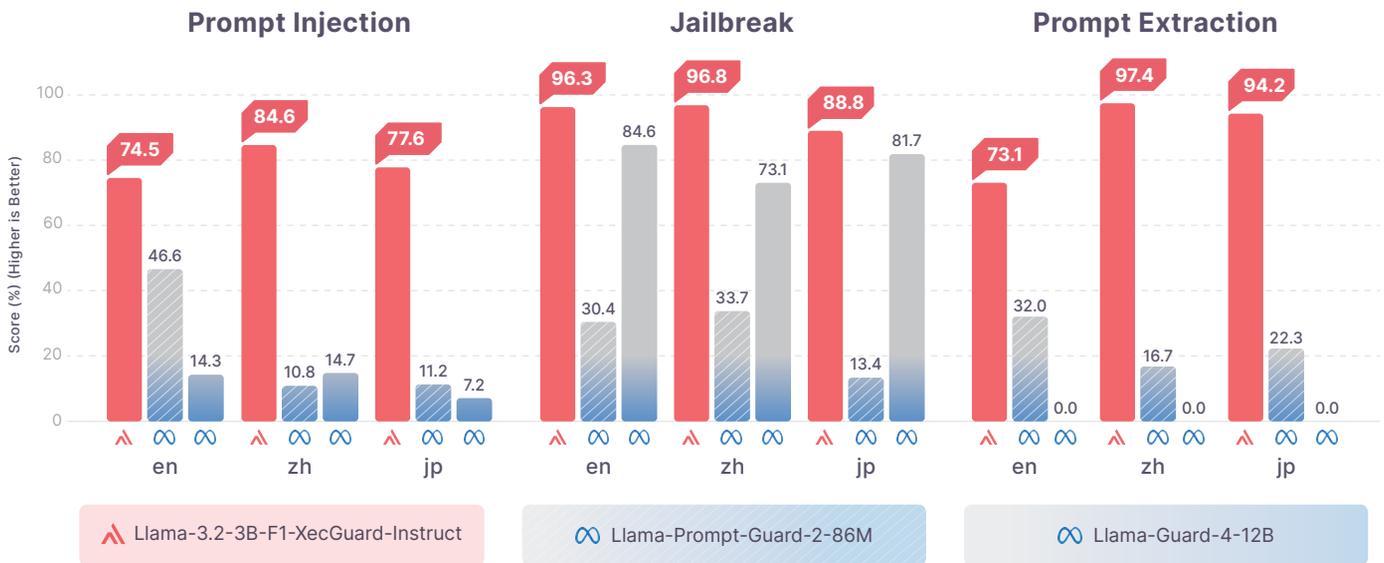


CyCraft Guards Trustworthy LLMs



AI Blue Teaming Next-Generation AI Firewall

Outstanding Performance: XecGuard's Industry-Leading Security Protection



Reinforce System Prompt Compliance in Normal Contexts Reinforce Malicious User Prompt Detection



Prompt Injection Defense

Detect malicious instructions, output manipulation, instruction obfuscation, logical/emotional appeals, and role-playing attacks, preventing deviations from application-defined System Prompts to ensure compliance.



Prompt Extraction Defense

Prevent attackers from guiding inputs to exfiltrate application System Prompts, safeguarding internal logic and confidential information.



Jailbreak Attack Defense

Block safeguard-bypassing context manipulation, preventing outputs that violate ethical norms or security standards to preserve AI model's safety and reliability.