Protocolo de Prevenção ao Uso Indevido de Dados Pessoais em Projetos de Machine Learning e Inteligência Artificial

**Time de Inteligência Artificial**

**Data**: dezembro de 2024

**Responsável**: Maelson Marques de Lima

# 1. Objetivo

Este protocolo tem como objetivo estabelecer diretrizes práticas para garantir que nenhum dado pessoal identificável seja utilizado de forma indevida em projetos de ciência de dados, aprendizado de máquina (ML) ou inteligência artificial (IA), conforme as exigências da LGPD (Lei Geral de Proteção de Dados), Projeto de Lei 2338/2023, diretrizes da ANPD e o NIST Privacy Framework.

# 2. Escopo

Aplica-se a todos os projetos conduzidos pela organização que envolvam o uso de bases de dados para treinamento, validação ou teste de modelos de IA e ML, tanto em ambientes de desenvolvimento quanto de produção.

# 3. Diretrizes do Protocolo

* Nenhum dado deve ser processado sem definição prévia do caso de uso e aprovação da área de compliance.
* Somente dados estritamente necessários devem ser utilizados, seguindo o princípio da minimização.
* A extração inicial de dados deve passar por um pipeline de inspeção automática que utilize LLMs e regras heurísticas para detecção de dados pessoais e sensíveis.
* Campos contendo nomes, documentos, telefones, e-mails ou quaisquer identificadores únicos devem ser automaticamente excluídos ou anonimizados.
* É obrigatório o uso de dicionário de dados com classificação de sensibilidade (pessoal, sensível, público).
* Nenhum dado rotulado como pessoal deve ser utilizado para treinar modelos, salvo quando for aplicável técnica de anonimização robusta e houver respaldo legal explícito.
* Todos os dados utilizados devem ser logados com hash de verificação, versão de schema e justificativa de uso.
* Auditorias devem ser periódicas e realizadas em pipelines de dados e modelos já implantados.

# 4. Ferramentas Recomendadas

* SQLAlchemy para inspeção de schemas
* Pandas para manipulação de dados
* OpenAI ou LLMs equivalentes para identificação contextual
* Great Expectations ou SodaSQL para validação de qualidade
* DVC ou MLFlow para rastreabilidade e versionamento de dados
* Git e GitHub/GitLab para versionamento de código e notebooks

# 5. Fluxo Operacional Resumido

1. Definir caso de uso e validar com compliance
2. Conectar ao banco e inspecionar estrutura com SQLAlchemy
3. Coletar amostra de até 100 registros por tabela
4. Usar LLM para avaliar presença de dados pessoais
5. Remover ou anonimizar colunas identificadas
6. Documentar e versionar schema aprovado
7. Realizar EDA e preparar dados para modelagem
8. Validar qualidade e registrar rastreabilidade
9. Submeter modelo a auditoria antes de ir à produção

# 6. Considerações Finais

O cumprimento deste protocolo visa mitigar riscos legais, éticos e técnicos associados ao uso indevido de dados pessoais. A proteção da privacidade e o uso responsável da IA são compromissos fundamentais da organização.