting AI-enabled fraud.

Use Case 1: Deepfakes - The use of deepfake-enabled fraud and scams in traditional finance and crypto/DeFi markets.

Winning Team: DAO-or-DAI (😈Fraudsters)

This scenario demonstrated a sophisticated AI-driven crypto pump-and-dump scam called DAO-or-DAI. The fraudsters used deepfake voice synthesis to conduct multilingual phishing outreach, tricking victims into revealing credentials. The scam then drained victims' wallets through malicious POAPs (Proof of Attendance Protocol tokens). To add legitimacy, the scheme fabricated a fake DAO infrastructure and leveraged celebrity promotions generated by AI deepfakes, making the scam highly convincing and difficult to detect.

Use Case 2: Generative AI & Social Engineering - The evolution of social engineering schemes, ransomware, and Trojan attacks through Generative AI

Winning Team: EVA (🛡️Defenders)
This scenario focused on AI-driven malware that adapts dynamically to exploit remote tools and security gaps in target systems. To counter this evolving threat, the defense solution Evolving Vigilant AI (EVA) was showcased. EVA combines emotional intelligence with advanced technical detection to identify and disarm sophisticated scams in real time. It provides proactive protection through SDKs, browser plugins, and standalone apps designed for banks, fintech companies, and vulnerable communities, enhancing fraud defense across multiple platforms.

Use Case 3: Synthetic IDs & Public Sector Fraud - The threat of AI-generated synthetic identities and documents in large-scale public sector fraud

Winning Team: Team Mission Impossible (😈Fraudsters)
This use case simulated a large-scale synthetic identity fraud campaign called Team Mission Impossible, targeting government emergency financial relief programs. The attack used deepfake technology to create AI-generated companies, synthetic voices, and forged documents to impersonate real individuals. By overwhelming IRS analysts with sophisticated, high-volume fraudulent claims, the scheme looked to siphon off millions of dollars undetected, exposing critical vulnerabilities in public sector identity verification systems.

Use Case 4: GenAI Defense - Consumer-facing GenAI bots for decentralized fraud defense

Winning Team: DataSentinel.AI (🛡️Defenders)
This scenario highlighted fraud schemes using generative AI bots equipped with voice cloning capabilities to exploit stolen personal data. Attackers leveraged these technologies to gain unauthorized credit, loans, and purchases by impersonating victims convincingly. The defense solution, DataSentinel.AI, is a personalized GenAI-powered bot that proactively safeguards individuals through multi-layered monitoring, real-time alerts, and rapid responses to voice cloning, data breaches, and financial fraud attempts, helping to prevent losses before they occur.

Use Case 5: Redress - Agentic/Agentive AI-driven victim assisted scams using social engineering and defenses to aid in prevention and recovery

Winning Team: IdentifAI Labs (😈Fraudsters)
This scenario highlighted a fraud scheme in which a rogue AI identity provider used agentic AI and social engineering to infiltrate onboarding systems and exploit victims at scale. Attackers harvested personal and financial data, created synthetic identities, and laundered funds through shell companies and synthetic accounts. The fraudster team, IdentifAI Labs, demonstrated how adaptive, agentic AI could mimic legitimate onboarding behavior to evade detection and transform victim-assisted scams into a scalable and highly evasive fraud model.

Common Strengths of Winning Solutions

- Combined generative AI and agentic or agentive AI in an autonomous manner to act proactively rather than reactively.
- Integrated cross-sector data sources, technologies and platforms for holistic defense.
- Balanced innovation with practical usability for rapid deployment in real-world scenarios.
- Demonstrated adaptability to evolving scenarios through continuous learning.

Partnerships

During the event, new partnerships and collaborations were forged, linking diverse stakeholders in a collective effort against AI-driven fraud across the finance, regulatory, technology and policy space.

"The collaboration across sectors was invaluable for designing realistic and effective defense strategies."
Cybersecurity Expert

Policy & Behavioral Influence

The event underscored the urgency of updated policies and cooperative frameworks to manage AI-related fraud risks, as well as the importance of technology showcases to stay ahead of emerging threats. It also raises the need for information-sharing frameworks to enable rapid responses to fraud and scams.

TOOLS UTILIZED - BATTLEFRONT AI ARSENAL

The event equipped participants with a curated AI arsenal designed to allow them to replicate both fraud and defense scenarios across all the use-cases. Details of this toolkit are as follows:

AWS Workshop Studio

AWS provided a dedicated Workshop Studio platform that allowed participants to test prompt-based AI-fraud detection tools. This was particularly effective for deepfake, synthetic ID, and social engineering use-cases. Users could input data and experiment with the tools in real time. Advanced participants leveraged the AWS Generative AI toolkit, including Amazon Bedrock and V0, to specifically explore the Agentive AI for fraud redress and recovery as well as the GenAI bots for defense use-cases.

Open Source Tools

Participants were provided with a curated list of open-source tools that could be leveraged to simulate both fraudster and defender tactics. The following table highlights the use of these tools across different use-cases.

Use Case	Fraudster (Prompt-based)	Fraudster (Advanced)	Defender (Prompt-based)	Defender (Advanced)
Deepfake Fraud	–	DeepFaceLab	Deepware.ai	FaceForensics++
Social Engineering	Emkei's Fake Mailer	GPT-Neo	OpenPhish	ELK Stack
Synthetic ID Fraud	–	StyleGAN	Deepware.ai	OpenCV
GenAI Bots for Defense	–	DialoGPT	Botpress	ChatAnalytics
Agentive AI for Redress/Recovery	–	Faker	–	GraphSense

Additional AI and open source tools used by the teams include: Lovable, Stagehand API, Open AI, Philanthropic API, Deepseek, Vertex AI, ChatGPT, DeepL, Bing Image Creator Grammarly, Quill bot, OpenVoice, React, node.js, Vertex AI, V0, Anthropic.

Sardine's Device SDK

Complementing AWS and open-source recommendations, Sardine shared its proprietary Device SDK tool for deepfake detection. This provided insights into advanced fraud detection approaches and illustrated how specialized tools can be integrated into broader defense frameworks.

CHALLENGES & LESSONS LEARNED

Challenges

- Generative and agentic AI are accelerating fraud and scams at an exponential rate, with attacks increasingly personalized and automated at scale.
- As fraud tactics evolve, it is becoming more difficult for defenses to keep pace.
- Managing the complexity of diverse AI tools within a time-constrained simulation posed challenges for some participants, particularly those without a technical or AI prompt-engineering background.
- Criminals operate across continents and financial systems, using AI tools to localize language, making fraud and scams increasingly difficult to detect and investigate.
- There is a gap between social media and communications platforms, where fraud and scams often begin; financial systems, where losses occur; and law enforcement and regulators that needs to be addressed.

Lessons Learned

- AI is not just a tool for good but also a force multiplier for fraud.
- Ensuring diverse representation across sectors to mitigate AI bias requires greater inclusion of public-sector and consumer-advocacy voices from a range of ages and backgrounds.
- Balancing openness in simulations with confidentiality and security concerns requires refined protocols.
- The psychological toll of fraud and scams may necessitate greater personalization of AI defenses, while still balancing privacy and data-protection concerns.
- A human must remain in the loop for any AI-related activity.
- More effective coordination and enhanced data and information sharing are required to address AI-enabled fraud.
- Defenses must be multi-modal and multi-layered to effectively address frauds and scams and can no longer simply be defensive in nature.

"There is a need to make the financial ecosystem Return on Investment (ROI) negative for fraudsters."
Fintech Participant

- Realistic simulation events serve as powerful teaching tools.
- Greater digital literacy and public awareness are needed to protect against fraud, especially in vulnerable or targeted communities such as the elderly or teens.
- Fraud defenses must be as creative as fraud attacks, monitoring not only patterns, amounts, and transaction data but also persuasion tactics, language, tone, and behaviors.
- Fraud prevention is not solely a technical or regulatory challenge - it must keep humans at the center of all activity.

NEXT STEPS & RECOMMENDATIONS

To effectively combat AI-driven fraud, sustained multi-sector collaboration is crucial, focusing on both proactive detection of fraudulent attacks and robust victim protection. Stakeholders must continue to prioritize investments in AI-based prevention and detection tools, comprehensive fraud education, and unified intelligence-sharing frameworks

To that end, AIR will continue refining its approach to convene the ecosystem and identify solutions to combat global financial fraud. Activities may include co-hosting a virtual roundtable, podcast, and international events. AIR will also continue to work closely with partners to explore fraud policy and regulation, and to host Battlefront-style events and technology showcases to ensure regulators, industry, and global stakeholders remain informed about emerging technologies and better equipped to counter evolving fraud threats. Organizations interested in partnering with AIR are encouraged to get in touch.