Curso: Auditor de IA

Duração: 21h
Área formativa: Cursos

Sobre o curso

Aprende a auditar sistemas de IA, modelos generativos e agentes autónomos segundo o AI Act, RGPD e frameworks internacionais.

A adoção de Inteligência Artificial está a crescer rapidamente nas organizações. Com essa expansão surge uma exigência crítica: garantir que os sistemas de IA são seguros, éticos, transparentes e conformes com a regulamentação europeia.

Neste curso Auditor de AI, aprendes a avaliar, validar e auditar modelos generativos, modelos de decisão, agentes autónomos e automações inteligentes. Desenvolves competências para identificar riscos, aplicar controlos, validar conformidade legal e produzir documentação auditável.

Ao longo de 21 horas, trabalhas metodologias, métricas e frameworks reconhecidas internacionalmente, culminando numa auditoria aplicada em contexto simulado.

Metodologia

:: Sessões baseadas em casos reais

:: Exercícios práticos e análise de cenários

:: Aplicação de ferramentas e métricas de auditoria

Auditoria final aplicada em contexto simulado

Objectivos

Ao longo deste curso vais:

:: Compreender o ciclo de vida e os riscos de sistemas de IA;

:: Avaliar modelos generativos, modelos preditivos, agentes e automações inteligentes;

:: Aplicar metodologias de avaliação de qualidade, fairness, segurança e ética;

:: Classificar sistemas segundo o AI Act;

:: Validar conformidade com AI Act, RGPD e políticas internas;

:: Analisar datasets, outputs e logs de forma crítica;

:: Definir controlos e medidas de mitigação;

:: Produzir relatórios e documentação de auditoria;

:: Realizar uma auditoria completa em contexto simulado.

Pré-requisitos

:: Competências digitais básicas;

:: Compreensão geral de processos de negócio;

:: Não é necessário saber programar.

Destinatários

:: Auditores internos e externos;

:: Profissionais de compliance, risco e governance;

:: Gestores de projeto envolvendo IA;

:: Analistas técnicos e funcionais;

:: Profissionais de TI, qualidade ou segurança da informação;

:: Consultores e responsáveis por transformação digital.

Programa

Introdução à Auditoria de IA, Regulação e Ciclo de Vida dos Sistemas

Este módulo estabelece a base conceptual da auditoria de IA, enquadrando o papel do auditor, o ciclo de vida completo de sistemas de Inteligência Artificial e a aplicação do AI Act e do RGPD. Sabe como analisar quando e por que razão um sistema deve ser auditado, classificando soluções segundo níveis de risco e obrigações legais.

- O que é auditar IA? Diferença entre auditoria, avaliação e monitorização
- Tipos de sistemas auditáveis: LLMs, modelos preditivos, agentes, automações inteligentes
- Ciclo de vida da IA: design → data → model → deploy → monitorização
- AI Act: classificação por risco, obrigações, documentação exigida
- Relação entre AI Act, RGPD e políticas internas
- Responsabilidade organizacional e accountability

Riscos e Impactos em Sistemas de IA

Com foco na identificação e classificação de riscos técnicos, éticos, operacionais e jurídicos associados a modelos generativos, agentes e automações inteligentes, neste módulo desenvolves uma matriz de risco estruturada e aplicas metodologias como AI Impact Assessment para avaliar impactos reais e potenciais.

- Tipos de risco: técnico, ético, operacional, reputacional, jurídico
- Hallucinations, decision instability, model drift, prompt injection
- Riscos nos dados: enviesamentos, incompletude, dados sensíveis, fuga de informação
- Riscos em agentes e automações (loops, erros de execução, decisões incorretas)
- AI Impact Assessment (AIA): metodologia base
- Identificação de riscos em outputs e logs

Auditoria Técnica: Dados, Modelos e Outputs

Aprende a avaliar tecnicamente sistemas de IA sem necessidade de programação, utilizando métricas de desempenho, testes e análise crítica de datasets. Neste módulo abordas os requisitos técnicos do AI Act, robustez, segurança e validação de resultados.

- Métricas essenciais: accuracy, precision, recall, F1-score, perplexity
- Testes A/B e testes de stress
- Avaliação de qualidade de datasets (completude, equilíbrio, noise, skew)
- Validação de outputs: consistência, coerência, confiabilidade
- Logs, tracking e análise de comportamento
- Requisitos técnicos do AI Act: performance, robustness, cybersecurity

Fairness, Ética e Não-Discriminação

Explora métodos para identificar e mitigar bias em sistemas de IA, garantindo equidade e alinhamento com princípios éticos. Aplica neste módulo testes de fairness e define estratégias de mitigação em modelos generativos e preditivos.

- O que é fairness? Tipos de bias e impacto societal
- Testes de fairness: equal opportunity, demographic parity, equalised odds
- Riscos éticos em LLMs, modelos preditivos e automações
- Mitigação: data balancing, rule-based overrides, guardrails
- Ética aplicada: explainability, transparência, responsabilidade
- Erros éticos comuns em IA generativa

Segurança, Privacidade e Controlo

Aborda vulnerabilidades técnicas, proteção de dados e requisitos do RGPD aplicados à Inteligência Artificial. Descobre como avaliar pipelines, acessos e controlos operacionais para garantir segurança e conformidade.

- Vulnerabilidades: adversarial attacks, prompt injection, data exfiltration
- Segurança de pipelines de IA: tokens, acessos, ambientes, logs
- Privacidade em IA: dados pessoais, minimização, retention policies
- RGPD aplicado a IA: bases legais, consentimento, DPIAs
- Guardrails, validação humana e controlo operacional
- Padrões técnicos do AI Act para segurança

Auditoria a Agentes de IA e Automações Inteligentes

Focado na auditoria de ecossistemas distribuídos e sistemas multi-agente, este módulo analisa decisões automatizadas, integrações com APIs e fluxos complexos. Sabe como avaliar quando é obrigatório implementar human-in-the-loop e como validar exceções.

- Avaliação de agentes individuais e multi-agent systems
- Riscos em automações inteligentes (Make, Power Automate, Zapier, n8n)
- Validação de decisões, ações e exceções
- Human-in-the-loop: quando é obrigatório
- Testes de comportamento e limites
- Auditoria de integrações com APIs e dados internos

Frameworks, Metodologias e Evidências

Explora a aplicação prática de frameworks internacionais como NIST AI RMF e ISO/IEC 42001 para estruturar auditorias de IA com rigor e consistência. Aprende a identificar evidências obrigatórias e preparar auditorias internas e externas.

- NIST AI Risk Management Framework
- ISO/IEC 42001 – AI Management System
- Estrutura de auditorias de IA segundo best practices
- Evidências obrigatórias e não obrigatórias
- Checklists de conformidade
- Preparação para auditorias internas/externas