

Aplicando o Método de Análise (Framework)

Viés e Justiça

- **Tipos de viés presentes:**
 - **Viés de dados:** o modelo foi treinado com currículos de 10 anos, em sua maioria de homens, refletindo desigualdades históricas de gênero no setor de tecnologia.
 - **Viés algorítmico:** o sistema aprendeu a penalizar termos associados a mulheres (como “women’s chess club”) e a favorecer padrões masculinos.
 - **Grupos afetados:** mulheres foram desproporcionalmente prejudicadas no processo de seleção.
 - **Distribuição de benefícios e riscos:** os homens tiveram maior chance de aprovação; mulheres sofreram desvantagem estrutural, o que perpetua desigualdade no mercado de trabalho.
-

Transparência e Explicabilidade

- **Transparência:** limitada. A lógica do sistema não era acessível a candidatos nem totalmente clara para recrutadores.
 - **Explicabilidade:** difícil compreender os critérios de rejeição ou aprovação. O modelo funcionava como uma “black box”.
 - **Problema:** falta de auditabilidade clara compromete a confiança no processo seletivo.
-

Impacto Social e Direitos

- **Impacto no mercado de trabalho:** perpetuação da exclusão feminina em tecnologia.
- **Autonomia das pessoas:** candidatas não tinham como contestar ou entender decisões da IA.
- **Direitos fundamentais:** risco de discriminação de gênero, o que fere princípios de igualdade (Constituição Federal) e direitos assegurados pela **LGPD**, que garante transparência e não discriminação em sistemas automatizados

Responsabilidade e Governança

- **Como a equipe poderia ter agido diferente:**
 - Revisando criticamente os dados históricos antes de treinar o modelo.
 - Implementando testes de fairness (justiça) antes de colocar o sistema em uso.
 - Garantindo revisões humanas (human-in-the-loop) nas decisões.
- **Princípios de Ethical AI by Design:**
 - Justiça e não discriminação.
 - Transparência e explicabilidade.
 - Accountability (responsabilidade clara da equipe de desenvolvimento).

- **Leis/Regulações aplicáveis:**
 - **LGPD (Brasil):** artigos que tratam de decisões automatizadas e não discriminação.
 - **Direito trabalhista:** proibição de discriminação por gênero.
 - **Direitos Humanos:** igualdade de acesso a oportunidades.

Posicionamento

Minha posição: O sistema, como foi concebido, **não deve ser utilizado**. No entanto, não é necessário bani-lo totalmente: ele pode ser **redesenhado** com bases éticas sólidas e maior supervisão.

Recomendações práticas:

1. **Auditoria de dados** – revisar conjuntos de dados para identificar e mitigar vieses antes do treinamento.
2. **Testes de equidade contínuos** – implementar métricas de fairness (ex.: igualdade de oportunidades entre gêneros e raças) e realizar monitoramento periódico.
3. **Explicabilidade e supervisão humana** – garantir que candidatos possam solicitar justificativas claras e permitir que decisões finais sejam sempre revisadas por pessoas.