SANJAY PUNDAREKH

Offensive Security & AI Security Manager | Penetration Testing · Red Teaming · AI Red Teaming · LLM & Agentic Exploitation

Bangalore, India | +91 9538952828 | sanjaypundarekh@gmail.com | linkedin.com/in/sanjay-pundarekh

PROFESSIONAL SUMMARY

Cybersecurity and AI Security Manager with **5+ years at EY India**, where I **lead the firm's AI Security capability**. I pair deep red teaming and penetration testing experience with AI security - AI red teaming, LLM/SLM and agentic-system exploitation, prompt injection, guardrail and alignment manipulation, and AI threat modelling (MITRE ATLAS, NIST AI RMF, MAESTRO). I operate **AI-natively**, embedding LLMs and custom agentic tooling directly into offensive workflows both as assessment targets and to accelerate reconnaissance, exploitation and reporting across multi-million-dollar engagements.

PROFESSIONAL EXPERIENCE

EY India LLP - Big 4 Consulting Firm **Bangalore, India**

Manager - Cybersecurity and AI Security *September 2026 - Present*

- **Built and now lead EY India's AI Security competence within the Bangalore office**, delivering AI red teaming, AI threat modelling, and AI security gap assessments; contributed to new client acquisition generating USD 20M+ in revenue.
- Spearhead R&D for the AI Security capability, embedding AI risk assessments into core security services to enable secure client adoption of AI systems.
- **Conduct Third-Party Vendor Review (TPRM) assessments across multiple clients**, validating prospective vendors' security posture and control maturity prior to onboarding to quantify supply-chain risks.
- Lead the Network Penetration Testing team; train analysts on current adversary TTPs and deliver 300+ critical and high-risk findings supporting multi-million-dollar engagements.

Senior Consultant - Cybersecurity *April 2024 - August 2026*

- **Delivered multiple AI red teaming assessments**, surfacing 25+ Critical and High-severity issues spanning AI security and AI safety, helping clients safely scale their adoption of AI.
- **Operated AI-natively across delivery**, built and operationalized custom MCP servers, agentic harnesses (Claude Code, Gemini CLI, Hermes Agent), and locally deployed LLMs across red teaming and penetration testing to automate reconnaissance and reporting and shorten delivery timelines.
- Performed application and AI threat-modelling exercises aligned to MITRE ATLAS, NIST AI RMF, and MAESTRO, identifying architectural and security flaws and improving client security posture.
- Led Red Team and Assumed Breach assessments across enterprise and cloud environments, consistently achieving domain compromise and identifying critical architectural weaknesses.
- Core team member empanelling EY India with CERT-In (90% in the practical evaluation); design CTF challenges for DSCI's Annual Information Security Summit (2023–2025); lead Bangalore's offensive-security training and contribute winning client proposals.

Consultant - Cybersecurity *October 2022 - March 2024*

- Led a Red Team assessment for a multinational IT services firm, compromising the internal network and Azure cloud assets to achieve Domain Administrator in 36 days.
- Conducted security assessments across on-premise and cloud infrastructure, uncovering critical threats and misconfigurations and driving a 40% reduction in potential breach exposure.
- Directed cyber war-gaming simulations and tabletop exercises, strengthening clients' incident response capabilities and resilience against emerging threats.
- Ranked in the top 25% within EY's Advanced Security Centre (ASC) on Hack the Box.

Associate Consultant - Cybersecurity *January 2022 - September 2022*

- Led a team of 8 consultants on a key-account penetration test, identifying 200+ critical/high and 800+ medium/low vulnerabilities across web, SAP, and hybrid cloud/on-prem infrastructure.
- Penetration tested a hybrid Active Directory environment, identifying tens of thousands of domain users with local administrator privileges on a critical server, enabling lateral compromise of subsequent servers.

Cybersecurity Analyst - Cybersecurity *January 2021 - January 2022*

- Performed penetration testing and secure code review on Android and iOS mobile applications; drafted CIS-aligned hardening checklists for network devices.

SKILLS & TOOLS

AI Security: AI Red Teaming, AI Threat Modelling, LLM/SLM & agentic-system security assessments, prompt injection, guardrail bypass, alignment manipulation, model robustness evaluation; MITRE ATLAS, NIST AI RMF, MAESTRO.

AI-Native Tooling: Claude Code (with custom offensive-security skills), Gemini CLI, Hermes Agent, custom Model Context Protocol (MCP) server development, locally deployed LLMs, n8n workflow automation, AI-driven pipelines integrating models with security tooling.

Offensive Security: Red Teaming, Assumed Breach Assessments, Network / Web / Mobile (iOS & Android) / SAP / Cloud Penetration Testing, Active Directory & Entra ID exploitation, AWS & Azure offensive operations, secure code review, threat modelling, security architecture review.

Third-Party & Risk: Third-Party Vendor Review (TPRM), vendor security-posture validation, pre-onboarding and supply-chain risk assessment, control-maturity evaluation.

Frameworks & Standards: MITRE ATT&CK, MITRE ATLAS, OWASP Top 10, CIS Benchmarks, NIST AI RMF, MAESTRO.

Programming & Tooling: Python, PowerShell, Bash; Windows & Linux.

Leadership & Communication: Team leadership across remote and on-site engagements, senior-stakeholder advisory, proposal development, technical training delivery, and translating complex security risk for technical and non-technical audiences.

SECURITY TOOLING & RESEARCH

- **Jörmungandr (Model Robustness Evaluation) -** An AI robustness framework that stress-tests LLMs and SLMs against real-world failure scenarios including alignment manipulation, guardrail bypass, narrative reframing, and reliability degradation.
- **AI-Integrated Black Box Assessment Tool -** An AI-enabled external attack-surface framework integrating Claude Code with automated reconnaissance pipelines to identify, fingerprint, and correlate internet-facing assets, with centralized dashboarding for exposure tracking.

CERTIFICATIONS

Offensive Security Certified Professional (OSCP+) · Certified Red Team Expert (CRTE) · Offensive Cloud Specialist - AWS (HTB) · Penetration Tester Level 2 - Dante (HTB)

EDUCATION

Dayananda Sagar College of Engineering **Bangalore, India**

B.E., Electronics & Instrumentation Engineering *August 2016 - August 2020*

- CGPA 9.27/10.0 (3.89/4.0); First Class with Distinction; ranked in the top 5% of the class. Relevant coursework: Computer Networks, Process Automation, Microcontrollers, Control Systems, Artificial Intelligence.