COURSE TABLE OF CONTENTS

RAG Security

OEM: Koenig Original • Duration: 2 Days (16 hrs)

COURSE MODULES & TOPICS

Module 1: Architecture and Threat Modelling

- Documents, chunking, embeddings, vector stores, retrieval, generation
- Trace training application's data flow and trust boundaries (LAB)

Module 2: Ingestion and Document Integrity

- Source trust, document parsing, provenance, ingestion permissions
- Ingest synthetic documents and test source and metadata validation (LAB)

Module 3: Poisoning and Retrieval Manipulation

- Untrusted knowledge entries, misleading context, retrieval influence
- Introduce poisoned test document and observe retrieved context (LAB)

Module 4: Prompt Injection via Retrieved Content

- Instructions in documents, answer manipulation, disclosure attempts
- Test indirect prompt injection through a retrieved document (LAB)

Module 5: Access Control and Data Leakage

- Document permissions, tenant isolation, sensitive context exposure
- Test document access across roles and tenant boundaries (LAB)

Module 6: Embedding and Vector Store Risks

- Embedding privacy, inversion concepts, vector store exposure
- Inspect vector store permissions and exposed document metadata (LAB)

Module 7: Defences and Security Evaluation

- Permission aware retrieval, content controls, output handling, logging
- Apply controls and rerun attack and benign test cases (LAB)

Module 8: Integrated Assessment

- Evidence collection, business impact, mitigations, residual risk
- Assess RAG application and present concise findings report (LAB)

© Koenig Solutions