FireTail

APIs Are the Attack Surface That Matters in AI-Powered Apps.

Lessons from the real world.

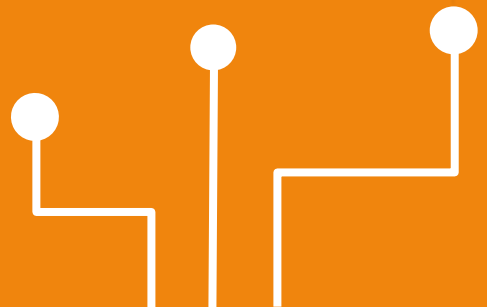
Overview.

What we'll cover today.

- **Lessons learned:**
  How to approach securing new technology platforms
- **Context:** Where we are with AI today
- **AI Security:**
  Key requirements
- **About us:** FireTail overview
- **Closing:** Get these slides

FireTail

History...



FireTail

The Evolution of Digital Innovation.

Phase	Start	Lift & shift	Cloud 2.0	Cloud + AI
Operations	Server-based	Virtual server	Cloud-native	CloudOps/ AIOps
Access	Controlled	Loosening	Distributed	Distributed
Challenges	Perimeter	Shadow IT Data exposures Misconfigurations	Complex attacks Cloud identity API attack surface	Complex attacks API breaches Data governance
Tools	Firewall, EDR, SIEM, DLP	CSPM	CNAPP, CIEM API Security	CNAPP, CIEM AAI + PI Security
Stakeholders	IT	IT, developers	IT, developers	Everyone

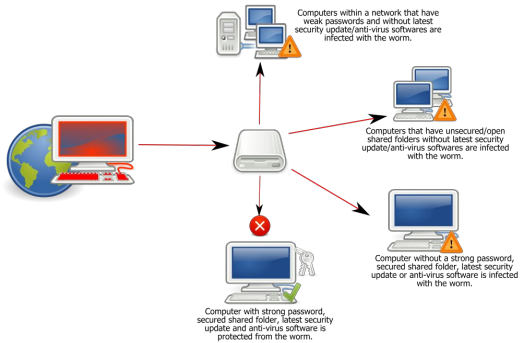
The Evolution of Digital Innovation.

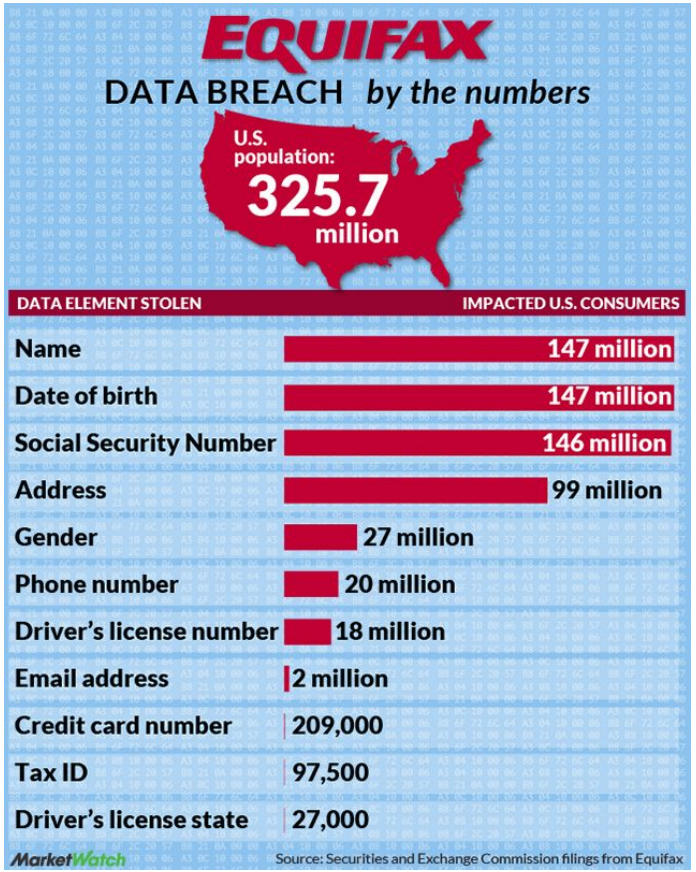
	On Prem	Cloud	APIs	AI
Phase	Start	Lift & shift	Cloud 2.0	Cloud + AI
Operations	Server-based	Virtual server	Cloud-native	CloudOps/ AIOps
Access	Controlled	Loosening	Distributed	Distributed
Challenges	Perimeter	Shadow IT Data exposures Misconfigurations	Complex attacks Cloud identity API attack surface	Complex attacks API breaches Data governance
Tools	Firewall, EDR, SIEM, DLP	CSPM	CNAPP, CIEM API Security	CNAPP, CIEM AI + API Security
Stakeholders	IT	IT, developers	IT, developers	Everyone

Initial virus / firewall breaches.

Cyber attacks 1.0 / 1.5.

Worm:Win32 Conficker

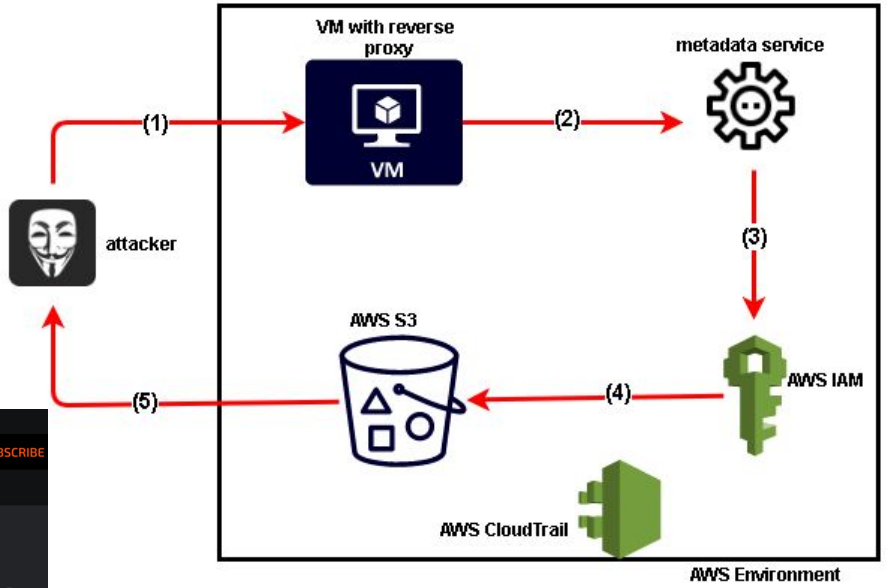




The Evolution of Digital Innovation.

Phase	Start	**Lift & shift**	Cloud 2.0	Cloud + AI
Operations	Server-based	**Virtual server**	Cloud-native	CloudOps/ AIOps
Access	Controlled	**Loosening**	Distributed	Distributed
Challenges	Perimeter	**Shadow IT** **Data exposures** **Misconfigurations**	Complex attacks Cloud identity API attack surface	Complex attacks API breaches Data governance
Tools	Firewall, EDR, SIEM, DLP	**CSPM**	CNAPP, CIEM API Security	CNAPP, CIEM AI + API Security
Stakeholders	IT	**IT, developers**	IT, developers	Everyone

Cloud attacks 1.0 / 1.5.



ars TECHNICA AI BIZ & IT CARS CULTURE GAMING HEALTH POLICY SCIENCE SECURITY SPACE TECH FORUM | SUBSCRIBE

SECURITY THROUGH...WHAT EXACTLY?

Defense contractor stored intelligence data in Amazon cloud unprotected [Updated]

Booz Allen Hamilton engineer posted geospatial intelligence to Amazon S3 bucket.

SEAN GALLAGHER – MAY 31, 2017 4:00 PM | 129

CloudSploit
Aug 2, 2019 · 6 min read ★ · Listen

A Technical Analysis of the Capital One Hack

. . .

The recent disclosure of yet another cloud security misconfiguration leading to the loss of sensitive personal information made the headlines this past week. This particular incident came with a bit more information from the indictment of the accused party, allowing us to piece together the revealed data and take an educated guess as to what may have transpired leading up to the loss of over 100 million credit card applications and 100 thousand social security numbers.

The Evolution of Digital Innovation.

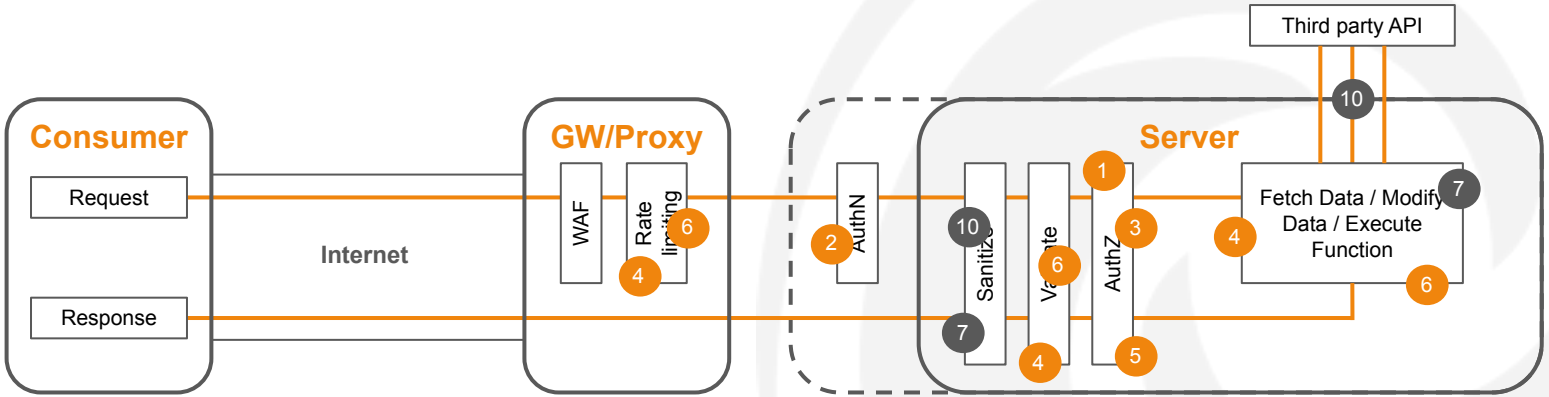
Phase	Start	Lift & shift	**Cloud 2.0**	Cloud + AI
Operations	Server-based	Virtual server	**Cloud-native**	CloudOps/ AIOps
Access	Controlled	Loosening	**Distributed**	Distributed
Challenges	Perimeter	Shadow IT Data exposures Misconfigurations	**Complex attacks** **Cloud identity** **API attack surface**	Complex attacks API breaches Data governance
Tools	Firewall, EDR, SIEM, DLP	CSPM	**CNAPP, CIEM** **API Security**	CNAPP, CIEM AI + API Security
Stakeholders	IT	IT, developers	**IT, developers**	Everyone

Verizon.

Tracfone Consent Decree.

- **Regulatory Compliance and Penalties:** Failure to secure APIs and protect customer data resulted in significant fines and regulatory scrutiny, as demonstrated by the $16 million Tracfone settlement with the FCC.
- **API Vulnerabilities as a Critical Risk:** TracFone's data breaches were linked to API vulnerabilities, underscoring that enterprises must prioritize securing APIs to prevent unauthorized access to sensitive data.
- **Impact on Customer Trust:** Breaches exposing customer information like PII and CPNI can damage customer trust and brand reputation, making robust API security a business imperative.
- **Comprehensive Security Measures:** The FCC's mandated security measures for TracFone include adopting standards such as NIST and OWASP, highlighting the need for enterprises to implement widely-accepted security frameworks.

Threat modeling.



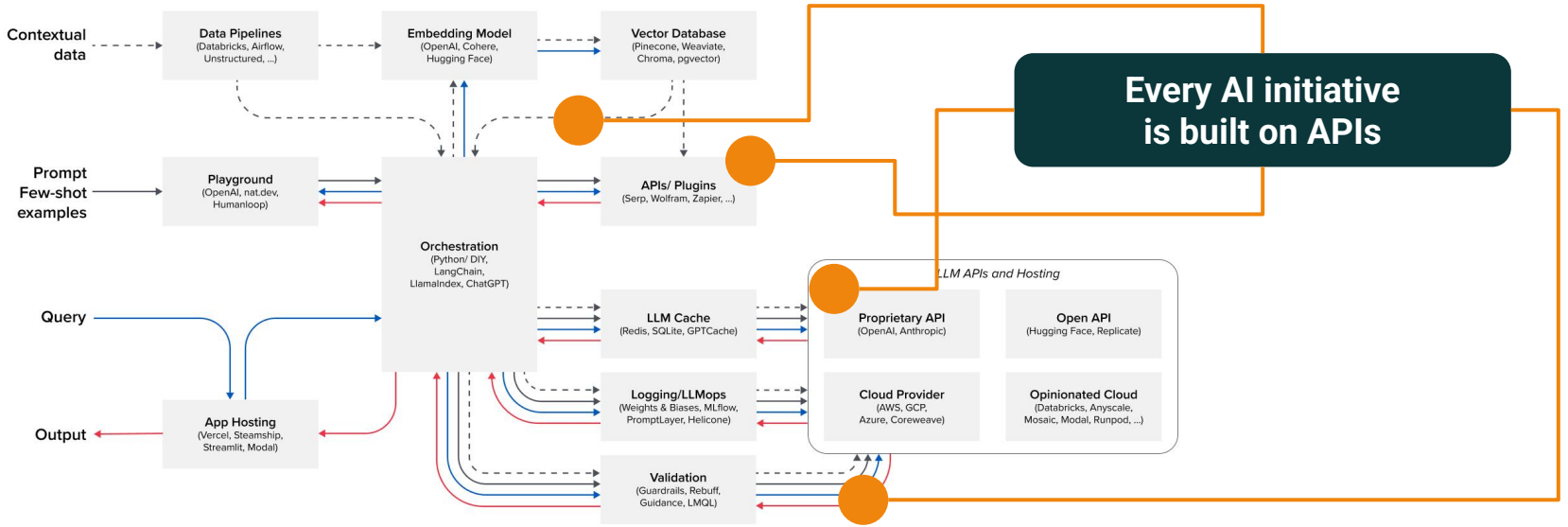
1. BOLA.
2. Broken AuthN.
3. BOPLA.
4. Unrestricted Resource Consumption.
5. BFLA.
6. Unrestricted Process Access.
7. SSRF.
8. Misconfiguration.
9. Improper Inventory Management.
10. Unsafe consumption of APIs.

FireTail

The Evolution of Digital Innovation.

Phase	Start	Lift & shift	**Cloud 2.0**	**Cloud + AI**
Operations	Server-based	Virtual server	**Cloud-native**	**CloudOps/ AIOps**
Access	Controlled	Loosening	**Distributed**	**Distributed**
Challenges	Perimeter	Shadow IT Data exposures Misconfigurations	**Complex attacks** **Cloud identity** **API attack surface**	**Data exfil** **API breaches** **AI governance**
Tools	Firewall, EDR, SIEM, DLP	CSPM	**CNAPP, CIEM** **API Security**	**CNAPP, CIEM** **AI + API Security**
Stakeholders	IT	IT, developers	**IT, developers**	**Everyone**

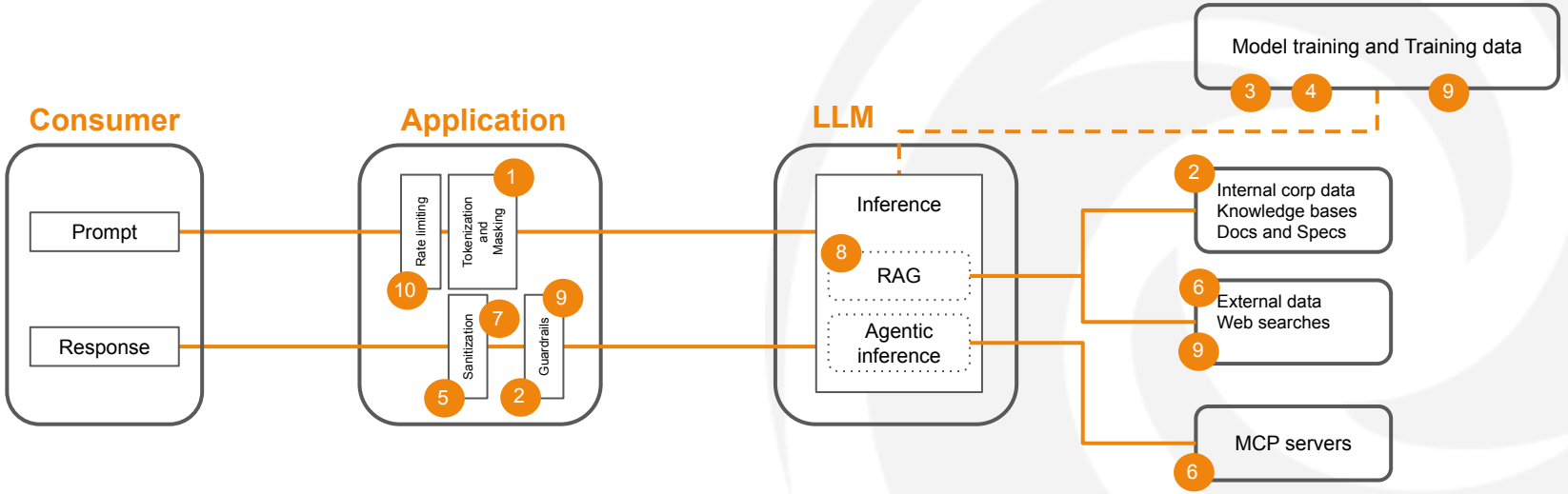
There is no AI without APIs.



Source: a16z Enterprise

FireTail

Threat modeling LLM Top 10.



1. Prompt Injection.
2. Sensitive Information Disclosure.
3. Supply Chain.
4. Data and Model Poisoning.
5. Improper Output Handling.
6. Excessive Agency.
7. System Prompt Leakage.
8. Vector and Embedding Weaknesses.
9. Misinformation.
10. Unbounded Consumption.

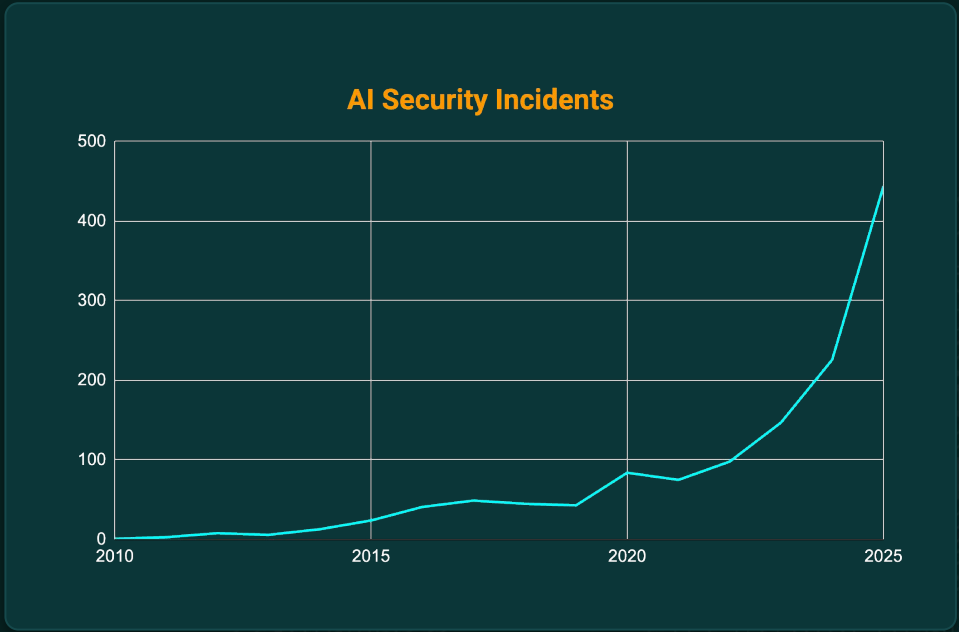
FireTail

THE EXPONENTIAL GROWTH OF AI SECURITY INCIDENTS

FireTail

AI security incidents are climbing fast

November 2022 was both an industry inflection point, and a security inflection point. The data doesn't lie. This will be the biggest security challenge for the decade ahead.



OpenClaw.

A Wakeup Call for the Enterprise.

1. **The Exposure:** Over 1,000 local developer machines (Macs/Linux) exposed to the public internet with root shell access due to a default "localhost" misconfiguration.
2. **The Leak:** 1.5 million API keys (OpenAI, AWS, Anthropic) and sensitive user records leaked via a related AI social network called "MoltBook."
3. **The Reality:** This wasn't just a consumer data breach; it was a live-fire demonstration of what can go wrong, illustrated at scale.
4. **The Takeaway:** "Shadow Agents" are part of the the new Shadow IT and they may be inside your organization already.

The problem with Securing new tech.

FireTail

Same story, but evolving .

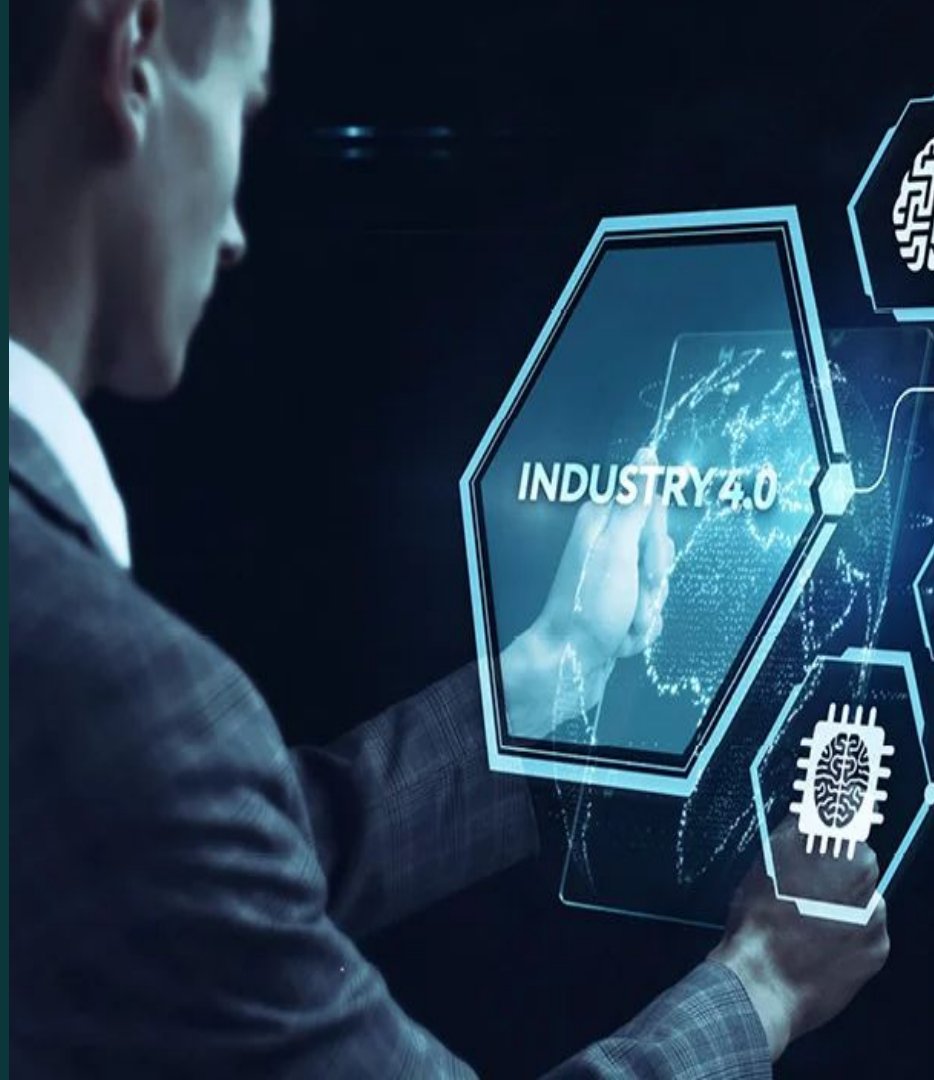
Threat-tool mismatch	Each new technology enables great new things, but also introduces new threats that existing tools don't cover.
Budget & org time lag	Defenders face organizational challenges in getting budget approved, gathering requirements, buying & deploying.
Security debt	The previous tech wave isn't fully secured.
Stakeholder expansion	Everything, everywhere, all at once.

FireTail

The AI revolution.

The current pace of change and potential for disruption presents both opportunities and existential threats to companies. Companies cannot afford to sit on the sidelines while this happens, yet they must also avoid the risks of loss of trust and costly data breaches.

FireTail

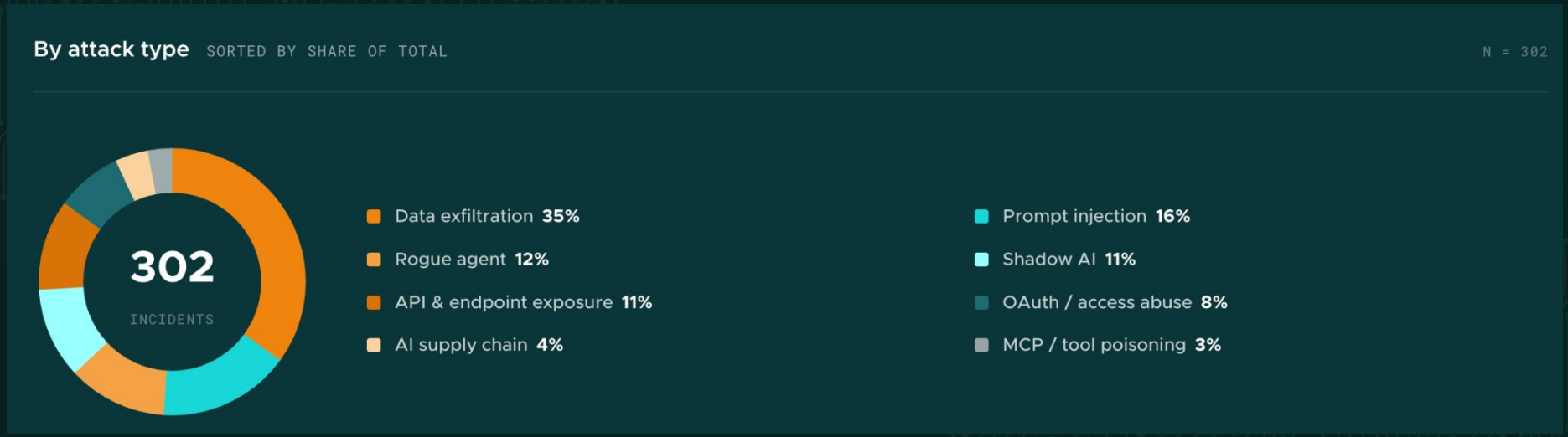


AI & the API attack surface.

FireTail

THE EXPONENTIAL GROWTH OF AI SECURITY INCIDENTS

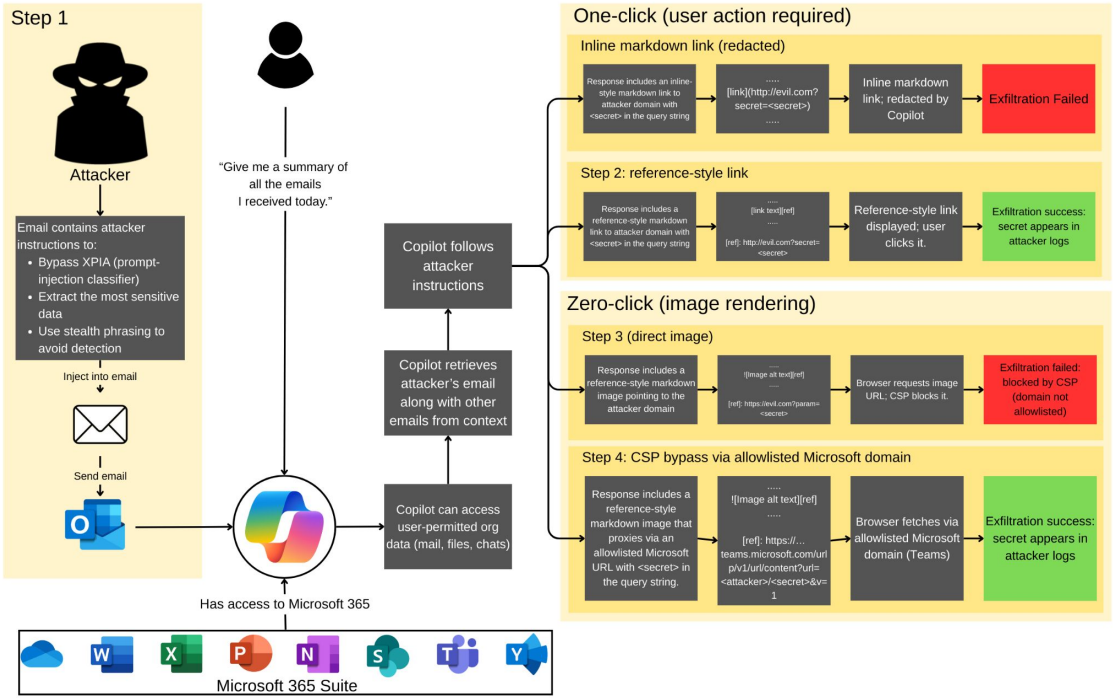
FireTail



FireTail's State of AI Security 2026

EchoLeak.

Prompt Injection & Data Exfil via API.

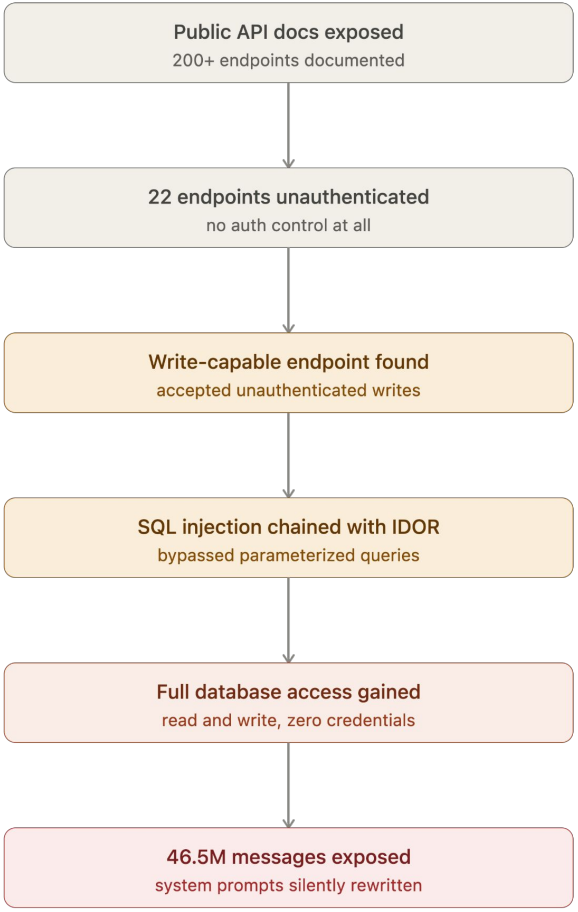


1. Copilot calls the Microsoft Graph API to fetch email
2. Auth check is performed but no content-layer inspection
3. Email content contains the prompt injection
4. EchoLeak's data theft happened through Copilot's own markdown/image-rendering pipeline. The injected instruction caused Copilot's response to include a markdown image tag pointing to an attacker-controlled URL, with stolen data appended as a parameter.

Source: https://arxiv.org/html/2509.10540v1

Lilli.

Internal AI; public APIs.



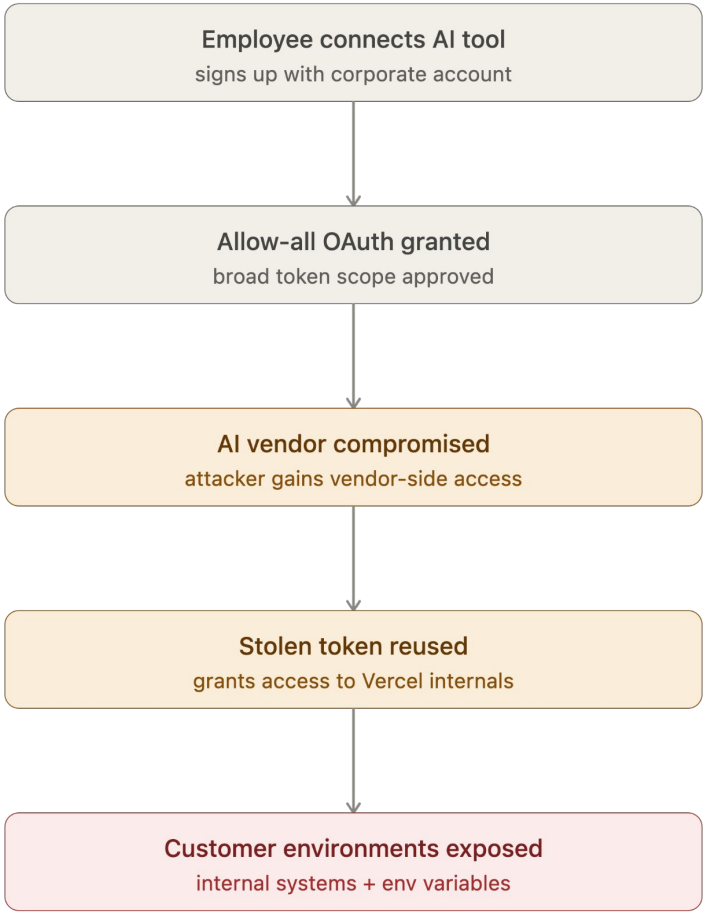
~2 hours, fully autonomous, $20 in compute

1. Public documentation
2. An LLM is exhaustive
3. Chaining is a particular skill of LLMs
4. LLM gains CRUD on DB via API

Source: Various

Vercel.

API oAuth.



Entry point: one OAuth scope, not a hijacked package or code exploit

1. Sign-up
2. Get token
3. Token isn't least privilege
4. Token -> vendor pwned
5. Vendor pwned -> customers

Source: Various

LiteLLM.

Supply chain API GW.

Entry point: an AI component in the supply chain nobody was tracking

- Every AI app needs an API
- An API needs a gateway
- Speed > security
- Ouch
- YMMV

Source: Various

AI & the API attack surface.

FireTail

In these examples...

- The models were never compromised.
- Every attacker can afford to be thorough. Any attack surface can and will (probably) be found.
- Any lack of authN (& maybe also /Z) evident from code, documentation or API spec will be found.
- There's no excuse for not doing basic SCA / SAST on API code.
- There's no excuse for not running a basic API GW or WAF.
- Secure configuration > new tools in these cases.
- Don't sacrifice security for speed.
- If you can avoid it. FOMO is real. Org inertia is real.

Secure APIs to secure AI.

Discovery.

Find all of your AI initiatives, from code to cloud with simple, continuous integrations.

Visibility.

Get a complete view of your AI integrations, with configurations & context.

Policy.

Runtime protection with call validation, authentication checks, injection prevention.

Secure Code.

APIs analyzed for misconfigurations and vulnerabilities. AI assessed for data sharing risks.

- $4.37M est. cost/breach
- API GW reduces risk 91%
  - authN/Z
  - Rate limits
  - Anomaly detection
- Securing APIs is the single most effective and cheapest action
- 1-3 wk ROI

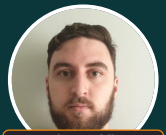
About FireTail.

FireTail

FOUNDERS

Jeremy Snyder
Co-Founder & CEO

- ex-AWS, DivvyCloud, Rapid7, REAN
- Led sales, biz dev, corp dev at DivvyCloud
- Rapid7 M&A



Riley Priddle
Co-Founder & CTO

- ex-SkyTV, Hewlett Packard, Otro
- 5x AWS certified
- 10+ yrs of web service development

ADVISORS

Mikko Hyppönen
Strategy

Sounil Yu
Security

Ted Julian
Commercial

Matt Peters
Product

Jordan McQueen
Engineering

Jonathan Leneghan
Engineering

INVESTORS

PALADIN CAPITAL GROUP

zscaler

SECURE OCTANE

GENERAL ADVANCE

FireTail

MODERN CYBER with Jeremy Snyder

CYBERTECH 100

TC STARTUP BATTLEFIELD 200

CIS

AI SECURITY INCIDENT TRACKER

API DATA BREACH TRACKER

AICPA SOC

ASIA 2025 FINALIST black hat STARTUP SPOTLIGHT

USA 2025 FINALIST black hat STARTUP SPOTLIGHT

Blackhat Startup Spotlight 2025.

Global finalist

FireTail customers include:

4.8 ★ on G2

Plerion

tidal .CLOUD

modelve

1upHealth

ION MOBILITY

"FireTail has been an absolute game-changer for us." ReFocus

THE COMPLETE AI SECURITY & GOVERNANCE PLATFORM

One platform. Complete coverage.

FireTail

Workforce .

All AI tools and services your employees use.

BROWSER • WORKSPACE • ENDPOINT

Workload .

All AI built into your applications and agents

CODE • CLOUD • AGENTS

UNIFIED AI OBSERVABILITY

- Continuous AI Discovery
- Centralized AI Logs with Detection & Response
- AI Security Posture Management
- AI Governance & Policy Engine

UNIFIED AI OBSERVABILITY PIPELINES

FireTail

WORKFORCE AI PIPELINE

All AI tools and services your employees use.

BROWSER • WORKSPACE • ENDPOINT

WORKLOAD AI PIPELINE

All AI built into your applications and agents.

CODE • CLOUD • AGENTS

Comprehensive Workforce Observability Pipeline

USE CASES

- Shadow AI Discovery
- Sensitive Data Protection
- Topic-Driven Guardrails
- Consumption & FinOps Insights
- GRC Reporting

Comprehensive Workload Observability Pipeline

USE CASES

- Agent & Application Inventory
- Agentic Workflow Tracing
- Completion & Failure Analysis
- Monitoring & Alerting
- Token & Cost Insights

FireTail captures, normalizes and centralizes data from all AI tools and systems, regardless of provider, to deliver complete AI observability across the enterprise.

Try FireTail.ai Free.

Get these slides

And get in touch with you. We'll also offer you access to a POC environment with 30 days of AI security access.

Email me: jeremy@firetail.ai

FireTail

Thank you!

FireTail