**Necessidade da proteção de dados da IA**

A IA é um recursos valioso para o desenvolvimento de modelos e algoritmos

para automatização de processos, usada para várias tarefas como previsão

reconhecimento de padrões, análise de dados e outros.

E sendo uma ferramenta que pode trazer benefícios ela também pode trazer

implicações de segurança que precisamos considerar

Vazamento de dados - Hackers pode ter acesso a dados pessoas

Viés - Ela pode perpetuar preconceitos e discriminação se não for usada corretamente

Vulnerabilidade - Identificar vulnerabilidades no sistema de segurança do modelo

Deve se ter Ética e privacidade na construção da IA

Como os dados que fornecemos sem ter consciência podem gerar riscos

Segundo Danilo Macedo, líder de Relações Governamentais e Assuntos Regulatórios, na IBM eles classificam em 2 categorias, alto ou baixo risco

Alto risco seriam usar os dados das pessoas com fonte de renda (monetização de dados externos), as pessoas que estão nessa categoria muitas vezes não sabem como são acessados e usados e qual o risco que podem fornecer.

E os de Baixo risco seriam usar os dados das pessoas para melhorar operações,

produtos ou serviços( monetização interna ), e nessa categoria as pessoas esperam que não passe disso.

"Esta distinção vai nos permitir procurar uma regulamentação mais adequada e mais equilibrada. Por exemplo, você pode ajustar a carga regulatória para ser proporcional aos riscos dos modelos comerciais baseados em dados, aumentar a transparência da revenda de dados, exigir que os compradores verifiquem se os dados foram tratados de forma legal e transparente, entre outras obrigações." - por Danilo Macedo

<https://www.privacidade.com.br/a-ia-no-contexto-da-protecao-de-dados/>

<https://www.ibm.com/blogs/ibm-comunica/protecao-de-dados-e-inteligencia-artificial/>

É necessário algumas medidas de segurança que devemos implementar para proteger nossos dados e ficar em conformidade com a LGPD ( Lei geral de proteção de dados ),

baseado no 3° whitepaper do google, de como eles usam a criptografia para proteger os dados do usuário, sendo 3 fatores essenciais: Autenticidade, Integridade e criptografia.

Verificar a fonte dos dados, se é um humano ou um processo, destino dos dados, e se os dados vieram inalterados, e não menos importante, a criptografia, que utiliza dados legíveis e converte para dados não legíveis, como exemplo o método de Diffie-Hellman, que pode ser aplicado para garantir o acesso às pessoas devidamente autorizadas

**Medidas de segurança física e lógica para proteger os dados**

Aplicar uma medida de controle de acesso ajuda a proteger os dados em repouso, isolando e versionando esse dados, controlando que tem acesso aos mesmo

Uma das ferramentas que podem ser utilizadas e o github, uma plataforma de hospedagem de código com controle de versão usando o git, sendo uma das ferramentas mais utilizadas por engenheiros de software

**Melhorar a segurança dos dados**



Usar o mongodb para fazer a importação do banco de dados

diferente de como era usado antes, de importar diretamente o arquivo csv no colab

garantindo flexibilidade e estabilidade para manipular nossos dados

ela ajuda a não precisar fazer a normalização

Seguindo o passo a passo do fórum do havodata: [Pandas MongoDB Connection: 101 Simplified - Learn | Hevo](https://hevodata.com/learn/pandas-mongodb/)

**Ferramentas para monitoramento de dados**

**Zabbix**: monitoramento de rede, servidor, nuvem, aplicativos e serviços

Devido a sua integração com vários tipos de sistemas e suas funcionalidades ela se torna uma das melhores escolhas.

**PRTG**: Monitore todos os sistemas, dispositivos, tráfego, aplicativos, sites, aplicativos, serviços. Tecnologias integradas, descoberta de rede automática, notificações, relatórios e muito mais.

## **Rotina de backup**

## O recomendado seria fazer um rotina de backup automaticamente diariamente ou quando houvesse uma mudança significativa

## por meio físico: salvar 2 backups, um em outro dispositivo físico e um no PC.

por meio digital: salvar em 2 servidores, principal e o reserva.

## Fontes de pesquisas

<https://cloud.google.com/docs/security/encryption-in-transit?hl=pt-br>

<https://www.privacidade.com.br/a-ia-no-contexto-da-protecao-de-dados/>

<https://www.ibm.com/blogs/ibm-comunica/protecao-de-dados-e-inteligencia-artificial/>

Tecnologias e Ferramentas utilizadas

Python e suas biblioteca - Sklearn, panda

Google Colab

Windows 10

Excel

Google Docs

Canvas