Matriz de Riscos e Controles - IA em Grandes Empresas

Atlantyx | Referência de governança | Data: 12/11/2025

1. Objetivo

Apresentar riscos típicos de IA em grandes empresas e controles recomendados para mitigação.

2. Premissas

Escopo inclui assistentes RAG, automações com LLM e modelos preditivos.

Os controles listados devem ser adaptados ao setor (energia, finanças, varejo) e à criticidade.

Tabela 1 - Matriz de riscos e controles

Risco	Impacto	Sinais	Controles mínimos
Alucinação em resposta crítica	Alto	Resposta sem fonte	Obrigar citações; bloquei
Vazamento de informação (exfiltração)	Alto	Pedidos por segredos	RBAC/ABAC no retrieval;
Prompt injection	Médio/Alto	Instruções no doc: 'ignore as regras'	Separar system prompt; v
Custo explosivo por uso	Médio	Tokens por pergunta cresce	Cache; top-k dinâmico; li
Dados desatualizados	Médio	Docs conflitantes	Versionamento; effective_

Tabela 2 - Política de retenção (recomendação)

Tipo de dado	Retenção sugerida	Observação
Logs de métricas (sem conteúdo)	90 dias	Suporte a auditoria e investigação de incidentes
Conteúdo de prompt/resposta	0 dias (não reter)	Somente em laboratório com aprovação formal
Feedback do usuário	180 dias	Anonimizar quando possível