UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
TRABALHO DE GRADUAÇÃO

Lucas Oliveira Cavalcanti

Infraestrutura para processamento consentido de dados

Recife

2025

Lucas Oliveira Cavalcanti

Infraestrutura para processamento consentido de dados

Trabalho apresentado ao Programa de Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Bacharel em Ciência da Computação.

Área de Concentração: Processamento e Proteção de Dados

Orientador: Robson do Nascimento Fidalgo

Recife

2025

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Cavalcanti, Lucas Oliveira.
Infraestrutura para processamento consentido de dados / Lucas Oliveira Cavalcanti. - Recife, 2025.
33, tab.

Orientador(a): Robson do Nascimento Fidalgo
Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Informática, Ciências da Computação - Bacharelado, 2025.
Inclui referências.

1. Processamento de dados. 2. Proteção de dados. 3. Consentimento. 4. Propósito. 5. Sistema de gerenciamento de banco de dados. I. Fidalgo, Robson do Nascimento. (Orientação). II. Título.

000 CDD (22.ed.)

Lucas Oliveira Cavalcanti

Infraestrutura para processamento consentido de dados

Monografia apresentada ao curso de Ciência da Computação, como requisito parcial para a obtenção do Título de Bacharel em Ciência da Computação, Universidade Federal de Pernambuco.

Aprovado em: 25 de março de 2025.

BANCA EXAMINADORA

_____

Prof. Robson do Nascimento Fidalgo
Universidade Federal de Pernambuco

_____

Prof. Kiev Santos da Gama
Universidade Federal de Pernambuco

Aos meus pais, Karina e Rodrigo, minhas maiores inspirações na jornada do aprendizado.

AGRADECIMENTOS

Este trabalho representa não apenas o resultado de estudo e dedicação, mas também a soma do apoio, incentivo e ensinamentos de muitas pessoas que estiveram ao meu lado nesta jornada.

Agradeço primeiramente à minha família – Rodrigo, Karina, Lorena e Darcinete – pelo amor incondicional, pelo suporte em todos os momentos, por sempre acreditarem no meu potencial e pela oportunidade de estudar nessa instituição.

Aos meus amigos do Centro de Informática da UFPE – Artur, Rodrigo, Matheus, Gustavo, José Lucas e Bruna – meu mais sincero agradecimento por compartilharem comigo essa caminhada, pelos aprendizados e principalmente pelas risadas.

Aos professores do CIn, sou grato pelo conhecimento transmitido, pelos desafios propostos e pelo compromisso com a excelência no ensino. Em especial, agradeço ao meu orientador, Robson, pela paciência e pelos valiosos direcionamentos.

Por fim, expresso minha gratidão ao Centro de Informática e à Universidade Federal de Pernambuco pela qualidade do ensino e pelo impacto positivo que essa instituição trouxe não apenas para minha trajetória profissional, mas também para a vida da minha família. O CIn e a UFPE foram fundamentais e abriram muitas portas.

A todos vocês, meu muito obrigado!

RESUMO

O avanço da informatização e a ampla disponibilidade de dados impulsionaram uma verdadeira revolução na era da informação. No entanto, esse crescimento também trouxe preocupações significativas quanto à legitimidade do uso de dados pessoais. A crescente demanda por privacidade e segurança levou à criação de regulamentações como a General Data Protection Regulation (GDPR) na Europa, que serviu de base para a Lei Geral de Proteção de Dados (LGPD) no Brasil. Promulgada em 2018, a LGPD estabelece diretrizes fundamentais para o tratamento de dados pessoais, conferindo aos indivíduos maior controle sobre o uso de suas informações.

Apesar de sua relevância, a implementação da LGPD ainda apresenta desafios técnicos, especialmente para desenvolvedores e arquitetos de sistemas que buscam adequar suas soluções às exigências legais. Nesse contexto, este trabalho tem como objetivo investigar as implicações tecnológicas da LGPD propondo uma abordagem para a modelagem de tabelas e a configuração de Sistemas de Gerenciamento de Banco de Dados (SGBD), de forma a garantir o processamento consentido e seguro dos dados pessoais.

Palavras-chaves: Processamento de dados. Propósito. Consentimento. SGBD. LGPD. GDPR.

ABSTRACT

The advancement of digitalization and the widespread availability of data have driven a significant transformation in the information era. However, this progress has also raised critical concerns regarding the legitimacy of personal data usage. Growing demands for privacy and security have led to the development of regulations such as the General Data Protection Regulation (GDPR) in Europe, which served as a foundation for Brazil's General Data Protection Law (LGPD). Enacted in 2018, the LGPD establishes fundamental guidelines for the processing of personal data, granting individuals greater control over how their information is used.

Despite its importance, the practical implementation of the LGPD presents technical challenges, particularly for developers and system architects seeking to align their solutions with legal requirements. In this context, this study aims to analyze the technological implications of the LGPD and propose a structured approach for database modeling and the configuration of Database Management Systems (DBMS) to ensure the lawful and secure processing of personal data.

Keywords: Data Processing. Purpose. Consent. DBMS. LGPD. GDPR.

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

GDPR	General Data Protection Regulation
LGPD	Lei Geral de Proteção de Dados
SGBD	Sistema de Gerenciamento de Banco de Dados

SUMÁRIO

1 INTRODUÇÃO

O avanço da tecnologia e a digitalização crescente de serviços têm impulsionado um aumento exponencial na coleta e processamento de dados pessoais. Empresas de tecnologia utilizam esses dados para alimentar sistemas inovadores, como algoritmos de aprendizado de máquina que operam em tempo real sobre informações de milhões — e possivelmente bilhões — de indivíduos. No entanto, esse progresso vem acompanhado de desafios críticos relacionados à privacidade e à regulamentação do uso dessas informações, a fim de proteger os direitos dos cidadãos. Como argumentam Solove e Citron, os danos causados por vazamentos de dados muitas vezes envolvem riscos futuros e impactos emocionais, como ansiedade e medo, que ainda são subestimados pelos sistemas legais (SOLOVE; CITRON, 2017). Para responder a essas preocupações, regulamentações como a General Data Protection Regulation (GDPR) na Europa e a Lei Geral de Proteção de Dados (LGPD) no Brasil foram estabelecidas com o objetivo de proporcionar maior transparência e segurança no tratamento de dados pessoais. Essas legislações definem diretrizes rigorosas para o acesso, coleta, armazenamento e processamento dessas informações, buscando equilibrar inovação tecnológica e proteção da privacidade.

No contexto organizacional, essas regulamentações impõem desafios operacionais significativos, especialmente para sistemas que lidam com dados sensíveis. Considere, por exemplo, um operador de banco de dados em uma instituição financeira que precisa acessar informações de um cliente para processar uma solicitação de revisão de crédito. Para isso, é necessário consultar dados como histórico de transações, salário, score de crédito e vínculos contratuais. No entanto, conforme estabelecido pela LGPD, esse acesso deve ser justificado por um propósito legítimo e, quando aplicável, contar com o consentimento explícito do titular. Esse cenário evidencia a necessidade de mecanismos técnicos que assegurem a conformidade legal no tratamento de dados dentro de um Sistema de Gerenciamento de Banco de Dados (SGBD). Garantir que apenas usuários autorizados acessem informações sensíveis, respeitando princípios de transparência, é essencial para evitar usos indevidos.

Diante da crescente complexidade dessas normativas e do volume massivo de dados processados, as organizações precisam adotar soluções eficazes para garantir uma conformidade contínua e automatizada. Nesse sentido, governos e empresas têm investido em estudos e revisões legislativas para atualizar normas e assegurar que os direitos dos cidadãos sejam protegidos em um cenário de transformação digital acelerada. Este trabalho busca contribuir com

essa discussão ao propor uma abordagem técnica para garantir o processamento consentido de dados, alinhando-se às exigências regulatórias vigentes.

1.1 MOTIVAÇÃO

As normativas supracitadas impõem diretrizes rigorosas para o tratamento de dados pessoais, exigindo que o acesso a essas informações esteja fundamentado em princípios como **consentimento** e **propósito legítimo** que serão explorados no próximo capítulo (European Union, 2016; Presidência da República - Brasil, 2018). Assim, surge uma nova questão regulatória: a necessidade de respeitar o propósito específico para o qual cada dado foi consentido. Garantir que um SGBD esteja em conformidade com tais regulamentações não é trivial. A maioria dos sistemas de gerenciamento de banco de dados não possui mecanismos nativos para validar, em tempo de execução, se uma consulta ao banco de dados está em conformidade com as informações de consentimento e propósito. Como consequência, a responsabilidade por essa conformidade recai sobre os desenvolvedores e administradores do banco de dados, tornando o processo suscetível a falhas e vulnerabilidades operacionais.

Uma alternativa que poderia ser considerada, utilizando os recursos já presentes em SGBDs comerciais, seria aplicar mecanismos de controle de acesso tradicionais, como as instruções GRANT e REVOKE da linguagem SQL. No entanto, essa abordagem se mostra limitada quando o objetivo é garantir a conformidade com os propósitos específicos de uso dos dados, como exigido pelas regulamentações. Os comandos GRANT e REVOKE são amplamente utilizados para definir quais usuários ou grupos (roles) podem acessar determinados objetos do banco de dados e quais operações podem realizar sobre eles. Esses mecanismos são eficazes para restringir quem pode acessar quais recursos dentro do sistema, entretanto, não contemplam a motivação subjacente ao acesso, ou seja, não consideram o propósito específico pelo qual os dados estão sendo consultados ou processados. Na prática, um mesmo usuário pode ter motivos diferentes para acessar o mesmo dado em contextos distintos. Isso significa que, mesmo que o acesso esteja autorizado, ele pode não estar em conformidade com o propósito para o qual o titular dos dados concedeu consentimento (ABREU, 2021). Esse argumento reforça a necessidade de mecanismos específicos e automatizados que incorporem as informações de consentimento e propósito diretamente no processamento das consultas SQL.

1.2 OBJETIVOS

Considerando o aumento contínuo do volume de dados processados e a variedade de propósitos aos quais estão vinculados, atribuir unicamente ao operador do banco de dados a responsabilidade de garantir a conformidade regulatória torna-se impraticável. Para reduzir esse risco e assegurar que as diretrizes de proteção de dados sejam cumpridas de forma eficiente, faz-se necessário adotar mecanismos automatizados de controle de acesso e filtragem de informações. Diante desse contexto, este trabalho propõe uma solução baseada em duas frentes principais:

1. A definição de uma **infraestrutura de metadados** que permita armazenar e gerenciar informações sobre consentimento e propósito de processamento de dados.

2. A implementação de um **mecanismo no parser do SGBD** para validar e reescrever consultas, em tempo de execução, garantindo que apenas informações compatíveis com os propósitos consentidos sejam acessadas.

Essa abordagem visa contribuir para a governança de dados pessoais em ambientes corporativos, promovendo maior segurança para os operadores dos bancos de dados, transparência para os titulares dos dados e conformidade com regulamentações vigentes. Além disso, ao integrar princípios de processamento consentido diretamente na camada de consulta, espera-se reduzir a dependência de controles externos e aprimorar a eficiência no cumprimento das normativas de proteção de dados.

2 FUNDAMENTAÇÃO TEÓRICA

No contexto deste trabalho, é fundamental compreender os conceitos que embasam o processamento consentido de dados, permitindo a análise e implementação de mecanismos compatíveis com as exigências normativas. Este capítulo apresenta os fundamentos teóricos necessários para o desenvolvimento da pesquisa, explorando os principais conceitos relacionados à proteção de dados, consentimento, propósito e governança em SGBDs. Esses elementos fornecem a base conceitual para a proposta metodológica deste estudo.

2.1 GENERAL DATA PROTECTION REGULATION (GDPR)

A GDPR, implementada em 25 de maio de 2018, é a regulamentação da União Europeia que define regras rigorosas para o tratamento de dados pessoais. Seu principal objetivo é garantir que indivíduos tenham maior controle sobre suas informações e que empresas sigam princípios éticos no uso desses dados. A legislação se aplica a qualquer organização que processe dados de cidadãos da União Europeia, independentemente da localização da empresa.

Os principais princípios estabelecidos pela GDPR incluem:

- **Legalidade, justiça e transparência**: Os dados devem ser coletados e processados de maneira legal, justa e transparente.
- **Finalidade específica**: A coleta de dados deve ter propósitos explícitos e legítimos, sem usos incompatíveis posteriores.
- **Minimização de dados**: Apenas os dados estritamente necessários devem ser coletados.
- **Exatidão**: Os dados devem ser mantidos corretos e atualizados.
- **Limitação de armazenamento**: As informações devem ser mantidas apenas pelo tempo necessário.
- **Integridade e confidencialidade**: Devem ser adotadas medidas para garantir a segurança e a proteção contra acessos não autorizados, perdas ou vazamentos (European Union, 2016).

A não conformidade com a GDPR pode resultar em penalidades severas, incluindo multas de até 20 milhões de euros ou 4% do faturamento anual global da organização, o que for

maior. Além disso, a regulamentação concede aos titulares de dados direitos fundamentais, como o direito ao acesso, retificação, exclusão (*direito ao esquecimento*) e portabilidade de suas informações.

2.2 LEI GERAL DE PROTEÇÃO DE DADOS (LGPD)

Inspirada na GDPR, a LGPD foi sancionada no Brasil pela Lei nº 13.709/2018 e entrou em vigor em setembro de 2020. Assim como sua contraparte europeia, a LGPD tem como objetivo proteger os direitos fundamentais de liberdade e privacidade dos cidadãos, estabelecendo regras para o tratamento de dados pessoais em meios físicos e digitais.

A legislação se aplica a qualquer operação de tratamento de dados realizada por empresas ou órgãos públicos que ofereçam serviços no Brasil, independentemente da localização da organização. Seus princípios fundamentais incluem:

- **Necessidade e adequação**: O tratamento de dados deve estar limitado ao mínimo necessário para alcançar sua finalidade.
- **Transparência**: Os titulares devem ser informados sobre como seus dados estão sendo utilizados.
- **Bases legais**: O processamento pode ocorrer com base no consentimento do titular, cumprimento de obrigações legais ou legítimo interesse do controlador.
- **Segurança e prevenção**: Medidas técnicas devem ser implementadas para evitar incidentes de segurança (Presidência da República - Brasil, 2018).

Empresas que não estiverem em conformidade com a LGPD podem ser penalizadas com advertências e multas que podem chegar a 2% do faturamento da organização, limitadas a R$ 50 milhões por infração. Além disso, a LGPD estabeleceu a Autoridade Nacional de Proteção de Dados (ANPD), responsável pela regulamentação e fiscalização do cumprimento da lei no Brasil.

2.2.1 Comparação entre GDPR e LGPD

Embora a LGPD tenha sido inspirada na GDPR, existem algumas diferenças importantes entre as duas regulamentações:

- **Escopo territorial**: A GDPR se aplica a qualquer organização que processe dados de cidadãos da União Europeia, independentemente da sua localização. Já a LGPD se aplica a qualquer operação de tratamento de dados realizada no Brasil ou que envolva dados de indivíduos localizados no país.
- **Multas**: A GDPR prevê penalidades mais severas (até 20 milhões de euros ou 4% do faturamento global), enquanto a LGPD limita as multas a R$ 50 milhões por infração.
- **Bases legais**: Embora ambas contemplem o consentimento do titular como um dos fundamentos para o tratamento de dados, a GDPR oferece um escopo mais detalhado sobre o interesse legítimo e regras para transferência internacional de dados.
- **Autoridade fiscalizadora**: A GDPR é aplicada por órgãos reguladores independentes em cada país membro da UE, enquanto a LGPD é regulamentada exclusivamente pela ANPD.

Ambas as legislações desempenham um papel essencial na proteção da privacidade e na promoção da segurança de dados, estabelecendo diretrizes claras para empresas e instituições. Neste trabalho, a conformidade com essas regulamentações será analisada no contexto de processamento consentido de dados em SGBDs, destacando desafios e soluções para garantir sua aplicação eficaz.

2.3 CONSENTIMENTO E PROPÓSITO NO PROCESSAMENTO DE DADOS

O processamento de dados pessoais deve seguir diretrizes que garantam transparência e proteção à privacidade dos indivíduos. No contexto das regulamentações de proteção de dados, como a GDPR e a LGPD, dois conceitos fundamentais estruturam a base legal para a coleta e utilização dessas informações: o **consentimento** e o **propósito**. Esses princípios asseguram que o processamento de dados ocorra de maneira controlada e alinhada aos direitos dos titulares. As regulamentações de proteção de dados estabelecem que dados pessoais são informações que podem identificar direta ou indiretamente um indivíduo. Isso inclui identificadores diretos, como nome e CPF, bem como dados indiretos, como histórico de navegação, preferências de consumo e registros financeiros (SCHNEIDER, 2019).

O **propósito** refere-se à finalidade específica para a qual um conjunto de dados pessoais será tratado. Exemplos comuns incluem personalização de publicidade, análise de crédito, ge-

ração de estatísticas e autenticação de usuários (TAYLOR, 2020). A conformidade com normas como a GDPR e a LGPD exige que toda coleta e processamento de dados sejam justificados por um propósito legítimo e previamente determinado (European Union, 2016; Presidência da República - Brasil, 2018).

O **consentimento**, por sua vez, representa a autorização explícita do titular para que um conjunto de dados pessoais seja tratado para um propósito específico (SOLOVE, 2020). Esse consentimento deve ser fornecido de maneira *livre, informada e inequívoca*, garantindo que o indivíduo compreenda exatamente como suas informações serão utilizadas. Além disso, deve ser possível ao titular revogar esse consentimento a qualquer momento, reforçando sua autonomia sobre os próprios dados.

2.4 GERENCIAMENTO DE PROPÓSITO E REFORÇO DE CONSENTIMENTO

Para garantir que os dados sejam processados dentro dos limites estabelecidos pelos titulares e pelas regulamentações, é necessário implementar mecanismos eficazes de governança. O **gerenciamento de propósito** consiste em definir, armazenar e monitorar os propósitos para os quais os dados foram consentidos, possibilitando a aplicação de restrições (AGGARWAL, 2021). Esse processo é essencial para que empresas e instituições demonstrem conformidade regulatória e evitem acessos indevidos.

Complementarmente, o **reforço de consentimento** engloba um conjunto de medidas técnicas e organizacionais destinadas a assegurar que os dados pessoais só sejam acessados quando houver um propósito previamente consentido (NISSENBAUM, 2019). Entre essas medidas propostas por Nissenbaum, destacam-se:

- **Arquitetura de privacidade**: Soluções desenvolvidas que incorporem os princípios de privacidade desde sua concepção.
- **Políticas de controle de acesso**: Restrições baseadas nos propósitos consentidos, impedindo acessos indevidos a informações sensíveis.
- **Auditorias regulares**: Monitoramento e verificação de conformidade com as regulamentações vigentes sobre processamento dos dados para garantir que sejam utilizados conforme os termos acordados.

- **Mecanismos de rastreamento**: Registro detalhado incluindo informações como propósito de acesso, data e hora da ação e operador responsável, permitindo maior transparência e controle sobre o uso dos dados.

A implementação desses conceitos não apenas fortalece a proteção da privacidade dos indivíduos, mas também se torna um aspecto central na arquitetura de segurança de sistemas de gerenciamento de dados. Essa implementação é um desafio para os SGBDs tradicionais, pois não possuem suporte nativo para verificação de consentimento e propósito, visto que foram arquitetados antes das regulamentações de proteção de dados entrarem em vigência. No próximo capítulo, serão exploradas as infraestruturas e metodologias que propõem a aplicação prática dessas diretrizes nos SGBDs.

2.5 PROCESSAMENTO DE CONSULTAS E PARSER

Considerando o escopo deste trabalho, é fundamental apresentar um detalhamento das etapas envolvidas no *processamento de consultas*, um dos módulos centrais de um SGBD, pois define como as requisições dos operadores são interpretadas e executadas. Esse processo converte as consultas SQL em operações internas compreendidas pelo mecanismo de execução do banco de dados (ULLMAN; WIDOM, 2008).

As principais etapas do processamento de consultas incluem:

1. **Parsing**: Verifica a sintaxe e converte a consulta SQL em uma estrutura interna, geralmente uma árvore abstrata.
2. **Análise semântica**: Confirma a existência de tabelas, colunas e permissões do usuário.
3. **Otimização**: Determina a melhor estratégia de execução da consulta, considerando índices e estatísticas do banco de dados.
4. **Execução**: O plano de consulta é processado e os resultados são retornados ao usuário.

2.6 TRABALHOS RELACIONADOS

Esta seção apresenta os principais trabalhos relacionados ao controle de acesso a dados com base em consentimento e propósito, tema central deste trabalho. São discutidas propostas

que abordam o processamento consentido de dados em Sistemas de Gerenciamento de Banco de Dados (SGBD), destacando suas contribuições, limitações e relação com a solução deste trabalho.

Abreu (ABREU, 2021) propõe uma extensão da linguagem SQL com operadores específicos para controle de acesso orientado a propósito. A proposta permite que o desenvolvedor declare explicitamente os propósitos na cláusula SELECT, vinculando-os ao consentimento do titular. Embora interessante, essa abordagem exige que os desenvolvedores modifiquem todas as consultas manualmente, o que pode ser custoso e propenso a erros. Aggarwal (AGGARWAL, 2021) discute o gerenciamento de acesso baseado em propósito em ambientes de Big Data. O autor propõe um modelo flexível para representar permissões dinâmicas de acesso. No entanto, a proposta se limita a uma modelagem conceitual, sem detalhar mecanismos de aplicação em SGBDs relacionais tradicionais. Mais recentemente, Pappachan et al. (PAPPACHAN; ENCK; ROCHE, 2020) definem um *middleware* para processamento de consultas orientadas a propósito. O sistema atua interceptando requisições e verificando, em tempo de execução, se o acesso está de acordo com os propósitos consentidos. Embora o enfoque também seja a privacidade, o middleware funciona de forma externa ao SGBD, diferentemente deste trabalho, que incorpora a lógica de verificação diretamente no *parser* do banco de dados.

Apesar dos trabalhos discutidos, ainda há limitações em termos de integração nativa com os mecanismos de consulta dos SGBDs. A abordagem proposta neste trabalho contribui ao automatizar a verificação de consentimento e propósito dentro do próprio *parser* do banco de dados, promovendo uma solução menos dependente da intervenção manual por parte do desenvolvedor.

3 METODOLOGIA

A crescente necessidade de conformidade com regulamentações de proteção de dados exige que os SGBDs implementem mecanismos que assegurem que consultas realizadas sobre dados pessoais respeitem os propósitos previamente consentidos. Quando um operador de banco de dados executa uma consulta para recuperar informações dos usuários, como no caso de um comando SQL que solicita todos os registros da tabela `Usuarios`, é fundamental que o sistema possua controles que limitem o acesso aos dados conforme os consentimentos e propósitos estabelecidos.

Mesmo que o titular dos dados tenha concedido consentimento prévio para o processamento de suas informações, surge um novo desafio regulatório: assegurar que a consulta esteja limitada ao propósito. Tanto a LGPD quanto a GDPR estabelecem que o consentimento deve ser concedido para propósitos específicos, impedindo usos incompatíveis dos dados coletados.

3.1 TABELAS PARA METADADOS DE PROPÓSITO

Para garantir que o acesso aos dados pessoais ocorra de forma controlada e em conformidade com as regulamentações vigentes, é necessário estruturar um modelo de metadados que permita gerenciar os propósitos para os quais as informações podem ser utilizadas. Essa abordagem viabiliza a implementação de mecanismos de restrição baseados em consentimento, assegurando que cada consulta ao banco de dados esteja alinhada aos consentimentos dos titulares das informações. A estrutura proposta é composta por tabelas que organizam os dados de maneira a possibilitar um controle granular sobre quais informações podem ser acessadas para cada propósito. Dessa forma, é possível definir, armazenar e aplicar regras de acesso de maneira automatizada, reduzindo a necessidade de verificações manuais e garantindo maior transparência e auditabilidade.

A primeira etapa da infraestrutura consiste na definição de uma estrutura de tabelas para gerenciar os propósitos de uso dos dados, os campos necessários para cada propósito e a relação entre os usuários e os propósitos aos quais forneceram consentimento.

A seguir, são apresentadas as tabelas que compõem essa infraestrutura.

A Tabela 1 armazena os dados dos titulares, incluindo informações pessoais e sensíveis sujeitas a diferentes níveis de consentimento.

Tabela 1 – Tabela de Usuários

id_usuario	Nome	Data_Cadastro	Salário	CPF	Telefone
U001	Ana Silva	2024-01-10	5000.00	123.456.789-00	(99) 99999-5678
U002	Carlos Souza	2023-12-15	7000.00	987.654.321-00	(99) 97654-3210
U003	Mariana Costa	2024-02-05	6000.00	456.789.123-00	(99) 99876-5432
U004	Ricardo Lima	2022-08-20	4500.00	321.654.987-00	(99) 96543-2109

A Tabela 2 define os propósitos para os quais os dados podem ser processados. Além disso, especifica quais atributos podem ser acessados, garantindo que apenas os dados essenciais sejam utilizados conforme o consentimento do titular.

Tabela 2 – Tabela de Propósitos

id_proposito	Propósito	Coleta_Salário	Coleta_CPF	Coleta_Telefone
P1	Pesquisa de Mercado	FALSE	TRUE	TRUE
P2	Análise de Crédito	TRUE	TRUE	FALSE
P3	Marketing Direcionado	FALSE	FALSE	TRUE

A Tabela 3 estabelece a relação entre cada usuário e os propósitos para os quais concedeu consentimento. Esse relacionamento permite que o SGBD valide automaticamente se determinada consulta atende às restrições impostas pela legislação vigente.

Tabela 3 – Tabela de Relacionamento Usuário-Propósito

id_usuario	id_proposito
U001	P1
U002	P2
U003	P1
U003	P3
U004	P2

Essa estrutura funciona como um conjunto de metadados de propósito, permitindo que o mecanismo de reescrita de consultas avalie automaticamente quais dados podem ser retornados com base nos propósitos e consentimentos previamente registrados. Dessa forma, evita-se acessos indevidos e se assegura a conformidade com as regulamentações de proteção de dados.

3.1.1 Relacionamento entre as Tabelas

As tabelas propostas seguem os princípios de normalização de banco de dados, garantindo um modelo relacional eficiente e organizado. A estruturação segue os seguintes critérios:

- **Tabela de Usuários**: Contém informações pessoais dos titulares de dados.
- **Tabela de Propósitos**: Define as finalidades autorizadas para o processamento dos dados.
- **Tabela de Relacionamento Usuário-Propósito**: Modela a relação entre usuários e propósitos concedidos.

Além disso:

- **id_usuario** é a chave primária na Tabela de Usuários.
- **id_proposito** é a chave primária na Tabela de Propósitos.
- A Tabela de Relacionamento utiliza uma chave composta (**id_usuario, id_proposito**), garantindo que um mesmo usuário possa conceder consentimento para múltiplos propósitos sem redundância.

Essa modelagem permite que a filtragem de acessos seja realizada de maneira estruturada, possibilitando que o mecanismo de reescrita de consultas atue diretamente sobre os dados consentidos, tornando o processo compatível com as exigências regulatórias.

3.2 REESCRITA DE CONSULTA NO PARSER DO SGBD

Nesta seção, considerando a modelagem de tabelas apresentada anteriormente, detalhamos o funcionamento do módulo de reescrita de consultas, responsável por garantir que apenas os dados compatíveis com os propósitos autorizados sejam acessados. Esse mecanismo atua interceptando consultas no parser do SGBD e modificando sua estrutura conforme os parâmetros de consentimento e propósito definidos. A implementação desse mecanismo de reescrita está alinhada com trabalhos recentes que exploram o aprimoramento da linguagem SQL para fins de conformidade regulatória. Abreu (ABREU, 2021), por exemplo, propõe uma extensão da linguagem SQL para incorporar controles de acesso baseados em propósito diretamente nas consultas, evidenciando a necessidade de adaptações na estrutura dos SGBDs para garantir o tratamento seguro de informações pessoais.

Considere a seguinte consulta original:

```
1 SELECT * FROM Usuarios;
```

e o parâmetro de filtro `proposito_id` com o valor `'P2'` (correspondente ao propósito de Análise de Crédito), definido a nível de sessão como uma variável de ambiente. Não está no escopo deste trabalho propor um formato específico para a definição desse parâmetro; assume-se que ele é corretamente passado ao módulo de reescrita de consulta.

A seguir, descrevemos o algoritmo de reescrita, que ajusta dinamicamente a consulta para garantir conformidade com os propósitos e consentimentos previamente registrados.

3.2.1 Algoritmo de Reescrita de Consulta

O pseudocódigo abaixo ilustra o fluxo do módulo de reescrita de consultas, que insere filtros apropriados e ajusta a estrutura da consulta antes de sua execução.

```
Funcao ProcessarConsulta(query, parametros):
    // 1. Parse da consulta: converte o SQL em uma AST (Arvore de
       Sintaxe Abstrata)
    ast <- ParseSQL(query)

    // 2. Verifica se a consulta envolve a tabela "Usuarios"
    se NAO ast.refere_a("Usuarios") entao
        retorne ExecutarPlano(GerarPlanoExecucao(ast))
    fim_se

    // 3. Identificacao de campos sensiveis na clausula SELECT
    lista_campos_sensiveis <- {"salario", "cpf", "telefone"}
    campos_sensiveis_encontrados <- []

    para cada campo in ast.clausula_SELECT faca:
        se campo.nome in lista_campos_sensiveis entao:
            campos_sensiveis_encontrados.adicionar(campo.nome)
        fim_se
    fim_para

    // 4. Se nao houver campos sensiveis na consulta, executa normalmente
    se campos_sensiveis_encontrados.Vazio() entao:
        retorne ExecutarPlano(GerarPlanoExecucao(ast))
    fim_se

    // 5. Insercao de JOINs necessarios para verificar proposito e
```

```
        consentimento
    se NAO ast.contemJoin("Usuario_Proposito") entao:
        ast.inserirJoin("Usuarios", "Usuario_Proposito",
                        "Usuarios.id_usuario =
                            Usuario_Proposito.id_usuario")
    fim_se

    se NAO ast.contemJoin("Propositos") entao:
        ast.inserirJoin("Usuario_Proposito", "Propositos",
                        "Usuario_Proposito.id_proposito =
                            Propositos.id_proposito")
    fim_se

    // 6. Aplicacao obrigatoria do filtro por proposito
    se NAO parametros.contem("proposito_id") entao:
        lancar Erro("Filtro de proposito obrigatorio: parametro
            'proposito_id' nao fornecido")
    senao:
        ast.adicionarCondicao("Propositos.id_proposito = :proposito_id")
    fim_se

    // 7. Reescrita dos campos sensiveis conforme os consentimentos
        definidos
    para cada campo in campos_sensiveis_encontrados faca:
        nova_expressao <- "CASE WHEN Propositos.coleta_" + campo +
                          " = TRUE THEN Usuarios." + campo +
                          " ELSE NULL END AS " + campo
        ast.substituirExpressao(campo, nova_expressao)
    fim_para

    // 8. Otimizacao e geracao do plano de execucao
    ast_otimizada <- OtimizarAST(ast)
    plano_execucao <- GerarPlanoExecucao(ast_otimizada)

    // 9. Execucao da consulta final
    resultado <- ExecutarPlano(plano_execucao)
    retorne resultado
FimFuncao
```

3.2.2 Passo a Passo da Reescrita da Consulta

A seguir, detalha-se o processo de reescrita da consulta, assegurando que a execução ocorra conforme os consentimentos concedidos e os propósitos definidos.

1. **Parsing da Consulta:** A consulta original:

```
SELECT * FROM Usuarios;
```

   é convertida em uma Árvore de Sintaxe Abstrata (AST) para análise e manipulação estrutural.

2. **Identificação de Dados Sensíveis:** O sistema verifica se a tabela `Usuarios` contém campos sensíveis. No exemplo, os seguintes campos são identificados:

   - `salario`
   - `cpf`
   - `telefone`

3. **Inserção de JOINs Necessários:** Como a consulta original não inclui associações com as tabelas de controle de consentimento, são adicionadas junções para vincular os usuários aos propósitos autorizados:

```
JOIN Usuario_Proposito
    ON Usuarios.id_usuario = Usuario_Proposito.id_usuario
JOIN Propositos
    ON Usuario_Proposito.id_proposito = Propositos.id_proposito
```

4. **Aplicação do Filtro por Propósito:** Para restringir os dados ao propósito solicitado, é adicionada a cláusula de filtragem correspondente ao parâmetro de sessão:

```
WHERE Propositos.id_proposito = 'P2'
```

5. **Reescrita dos Campos Sensíveis:** Cada campo sensível é modificado de forma que seu valor só seja retornado se o consentimento correspondente for válido. As reescritas seguem o padrão:

   - **Salário:**

```
        CASE WHEN Propositos.coleta_salario = TRUE
             THEN Usuarios.salario
             ELSE NULL
        END AS salario
```

- **CPF:**

```
        CASE WHEN Propositos.coleta_cpf = TRUE
             THEN Usuarios.cpf
             ELSE NULL
        END AS cpf
```

- **Telefone:**

```
        CASE WHEN Propositos.coleta_telefone = TRUE
             THEN Usuarios.telefone
             ELSE NULL
        END AS telefone
```

6. **Otimização e Execução da Consulta:** A AST modificada é otimizada para eficiência, convertida em um plano de execução e finalmente processada pelo SGBD.

3.3 CONSULTA REESCRITA FINAL

A consulta final, após a reescrita e a aplicação dos controles de propósito, tem a seguinte forma:

```
SELECT
    Usuarios.id_usuario,
    Usuarios.nome,
    Usuarios.data_cadastro,
    CASE WHEN Propositos.coleta_salario = TRUE
         THEN Usuarios.salario
         ELSE NULL
    END AS salario,
    CASE WHEN Propositos.coleta_cpf = TRUE
         THEN Usuarios.cpf
         ELSE NULL
    END AS cpf,
    CASE WHEN Propositos.coleta_telefone = TRUE
         THEN Usuarios.telefone
         ELSE NULL
    END AS telefone
FROM Usuarios
```

```
18 JOIN Usuario_Proposito
       ON Usuarios.id_usuario = Usuario_Proposito.id_usuario
20 JOIN Propositos
       ON Usuario_Proposito.id_proposito = Propositos.id_proposito
22 WHERE Propositos.id_proposito = 'P2';
```

Nesse caso, para o propósito P2 (Análise de Crédito), se os metadados indicarem que:

- coleta_salario = **TRUE**
- coleta_cpf = **TRUE**
- coleta_telefone = **FALSE**

então os campos salario e cpf serão retornados com seus valores reais, enquanto telefone será substituído por NULL, garantindo conformidade com os consentimentos fornecidos.

4 RESULTADOS

A consulta reescrita busca apenas os usuários associados ao propósito 'P2' (Análise de Crédito), garantindo que os dados retornados estejam em conformidade com as regras definidas na tabela de propósitos.

Tabela 4 – Resultado da consulta SQL para o propósito 'P2'.

id_usuario	nome	data_cadastro	salario	cpf	telefone
U002	Carlos Souza	2023-12-15	7000.00	987.654.321-00	NULL
U004	Ricardo Lima	2022-08-20	4500.00	321.654.987-00	NULL

Os dados retornados obedecem às permissões definidas no banco de dados, garantindo que apenas informações compatíveis com os propósitos autorizados estejam disponíveis.

- **Usuários filtrados:** Apenas os usuários `U002` e `U004` aparecem no resultado, pois são os únicos associados ao propósito `P2` na tabela `Usuario_Proposito`.
- **Dados Retornados:**
  - **Salário:** Exibido porque o propósito `P2` permite a coleta desse dado (`coleta_salario = TRUE`).
  - **CPF:** Também exibido, pois o propósito autoriza seu acesso (`coleta_cpf = TRUE`).
  - **Telefone:** Retornado como `NULL`, pois `coleta_telefone = FALSE` para `P2`.

4.1 IMPACTO DA SOLUÇÃO PROPOSTA

A reescrita da consulta e a implementação do mecanismo de controle de acesso baseado em propósitos demonstraram-se eficazes para garantir a conformidade com as regulamentações de proteção de dados, como a LGPD e a GDPR. Essa abordagem não apenas assegura a privacidade dos indivíduos, mas também contribui para a governança de dados, fornecendo um controle automatizado sobre o acesso às informações sensíveis.

Dentre os principais benefícios da solução proposta, destaca-se a automação do controle de acesso, reduzindo significativamente a possibilidade de acessos indevidos. Além disso, a rastreabilidade das consultas garante maior transparência e auditabilidade, permitindo que as organizações estejam em conformidade com exigências regulatórias sem comprometer a

eficiência operacional. A flexibilidade da abordagem permite que novos propósitos sejam facilmente incorporados sem exigir grandes mudanças na estrutura do banco de dados, garantindo escalabilidade e adaptabilidade a diferentes cenários empresariais.

No entanto, algumas limitações devem ser consideradas. A aplicação de filtros adicionais pode impactar o desempenho do banco de dados, especialmente em ambientes com grande volume de dados e múltiplos acessos simultâneos. Além disso, a compatibilidade da solução com diferentes SGBDs pode demandar ajustes específicos, sendo necessária uma adaptação para suportar múltiplas tecnologias sem comprometer a integridade das restrições de acesso.

Outro ponto crítico a ser abordado em futuras evoluções da solução é a prevenção contra possíveis ataques de injeção de propósito, onde um usuário mal-intencionado poderia tentar burlar as regras de controle para obter acesso não consentido a dados sensíveis. Mecanismos adicionais de verificação e autenticação de parâmetros podem fortalecer a segurança e evitar vulnerabilidades exploráveis.

Dessa forma, a solução proposta apresenta uma abordagem para o processamento consentido de dados, equilibrando privacidade e conformidade regulatória. Contudo, melhorias contínuas são necessárias para aprimorar a segurança, escalabilidade e eficiência da implementação.

5 CONCLUSÕES E TRABALHOS FUTUROS

A solução proposta permitiu a consulta de dados de usuários de acordo com as permissões definidas por propósitos específicos. A estrutura relacional adotada entre as tabelas `Usuarios`, `Propositos` e `Usuario_Proposito` viabilizou um controle granular sobre quais informações podem ser acessadas em cada contexto. Isso fortalece aspectos de segurança e conformidade com regulamentações como a **LGPD** e a **GDPR**, garantindo que apenas os dados necessários sejam expostos conforme o propósito da consulta.

Além disso, a abordagem utilizada assegura flexibilidade na definição de novos propósitos e regras de coleta de dados, sem a necessidade de reformular a estrutura do banco de dados. Essa escalabilidade permite que a solução evolua conforme novos requisitos regulatórios ou de negócio surjam.

No entanto, algumas oportunidades de otimização foram identificadas ao longo do processo, apontando direções para trabalhos futuros que visam aprimorar a eficiência, flexibilidade e aplicabilidade da solução proposta.

- **Otimização da Consulta SQL:** A atual consulta pode ser reestruturada para minimizar a sobrecarga computacional causada pelo alto número de junções (`JOINs`). Estratégias como a adoção de **CTEs (Common Table Expressions)** ou **Views Materializadas** podem ser exploradas para melhorar a eficiência na recuperação de dados, reduzindo a complexidade das operações e otimizando os tempos de resposta. Além disso, a aplicação de índices adequados nas colunas de junção pode contribuir significativamente para o desempenho do sistema.

- **Reestruturação do Armazenamento dos Metadados:** Atualmente, as permissões de coleta de dados são armazenadas na tabela `Propositos`. Uma abordagem mais flexível e escalável pode ser a adoção de um modelo baseado em **configuração dinâmica utilizando JSON**, permitindo ajustes nos níveis de acesso sem necessidade de alterações estruturais no banco de dados. O uso de **bancos NoSQL** também pode ser explorado para oferecer maior maleabilidade na manipulação dessas permissões, especialmente em cenários onde há frequentes mudanças nas regras de acesso.

- **Expansão da Análise de Dados Sensíveis para Outras Tabelas:** Embora o foco deste trabalho tenha sido a tabela `Usuarios`, outras tabelas do banco de dados po-

dem conter informações sensíveis que devem seguir as mesmas diretrizes de controle de acesso. Como proposta futura, sugere-se a **padronização da classificação de dados sensíveis** em todo o banco de dados, permitindo a implementação de restrições consistentes em múltiplas tabelas e assegurando a conformidade regulatória em diferentes contextos.

- **Implementação Prática em um SGBD:** Como continuação deste estudo, propõe-se a realização de uma implementação prática em um Sistema de Gerenciamento de Banco de Dados (SGBD) real. Essa aplicação permitirá validar a eficiência do mecanismo proposto em um ambiente de produção, analisando seu impacto no desempenho, a facilidade de integração e possíveis desafios operacionais. Testes em diferentes SGBDs, como PostgreSQL, MySQL e SQL Server, podem fornecer insights valiosos sobre a adaptabilidade da solução a diferentes arquiteturas.

Com essas melhorias, espera-se não apenas otimizar o desempenho das consultas, mas também garantir um modelo de dados mais flexível e escalável. A aplicação prática dessas soluções contribuirá para o aprimoramento da governança de dados, fortalecendo a conformidade regulatória e a eficiência no processamento consentido de informações.

REFERÊNCIAS

ABREU, I. C. d. *Processamento de Consulta Orientado a Propósitos: Uma Extensão da Linguagem SQL*. Trabalho de Conclusão de Curso (Graduação em Computação) — Universidade Federal do Ceará, Fortaleza, Brasil, 2021. Disponível em: <http://repositorio.ufc.br/handle/riufc/63747>.

AGGARWAL, C. C. Managing purpose-based data access control in big data environments. *International Journal of Data Governance*, v. 5, n. 1, p. 12–25, 2021.

European Union. *General Data Protection Regulation (GDPR)*. 2016. Accessed: Feb. 28, 2025. Disponível em: <https://gdpr-info.eu/>.

NISSENBAUM, H. *Contextual Integrity: A Framework for Privacy in the Digital World*. [S.l.]: Stanford University Press, 2019.

PAPPACHAN, P.; ENCK, W.; ROCHE, D. S. Towards purpose-based access control for privacy: Concepts and applications. In: *Proceedings of the 25th ACM Symposium on Access Control Models and Technologies (SACMAT)*. [S.l.: s.n.], 2020. p. 115–126.

Presidência da República - Brasil. *Lei Geral de Proteção de Dados Pessoais (LGPD) - Lei nº 13.709, de 14 de agosto de 2018*. 2018. Accessed: Feb. 28, 2025. Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>.

SCHNEIDER, F. B. *Privacy and Security in the Digital Age*. Springer, 2019. Disponível em: <https://link.springer.com/book/10.1007/978-3-030-25594-9>.

SOLOVE, D. J. *The Myth of the Privacy Paradox: Trusting the Law to Protect Privacy in the Digital Age*. [S.l.]: Oxford University Press, 2020.

SOLOVE, D. J.; CITRON, D. K. Risk and anxiety: A theory of data-breach harms. *Texas Law Review*, v. 96, p. 737–786, 2017. Disponível em: <https://texaslawreview.org/wp-content/uploads/2018/03/Solove.pdf>.

TAYLOR, L. Data privacy and ethics in the age of ai. *Journal of Data Protection & Privacy*, v. 3, n. 2, p. 45–58, 2020.

ULLMAN, J. D.; WIDOM, J. *A First Course in Database Systems*. 3. ed. [S.l.]: Pearson, 2008.