

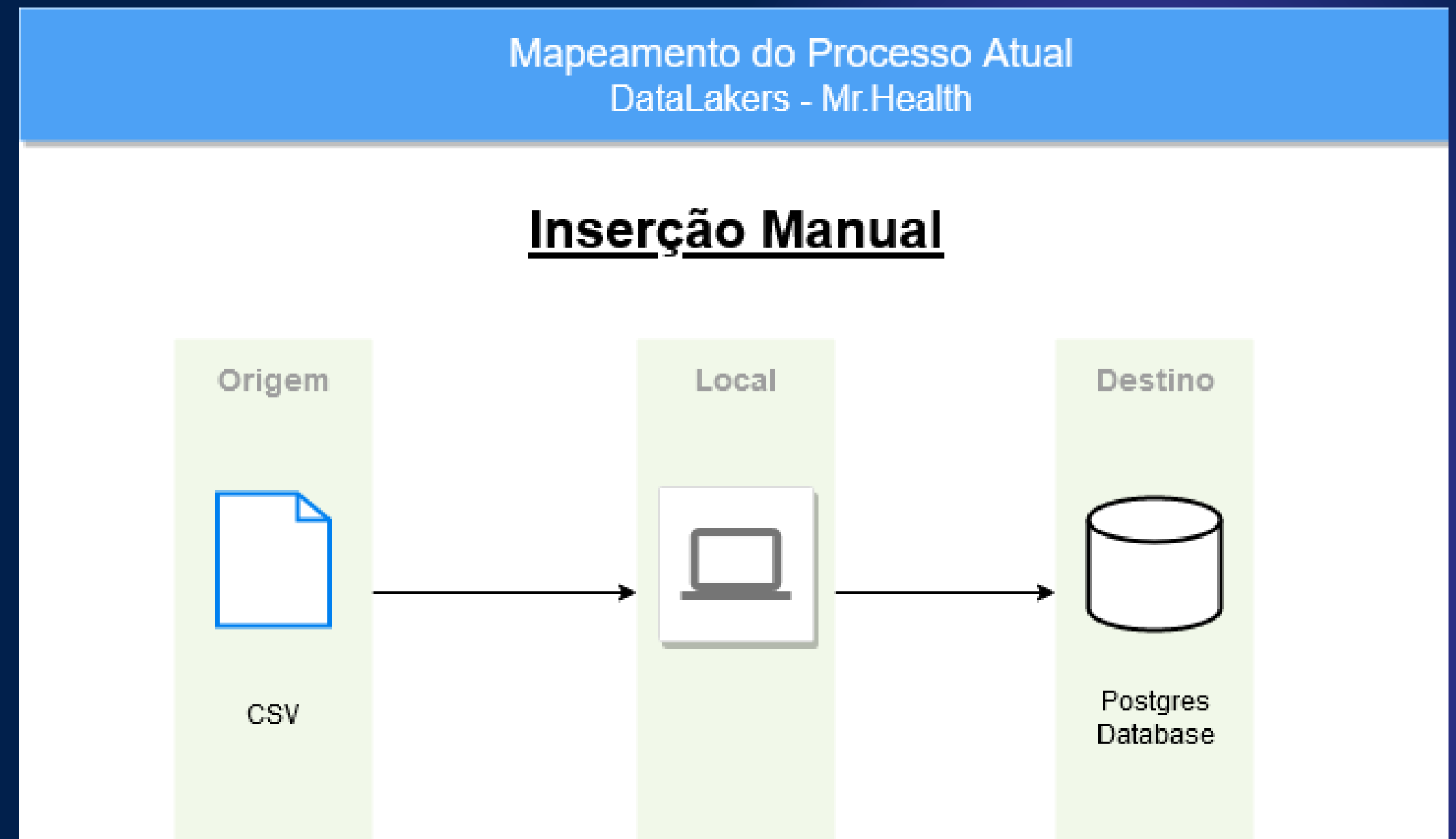
# Mr. Health Case

Pablo Rodrigo Massirer

# Atividade 1. Mapeamento do Processo Atual

## Pontos problematicos

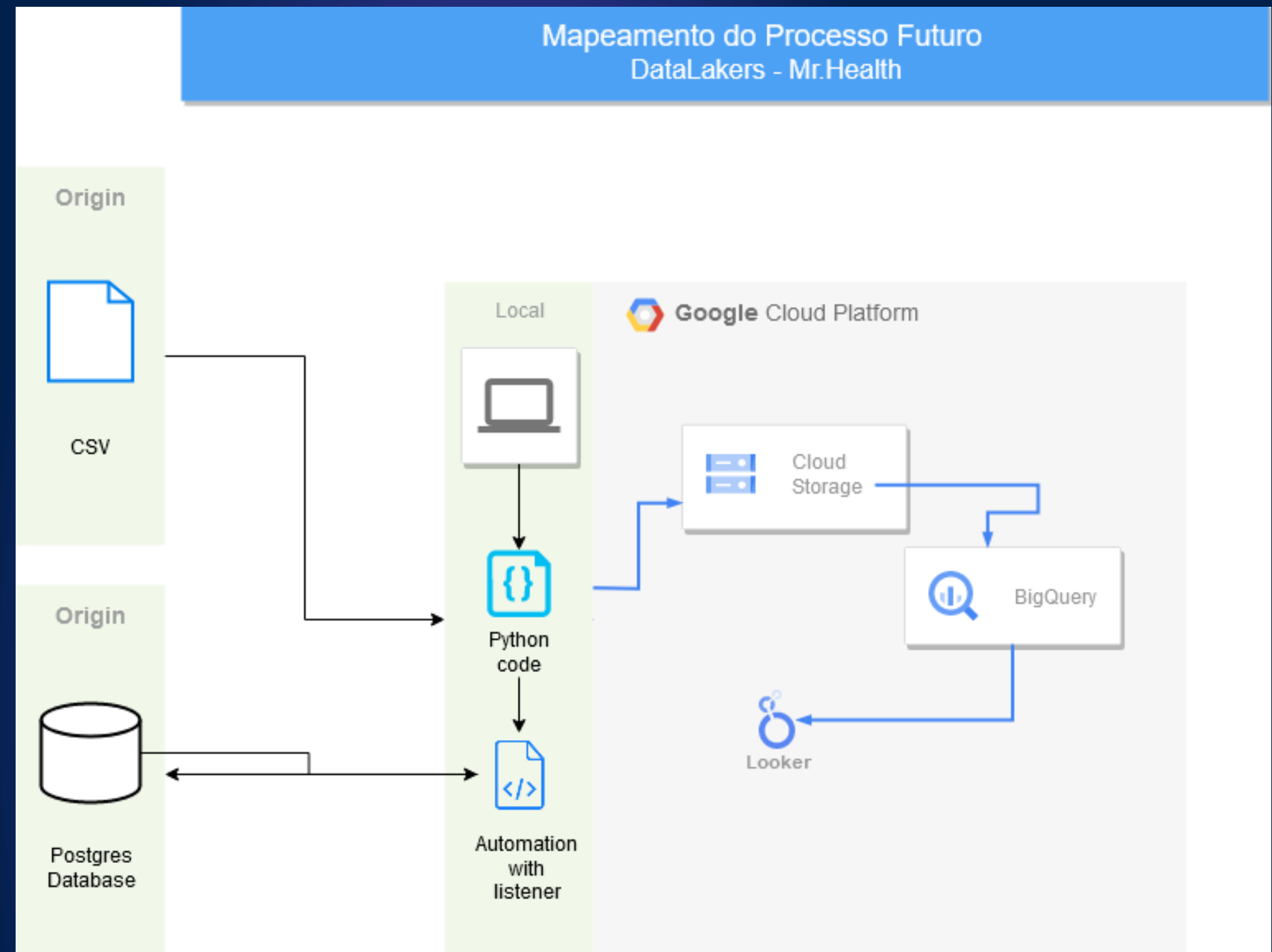
1. Erro humano;
2. Custo do tempo;
3. Falta de validação;
4. Escalabilidade limitada;
5. Riscos de segurança;
6. Dificuldade com Updates.



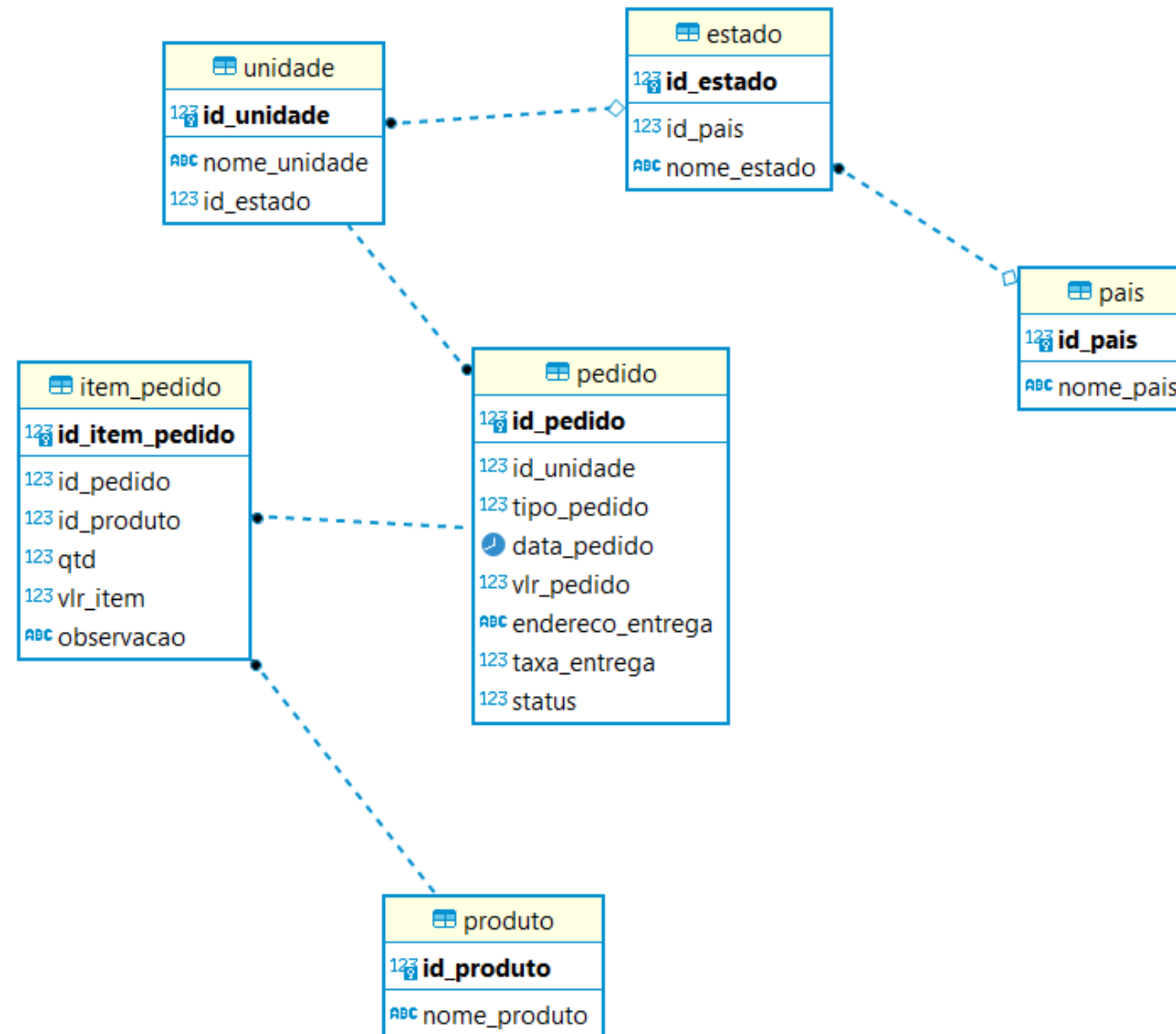
# Atividade 2. Mapeamento do Processo Futuro

## Pontos melhorados

1. Erro humano; ✓
2. Custo do tempo; ✓
3. Falta de validação; ✓
4. Escalabilidade limitada; ✓
5. Riscos de segurança; ✓
6. Dificuldade com Updates. ✓



# Atividade 3. Diagrama de Arquitetura



# Atividade 4. Descrição das camadas Datalake

## Camada 1. Ingestion Layer

Camada inicial que armazenará os dados raw no formato original

## Camada 2. Transformation layer

Camada que armazenará os dados para serem transformados e limpos

## Camada 3. Modeling Layer

Camada que armazenará os dados para serem modelados

## Camada 4. Presentation Layer

Camada que armazenará os dados para serem utilizados em sistemas BI

# 5. Relação de atividades

## Levantamento de Requisitos

Realizar reuniões com os stakeholders para entender os requisitos e necessidades do processo de ingestão e consolidação de dados. Documentar os requisitos em detalhes.

## Identificação de Fontes de Dados

Identificar todas as fontes de dados relevantes para o projeto, incluindo bancos de dados, arquivos, entre outros.

## Data Modeling

Definir o modelo de dados que será utilizado para a consolidação dos dados. Isso inclui a estrutura das tabelas, campos, chaves primárias, etc.



# 5. Relação de atividades

## Seleção de plataforma de nuvem

Avaliar e escolher a plataforma de nuvem apropriada (por exemplo, AWS, Azure, Google Cloud) com base nos requisitos, escalabilidade e considerações de custo do projeto. Nesse caso GCP.

## Configuração de infraestrutura

Configurar a infraestrutura de nuvem, incluindo IAM, serviços de armazenamento e componentes de rede necessários para armazenamento e processamento de dados.

## Desenvolvimento de ETL e seus testes

Criar os processos de ETL para extrair os dados das fontes, transformá-los conforme o modelo de dados definido e carregá-los no local de armazenamento, Data Lake/Data Warehouse. Desenvolver testes automatizados para validar a precisão e integridade dos dados

# 5. Relação de atividades

## Tratamento de Erros e Exceções

Definir estratégias para o tratamento de erros e exceções durante o processo de ingestão, garantindo a robustez e confiabilidade do sistema.

## Data Transformation

Desenvolver processos de transformação de dados para limpar, enriquecer e pré-processar os dados conforme necessário para consolidação.

## Monitoramento e Alerta

Implementar sistemas de monitoramento e alerta para rastrear o desempenho e a integridade dos processos de ingestão e consolidação de dados.



# 5. Relação de atividades

## Integração com ferramentas de análise de dados e BI

Integrar os dados ingeridos e consolidados com análise de dados e ferramentas de inteligência de negócios para análise de dados e relatórios.

## Documentação

Criar documentação detalhada do processo de ingestão e consolidação de dados, incluindo fluxos de trabalho, diagramas, e descrições dos componentes envolvidos.



# Obrigado

pablomassirer@gmail.com

<https://github.com/pablomassirer>

<https://www.linkedin.com/in/pablomassirer>