# AI SECURITY

## ENSURING RESPONSIBLE AND SECURE AI DEPLOYMENTS

AI is reshaping how organisations operate, offering new capabilities in automation, reasoning, and decision support.

But its adoption also introduces novel security risks. AI systems are susceptible to attacks - often in ways that traditional security tools don't detect. Without proper safeguards, these vulnerabilities can lead to unauthorised access, reputational damage, and misuse of sensitive data.

### COMMON CHALLENGES

These are some of the issues that AI deployments may face if not fully assessed and secured:

- Prompt injection and jailbreaks that subvert system instructions
- Unintentional exposure of proprietary or user-submitted data
- Gaps in governance, explainability, and responsible AI use
- Vulnerable model-serving APIs and insecure plugin/tooling environments
- Unclear compliance boundaries in regulated or sensitive domains

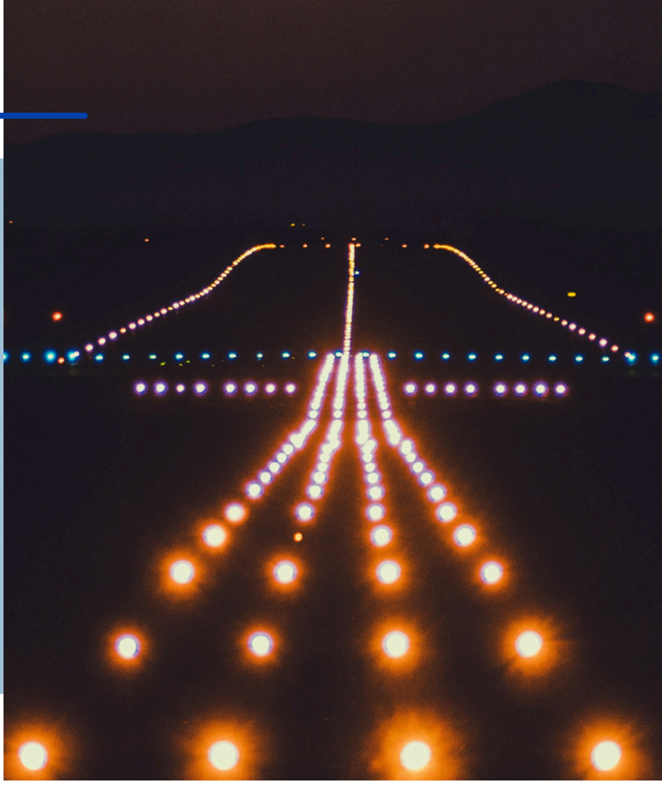
OUR SOLUTION →

# OUR SOLUTION

Our AI Security service enables organisations to deploy and integrate AI responsibly and securely.

We assess AI systems, APIs, and infrastructure to identify weaknesses, provide mitigation strategies, and establish best practices for safe and ethical use.

This includes threat modelling, security testing, and tailored guardrails for LLM deployments.



We provide:

- Security and **red-team testing** for models, APIs, cloud infrastructure, and databases
- Assessments for **prompt injection** and training data poisoning
- **Model behaviour audits** (e.g., adversarial prompt testing, jailbreak detection, hallucination analysis)
- Guidance on **safe model access**, audit trails, and usage monitoring
- Collaboration with internal security teams on **penetration testing** and bug bounty validation
- Secure reporting with executive summaries, technical findings, and **mitigation plans**

# WHY US?

Led by certified professionals with hands-on AI and security expertise

Deep knowledge in both offensive security and AI systems

Independent and transparent—no conflicts of interest, no technology or vendor lock-in

Trusted by organisations deploying AI in sensitive and high-impact environments

## ENSURE SECURE AND RESPONSIBLE AI ADOPTION

Get in touch to protect your systems and people against new and evolving AI threats.

THIRD
PLATFORM

www.thirdplatform.ai

hello@thirdplatform.ai

0131 210 0129