Building AI Agents with a practical deep dive into Microsoft's new AI Red Teaming Agent

Course code: WAIA

AI Red Teaming relies on the creative human expertise of highly skilled safety and security professionals to simulate attacks. The process is resource and time intensive and can create a bottleneck for many organizations to accelerate AI adoption. With the AI Red Teaming Agent, organizations can now leverage Microsoft's deep expertise to scale and accelerate their AI development with Trustworthy AI at the forefront.

Outline:

- Definition and types of AI agents
- Real-world applications and use cases
- Discussion: The role of AI agents in modern technology
- Understanding AI red teaming and its importance
- Overview of Microsoft's AI Red Teaming Agent
- Key features: automated scans, attack strategies, and reporting
- Supported risk categories and attack techniques
- Installing necessary tools and dependencies
- Configuring Azure AI Foundry and the AI Red Teaming Agent
- Running scans on a sample AI model
- Final Project: Build an AI Agent to perform Network Exploitation tasks

Prerequisites:

- Basic understanding of AI and machine learning concepts
- Azure MS account

Who Should Attend?

AI engineers, ML practitioners, security researchers, and technical decision-makers who want to integrate Trustworthy AI and proactive testing into their development pipeline.

GOPAS Praha
Kodaňská 1441/46
101 00 Praha 10
Tel.: +420 234 064 900-3
info@gopas.cz

GOPAS Brno
Nové sady 996/25
602 00 Brno
Tel.: +420 542 422 111
info@gopas.cz

GOPAS Bratislava
Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 248 282 701-2
info@gopas.sk