Google Developer Groups
Cloud Chennai

AI Prompt Attacks: Cloud Under Siege, Defenses Activated

The Battle for Cloud Integrity Starts Here

cloud
community
days 2025

Google Developer Groups
Cloud Chennai

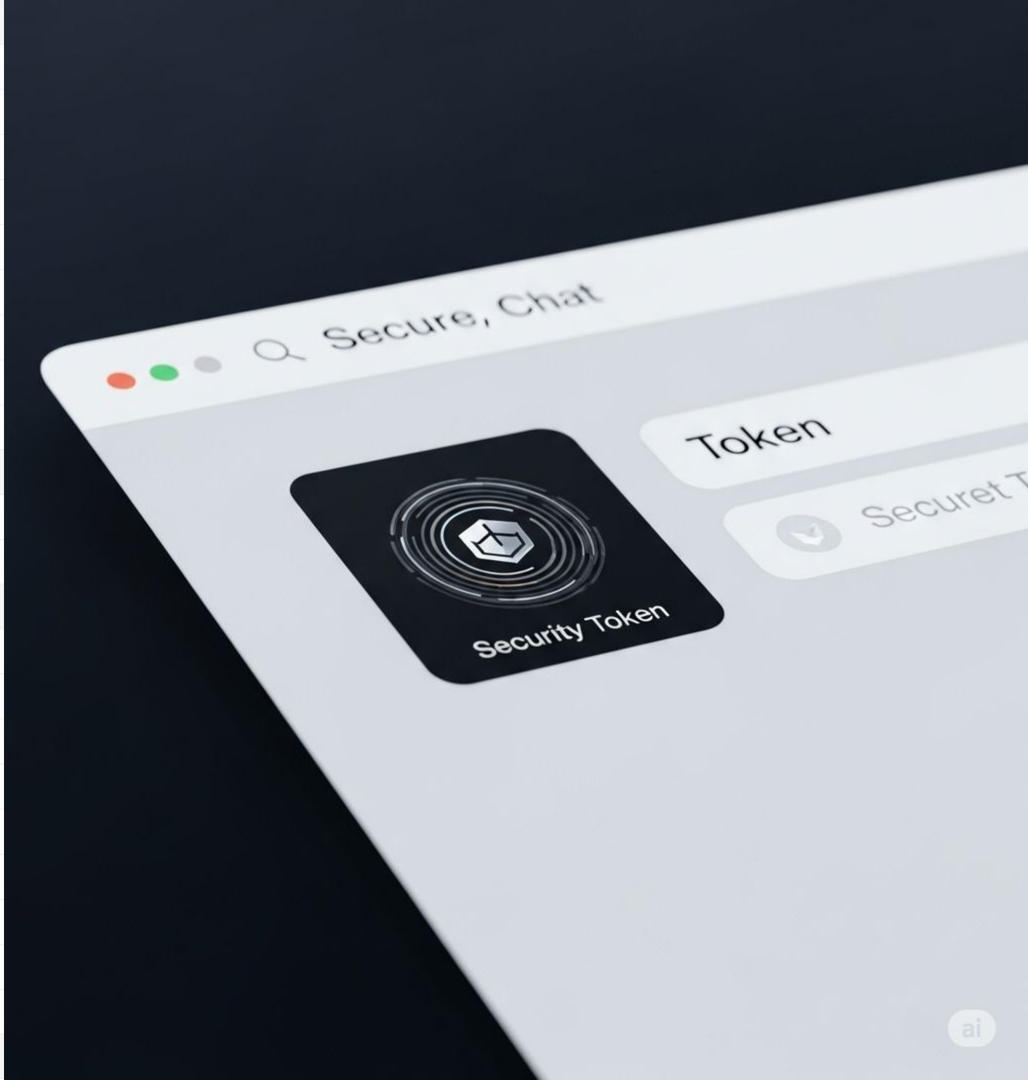
I am Sai, a cybersecurity professional with 16+ years of hands-on experience spanning red teaming, penetration testing, and cloud security architecture. Over the years, I have worked closely with organizations to strengthen their defenses, helped shape cybersecurity strategies for state government, and provided consulting across sectors. I regularly speak at industry forums, mentor aspiring security professionals, and teach at respected institutions, driven by a strong belief in building secure systems and empowering the next generation of talent

Sai Krishna Kumar Voootukuru
Senior Cloud Security Architect, Wiz

Agenda

- AI as a new threat surface
- Real attack: Prompt Injection → Token Theft → Privilege Escalation
- Demo walkthrough
- Mitigations
- Key takeaways

"What if your AI app exposed your Cloud secrets... live?"



Google Developer Groups
Cloud Chennai

Cloud misconfiguration

is one of the top risks in the cloud

https://www.business-standard.com/technology/tech-news/ransomware-attacks-human-error-main-cause-of-cloud-data-breac
hes-report-123041900719_1.html

81%

of organizations report that human error is the cause behind most cloud security breaches

https://www.business-standard.com/technology/tech-news/ransomware-attacks-human-error-main-cause-of-cloud-data-breac
hes-report-123041900719_1.html

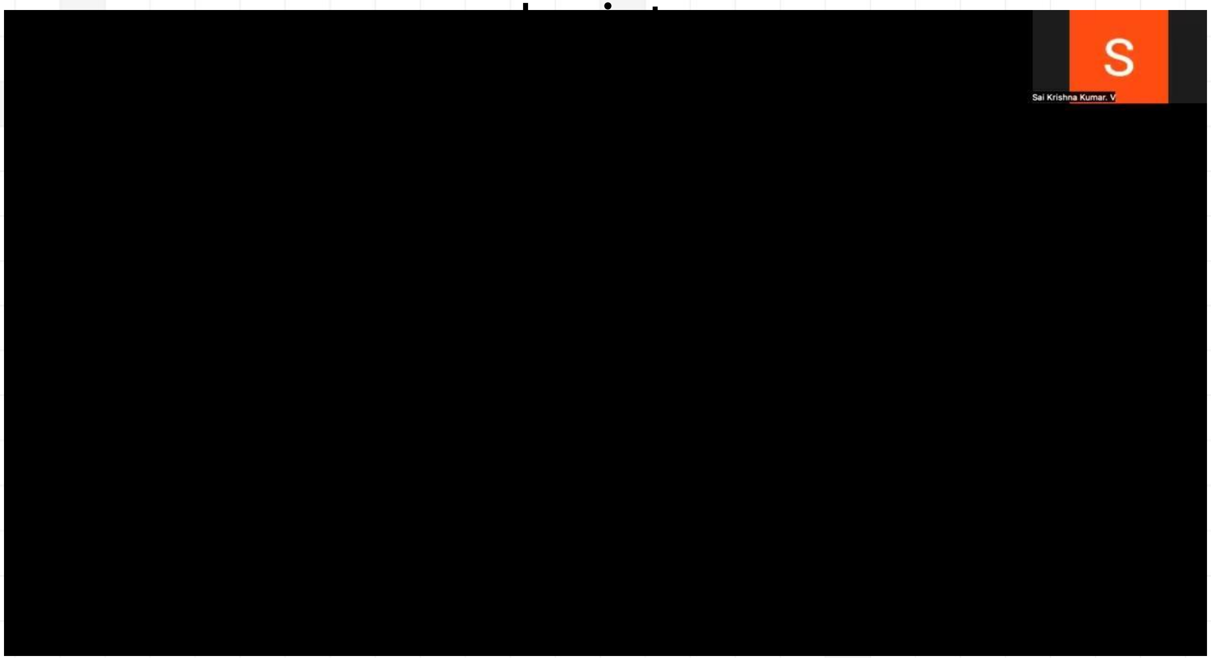
99%

of cloud users, roles, and service accounts are overly permissive

How many of you use AI?

The AI Supply Chain is the New Attack Surface

AI prompt → backend code → API call → metadata [illegible]



LLMs can be manipulated;if they connect to cloud resources, you expose your identity perimeter

Unmasking the Impersonators: Service Account Abuse

Excessive Privileges → Service Account Impersonation → Root Access

Privilege escalation thrives on unmanaged identities, exposing organizations.

Google Developer Groups
Cloud Chennai

MITRE ATT&CK Mapping

Tactic	Technique	Description
Initial Access	T1190 - Exploit Public-Facing App	Prompt injection via public AI endpoint
Execution	T1609 - Cloud Infrastructure Discovery	AI makes internal metadata API call
Credential Access	T1552.001 - Unsecured Credentials	Access token retrieved from metadata server
Privilege Escalation	T1530 - Impersonation	Used token to impersonate higher-privileged SA
Lateral Movement	T1570 - Lateral Tool Transfer	Used impersonated SA to access other services

Defensive Measures That Work

Defend Against Prompt Injection

- **Sanitize All Inputs**: Never trust user-supplied prompts, enforce strict validation.
- **Isolate Prompt Context**: Keep system prompts and user inputs separate.
- **Control LLM Outputs**: Strip sensitive content (e.g., URLs, commands, secrets).

Lock Down Service Accounts

- **Enforce Least Privilege**: Avoid wildcard roles (Editor, Owner) - assign only what's necessary.
- **Limit Token Scope & Lifetime**: Use access boundaries and TTLs for service account tokens.

Block Metadata Token Theft

- **Restrict Metadata Access**: Disable 169.254.169.254 in container runtimes.
- **Use Workload Identity Federation**: Replace metadata-based tokens with short-lived, auditable identities.

Detect & Respond Early

- **Enable Cloud Audit Logs**: Track prompt usage, token issuance, and IAM escalations.
- **Deploy Cloud Threat Detection**: Use Cloud IDS/SCC to catch anomalous activity.

Key Takeaways

- AI interfaces can lead to **real cloud identity theft**
- Metadata server is the **Achilles' heel** of cloud apps
- Over-permissive IAM = lateral movement
- Prevention needs **AI + cloud security synergy**

Thank you GDG Chennai !!

நன்றி !

Sai Krishna Kumar Vootukuru

Cybersecurity Leader | Speaker | Trainer | Empowering Secure Digital Transformations...

Google Developer Groups
Cloud Chennai

Sai Krishna Kumar Vootukuru

Senior Cloud Security Architect

WIZ