AI-SEC

Secure GenAI und Agent Threat Modeling

Zwei Tage, nach denen Sie ein Threat Model, ein Datenflussbild und einen Testplan für eine echte Anwendung haben statt einer Liste theoretischer Risiken.

DAUER	GRUPPE	FORMAT
2 Tage	bis 12 Personen	Inhouse · Remote · Im Lab · Deutsch oder Englisch

DAS EXISTIERT DANACH

- Ein Datenflussbild Ihrer Anwendung von der Eingabe über Modell, Wissensquellen und Werkzeuge bis zur Ausgabe
- Ein Threat Model, das eine STRIDE-artige Systematik auf GenAI und Agenten überträgt
- Ein Katalog konkreter Schutzmaßnahmen: Guardrails, Berechtigungen und Ausgabefilter, je Bedrohung zugeordnet
- Ein Testplan, der prüfbar macht, ob eine Maßnahme wirkt
- Eine priorisierte Restrisikoliste mit Verantwortlichen

AGENDA

Tag 1 · 09:00	Warum GenAI andere Bedrohungen hat: Prompt Injection, unsichere Ausgaben, Datenabfluss über Wissensquellen und Werkzeuge. Kernfrage: Wo unterscheidet sich das von klassischer Anwendungssicherheit und wo nicht?
Tag 1 · 10:30	Datenflussbild Ihrer Anwendung: Wir zeichnen den vollständigen Fluss: Eingabe, Modell, RAG-Quellen, Tool-Aufrufe, Ausgabe. Übung: Jede Vertrauensgrenze wird markiert und benannt.
Tag 1 · 13:30	STRIDE-artige Systematik auf GenAI übertragen: Wir gehen die Bedrohungskategorien durch und übersetzen sie auf Prompt, Kontext, Werkzeuge und Ausgabe. Der OWASP-LLM-Referenzrahmen dient als Checkliste, damit keine Kategorie vergessen wird.
Tag 1 · 15:30	Bedrohungen an Ihrer Anwendung sammeln: Übung an Ihrem Datenflussbild: Je Vertrauensgrenze benennen wir plausible Angriffe. Tagesergebnis: ein erstes Threat Model mit priorisierten Bedrohungen.
Tag 2 · 09:00	Live-Demonstration im Lab: Wir zeigen an einer präparierten Umgebung, wie eine Prompt Injection über eine RAG-Quelle wirkt und was ein zu weit berechtigtes Werkzeug preisgibt. Keine Kundendaten, nur zum Verständnis der Wirkung.
Tag 2 · 11:00	Schutzmaßnahmen je Bedrohung: Guardrails, Berechtigungen nach geringstem Recht, Ausgabefilter, Quellenprüfung. Entscheidung je Bedrohung: welche Maßnahme, welcher Aufwand, welches Restrisiko bleibt?
Tag 2 · 14:00	Testplan bauen: Zu jeder Maßnahme ein Test, der ihre Wirkung nachweist. Übung: Wir formulieren pro wichtiger Bedrohung einen konkreten Testfall mit erwartetem Ergebnis.
Tag 2 · 16:00	Restrisiken und Übergabe: Wir tragen zusammen, was bleibt, und benennen Verantwortliche. Tagesergebnis: Threat Model, Maßnahmenkatalog, Testplan und priorisierte Restrisikoliste für Ihre Anwendung.

IHR TRAINER

Dino Bordonaro

8x Microsoft MVP in Folge (Microsoft Azure · Cloud and Datacenter Management), Microsoft Certified Trainer, fast 30 Jahre IT. Baut souveräne KI-Systeme auf eigener Hardware (Azure Local, Foundry Local, NVIDIA-GPUs) und schult genau darauf. Über 800 Professionals ausgebildet, jedes Konzept im eigenen Enterprise-Lab im TÜV-geprüften Rechenzentrum erprobt.

15.800 € pauschal, zzgl. MwSt.

Der Track liefert Modell, Maßnahmen und Testplan als Entwurf. Er ersetzt keine Implementierung, kein Penetrationstest-Mandat und keinen betriebenen Sicherheitsdienst.

AI-SEC

Dino Bordonaro · BORDONARO IT GmbH & Co. KG · info@bordonaro-it.de · +49 6232 298 53 - 0 · bordonaro.ai