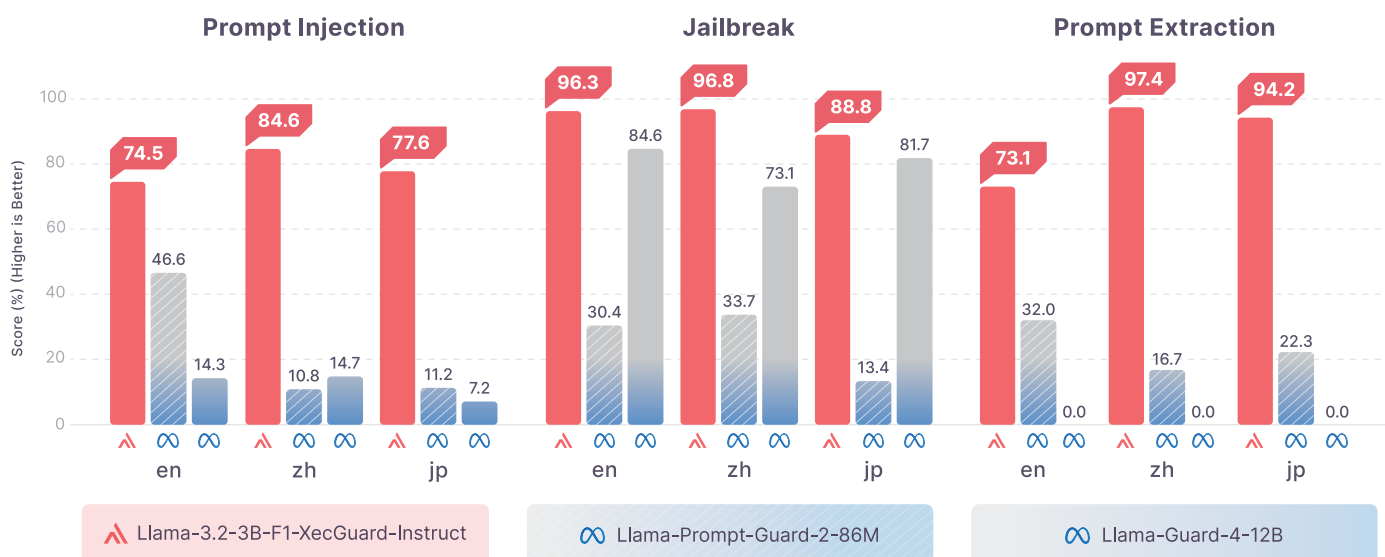


AI Blue Team
新世代 AI 的標配防火牆

超越防護基準：XecGuard 全面領先業界安全表現



強化 System Prompt 正常語境指令遵循 | 強化 User Prompt 惡意語境識別能力



防禦提示詞注入 Prompt Injection

識別惡意指令、輸出格式操控、指令混淆編碼、邏輯與情感說服、角色扮演，避免輸出結果偏離應用程式自訂的 System Prompt，維持正常合規運作。



防禦提示詞竊取 Prompt Extraction

防止攻擊者透過引導輸入竊取應用程式的 System Prompt，確保內部邏輯與機密安全不外洩。



防禦越獄攻擊 Jailbreak Attack

阻擋繞過防護的語境操控，避免模型輸出違反善良風俗或安全規範的內容，維護 AI 的信任與可靠性。