

Ciclo de vida dos dados - Empresa Gigamart

1)Coleta

Através da base de dados recebida não temos como identificar o registro de origem (data lineage) desses dados. É possível que eles sejam coletados através da plataforma de vendas da Gigamart ou do CRM utilizado pela organização.

Alguns dos riscos comuns desta etapa, se ela não for bem estabelecida e organizada, é a existência de dados imprecisos e incompletos, coleta de dados sem consentimento, ausência de padrões e coleta de dados que não são necessários para gerar valor ao negócio.

Para conter esses riscos devemos:

- Garantir o consentimento da coleta por parte do usuário;
- Estabelecer regras de negócio claras para que saibamos exatamente quais dados devemos coletar;
- Quando possível definir respostas categorizadas para campos, de forma que o usuário não se sinta livre para escrever o que quiser, em um campo como por exemplo "Estado Civil";
- Validações automatizadas de entrada

2)Armazenamento

A fase de armazenamento é onde vai ocorrer o tratamento dos dados, esses dados provavelmente encontram-se armazenados em um Data Warehouse, já que são tabulares e este é otimizado para consultas e análises.

Alguns dos riscos evidenciados nessa etapa são, dados inconsistentes, desorganizados, vazamentos e uso não autorizado. Isto pode acontecer devido a ausência de papéis bem definidos e políticas de acesso mal estabelecidas.

Para conter esses riscos devemos:

- Estabelecer políticas de controle de acesso a dados, utilizando ferramentas para conter o acesso não autorizado;
- Backup;
- Versionamento;
- Utilização de ferramentas que mantenham o ambiente digital seguro, como antivírus e firewall.
- Respeitar normas da LGPD

3)Uso

Nesta fase é onde iremos de fato fazer uso dos dados coletados e tratados. Seja para geração de csvs ou dashboards, vamos começar a entender o que de fato está acontecendo, gerando informação e valor. Neste caso, estamos fazendo uso dos dados

através de csvs e analisando a sua qualidade para gerar informação e como resultado apoiar na decisão dos times de vendas

Os riscos evidenciados nesta fase são, uso indevido, interpretação incorreta e tomada de decisão baseada em dados errôneos. É importante lembrar que esta fase depende diretamente da qualidade dos dados.

Para conter esses riscos devemos:

- Documentar atividades;
- Estabelecer permissões;
- Rastreamento do uso.

4) Compartilhamento

Nesta fase, os dados são enviados para outros times e parceiros que vão usufruir destas informações para tomar decisões com embasamento. Neste caso da empresa Gigamart, é possível que os dados sejam compartilhados com o time da área comercial para apoiar na sua análise de vendas por região.

Alguns dos riscos comuns desta fase são, vazamento de dados, envio de dados incompletos e violação da LGPD.

Para conter esses riscos devemos fazer:

- Mascaramento de dados sensíveis;
- Criptografia;
- Contratos e termos de uso.

5) Arquivamento

Na fase de arquivamento, vamos identificar dados mais antigos e que não estão sendo usados com a mesma frequência para formar um histórico. No caso da Gigamart, podemos arquivar dados de vendas que foram coletados a mais de 1 ano ou de clientes que encontram-se inativos.

Alguns dos riscos comuns desta etapa são, armazenamento de dados desnecessários, retenção de dados além do permitido, custo de retenção e dificuldade de recuperação.

Para conter esses riscos devemos realizar:

- Separação entre dados ativos e arquivados
- Políticas BEM definidas
- Armazenamento seguro
- Organização e rastreabilidade

6) Exclusão

A exclusão serve para eliminar de uma vez por todas aqueles dados que não são mais utilizados e nem necessitam ser armazenados no histórico. A exclusão deve ser feita de forma regular e cautelosa. Ela nos ajuda a disponibilizar recursos e baixar o custo de

armazenamento de informações que não podem mais ser armazenadas ou não são mais tão valiosas assim.

Alguns dos riscos que envolvem esta etapa são, exclusão acidental de dados úteis, falhas na remoção e retenção ilegal.

Para conter os riscos devemos:

- Estabelecer uma política de descarte clara
- Uso de ferramentas de limpeza definitiva
- Registro do processo de exclusão, através de logs

Papéis e Responsabilidades

Data Owner:

O Data Owner da Gigamart é o **Diretor Comercial e líder da área de vendas** da empresa. Ele/a terá como responsabilidades garantir que os dados estejam alinhados com os objetivos do negócio, definir políticas de acesso, como estes dados serão protegidos e qual a qualidade esperada.

Data Steward:

Na Gigamart quem será responsável por assumir esse papel será um **Analista de Dados Sênior**. O Data Steward terá como responsabilidades garantir que os dados estejam corretos, atualizados, completos e monitorar a qualidade geral dos dados utilizados. Ele/a será responsável por impor a política estabelecida pelo Data Owner.

Data Custodian:

Quem vai assumir o papel de Data Custodian é um **Engenheiro de Dados Sênior**. Este será responsável pela infraestrutura, backup e configurar perfis de acesso aos dados.

Data Consumer:

O Data Consumer é aquele que vai fazer uso direto dos dados para tomar decisões. Quem vai assumir esse papel na Gigamart é o **Gerente de Vendas**. Suas responsabilidades incluem fazer análises, relatórios, gerar dashboards e tomar decisões.