

Relatório – Análise Ética em Inteligência Artificial

Caso Escolhido

O caso analisado é o do viés algorítmico no recrutamento da Amazon. Em 2018, a empresa suspendeu o uso de um sistema de inteligência artificial para triagem de currículos após identificar que o modelo apresentava viés contra mulheres. O algoritmo foi treinado com dados históricos de contratações majoritariamente compostas por homens, o que levou a decisões discriminatórias.

Análise com Framework Ético

Viés e Justiça

- Tipo de viés:

Viés de dados: modelo treinado com histórico desigual (predomínio de homens).

Viés algorítmico: penalização de termos associados a gênero feminino.

- Grupos afetados: mulheres candidatas a vagas de tecnologia.

- Distribuição: injusta, já que candidatos homens tinham vantagem desproporcional.

Transparência e Explicabilidade

O sistema funcionava de forma vedada: os recrutadores não entendiam claramente por que certos currículos eram descartados. Não havia explicabilidade para justificar as decisões automáticas, comprometendo a confiança no processo.

Impacto Social e Direitos

- Mercado de trabalho: perpetuava desigualdades históricas de gênero na tecnologia.

- Autonomia das pessoas: candidatas mulheres eram eliminadas sem saber o motivo.

- Direitos fundamentais: violação indireta do princípio da não discriminação e conflito com a LGPD, que exige transparência e uso ético de dados pessoais.

Responsabilidade e Governança

- Ações esperada da equipe de desenvolvimento:

Revisar os dados de treinamento para corrigir vieses históricos.

Implementar auditorias contínuas no algoritmo.

Garantir diversidade na equipe de desenvolvimento.

Leis/regulações aplicáveis:

LGPD (Brasil): princípio da não discriminação e direito à explicação de decisões automatizadas.

Leis trabalhistas antidiscriminatórias.

Diretrizes internacionais de IA responsável (OCDE, União Europeia).

Posicionamento

O sistema não deve ser banido completamente, mas precisa ser redesenhado. O uso de IA em recrutamento pode trazer benefícios como agilidade e redução de custos, desde que os vieses estruturais sejam corrigidos e haja maior transparência.

Recomendações Práticas

Auditoria de Dados e Algoritmos: aplicar técnicas de mitigação de viés, balancear os dados de treinamento e revisar periodicamente o modelo.

Explicabilidade das Decisões: garantir que candidatos tenham acesso às razões pelas quais seu currículo foi aceito ou rejeitado.

Governança Ética: criar um comitê multidisciplinar (jurídico, RH, tecnologia, diversidade) para revisar continuamente os impactos da IA.