

# A COLLABORATION ACROSS INDUSTRY, ACADEMIA, AND GOVERNMENT



**MITRE Adversarial Threat Landscape for AI Systems (ATLAS™) is a globally accessible, living knowledge base of adversary tactics and techniques based on real-world attack observations and realistic demonstrations from artificial intelligence (AI) red teams and security groups.**

There are a growing number of vulnerabilities in AI-enabled systems as the incorporation of AI increases the attack surfaces of existing systems beyond those of traditional cyberattacks. We developed ATLAS to raise community awareness and readiness for these unique threats, vulnerabilities, and risks in the broader AI assurance landscape.

ATLAS is modeled after the MITRE ATT&CK® framework and its tactics, techniques, and procedures (TTPs) are complementary to those in ATT&CK.

#### **ATLAS can be used to:**

- Inform security analysts and AI developers/implementers of realistic threats to AI-enabled systems
- Enable threat assessments and internal red teaming
- Understand real-world adversary behaviors and mitigation pathways
- Report unique real-world adversary attacks on AI-enabled systems

“

The ATLAS framework has become the de facto Rosetta Stone for security professionals to make sense of this ever-shifting AI security space. Today's latest ATLAS evolution—to include more LLM attacks and case studies—underscores the framework's incredible relevance and utility.

Microsoft, an ATLAS and Arsenal Collaborator

”

## Why AI Security?

Real world attacks on AI-enabled systems are happening and our ATLAS community is dedicated to understanding and mitigating these threats. For example, an attack submitted to ATLAS as a case study showed losses of over \$77 million through an adversarial attack on a facial recognition system used by a foreign tax authority.

AI security threats and risks can also impact our allies and collaborators across the globe. In March 2023, the MITRE ATLAS Team kicked off a [NATO Exploratory Team \(ET\)](#) to address the evolving landscape of threats and assurance concerns for AI-enabled systems.

## Incident Sharing through ATLAS

Operating at the intersection of government, industry and academia, MITRE is uniquely positioned to create mechanisms for timely, relevant, and secure AI incident and vulnerability reporting. Our public ATLAS website documents real-world reported case studies and we are building on MITRE's historical strength in enabling protected or anonymized threat reporting within our community. The ATLAS team is also collaborating with industry and academia on open-source tools like the [AI Risk Database](#), a tool for discovering vulnerabilities associated with public AI models.

## Threat Emulation and Red Teaming

In early 2023, the ATLAS team released the collaboratively developed [Arsenal and Almanac plugins](#) to add implementations of ATLAS techniques and new adversary profiles that target AI-enabled systems to [CALDERA™](#), an existing MITRE threat emulation tool largely leveraged by the cyber world.

## Mitigations

The ATLAS team released a draft set of [mitigations](#) and is continually incorporating community techniques for mitigating AI security threats. Mitigations include both security concepts and classes of technologies that can be used to prevent an attack from being successfully executed.



There are a growing number of vulnerabilities in AI-enabled systems as the incorporation of AI increases the attack surfaces of existing systems beyond those of traditional cyberattacks.

Join our collaborative community, rapidly report threats, and help shape future tools for AI security, threat mitigation, bias, privacy, and other critical aspects of AI assurance.

- email: [atlas@mitre.org](mailto:atlas@mitre.org)
- slack: [mitreatlas.slack.com](https://mitreatlas.slack.com)
- github: [github.com/mitre-atlas](https://github.com/mitre-atlas)
- webpage: [atlas.mitre.org](https://atlas.mitre.org)

*MITRE's mission-driven teams are dedicated to solving problems for a safer world. Through our public-private partnerships and federally funded R&D centers, we work across government and in partnership with industry to tackle challenges to the safety, stability, and well-being of our nation.*

**MITRE** | SOLVING PROBLEMS  
FOR A SAFER WORLD®