Capítulo 3

LGPD em Ambientes de Bancos de Dados nas Organizações

Ana Carolina Brito de Almeida, Letícia Dias Verona, Maria Luiza Machado Campos e Fernanda Araujo Baião

Abstract

Information Security has been an especially relevant topic for the development of Information Systems (IS) in organizations. In 2018, it became even more crucial as the Law 13.709 (LGPD) was passed in Brazil, giving organizations more responsibility for the collection, processing and protection of personal data. It is well known that much of the corporate information is stored in repositories under the management of Database Management Systems (DBMS). In this context, this course addresses the theme of LGPD in database systems (DB) environments within organizations. The course includes a discussion of the two crucial concepts of Data Security and Data Privacy; presents an overview of LGPD, exemplifying its ten principles; and discusses the operational support of some DBs to the principles advocated by such Law.

Resumo

A Segurança da Informação é um tema especialmente relevante para desenvolvimento de Sistemas de Informação (SI) nas organizações. Em 2018, tornou-se ainda mais crucial, pois foi sancionada a Lei 13.709 (LGPD), no Brasil, que atribui mais responsabilidade às organizações quanto à coleta, ao tratamento e à proteção dos dados pessoais. É notório que grande parte das informações corporativas estão armazenadas em repositórios sob a gestão de Sistemas Gerenciadores de Bancos de Dados (SGBD). Neste contexto, o presente minicurso aborda, o tema de LGPD em ambientes de bancos de dados (BDs) nas organizações. O minicurso discute os conceitos de Segurança e Privacidade; apresenta uma visão geral da LGPD, exemplificando seus dez princípios; e discute o suporte operacional em BDs aos princípios preconizados na tal Lei.

3.1. Introdução

A Segurança da Informação é uma preocupação constante durante o desenvolvimento de Sistemas de Informação. Tal área obteve ainda mais visibilidade com a sanção da LGPD,

a Lei Geral de Proteção de Dados [Brasil 2018]. Essa Lei é baseada no Regulamento Geral sobre a Proteção de Dados 2016/679 (RGPD, ou, como é mais conhecida em inglês, GDPR - General Data Protection Regulation), elaborado pela União Europeia. A Lei obriga organizações a seguirem uma série de itens quanto à coleta, ao tratamento e à proteção dos dados pessoais. A RGPD tem como foco proteger direitos fundamentais de liberdade e privacidade dos indivíduos, complementando regulamentações previamente existentes na Convenção Europeia de Direitos Humanos e impondo maiores obrigações aos agentes privados detentores de dados pessoais. A lei brasileira foi aprovada em 2018 e ainda necessita ser regulamentada para sua entrada em vigor em 2020. O seu espectro de atuação, bastante amplo e ainda muito subjetivo, contempla o direito ao cidadão de impedir a divulgação ou posse de qualquer dado pessoal a seu respeito, isolado ou agregado estatisticamente. Em ambas, é destacada a necessidade de consentimento explícito para coletar os dados pessoais e transparência total sobre o que será feito a partir deles.

Em tempos de internet das coisas (IoT), redes sociais e aplicativos móveis, é importante que haja atenção tanto por parte dos usuários quanto ao cadastro e à transferência de seus dados quanto pelas empresas que detém tais dados, de forma a protegê-los de vazamentos. Por exemplo, quando um usuário se cadastra em um aplicativo móvel de corrida, que registra o tempo e a distância que ele percorre, a empresa desenvolvedora do aplicativo não pode enviar os dados coletados para uma empresa de plano de saúde, suplemento alimentar ou de marcas esportivas sem o consentimento do usuário.

Usualmente, os dados de usuários são mantidos em repositórios corporativos nas organizações. Além disso, a maioria das organizações armazena os dados pessoais coletados dos clientes em Sistemas de Gerenciamento de Banco de Dados (SGBD). Dessa forma, é necessário saber como e o quanto as empresas detentoras dos principais SGBDs comerciais estão preparadas para dar suporte à implantação de estratégias da LGPD de forma eficaz e eficiente.

Tradicionalmente, os SGBDs dos principais fornecedores de mercado dispõem de recursos para prover segurança da informação em seus produtos. Mais recentemente, no entanto, este aspecto vem se ampliando para tratar da privacidade, também motivado por este cenário recente em que a RGPD se insere. Neste sentido, a Oracle disponibiliza uma série de pacotes para aumentar a segurança dos dados armazenados[1], tais como: *Oracle Advanced Security, Oracle Key Vault, Oracle Data Masking and Subsetting* etc. Tais pacotes oferecem suporte à: criptografia de dados transparente; gerenciamento de chave de criptografia; controle de acesso multifatores e usuários com privilégios; classificação e descoberta de dados; monitoramento e bloqueio de atividades de banco de dados (BD); auditoria e relatórios consolidados; e mascaramento de dados. Já a Microsoft disponibilizou um *guia*[2] de conformidade com o RGPD, mencionando como ferramentas do SGBD SQL Server 2017 podem auxiliar neste sentido, visando auxiliar estratégias de implantação da LGPD. Alguns exemplos de ferramentas são o *Microsoft Compliance Manager*, uma solução que permite aos clientes que trabalham com nuvem gerenciar sua

[1] https://www.oracle.com/database/security/

[2] https://info.microsoft.com/sql-server-gdpr-ebook-registration.html

própria conformidade e o *Data Discovery and Classification*, uma ferramenta para descobrir, classificar, rotular e relatar os dados sensíveis nos BDs dos usuários.

O objetivo do minicurso é possibilitar uma visão geral da LGPD tanto com a perspectiva de usuário quanto com a perspectiva de um administrador de BD, exemplificando recursos de alguns dos principais SGBDs de mercado. O conteúdo está estruturado em tópicos principais, descritos a seguir:

- ✓ **Segurança e Privacidade:** apresentação e discussão dos conceitos de segurança e de privacidade.

- ✓ **Introdução e visão geral da Lei Geral de Proteção de Dados Pessoais (LGPD):** apresentação dos principais conceitos envolvidos na Lei, as boas práticas indicadas e exemplos de aplicação dos dez princípios sobre o tratamento de dados pessoais com o uso de aplicativos de celular, incluindo redes sociais.

- ✓ **LGPD e mecanismos de segurança nos SGBDs:** apresentação das funcionalidades encontradas nas principais empresas desenvolvedoras de SGBD do mercado para contemplar os tópicos da LGPD: anonimização de dados e criptografia, notificação de vazamento de dados, recursos de auditoria. Apresentação de exemplos práticos, ilustrando diversas situações reais e frequentemente encontradas no dia a dia de uma empresa, explorando os mecanismos de proteção apresentados e como eles atendem essas situações. Discussão do estado da arte, oportunidades e desafios.

3.2. Segurança e Privacidade

É importante destacar que, embora estejam muito relacionados e exista uma sobreposição considerável entre questões relacionadas ao acesso a recursos (segurança) e questões relacionadas ao uso de informações (privacidade), existem diferenças importantes entre os conceitos de segurança e privacidade.

Elmasri e Navathe (2019) diferenciam bem esses conceitos, definindo que Segurança na Tecnologia da Informação refere-se a muitos aspectos da proteção de um sistema do uso não autorizado, incluindo autenticação de usuários, criptografia de informações, controle de acesso, políticas de *firewall* e detecção de intrusões. Para o propósito do minicurso, limitaremos nosso tratamento de Segurança aos conceitos associados a quão bem um sistema pode proteger o acesso às informações que ele contém, incluindo a integridade e disponibilidade dessas informações.

O conceito de privacidade, por sua vez, transcende o aspecto de segurança e de fato vem sendo tratado como parte de uma discussão mais ampla sobre Transparência em cenários recentes da 4a Revolução Industrial (*Fourth Industrial Revolution*, ou 4IR), caracterizados por uma fusão de tecnologias e eliminando fronteiras entre os meios físico, digital e biológico [Teixeira et al, 2019]. A privacidade examina até que ponto o uso de informações pessoais que o sistema detém sobre um usuário está em conformidade com as suposições explícitas ou implícitas em relação a esse uso. Do ponto de vista do usuário final, a privacidade pode ser considerada a partir de duas perspectivas diferentes: impedir o armazenamento de informações pessoais e garantir o uso apropriado de informações

pessoais. Para o presente minicurso, a ideia básica é discutir os mecanismos que os Sistemas de Gerenciamento de Banco de Dados oferecem para garantir o uso apropriado de informações pessoais por eles armazenadas.

Na indústria, a privacidade realmente se concentra nos seguintes conceitos [Dean 2017]:

- ✔ Quais dados devem ser coletados?
- ✔ Quais são os usos permitidos?
- ✔ Com quem isso pode ser compartilhado?
- ✔ Por quanto tempo os dados devem ser retidos?
- ✔ Qual modelo de controle de acesso granular é apropriado?

De uma forma resumida, os controles de segurança são criados para controlar quem pode acessar as informações, enquanto a privacidade é mais granular, controlando quais dados específicos eles podem acessar, e quando. Dean [2017] exemplifica os dois conceitos em um cenário: Se você deposita em uma instituição financeira nacional, todos os caixas do país podem ser provisionados (ou seja, acesso de segurança concedido) para acessar os detalhes da sua conta. Isso fornece a flexibilidade para um cliente visitar uma filial em sua cidade natal, uma filial na costa oeste durante uma viagem de negócios ou uma filial da Flórida durante as férias. Mas a privacidade é outra camada. Embora o caixa possa ser provisionado para exibir todos os detalhes da conta dos clientes, a privacidade só permite acesso quando existe uma necessidade comercial; como um cliente entrando em uma filial em outra cidade para acessar suas contas. Mas a privacidade não permite que o mesmo caixa veja o saldo da conta de seus vizinhos ou talvez o saldo de uma pessoa famosa, apenas porque eles estão interessados - apesar de seus privilégios de acesso lhes concederem acesso.

Portanto, a aplicação comercial dos termos privacidade e segurança é muito diferente, mas com uma certa sobreposição. Segundo Dean (2017), "Você não pode ter privacidade sem segurança, mas pode ter segurança sem privacidade".

3.3. Lei Geral de Proteção de Dados Pessoais (LGPD)

A Lei Geral de Proteção de Dados (Lei 13.709/18), alterada pela Lei 13.853/19, é uma lei brasileira fortemente inspirada pelo RGPD europeu, e que dispõe sobre o tratamento de dados pessoais com objetivo de proteger a liberdade e a privacidade dos cidadãos [Brasil 2018]. A lei abrange toda atividade que envolve coleta, tratamento e armazenamento de dados pessoais, seja ela praticada por entes privados ou públicos no Brasil, inclusive empresas internacionais com atividade no país.

3.3.1. Principais conceitos e princípios

O conceito de dado pessoal, conforme descrito na lei, significa qualquer informação, que individualmente ou combinada com outras, possa identificar uma pessoa ou submetê-la a

um tratamento específico. Como exemplos de dados pessoais podem ser citados nome, CPF, endereço, *cookies* gravados em computadores pessoais, informações partilhadas em redes sociais, dados financeiros ou qualquer informação que permita a identificação de um indivíduo. A lei estabelece ainda o conceito de dado pessoal sensível como sendo dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural.

A LGPD determina que o titular do dado, ou seja, a pessoa natural a quem se referem os dados objeto da coleta, tratamento e armazenamento, tem direito a saber como seus dados estão sendo tratados, a ter acesso aos mesmos, corrigi-los, pedir a sua exclusão, correção e revogar o consentimento ao seu uso. Pode ainda solicitar a informação de quem teve acesso aos seus dados através de atividades compartilhadas com o controlador, que vem a ser a entidade pública ou privada a cujos interesses o processamento dos dados é submetido.

Segundo a LGPD, toda a atividade de tratamento de dados deve obedecer aos seguintes princípios:

- ✓ Princípio 1 - finalidade: realização do tratamento para propósitos legítimos, específicos, explícitos e informados ao titular, sem possibilidade de tratamento posterior de forma incompatível com essas finalidades;
- ✓ Princípio 2 - adequação: compatibilidade do tratamento com as finalidades informadas ao titular, de acordo com o contexto do tratamento;
- ✓ Princípio 3 - necessidade: limitação do tratamento ao mínimo necessário para a realização de suas finalidades, com abrangência dos dados pertinentes, proporcionais e não excessivos em relação às finalidades do tratamento de dados;
- ✓ Princípio 4 - livre acesso: garantia, aos titulares, de consulta facilitada e gratuita sobre a forma e a duração do tratamento, bem como sobre a integralidade de seus dados pessoais;
- ✓ Princípio 5 - qualidade dos dados: garantia, aos titulares, de exatidão, clareza, relevância e atualização dos dados, de acordo com a necessidade e para o cumprimento da finalidade de seu tratamento;
- ✓ Princípio 6 - transparência: garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial;
- ✓ Princípio 7 - segurança: utilização de medidas técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou difusão;
- ✓ Princípio 8 - prevenção: adoção de medidas para prevenir a ocorrência de danos em virtude do tratamento de dados pessoais;
- ✓ Princípio 9 - não discriminação: impossibilidade de realização do tratamento para fins discriminatórios ilícitos ou abusivos;

- ✓ Princípio 10 - responsabilização e prestação de contas: demonstração, pelo agente, da adoção de medidas eficazes e capazes de comprovar a observância e o cumprimento das normas de proteção de dados pessoais e, inclusive, da eficácia dessas medidas.

O acesso aos dados pessoais é normatizado e somente pode ser realizado se o titular dos dados der seu consentimento explícito para realização de atividades específicas, a não ser que o tratamento se enquadre em: cumprimento de leis, tanto da atividade privada do controlador quanto da administração pública; realização de estudos por órgão de pesquisa; exercício de direitos contratuais ou judiciais; quando necessário para a execução de contrato ou de procedimentos preliminares relacionados a contrato do qual seja parte o titular; proteção da vida ou integridade física do titular ou terceiros; tutela da saúde do titular; garantia de prevenção à fraude ou segurança do titular e para a proteção do crédito.

3.3.2. A genealogia da LGPD

Em 2010, foi realizada no Brasil a primeira consulta pública a respeito da proteção de dados. Em 2014, entra em vigor o decreto do Marco Civil da Internet que, dentre os seus objetos, inclui a privacidade do usuário da Internet no Brasil. De 2014 a 2018, diversos projetos de lei se referiram ao tema. Destes, a criação do cadastro positivo que previu a criação de um banco de dados para que as instituições financeiras facilitem o acesso ao crédito a bons pagadores englobou uma ampla discussão sobre privacidade e por fim permitiu que o titular dos dados solicite sua exclusão e esquecimento. Em agosto de 2018, estes projetos de lei foram consolidados na LGPD. No fim do mesmo ano, uma Medida Provisória (869/2018) vetou alguns artigos e criou a Autoridade Nacional de Proteção de Dados (ANPD) com o objetivo de fiscalizar e regulamentar a aplicação da LGPD. As funções e procedimentos desta agência e como se dará sua atuação ainda são desconhecidos para o grande público.

Os fatores decisivos para a concretização e publicação da lei foram o interesse do governo brasileiro de ingressar na OCDE. Um dos requisitos para a entrada no grupo é que o país possua uma lei geral de proteção de dados que permita discutir questões comerciais entre países e a publicação do RGPD europeu, em cujas bases a lei brasileira se assenta. Empresas brasileiras com subsidiárias fora do país, com clientes e fornecedores europeus, e mesmo as que poderiam ter dados de um cidadão europeu na sua base, passaram a se preocupar com o cumprimento do regulamento e com a necessidade de segurança jurídica interna.

Existem, entretanto, diferenças importantes entre a lei brasileira e o regulamento europeu. Em relação à aplicação da lei, as sanções europeias são ágeis e as multas substanciais. Empresas multinacionais de *software*, de serviços *on-line* e de transporte foram multadas por autoridades europeias em dezenas de milhões de dólares por expor indevidamente dados dos seus usuários. As multas brasileiras são limitadas a 2% do faturamento bruto da empresa, o que pode causar uma desproporcionalidade entre o lucro obtido pelo controlador dos dados e o dano causado ao seu titular.

Em termos conceituais, a RGPD afirma que o dado só deve ser usado para o propósito limitado para o qual foi coletado. Já a legislação brasileira admite o uso dos

dados para o legítimo interesse do controlador, o que pode ser conflitante com os direitos do titular e múltiplas circunstâncias. Além disto a LGPD considera a possibilidade da transferência dos dados para outro controlador, o que não é previsto no regulamento europeu sem o consentimento explícito do titular. A redação da lei brasileira permite um amplo espectro de interpretações e questionamentos e alguns são apontados na seção a seguir.

3.3.3. Questões importantes a serem respondidas

A anonimização de dados ou a solicitação de consentimento é uma prática comum na realização de pesquisas científicas e outros trabalhos baseados em dados. Ainda assim, os termos da lei, se não esclarecidos, podem causar uma insegurança jurídica em instituições de pesquisa e entidades governamentais. A definição de dado pessoal como todo dado que pode identificar unicamente uma pessoa natural faz com que, se considerarmos sua combinação com outro dado, possa abranger qualquer informação, ainda que para o controlador original e com interesses legítimos de pesquisa seja um dado anonimizado.

No campo da saúde, a questão é agravada pela definição da lei considerar dados biológicos, genéticos e relativos à saúde como dados pessoas sensíveis, que possuem um grau maior de severidade na aplicação da lei. As pesquisas relacionadas à saúde, em muitos estudos, envolvem a ampla discussão de um caso específico, os protocolos de tratamento adotados e os resultados obtidos. Esses dados podem alegadamente identificar unicamente uma pessoa natural e impedir o seu uso para fins de avanço da ciência.

Do outro lado do espectro de interesses, a inclusão de proteção ao crédito na área financeira, como uma das possibilidades de viabilização de coleta e tratamento de dados sem consentimento, cria uma vulnerabilidade ao titular, pois esses dados, a exemplo de adimplência e contratos de empréstimos, são itens de privacidade importantes para o cidadão.

Ainda sob a ótica das lacunas da lei que podem ser utilizadas em prol das empresas está o conceito de legítimo interesse. Em uma sociedade de livre mercado, o lucro é um interesse legítimo de uma empresa e esse argumento pode justificar a coleta sem consentimento - e sem conhecimento - de dados pessoais com o intuito de direcionar publicidade, o que em última instância implica também em manipulação de emoções, desejos e orientações políticas.

Por fim, a privacidade de agentes do poder público deve ser equilibrada com o interesse civil em fiscalizar suas ações e os limites entre a lei de transparência e a LGPD podem ser fluidos e, por essa razão, devem ser explicitados.

A LGPD, no momento da publicação deste capítulo, ainda carece de regulamentação e muitos conceitos e aplicações possuem lacunas de entendimento. A necessidade de adaptação de empresas e entes públicos, entretanto, é notória e urgente para que possam atender os requisitos mínimos da lei. Esta necessidade gera uma demanda de entendimento e conhecimento dos profissionais e cientistas da área da Ciência da Computação e áreas correlatas, bem como uma adaptação administrativa da maioria das empresas com negócios no país.

As seções a seguir objetivam fornecer um panorama das ferramentas e possibilidades existentes em alguns dos principais SGBDs existentes, com relação aos aspectos de segurança e seus impactos nas questões de privacidade.

3.4. LGPD e mecanismos de segurança nos SGBDs

A presente seção descreve como as principais empresas desenvolvedoras de ambientes de Bancos de Dados, e seus respectivos SGBDs, buscam auxiliar as organizações na adaptação à LGPD.

Com base em um estudo sobre a RGPD [Rajasekharan 2017], levantamos que os principais requisitos de segurança de dados da LGPD podem ser amplamente classificados em três categorias: **avaliação**, **prevenção** e **monitoramento/detecção**.

A **avaliação** está relacionada ao impacto na proteção de dados quando certos tipos de processamento de dados pessoais provavelmente apresentarão um "alto risco" para o titular dos dados. A avaliação deve incluir uma avaliação sistemática e abrangente dos processos, perfis da organização e como essas ferramentas salvaguardam os dados pessoais (LGPD - Capítulo VII - Seção II - Art. 50. § 2º - Letra d).

Em relação à **prevenção** de brechas de segurança, a própria LGPD recomenda algumas técnicas para prevenir os ataques. São elas: anonimização e pseudonimização, controle de acesso de usuário privilegiado, controle de acesso refinado e minimização de dados. A anonimização de dados é a técnica de embaralhar ou ofuscar completamente os dados e a pseudonimização refere-se à redução da vinculação de um conjunto de dados com a identidade original de um titular de dados. A LGPD afirma que as técnicas de anonimização e pseudonimização podem reduzir o risco de divulgação acidental ou intencional de dados, tornando as informações não identificáveis para um indivíduo ou entidade (LGPD - Capítulo II - Seção II - Art. 13.). O controle de acesso de usuário privilegiado que têm acesso aos dados pessoais deve impedir ataques de informações privilegiadas e contas de usuário comprometidas (LGPD - Capítulo VI - Seção I). Além do controle privilegiado do usuário, a LGPD recomenda a adoção de uma metodologia refinada de controle de acesso para garantir que os dados pessoais sejam acessados seletivamente e apenas para uma finalidade definida. Esse tipo de granulação fina do controle de acesso pode ajudar as organizações a minimizar o acesso não autorizado aos dados pessoais (LGPD - Capítulo II - Seção I - Art. 10 - § 1º). A minimização de dados diz respeito à recomendação de minimizar a coleta e retenção de dados pessoais para reduzir o limite de conformidade. Ao coletar, processar ou compartilhar dados pessoais, é necessário ser frugal e limitar a quantidade de informações às necessidades de uma atividade específica (Capítulo I - Art. 6º - III).

O **monitoramento/detecção** de brechas é necessário porque nenhuma organização, mesmo com a adoção de medidas preventivas de segurança, consegue eliminar totalmente a possibilidade de uma violação de dados. A LGPD recomenda esse monitoramento e o alerta para detectar tais violações através dos seguintes mecanismos: dados de auditoria e monitoramento e alerta oportuno. A LGPD exige não apenas o registro ou a auditoria das atividades nos dados pessoais, mas também recomenda que esses registros devem ser mantidos centralmente sob a responsabilidade do controlador (LGPD - Capítulo VI - Seção I - Art. 37.). Por fim, o monitoramento constante das atividades de dados pessoais é fundamental para detectar anomalias. A LGPD também

exige notificações oportunas em caso de violação (LGPD - Capítulo VII - Seção I - Art. 48.).

Além disso, a LGPD também exige conformidade com os princípios de proteção de dados para aprimorar a qualidade e o rigor da proteção dos dados. Entre os dez princípios da LGPD, destacam-se três deles relacionados aos ambientes de SGBDs:

- ✓ Princípio 7 – Segurança: proteger os dados armazenados;
- ✓ Princípio 8 – Prevenção: coibir vazamento de dados;
- ✓ Princípio 10 – Responsabilização: o agente de tratamento de dados pessoais deve demonstrar quais medidas foram adotadas para evitar o vazamento de dados (auditoria).

Diante da categorização ampla dos requisitos de segurança da LGPD e desses três princípios que impactam diretamente os ambientes de SGBDs, investigam-se soluções de diversas naturezas que possam diminuir a vulnerabilidade dos dados armazenados nas organizações. Essas soluções envolvem tanto funcionalidades diretas do SGBD e ferramentas associadas, quanto serviços, disponibilizados em plataformas na nuvem, que atuam como uma camada, abrangendo não só os SGBDs como também as demais aplicações executadas na mesma plataforma.

Baseado no quadrante mágico da Gartner de 2018 (Figura 3.1), decidiu-se investigar três principais empresas (Microsoft, Oracle, Amazon Web Services), além de um SGBD gratuito, o PostgreSQL.

Figura 3.1. Quadrante mágico para Sistemas Gerenciadores de Bancos de Dados Operacionais [Feinberg et al, 2018]

3.4.1. Microsoft

A Microsoft dispõe de diversos aspectos de segurança para auxiliar os usuários de SGBDs na adaptação à LGPD, controlando o acesso, prevenindo e detectando intrusos e vulnerabilidades e gerando relatórios de auditoria.

Seguindo a categorização ampla dos requisitos de segurança da LGPD, alguns aspectos propostos pela Microsoft são [Microsoft 2018a]:

- Avaliação: *data discovery and classification* e *sql vulnerability assessment*.
- Prevenção: *dynamic data masking (DDM), static data masking, sql server authentication, object-level permissions, role-based security, row-level security, transport layer security (TLS), transparent data encryption (TDE)* e *always encrypted*.
- Monitoramento/detecção: *sql server audit, sql server temporal tables* e *sql vulnerability assessment*.

3.4.1.1 Microsoft - Categoria de Avaliação

Toda organização possui um grande volume de dados, incluindo dados pessoais (e, particularmente, também dados sensíveis que devem ser protegidos segundo a LGPD). Diante do grande volume de dados, é importante que o controlador tenha o auxílio de uma ferramenta que o ajude a avaliar, identificar e categorizar os dados pessoais.

A *feature* ***data discovery and classification*** ajuda a organizar e classificar os dados para garantir o manuseio adequado e o melhor gerenciamento de informações pessoais [Microsoft 2019a]. Essa *feature* é acessada através da ferramenta *SQL Server Management Studio* (SSMS), ao selecionar um banco de dados, e já disponibiliza uma lista com as recomendações de classificação dos dados (Figura 3.2). Em seguida, o administrador do banco pode selecionar as recomendações com as quais concorda. Além disso, o administrador do banco de dados pode adicionar classificações de forma manual.

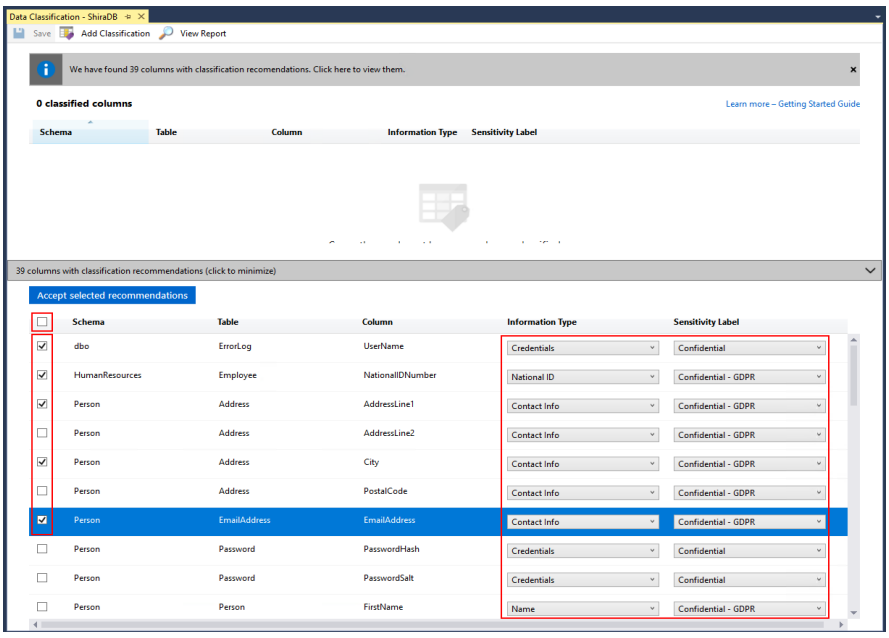

Figura 3.2. Lista de classificação dos dados proposta pelo Data Discovery and Classification [Microsoft 2019a]

Além da descoberta de dados pessoais, é importante avaliar as possíveis vulnerabilidades que existem no SGBD. A *feature* ***sql vulnerability assessment*** pode ajudar a detectar problemas de segurança e permissões. Quando um problema é detectado, pode-se fazer relatórios com uma busca detalhada no banco de dados para encontrar ações para resolução aos problemas através do SSMS [Microsoft 2017d]. Esse relatório (auditoria) também auxilia na categoria de monitoramento/detecção da LGPD. A Figura 3.3 apresenta um exemplo de relatório de verificação de vulnerabilidades. O relatório apresenta uma visão geral do seu estado de segurança, quantos problemas foram encontrados e suas respectivas gravidades.

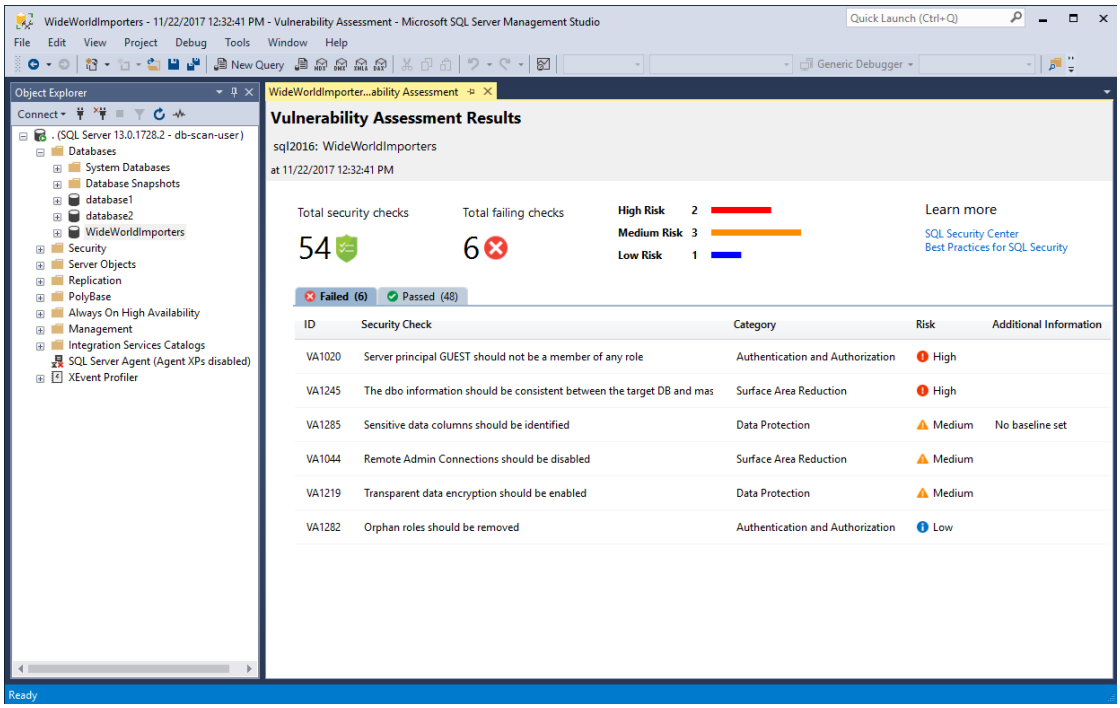

Figura 3.3. Exemplo de relatório de verificação de vulnerabilidades [Microsoft 2017d]

A Figura 3.4 apresenta uma solução possível para um problema de vulnerabilidade recomendada pela própria ferramenta de verificação. A recomendação é remover o membro `GUEST` de todos os papéis.

3.4.1.2 Microsoft - Categoria de Prevenção

A Microsoft disponibiliza *features*, de acordo com as recomendações da LGPD na categoria de prevenção, para a anonimização e pseudonimização através do mascaramento dinâmico e estáticos de dados, o controle de acesso de usuário privilegiado e a criptografia de dados, que apesar de não constar explicitamente na LGPD, é um meio de segurança para os dados, em caso de vazamento.

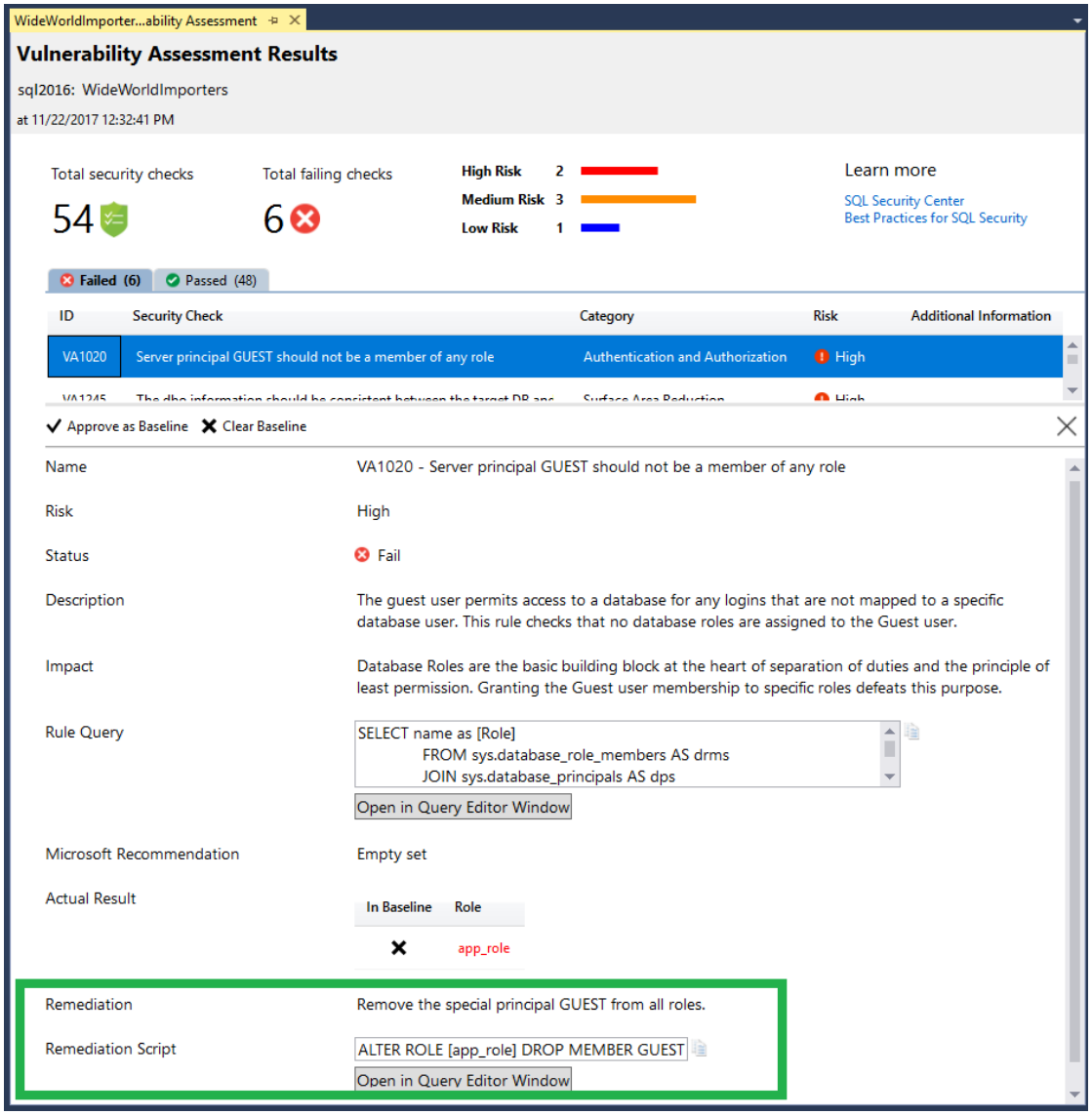

Figura 3.4. Exemplo de solução para vulnerabilidade detectada [Microsoft 2017d]

A *feature* ***dynamic data masking*** **(DDM)** limita a exposição aos dados pessoais através do mascaramento deles para usuários não privilegiados [Microsoft 2019c]. Ele mascara os dados em tempo de execução, facilitando a modelagem e o código de segurança nas aplicações. O mascaramento de dados pode ocorrer de forma total ou parcial e, no caso de dados numéricos, existe um tipo de máscara aleatória. Por exemplo, na Figura 3.5, tem-se um exemplo de comando que adiciona uma função de mascaramento (partial) para a coluna `LastName` da tabela `Membership`. O primeiro argumento da função é o prefixo, ou seja, a posição do primeiro caractere real a ser mostrado do dado, o segundo argumento possui os caracteres do meio e que irão mascarar o conteúdo no momento da exibição do resultado da consulta e o último argumento é o sufixo, ou seja, a posição do último caractere real a ser mostrado.

```
ALTER TABLE Membership
ALTER COLUMN LastName ADD MASKED WITH (FUNCTION = 'partial(2,"XXX",0)');
```

Figura 3.5. Comando de inclusão de mascaramento de dados a uma coluna existente [Microsoft 2019c]

A *feature* ***static data masking***, disponível no Sql Server Management Studio 18.0 preview 5 e posterior, possibilita a criação de uma cópia do banco de dados que tenha os dados pessoais mascarados para que o usuário possa compartilhar tal cópia sem compartilhar os dados pessoais contidos no banco [Granet 2018]. Diferente do DDM, o mascaramento ocorre em nível de armazenamento e todos os usuários da cópia do banco de dados tem o mesmo dado mascarado. Esse mascaramento ocorre em nível de coluna como pode ser visto na Figura 3.6, no passo 1 (step 1), onde as colunas `AddressLine`, `DateOfBirth`, `EmailId`, `FirstName`, `LastName` e `SSN`, da tabela `dbo.Customers`, são selecionadas para mascaramento. As funções escolhidas para o mascaramento de dados são: *shuffle, single value, null, group shuffle* e *string composite*. A função *shuffle* embaralha os dados (`AddressLine`) para as novas linhas e não introduz nenhum valor novo. A função *single value* substitui todos os conteúdos da coluna (`DateOfBirth`) pelo único valor inserido no momento da configuração. A função *null* substitui o conteúdo da coluna (`EmailId`) pelo valor null. Nesse caso, a coluna precisa ser opcional para poder usar essa função. A função *group shuffle* vincula mais de uma coluna (`FirstName` e `LastName`) no mascaramento aleatório, ou seja, usa o conteúdo de mais colunas para o embaralhamento. A função *string composite* permite o mascaramento da coluna inteira ou de parte dela. Por exemplo, o `SSN` pode ser parcialmente mascarado, mantendo-se apenas os seus quatro últimos dígitos. Ainda na Figura 3.6, no passo 2 (step 2), seleciona-se a localização do arquivo destino da cópia mascarada do banco de dados e no passo 3 (step 3) detalha-se o nome do banco de dados destino e o arquivo onde constará o log do mascaramento. Na Figura 3.7 tem-se um exemplo dos dados não mascarados (*Unmasked Data*) e dos dados após o mascaramento (*Masked Data*).

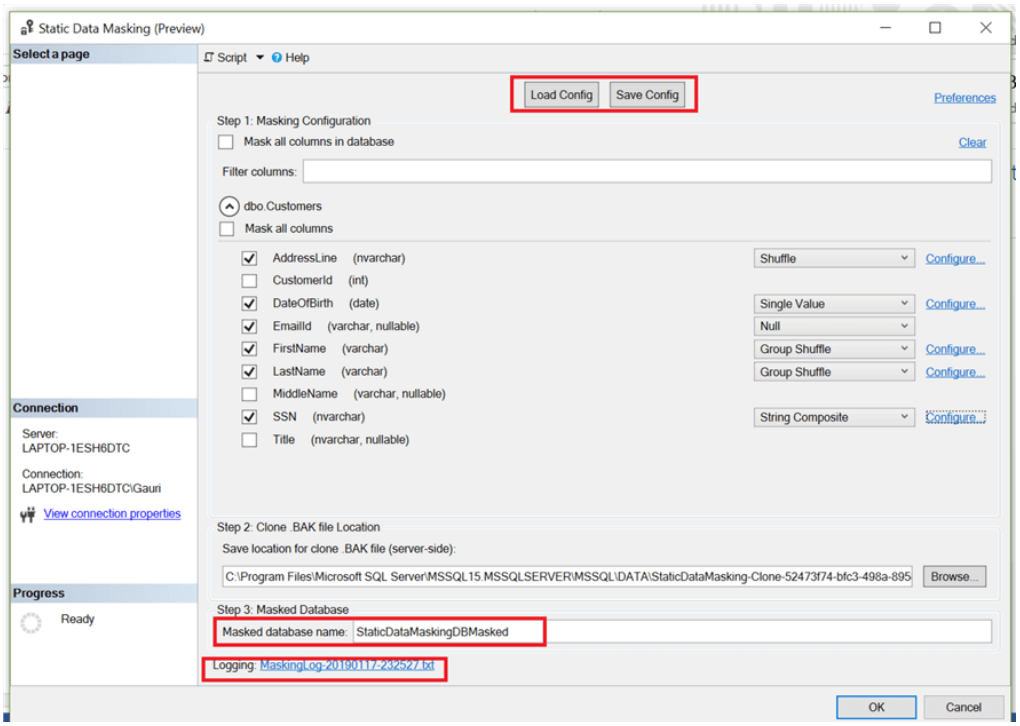

Figura 3.6. Seleção de colunas a serem mascadas fisicamente na cópia do banco de dados [Mahajan 2019]

Unmasked Data

FirstName	MiddleName	LastName	DateOfBirth	SSN	AddressLine	EmailId
Mihail	U	Frintu	1969-01-05	364-95-1699	1970 Napa Ct.	mihail@adventure-works.com
Ken	M	Ray	1996-05-29	150-85-5065	9833 Mt. Dias Blv.	terri0@adventure-works.com
Terri	A	Selikoff	1991-03-17	549-82-9234	7484 Roundtree Drive	roberto0@adventure-works.com
Roberto	N	Poland	1977-03-29	281-65-5849	9539 Glenside Dr	rob0@adventure-works.com
Rob	W	Rettig	1955-05-22	413-25-8713	1226 Shoe St.	gail0@adventure-works.com
Gail	V	Osada	1980-01-29	204-67-1050	1399 Firestone Drive	jossef0@adventure-works.com
Jossef	J	Philips	1991-03-17	764-92-9954	5672 Hale Dr.	dylan0@adventure-works.com
Dylan	C	Netz	1977-03-29	582-55-5002	6387 Scenic Avenue	diane1@adventure-works.com
Diane	M	Keyser	1955-05-22	798-17-7390	8713 Yosemite Ct.	gigi0@adventure-works.com
Gigi	M	Brown	1980-01-13	868-43-4288	250 Race Court	michael6@adventure-works.com
Michael	T	Kaliyath	1972-01-26	864-70-1391	1318 Lasalle Street	ovidiu0@adventure-works.com
Ovidiu	S	Frintu	1966-01-29	381-02-3744	5415 San Gabriel Dr.	thierry0@adventure-works.com
Thierry	N	Creasey	1980-01-29	755-88-1537	9265 La Paz	janice0@adventure-works.com
Janice	R	Cook	1972-01-14	732-84-0387	8157 W. Book	michael8@adventure-works.com
Michael	A	Martinez	1989-01-29	207-57-9704	4912 La Vuelta	sharon0@adventure-works.com
Sharon	Z	Goldstein	1969-01-29	045-14-7883	40 Ellis St.	david0@adventure-works.com
David	A	Cornelsen	1969-01-29	732-05-0120	6696 Anchor Drive	kevin0@adventure-works.com
Kevin	J	Petculescu	1976-12-18	126-77-2761	1873 Lion Circle	john5@adventure-works.com
John	R	Stadick	1969-01-29	822-08-3794	3148 Rose Street	mary2@adventure-works.com
Mary	R	Wedge	1973-01-29	586-80-9583	6872 Thornwood Dr.	wanida0@adventure-works.com

Masked Data

FirstName	MiddleName	LastName	DateOfBirth	SSN	AddressLine	EmailId
Patrick	U	Earls	2000-01-01	XXX-YY-1699	40 Ellis St.	NULL
Gigi	M	Brown	2000-01-01	XXX-YY-5065	2115 Passing	NULL
Mark	A	Yu	2000-01-01	XXX-YY-9234	9537 Ridgewood Drive	NULL
Ryan	N	Sacksteder	2000-01-01	XXX-YY-5849	4948 West 4th St	NULL
James	W	Scardelis	2000-01-01	XXX-YY-8713	7511 Cooper Dr.	NULL
Pete	V	Caron	2000-01-01	XXX-YY-1050	1285 Greenbrier Street	NULL
Danielle	J	Hay	2000-01-01	XXX-YY-9954	Pascalstr 951	NULL
Sandeep	C	Vanderhyde	2000-01-01	XXX-YY-5002	2354 Frame Ln.	NULL
Kevin	M	Koch	2000-01-01	XXX-YY-7390	7726 Driftwood Drive	NULL
Roberto	M	Poland	2000-01-01	XXX-YY-4288	2466 Clearland Circle	NULL
Bonnie	T	Ralls	2000-01-01	XXX-YY-1391	34 Waterloo Road	NULL
Mary	S	Martin	2000-01-01	XXX-YY-3744	10203 Acorn Avenue	NULL
Gail	N	Osada	2000-01-01	XXX-YY-1537	6387 Scenic Avenue	NULL
Denise	R	Bischoff	2000-01-01	XXX-YY-0387	2059 Clay Rd	NULL
Thomas	A	Barbariol	2000-01-01	XXX-YY-9704	5669 Ironwood Way	NULL
Houman	Z	Yalovsky	2000-01-01	XXX-YY-7883	5666 Hazelnut Lane	NULL
Sidney	A	Lertpiriyasuwat	2000-01-01	XXX-YY-0120	8463 Vista Avenue	NULL
Reuben	J	Watters	2000-01-01	XXX-YY-2761	1061 Buskirk Avenue	NULL
Peter	R	Keyser	2000-01-01	XXX-YY-3794	502 Alexander Pl.	NULL
Michael	R	Gubbels	2000-01-01	XXX-YY-9583	9784 Mt Etna Drive	NULL

Figura 3.7. Dados antes e após o mascaramento estático de dados [Mahajan 2019]

A partir daqui as *features* de prevenção lidam com o controle de acesso dos usuários. A *feature* ***sql server authentication*** ajuda a gerenciar as identidades dos usuários que acessam os bancos de dados e os servidores, impedindo o acesso não autorizado e pode ser configurado no SSMS [Microsoft 2018b]. Existem duas formas de autenticação no SGBD sql server: a autenticação do Windows e o modo misto. A autenticação do Windows é a forma padrão, onde as contas de usuário e grupos específicas do Windows são confiáveis para conectarem ao sql server. Já o modo misto suporta tanto a autenticação pelo Windows quanto pelo próprio sql server. Os pares de nome de usuário e senha são mantidos no sql server. A recomendação da Microsoft é utilizar a autenticação do Windows sempre que for possível, visto que a autenticação do Windows usa diversas mensagens criptografadas para autenticar os usuários no sql server, enquanto as credenciais do sql server trafegam pela rede, tornando-as menos seguras.

A *feature object-level permissions* permite a concessão de permissões em um nível excepcionalmente granular – até visualização de tabela, procedimento armazenado,

função escalar ou serviço de fila [Microsoft 2014]. Na Figura 3.8 tem-se um exemplo de consulta sobre as permissões que os usuários possuem sobre os objetos do banco de dados.

Results | Messages

	UserName	UserType	DatabaseUserName	Role	PermissionType	PermissionState	ObjectType	ObjectName	ColumnNa
1	NULL	Windows User	dbo	NULL	CONNECT	GRANT	DATABASE	NULL	NULL
2	NULL	SQL User	guest	NULL	NULL	NULL	NULL	NULL	NULL
3	Test	SQL User	Test	NULL	CONNECT	GRANT	DATABASE	NULL	NULL

Figura 3.8. Permissões de usuários aos objetos do banco de dados [Microsoft 2014]

A *feature* ***role-based security*** permite conceder permissões baseadas em papéis ou grupos de usuários ao invés de usuários individuais, reduzindo o ataque ao banco de dados e simplificando a administração de segurança [Microsoft 2017a]. O sql server disponibiliza papéis, em nível de servidor, para a administração do SGBD e as permissões atribuídas a eles não podem ser alteradas. O papel *sysadmin* abrange todos os outros papéis e tem escopo ilimitado, devendo ser atribuído somente a usuários altamente confiáveis. Além disso, existem papéis em nível de banco de dados, tendo um conjunto pré-definido de permissões. Os usuários do banco de dados podem ser adicionados aos papéis do banco de dados ou do servidor.

A *feature* ***row-level security*** restringe o acesso, de acordo com os direitos do usuário, limitando o acesso a linhas em uma tabela baseado no relacionamento entre o usuário e o dado [Microsoft 2019b]. Essa *feature* é implementada no sql server através da instrução `CREATE SECURITY POLICY` e predicados criados como funções com valores embutidos da tabela. Na Figura 3.9 tem-se um exemplo de criação de uma função que retorna o valor 1 (um) quando o conteúdo de uma linha da coluna do representante de vendas (@SalesRep) é o mesmo que o usuário que executa a consulta (@SalesRep = USER_NAME()) ou se o usuário que está executando a consulta for o gerente (USER_NAME() = 'MANAGER').

```
CREATE FUNCTION Security.fn_securitypredicate(@SalesRep AS sysname)
    RETURNS TABLE
WITH SCHEMABINDING
AS
    RETURN SELECT 1 AS fn_securitypredicate_result
WHERE @SalesRep = USER_NAME() OR USER_NAME() = 'Manager';
```

Figura 3.9. Função que realiza o cruzamento do usuário que realiza a consulta e a linha que está sendo consultada [Microsoft 2019b]

Na Figura 3.10 é apresentado o comando para a criação de uma política de segurança que adiciona a função anterior (Figura 3.9) como um predicado de filtro sobre a tabela de vendas (dbo.Sales). O estado (STATE=ON) precisa ser definido como ON para habilitar a política.

```
CREATE SECURITY POLICY SalesFilter
ADD FILTER PREDICATE Security.fn_securitypredicate(SalesRep)
ON dbo.Sales
WITH (STATE = ON);
```

Figura 3.10. Comando de criação de política de segurança [Microsoft 2019b]

Na Figura 3.11 tem-se as permissões de consulta (`GRANT SELECT ON`) na função para os usuários `Manager, Sales1` e `Sales2`. Dessa forma, quando o gerente consultar os dados, a função retornará 1 e o mesmo terá acesso a todas as linhas contidas

na tabela de vendas. Já os vendedores `Sales1` e `Sales2` só terão acesso às linhas de suas próprias vendas, pois, como ele não possui o papel de gerente, a função só retornará 1 quando o usuário que estiver consultando for igual ao conteúdo da linha da coluna representante de vendas.

```
GRANT SELECT ON security.fn_securitypredicate TO Manager;
GRANT SELECT ON security.fn_securitypredicate TO Sales1;
GRANT SELECT ON security.fn_securitypredicate TO Sales2;
```

Figura 3.11. Exemplo de permissões aos usuários [Microsoft 2019b]

Finalizando as *features* de prevenção, tem-se àquelas ligadas à criptografia. Na camada de transporte, a *feature* **TLS** é um protocolo de comunicação que garante comunicações altamente seguras, onde os dados são criptografados para ajudar a garantir que nenhum dado seja interceptado durante o tráfego entre o banco de dados e a aplicação cliente [Microsoft 2019d]. O TLS pode ser usado para validação do servidor quando uma conexão do cliente solicita criptografia. Se a instância do sql server estiver sendo executada em um computador ao qual foi atribuído um certificado de uma autoridade de certificação pública, a identidade do computador e a instância do SQL Server serão emitidas pela cadeia de certificados que leva à autoridade raiz confiável. Essa validação de servidor exige que o computador, no qual o aplicativo cliente está sendo executado, seja configurado para confiar na autoridade raiz do certificado que é usado pelo servidor.

A *feature* **TDE** protege os dados em repouso mesmo que a mídia física (cópias de segurança) seja perdida ou que os dados sejam descartados incorretamente [Microsoft 2019e]. Ele criptografa e descriptografa o banco de dados, as cópias de segurança e os logs de transações em tempo real, sem requerer qualquer mudança nas aplicações. A criptografia usa uma DEK (chave de criptografia do banco de dados), que é armazenada no registro de inicialização do banco de dados para disponibilidade durante a recuperação. A DEK é uma chave simétrica protegida por um certificado armazenado no banco de dados mestre do servidor ou uma chave assimétrica protegida por um módulo EKM (gerenciamento extensível de chaves). Na Figura 3.12 tem-se uma série de comandos para a utilização da TDE. Primeiro é necessário criar uma chave mestra (`CREATE MASTER KEY`), atribuindo-se uma senha. Em seguida, cria-se um certificado protegido pela chave mestra (`CREATE CERTIFICATE`). Após essa criação, cria-se uma chave de criptografia de banco de dados (`CREATE DATABASE ENCRYPTION KEY`), protegendo-a com o certificado anterior e por fim, define-se o banco de dados para usar a criptografia (`ALTER DATABASE ... SET ENCRYPTION ON`).

```
USE master;
GO
CREATE MASTER KEY ENCRYPTION BY PASSWORD = '<UseStrongPasswordHere>';
go
CREATE CERTIFICATE MyServerCert WITH SUBJECT = 'My DEK Certificate';
go
USE AdventureWorks2012;
GO
CREATE DATABASE ENCRYPTION KEY
WITH ALGORITHM = AES_128
ENCRYPTION BY SERVER CERTIFICATE MyServerCert;
GO
ALTER DATABASE AdventureWorks2012
SET ENCRYPTION ON;
GO
```

Figura 3.12. Comandos para uso do TDE [Microsoft 2019e]

A *feature* ***always encrypted*** é uma tecnologia que auxilia na proteção de dados pessoais enquanto eles estão em uso em nível de coluna [Microsoft 2017c]. Ela criptografa e descriptografa no computador cliente sem revelar a chave de criptografia para o servidor do banco de dados. Dessa forma, os dados ficam visíveis somente para as pessoas responsáveis por gerenciar tais dados e não para os administradores do banco de dados ou usuários altamente privilegiados que não tenham acesso. Como resultado, essa tecnologia fornece uma separação entre aqueles que possuem os dados (e podem exibi-lo) e aqueles que gerenciam os dados (mas que não devem ter acesso). Um driver é instalado no computador cliente e automaticamente, criptografa e descriptografa os dados confidenciais. O driver criptografa as colunas de dados confidenciais antes de passar os dados para o servidor de banco de dados e reconfigura automaticamente as consultas para que a semântica do aplicativo seja preservada. Um cenário em que esse tipo de tecnologia é útil é quando uma empresa deseja que um fornecedor externo administre o sql server. Dessa forma, eles não terão acesso aos dados confidenciais, pois os mesmos estarão criptografados no banco. Na Figura 3.13 apresenta-se um exemplo de comandos que cria os metadados de uma chave mestra (`CREATE COLUMN MASTER KEY`) de coluna, os metadados de chave de criptografia de coluna (`CREATE COLUMN ENCRYPTION KEY`) e uma tabela com colunas criptografadas (`CustName` e `SSN`). O valor de `ENCRYPTED_VALUE` foi cortado para não sobrecarregar a figura.

```
CREATE COLUMN MASTER KEY MyCMK
WITH (
     KEY_STORE_PROVIDER_NAME = 'MSSQL_CERTIFICATE_STORE',
     KEY_PATH = 'Current User/Personal/f2260f28d909d21c642a3d8e0b45a830e79a1420'
   );
-----------------------------------------------
CREATE COLUMN ENCRYPTION KEY MyCEK
WITH VALUES
(
    COLUMN_MASTER_KEY = MyCMK,
    ALGORITHM = 'RSA_OAEP',
    ENCRYPTED_VALUE = 0x01700000016C006F00630061006C006D0061006300680069006E0065002F006D0
);
-----------------------------------------------
CREATE TABLE Customers (
    CustName nvarchar(60)
        COLLATE  Latin1_General_BIN2 ENCRYPTED WITH (COLUMN_ENCRYPTION_KEY = MyCEK,
        ENCRYPTION_TYPE = RANDOMIZED,
        ALGORITHM = 'AEAD_AES_256_CBC_HMAC_SHA_256'),
    SSN varchar(11)
        COLLATE  Latin1_General_BIN2 ENCRYPTED WITH (COLUMN_ENCRYPTION_KEY = MyCEK,
        ENCRYPTION_TYPE = DETERMINISTIC ,
        ALGORITHM = 'AEAD_AES_256_CBC_HMAC_SHA_256'),
    Age int NULL
);
GO
```

Figura 3.13. Comandos para dados sempre criptografados [Microsoft 2017c]

3.4.1.3 Microsoft - Categoria de Monitoramento/detecção

Quando um vazamento acontece, a organização precisa detectá-lo o mais rapidamente possível para minimizar seu impacto, além de entender quais registros foram afetados. É importante que toda a organização seja alertada imediatamente quando alguma atividade

fora do comum for detectada e, a partir daí, monitore comportamentos suspeitos. Além disso, o controlador deve ter o auxílio de ferramentas de auditoria que possam gerar relatórios sobre as atividades que ocorreram no banco de dados.

A *feature* ***sql server audit*** rastreia as atividades do banco de dados para ajudar no entendimento e identificação de possíveis ameaças, abusos suspeitos ou violação de segurança [Microsoft 2016a]. Os eventos auditados (ações atômicas) podem ser gravados nos logs de eventos ou nos arquivos de auditoria e ocorrem em nível de servidor ou de banco de dados. Por padrão, o sql server não habilita a auditoria. Ela pode ser habilitada pelo SSMS ou via linha de comando sql. Utilizando a linha de comando, na Figura 3.14, são mostrados exemplos de criação de auditoria se servidor (`CREATE SERVER AUDIT`), a habilitação da auditoria do servidor (`ALTER SERVER ... STATE=ON`) e a criação de uma auditoria de banco de dados que audita as instruções `SELECT` e `INSERT` realizadas por qualquer usuário `dbo` para a tabela `HumanResources.EmployeePayHistory` no banco de dados `AdventureWorks2012` [Microsoft 2017b].

```
USE master ;
GO
-- Create the server audit.
CREATE SERVER AUDIT Payrole_Security_Audit
    TO FILE ( FILEPATH =
linux file path) ;
GO
-- Enable the server audit.
ALTER SERVER AUDIT Payrole_Security_Audit
WITH (STATE = ON) ;
USE AdventureWorks2012 ;
GO
-- Create the database audit specification.
CREATE DATABASE AUDIT SPECIFICATION Audit_Pay_Tables
FOR SERVER AUDIT Payrole_Security_Audit
ADD (SELECT , INSERT
     ON HumanResources.EmployeePayHistory BY dbo )
WITH (STATE = ON) ;
GO
```

Figura 3.14. Comandos para criar e habilitar auditorias de servidor e banco de dados [Microsoft 2017b]

A *feature* ***sql server temporal tables*** são tabelas com versão do sistema, sendo do tipo de tabela de usuário modeladas para manter um histórico completo da mudança dos dados em qualquer ponto no tempo e que podem ser usadas para gerar relatórios sobre os dados auditados [Microsoft 2016b]. A tabela temporal é dita com versão do sistema porque o período de validade para cada linha é gerenciado pelo sistema (SGBD). Cada tabela temporal possui duas colunas do tipo datetime2, que armazena o período de validade para cada linha sempre que uma linha é modificada. Além disso, a tabela temporal possui uma referência a outra tabela com um esquema espelhado (histórico), que é usada para armazenar a versão anterior da linha, automaticamente, sempre que uma linha na tabela temporal é atualizada ou excluída. Na Figura 3.15 é apresentado um exemplo de consulta à tabela de histórico (`FOR SYSTEM_TIME`), onde são retornadas as versões de linhas que satisfaçam a condição de `EmployeeID = 1000` e que estavam ativas durante o período de 01/01/2014 e 01/01/2015, inclusive.

```
SELECT * FROM Employee
    FOR SYSTEM_TIME
      BETWEEN '2014-01-01 00:00:00.0000000' AND '2015-01-01 00:00:00.0000000'
        WHERE EmployeeID = 1000 ORDER BY ValidFrom;
```

Figura 3.15. Consulta a tabela de histórico [Microsoft 2016b]

A *feature* ***sql vulnerability assessment***, conforme já descrito no item de avaliação, também auxilia com relatórios sobre as atividades ocorridas no banco de dados.

3.4.2. Oracle

A Oracle disponibiliza diversos aspectos de segurança para apoiar a privacidade dos dados. Para encolher a superfície de ataque e reduzir o número de maneiras pelas quais os invasores podem acessar os bancos de dados, é extremamente importante impor a segurança o mais próximo possível dos dados. Considerando as três categorias de requisitos da LGPD, descrevemos a seguir alguns dos diversos produtos que a Oracle disponibiliza para auxiliar na proteção dos dados armazenados e no controle de acesso a esses dados.

- Avaliação: *oracle data safe* (avaliação de segurança e risco) e *oracle database vault*.
- Prevenção: *oracle data safe* (mascaramento de dados), *oracle data masking and subsetting*, *oracle advanced security*, *oracle key vault*, *oracle database vault* e *oracle label security*.
- Monitoramento/detecção: *oracle data safe* (auditoria de atividades), *oracle audit vault* e *database firewall*.

3.4.2.1. Oracle - Categoria de Avaliação

Um dos desafios ao avaliar a natureza dos riscos é determinar o que avaliar, porque os aplicativos de banco de dados normalmente contêm vários pontos de entrada e têm dados pessoais espalhados por várias colunas e tabelas com controle de acesso vagamente definido. A Oracle auxilia nesse desafio da LGPD provendo ferramentas para ajudar na avaliação de segurança e risco.

Oracle data safe

O *oracle data safe* é um serviço na nuvem que garante segurança para os bancos de dados residentes na nuvem [Oracle 2019a]. Essa segurança é obtida através de avaliações de segurança e risco do usuário, auditoria de atividades, descoberta de dados confidenciais e mascaramento de dados.

A avaliação de segurança auxilia na identificação da existência de lacunas na estratégia de configuração do banco de dados e sugere maneiras de corrigir essas lacunas. Dessa forma, é possível identificar vulnerabilidades de segurança, por exemplo, quando senhas padronizadas estão sendo utilizadas ou quando os usuários possuem mais privilégios do que eles deveriam. Por exemplo, a Figura 3.16 apresenta um alerta do serviço para a quantidade de usuários que possuem o privilégio de DBA (DBA Role) no banco de dados, visto que esse tipo de privilégio deve ser restrito apenas aos administradores da base de dados e que possam ter acesso total aos dados, incluindo os dados sensíveis.

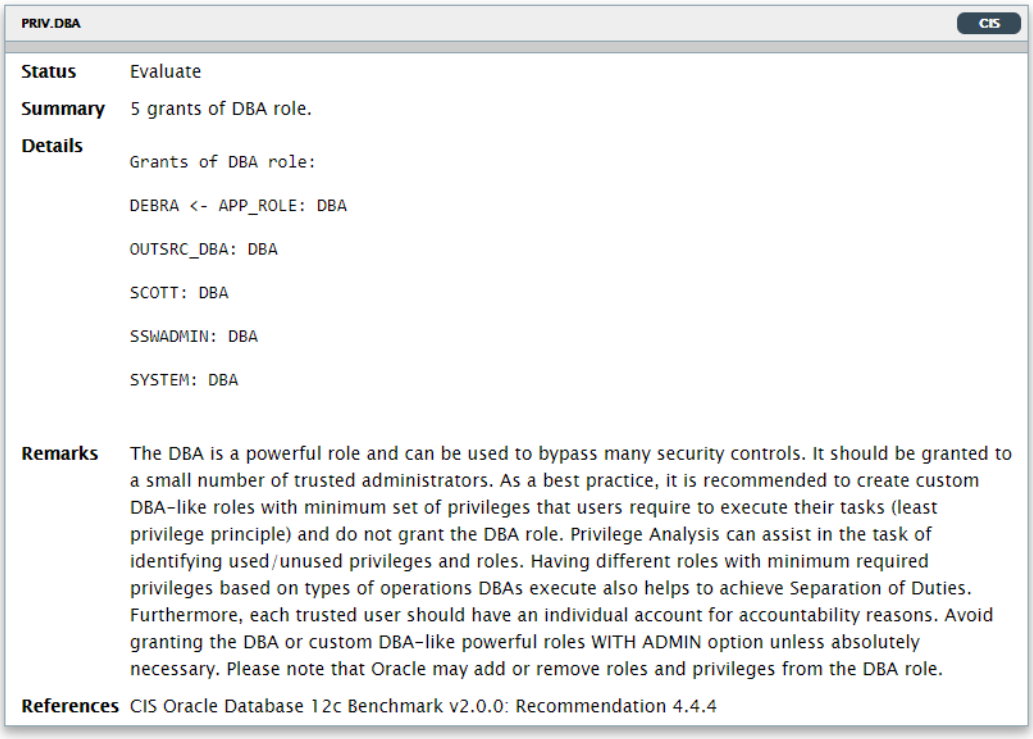

Figura 3.16. Exemplo de uso do serviço de avaliação de segurança[3]

A avaliação de risco do usuário permite avaliar e monitorar o usuário de forma a identificar possíveis riscos associados a contas privilegiadas. A Figura 3.17 apresenta um exemplo de análise realizada pelo serviço, onde algumas contas de usuários apresentam alto nível de exposição a riscos.

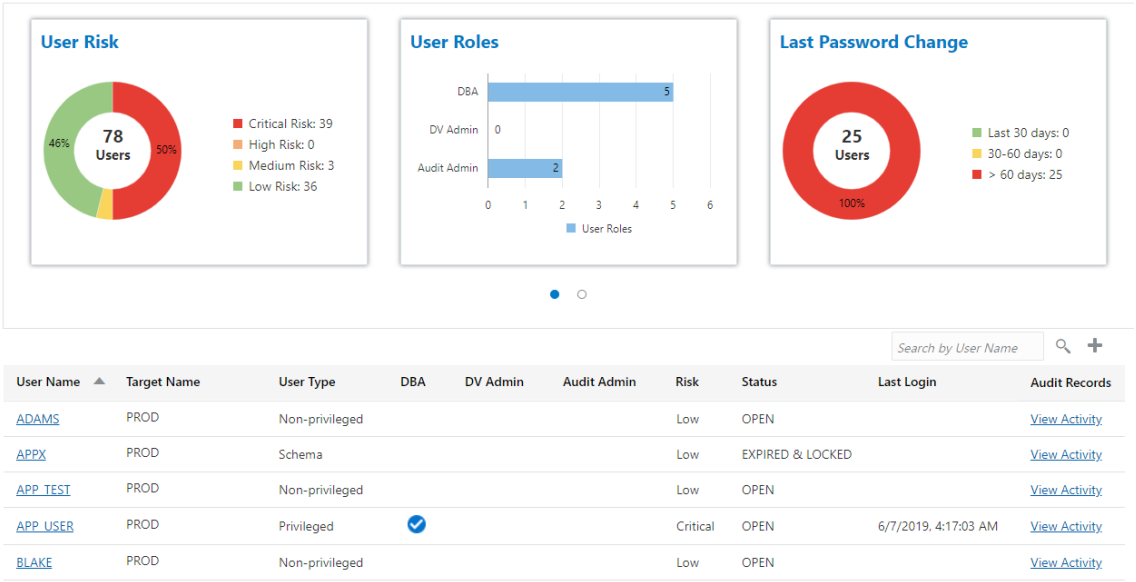

Figura 3.17. Exemplo de avaliação de riscos de usuários [Oracle 2019a]

[3] https://blog.bronto.com/bg/database/technologies/security/dbsat.html

As demais funcionalidades do *oracle data safe* são detalhadas nas suas respectivas categorias.

Oracle database vault

O *oracle database vault* trabalha em conjunto com o banco de dados oracle para evitar ameaças que exploram credenciais roubadas, usuários que usam contas privilegiadas para acessar dados confidenciais, criar novas contas, conceder privilégios adicionais, usuários que ignoram as políticas de uso dos dados das organizações, ameaças aos dados confidenciais durante a janela de manutenção das aplicações entre outras [Oracle 2019c]. Devido às suas diversas funções, o *oracle database vault* auxilia tanto na avaliação quanto na categoria de prevenção de ataques.

Após a identificação dos dados pessoais, torna-se importante identificar usuários (titulares de dados, terceiros, autoridades de supervisão e destinatários), incluindo usuários e administradores privilegiados (controladores, operadores), que não podem apenas acessar, mas também processar os dados pessoais [Rajasekharan 2017]. Durante a modelagem e a manutenção do sistema, privilégios adicionais podem ser concedidos inadvertidamente aos usuários. A análise de privilégios do *oracle database vault* ajuda a aumentar a segurança dos aplicativos, identificando os privilégios reais usados no tempo de execução. Os privilégios identificados como não utilizados, podem ser avaliados para uma possível revogação, ajudando a obter um modelo de privilégios mínimos.

3.4.2.2. Oracle - Categoria de Prevenção

Conforme já discutido, a LGPD recomenda diversas técnicas preventivas, tais como: pseudonimização, anonimização, controle de usuário privilegiado entre outras. Um dos desafios de qualquer controle de proteção de dados preventivo é a possível sobrecarga que ele cria nos sistemas e nas operações diárias de TI (tecnologia da informação). Esta sobrecarga pode vir em termos de mudança de processos; alterações necessárias no código-fonte do sistema, sobrecarga de desempenho e preocupações com escalabilidade. No entanto, a Oracle descreve que aborda tais desafios através de controles preventivos transparentes para a maioria dos sistemas e com um impacto mínimo no desempenho e nas operações contínuas de TI [Rajasekharan 2017].

Oracle data safe

O *oracle data safe* auxilia no item de anonimização dos dados, através do seu mascaramento, para que caso ocorra vazamento de dados pessoais, esses dados não sejam vinculados às pessoas reais. Para isso, é importante identificá-los. A descoberta de dados confidenciais ajuda a decidir quais dados devem ser protegidos. Esse serviço identifica e classifica mais de 125 tipos de dados sensíveis, tais como: dados de tecnologia da informação, dados financeiros, dados de saúde entre outros. Esse serviço é particularmente útil para empresas que possuem várias equipes de desenvolvimento e seus dados estejam distribuídos sobre vários bancos de dados, sendo difícil a identificação dos dados sensíveis e onde os mesmos estão localizados. A Figura 3.18 mostra algumas das categorias pré-definidas de dados sensíveis e a partir daí, o usuário pode selecionar a categoria que ele deseja descobrir quais seriam os dados sensíveis nos seus bancos de dados.

Figura 3.18. Exemplo de categorias de descoberta de dados sensíveis [Oracle 2019a]

O mascaramento de dados substitui os dados sensíveis do ambiente de produção por dados fictícios, mas realistas. Esse mascaramento pode ser usado para os ambientes de desenvolvimento e homologação das organizações, onde o desenvolvedor não precisa ter acesso, por exemplo, ao número do cartão de crédito de um usuário para realizar testes mais próximos da realidade. Dessa forma, o conjunto de dados de teste passa a ser realista, mas sem expor os dados sensíveis. A Figura 3.19 apresenta um exemplo de mascaramento do identificador do cliente (SSN) e do seu cartão de crédito (Credit Card).

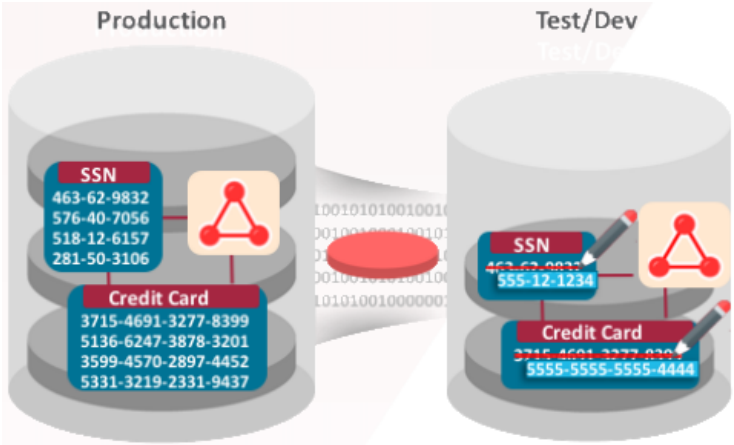

Figura 3.19. Exemplo de mascaramento de dados [Oracle 2019a]

Oracle data masking and subsetting

O *oracle data masking and subsetting* também auxilia na anonimização de dados e é um *plugin* acoplado ao banco de dados Oracle que cria um ambiente (desenvolvimento ou homologação) com um subconjunto dos dados de produção mascarados, mas realistas [Oracle 2017]. O mascaramento de dados segue um conjunto de regras pré-definidas para mapeamento. Esse plugin possui o objetivo similar ao mascaramento de dados contido no *oracle data safe*.

Oracle advanced security

Embora a LGPD não indique, explicitamente, a criptografia como uma forma de prevenção, a RGPD recomenda tal técnica, sendo ela extremamente importante para prevenção. O desafio para as organizações é a implementação da criptografia não só para os dados pessoais em tabelas criptografadas, mas também para *backups*, *data dumps* e

arquivos de *log*. O *oracle advanced security* é uma opção do banco de dados Oracle 19c e auxilia na criptografia de dados através do *Oracle advanced security transparent data encryption* (TDE), mas também na pseudonimização através do *Oracle advanced security data redaction* [Oracle 2019b].

O TDE criptografa os dados sensíveis na camada de banco de dados e, com isso, auxilia na prevenção de ataques que tentam ignorar o banco de dados e ler informações confidenciais de arquivos de dados no nível do sistema operacional, de *backups* de banco de dados ou de exportações de banco de dados. Os aplicativos e usuários autenticados no banco de dados continuam tendo acesso aos dados de forma transparente, enquanto usuários não autenticados que tentam burlar o banco de dados têm acesso negado para descriptografar os dados.

Na Figura 3.20 apresenta-se um exemplo de ataque que o banco de dados pode sofrer de pessoas que tenham acesso ao usuário do sistema operacional que tenha privilégios sobre os arquivos do banco de dados. No exemplo citado, o usuário do sistema operacional pode buscar por conteúdos com números, no arquivo que possui dados financeiros (*tablespace*) do banco de dados. Com isso, o usuário consegue obter a informação limpa com os números dos cartões de créditos registrados no banco de dados.

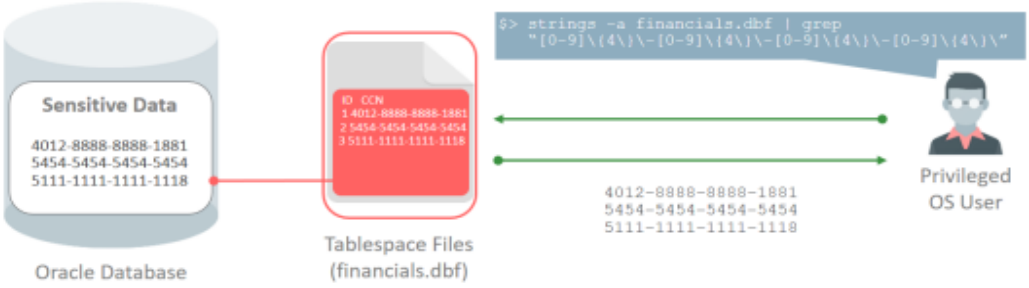

Figura 3.20. Exemplo de ataque ao banco de dados [Oracle 2019b]

Já na Figura 3.21 consegue-se ver o mesmo cenário com o TDE em ação. O TDE pode criptografar *tablespaces* ou até mesmo bancos de dados inteiros, incluindo as *tablespaces* SYSTEM, SYSAUX, TEMP e UNDO. Todo esse processo é transparente para as aplicações porque os processos de criptografia e descriptografia não requerem qualquer mudança na aplicação e os usuários das aplicações não conseguem lidar diretamente com os dados criptografados. Além dessas opções, o DTE também permite criptografar somente algumas colunas na tabela, desde que o usuário saiba quais seriam os dados sensíveis. Essa opção é relevante para bancos de dados enormes, tal como o *datawarehouse*.

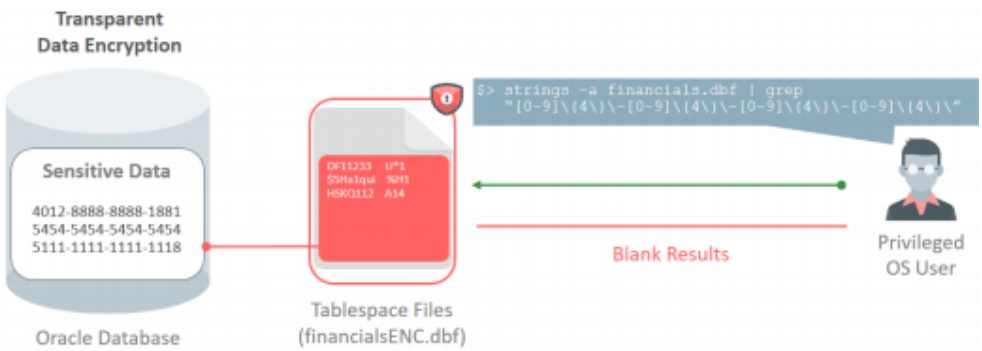

Figura 3.21. Exemplo de intervenção do TDE ao ataque no banco de dados [Oracle 2019b]

O *Oracle advanced security data redaction* fornece uma redação seletiva e dinâmica de dados sensíveis durante a apresentação dos resultados de uma consulta ao banco de dados, antes da exibição pelos aplicativos, para que os usuários não autorizados não possam visualizar tais dados. Os dados armazenados permanecem inalterados, enquanto os dados exibidos são transformados e editados antes de sair do banco de dados. O tipo de cenário em que essa redação pode ser relevante é o uso de aplicações de *call center*. O atendente não deve possuir acesso às informações confidenciais de clientes, como o número do cartão de crédito, visto que isso pode violar alguns regulamentos de privacidade (ex.: LGPD) e expor dados confidenciais sem necessidade.

A Figura 3.22 apresenta alguns exemplos de transformações que podem ser realizadas antes de serem exibidas nas aplicações.

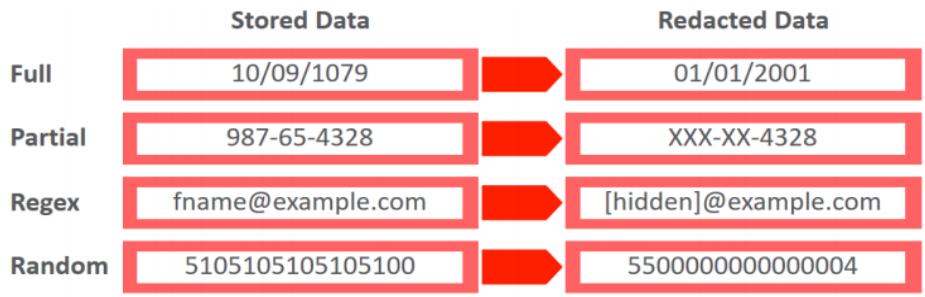

Figura 3.22. Exemplo de redação de dados para as aplicações [Oracle 2019b]

Oracle key vault

O *oracle key vault* também auxilia na prevenção através do controle centralizado sobre dados criptografados com o TDE. Ele possui a capacidade de suspender o acesso à chave mestra e renderizar os dados criptografados de forma ininteligível em caso de violação de dados ou atividade suspeita [Rajasekharan 2017]. O *oracle key vault* é um sistema de segurança para armazenar, centralizar e gerenciar chaves mestras do TDE (usadas para criptografia e descriptografia de dados) de vários bancos de dados Oracle e outros aplicativos de segurança [Oracle 2018a]. Esse sistema elimina alguns desafios operacionais de gerenciamento de chaves tais como: rotação periódica de senhas, realização de cópias de segurança e recuperação de senhas perdidas.

Além disso, conforme mostrado na Figura 3.23, as chaves de criptografia são armazenadas fisicamente e gerenciadas em um local separado de onde os dados criptografados residem, atendendo a uma regra frequente em regulamentos de segurança.

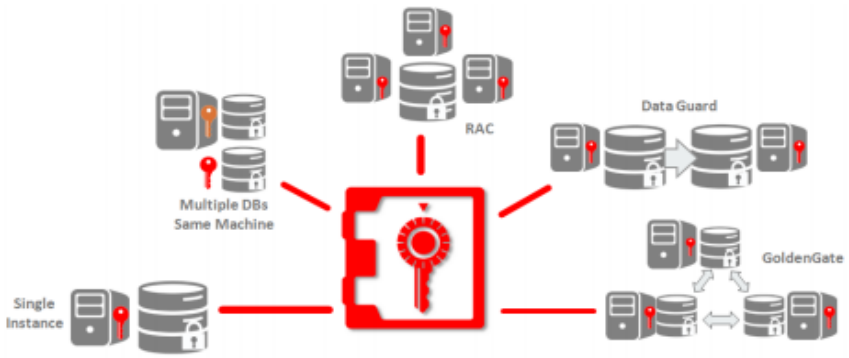

Figura 3.23. Cenário de uso do oracle key vault [Oracle 2018a]

Oracle database vault

O *oracle database vault* auxilia na prevenção através do controle de usuário privilegiado (papel de DBA). Esse tipo de conta, normalmente, possui acesso completo aos dados armazenados no banco de dados. No entanto, com o *oracle database vault*, é criado um ambiente de aplicação restrito ("Realm") dentro do banco de dados que previne o acesso aos dados da aplicação a partir de contas privilegiadas enquanto continua permitindo as atividades administrativas autorizadas regulares no banco de dados. Na Figura 3.24 tem-se um exemplo onde o DBA não consegue recuperar os dados de uma determinada tabela (`hr.emp`), ou seja, somente a pessoa autorizada para visualizar esses dados que consegue.

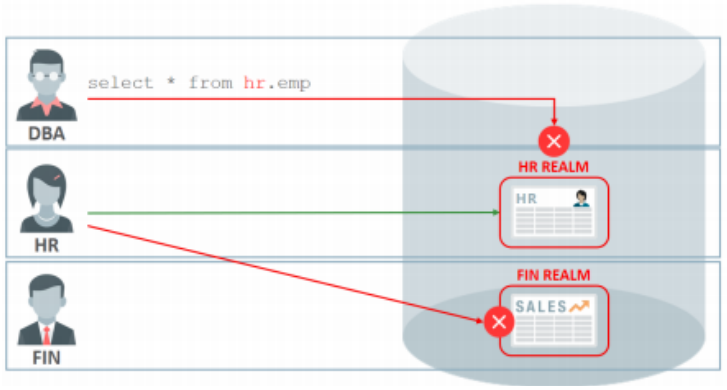

Figura 3.24. Oracle database vault atuando sobre contas privilegiadas [Oracle 2019c]

Outra função do *oracle database vault* é controlar a configuração do banco de dados de forma a impedir alterações no banco de dados que possam levar a configurações inseguras, desvios de configuração (alterações em estruturas de tabelas), reduzir a possibilidade de constatações de auditoria e melhorar a conformidade. Essa prevenção é adquirida através do controle do uso de comandos, tais como: ALTER SYSTEM, ALTER USER, CREATE USER, DROP USER entre outros. Por exemplo, na Figura 3.25, tem-se um cenário em que comandos do DBA são recusados, tais como: TRUNCATE TABLE e CONNECT de um IP desconhecido pelo banco de dados. Em resumo, é controlado o uso de comandos SQL que possam modificar o dicionário e a configuração do banco de dados e com isso, abrir o banco de dados para vulnerabilidades de segurança.

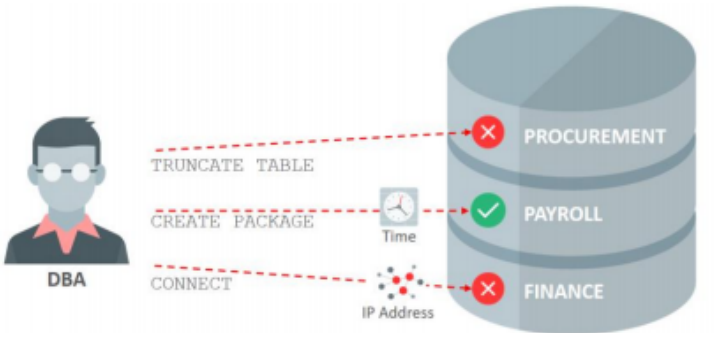

Figura 3.25. Oracle database vault atuando sobre a configuração do banco de dados [Oracle 2019c]

Oracle label security

O *oracle label security* também auxilia no controle de acesso, mas permitindo o controle de acessos multiníveis, requeridos por aplicações governamentais e militares [Oracle

2018b]. Ele é integrado ao *oracle enterprise manager* e está disponível junto com o banco de dados oracle – edição *enterprise*.

No controle de acesso multinível, tanto os dados quanto os usuários do banco de dados recebem uma classificação (*label* ou *classification*). A cada vez que um usuário tenta acessar um dado no banco, é verificado quais dados que o usuário possui acesso e o banco de dados só retorna aqueles dados que o usuário possui acesso. No exemplo da Figura 3.26, o usuário só possui acesso aos dados classificados como sensíveis (*sensitive*) do tipo *alpha* e *beta*. Sendo assim, somente duas linhas da tabela `locations` que são retornadas, mesmo a tabela possuindo cinco linhas (demais dados classificados como altamente sensíveis).

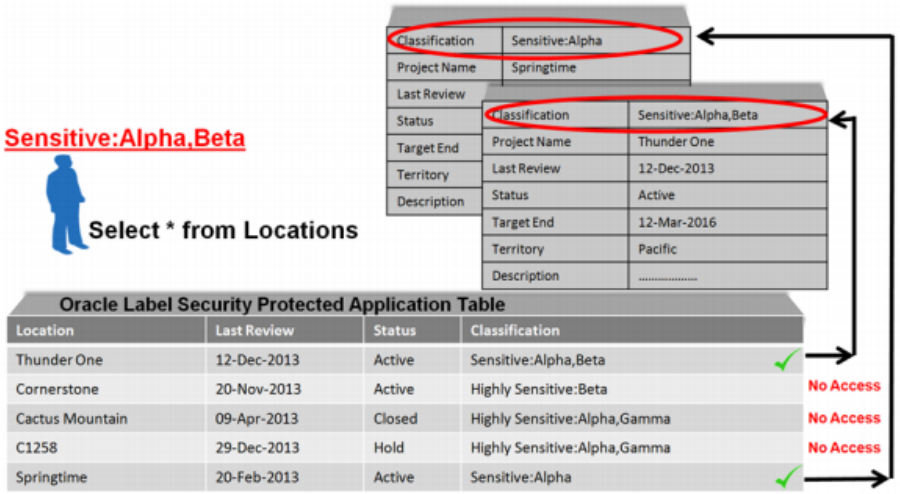

Figura 3.26. Oracle label security avaliando o acesso aos dados [Oracle 2018b]

3.4.2.3. Oracle - Categoria de Monitoramento/detecção

A LGPD determina que as organizações devem manter um registro de suas atividades de processamento. Esse registro só pode ser alcançado através do monitoramento e da auditoria constante das atividades sobre os dados pessoais. Esses dados de auditoria podem ser usados para notificar oportunamente as autoridades, em caso de violação. Além de exigir auditoria e alertas oportunos, a LGPD também exige que as organizações mantenham os registros de auditoria sob seu controle. Um controle centralizado dos registros de auditoria evita que invasores ou usuários mal-intencionados cubram os rastros de suas atividades suspeitas, excluindo registros da auditoria local [Rajasekharan 2017].

Oracle data safe

O *oracle data safe* auxilia na auditoria do banco de dados. A auditoria de atividades monitora as atividades dos usuários nos bancos de dados na nuvem, coletando e mantendo registros de auditoria por indústria e requisitos de conformidade regulamentar, acionando alertas para atividades não usuais. Por exemplo, a mudança em dados sensíveis pode ser auditada, caso ocorra falha no login de um administrador do banco de dados, pode ser gerado um alerta entre outros avisos possíveis. Na Figura 3.27 encontra-se um exemplo das atividades que podem ser monitoradas por esse serviço (ex.: quando o DBA desconecta do banco de dados (Event = LOGOFF); quando o DBA confirma uma transação (Event = COMMIT)).

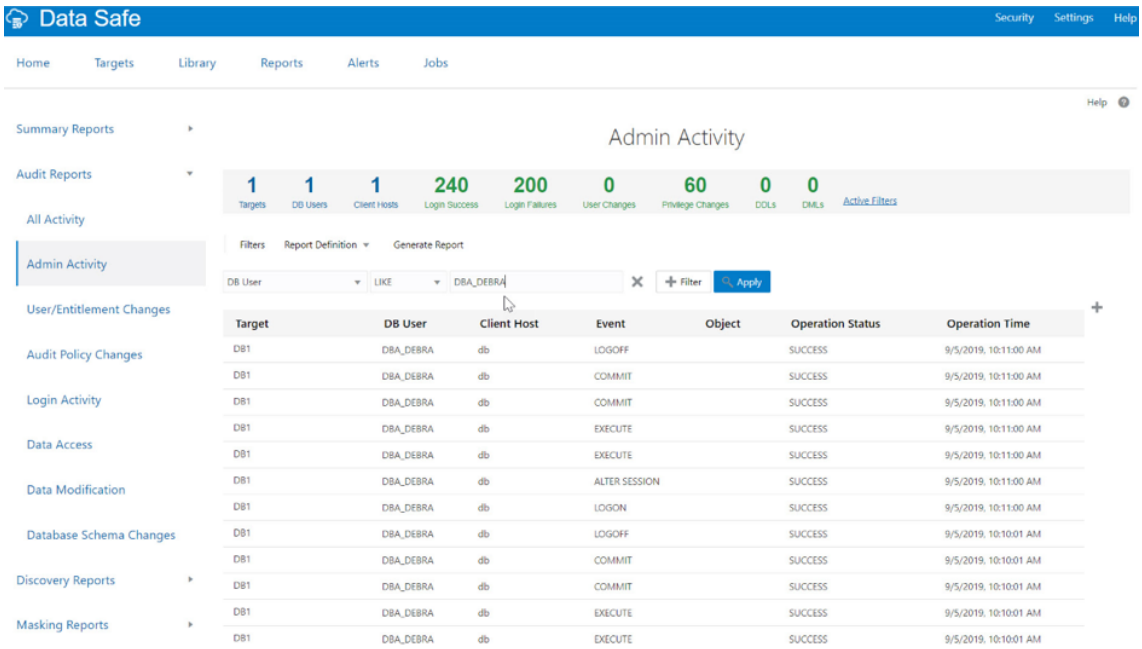

Figura 3.27. Exemplo de atividades monitoradas [Oracle 2019a]

Oracle audit vault e database firewall

O *oracle audit vault and database firewall* é uma plataforma de auditoria e proteção centrada em dados, que fornece monitoramento abrangente e flexível através de consolidação de dados de auditoria de bancos de dados Oracle e não Oracle, sistemas operacionais, sistemas de arquivos e aplicativos específicos [Rajasekharan 2017]. Ao mesmo tempo, o *oracle database firewall* pode atuar como a primeira linha de defesa na rede, impor o comportamento esperado do aplicativo, ajudando a impedir a injeção de SQL, o desvio do sistema e outras atividades que possam alcançar o banco de dados. O *oracle audit vault e o database firewall* podem consolidar os dados de auditoria de vários bancos de dados (Oracle, SQL Server, MySQL, DB2 entre outros) e monitora o tráfego SQL procurando, alertando e impedindo SQL não autorizado ou fora da política de segurança. Os responsáveis pela proteção de dados e os controladores podem especificar as condições sob as quais os alertas podem ser gerados em tempo real, tentando capturar os intrusos com as atividades anormais.

3.4.3. Amazon Web Services (AWS)

A AWS disponibiliza mais de 500 recursos e serviços com foco na segurança e compatibilidade. Segundo a categorização da LGPD, alguns dos serviços disponíveis são:

- Avaliação: *amazon macie* e *amazon inspector*.
- Prevenção: *AWS identity and access management (IAM)* e *aws key management service (KMS)*.
- Monitoramento/Detecção: *amazon guardduty e aws config*.

3.4.3.1. AWS - Categoria de Avaliação

O *amazon macie* é um serviço de segurança que usa aprendizado de máquina para descobrir, classificar e proteger, automaticamente, dados confidenciais na AWS [aws

2019a]. A Figura 3.28 apresenta um gráfico gerado a partir deste serviço sobre o comportamento de um usuário.

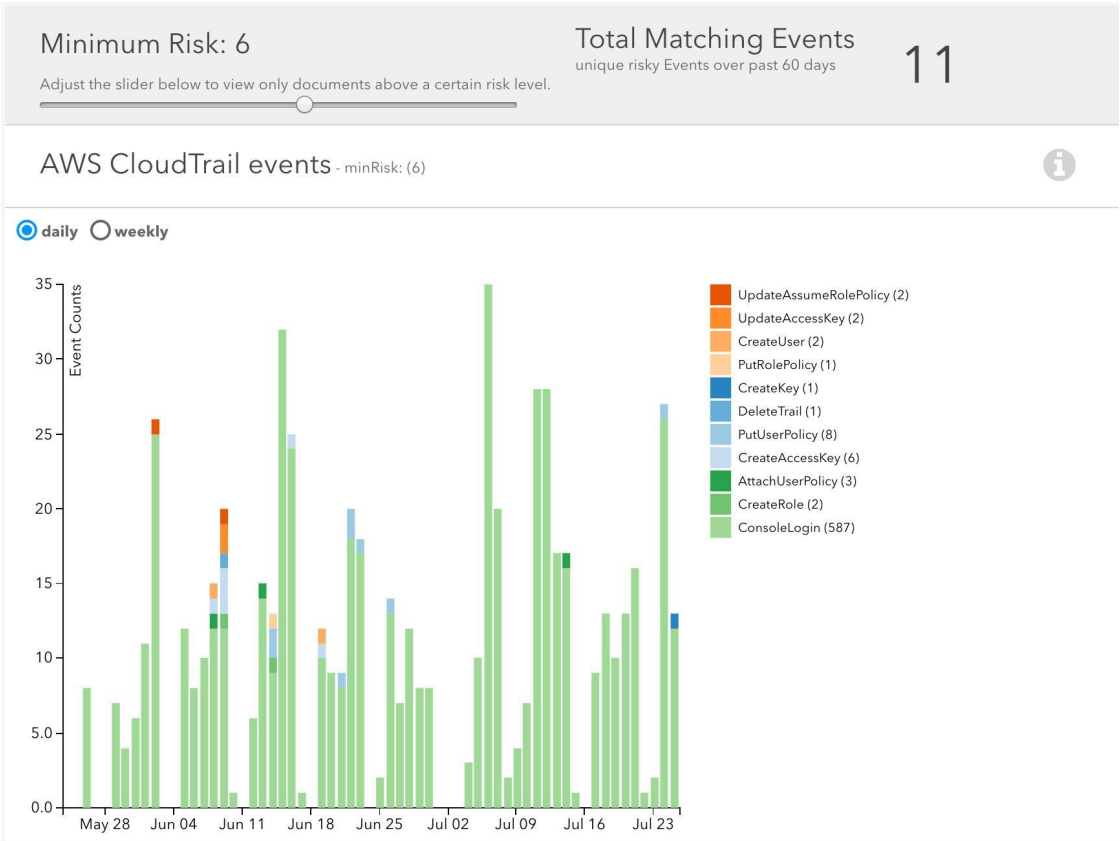

Figura 3.28. Análise do comportamento do usuário [aws 2019a]

O *amazon inspector* é um serviço de avaliação de segurança automático que ajuda a melhorar a segurança e a conformidade das aplicações implantadas na AWS [aws 2019b]. O *amazon inspector* avalia automaticamente as aplicações em busca de exposições, vulnerabilidades ou discrepâncias em relação às melhores práticas. Após realizar uma avaliação, o *amazon inspector* produz uma lista detalhada de descobertas de segurança priorizadas de acordo com o nível de severidade. A Figura 3.29 mostra um exemplo dessa lista. Caso o usuário selecione um dos itens que apresenta vulnerabilidade, é exibido um detalhamento sobre ele (Figura 3.30), com a descrição da vulnerabilidade (sujeito a ataques remotos por causa da função `bergetnext`) e a recomendação do serviço para o reparo da vulnerabilidade (atualizar o sistema operacional).

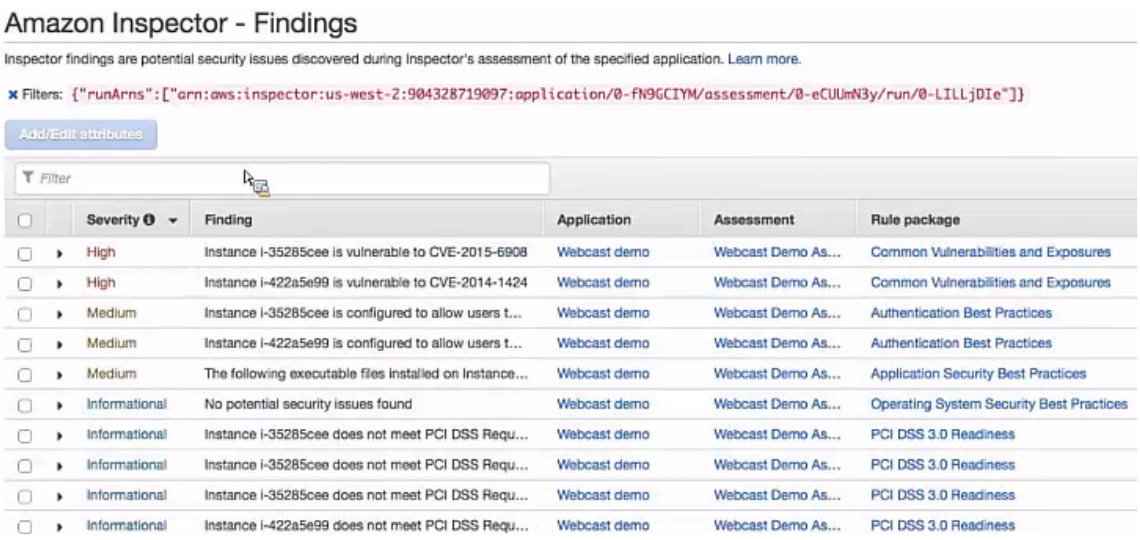

Figura 3.29. Lista de vulnerabilidades encontradas em uma aplicação [aws 2015]

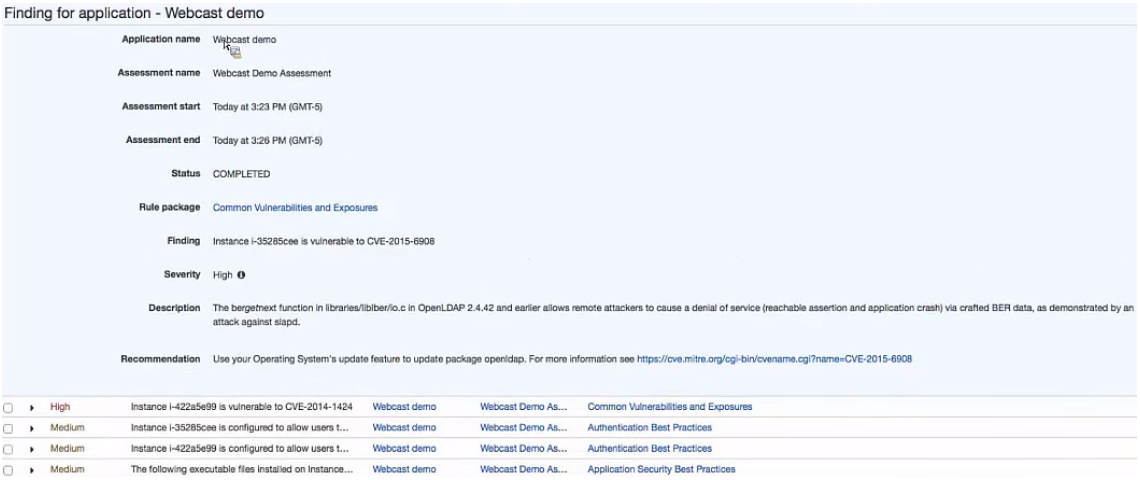

Figura 3.30. Detalhamento de uma das vulnerabilidades encontradas [aws 2015]

3.4.3.2. AWS - Categoria de Prevenção

O serviço *aws identity and access management (IAM)* permite a gerência, com segurança, do acesso aos serviços e recursos da AWS [aws 2019c]. Usando o IAM, pode-se criar e gerenciar usuários e grupos da AWS e usar permissões para conceder e negar acesso a recursos da AWS. A Figura 3.31 mostra um exemplo de atribuição de permissão ao usuário `testIAMuser`. O serviço permite adicionar o usuário a um grupo/papel pré-existente (ex.: `administrators`), copiar as permissões de um usuário existente ou atribuir políticas de segurança existentes de forma direta ao usuário.

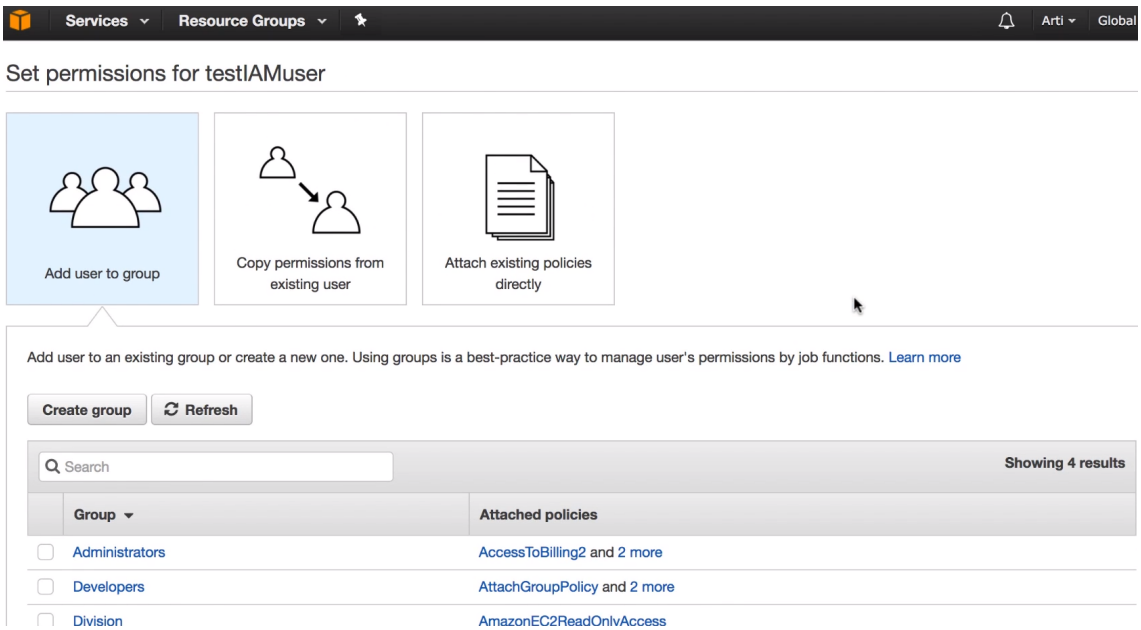

Figura 3.31. Exemplo de atribuição de permissão a um usuário [aws 2019c]

O serviço *aws key management service (KMS)* facilita a criação e o gerenciamento de chaves e o controle do uso de criptografia em uma ampla variedade de serviços da AWS e em seus aplicativos [aws 2019f]. A Figura 3.32 resume o comportamento desse serviço [stackoverflow 2018]. Existem dois tipos de chaves KMS: chaves mestras do cliente (CMKs) e chaves de dados (DKs). As chaves mestras do cliente nunca saem da infraestrutura da AWS e são geradas por chamada da API `CreateKey`. As chaves de dados são geradas por chamada da API `GenereateDataKey` que retorna uma versão "simples" e uma versão criptografada da chave. Essa criptografia é feita usando um CMK.

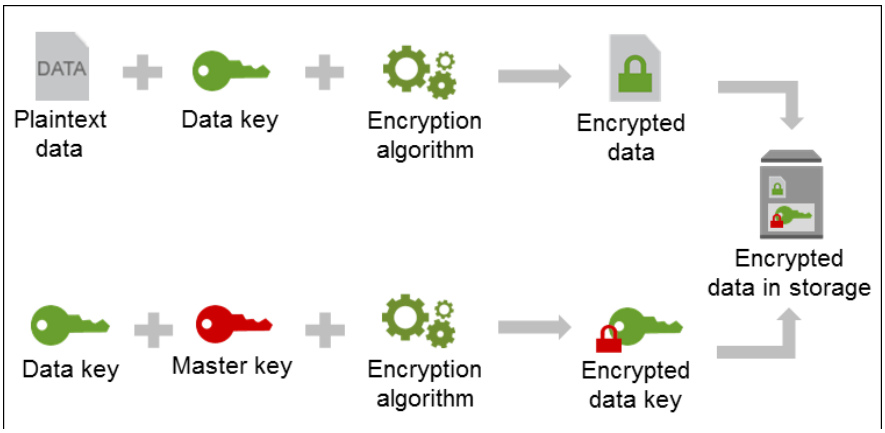

Figura 3.32. Visão geral do aws key management service [stackoverflow 2018]

3.4.3.3. AWS - Categoria de Monitoramento/Detecção

O *amazon guardduty* é um serviço de detecção de ameaças que monitora continuamente atividades mal-intencionadas ou e comportamentos não autorizados para proteger suas contas e cargas de trabalho da AWS [aws 2019d]. O serviço usa *machine learning*, detecção de anomalias e inteligência integrada contra ameaças para identificar e priorizar possíveis ameaças. A Figura 3.33 relata algumas ameaças encontradas e ordenadas por prioridade e a Figura 3.34 detalha uma das ameaças.

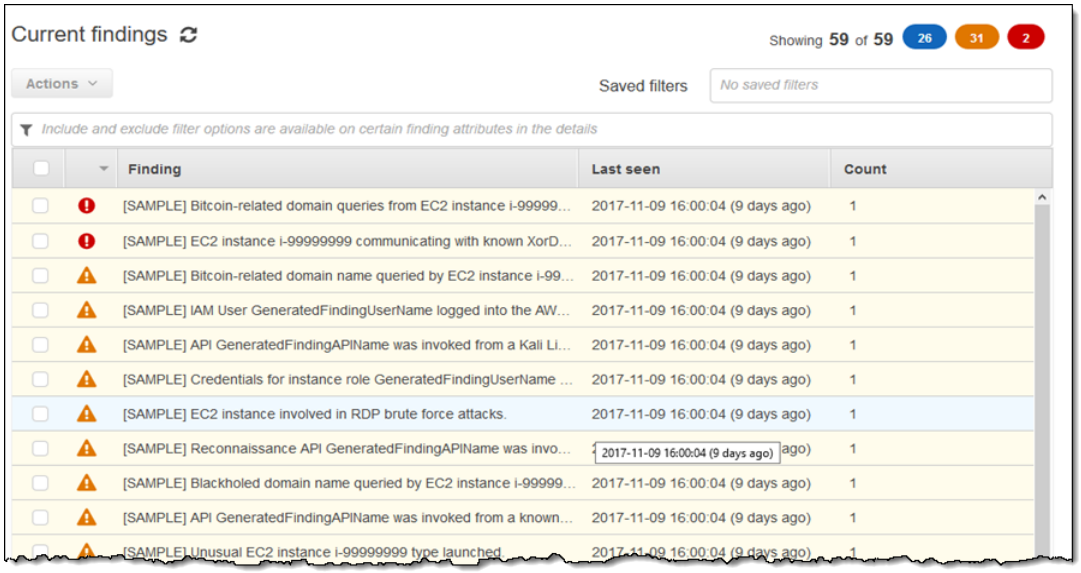

Figura 3.33. Ameaças detectadas [Barr 2017]

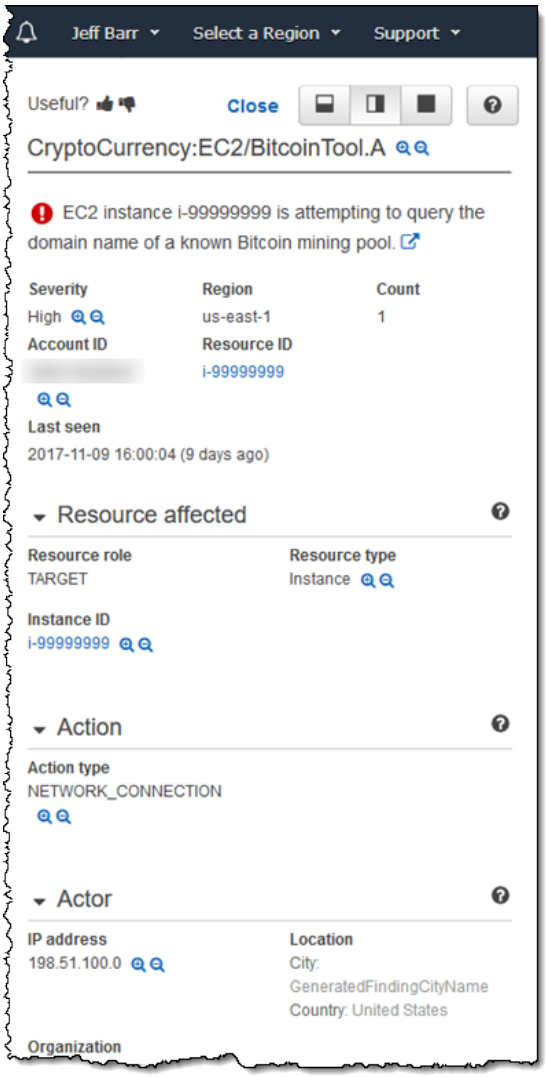

Figura 3.34. Detalhamento de uma das ameaças encontradas [Barr 2017]

O *aws config* é um serviço que permite acessar, auditar e avaliar as configurações dos recursos da AWS [aws 2019e]. O *aws config* monitora e grava continuamente os registros das configurações de recursos da AWS e lhe permite automatizar a avaliação das configurações registradas com base nas configurações desejadas. A Figura 3.35 demonstra uma visão geral da tela principal (*Dashboard*) do serviço *aws config*. Nessa tela, é possível (A) visualizar o número total de recursos que o *aws config* está registrando; (B) ver os tipos de recursos que o *aws config* está registrando, em ordem decrescente (por número de recursos). Caso o usuário selecione um tipo de recurso, o serviço abre a página de inventário de recursos; (C) escolher a exibição de todos os recursos também abre a página de inventário de recursos; (D) ver o número de regras não compatíveis; (e) ver o número de recursos não compatíveis; (f) ver as principais regras não compatíveis, em ordem decrescente (por número de regras); (G) escolher a exibição de todas as regras incompatíveis abre a página de regras.

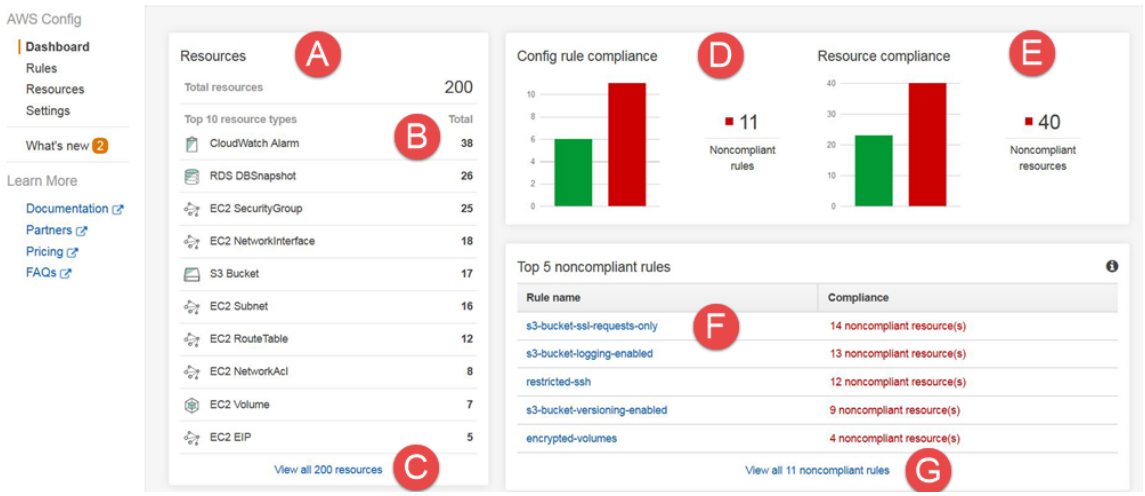

Figura 3.35. Dashboard do aws config [aws 2019e]

3.4.4. PostgreSQL

Existem algumas iniciativas de empresas e comunidades para auxiliar os usuários do SGBD relacional e gratuito PostgreSQL a se adaptarem às leis de proteção de privacidade. Entre as funcionalidades disponibilizadas, descreve-se neste capítulo, segundo as categorias da LGPD:

- Avaliação: até a escrita deste capítulo, não foram encontradas contribuições específicas dessa categoria.
- Prevenção: redação de dados e anonimização (*EnterpriseDB* e *PostgreSQL Anonymizer*); e criptografia de dados (*FUJITSU Enterprise Postgres* e *pgcrypto*).
- Monitoramento/Detecção: *pgAudit*.

3.4.4.1. PostgreSQL - Prevenção

Em termos de redação de dados, a empresa ***EnterpriseDB*** descreve uma forma, usando funções, visões, papéis e esquema padrão, para realizá-la de uma maneira que somente os usuários privilegiados consigam visualizar os dados no seu formato original [Linster 2018]. Os demais usuários, ao consultar os dados, visualizam de forma mascarada. Na Figura 3.36 tem-se um exemplo de função para mascarar a coluna `ssn` da tabela de

empregado (`employees`). Ela substitui todos os cinco primeiros números do `ssn` por 'x'. Adicionalmente, é incluído, ao final da função, o "`SECURITY DEFINER`" para especificar que a função deve ser executada com os privilégios do usuário que a possui.

```
CREATE OR REPLACE FUNCTION redact_ssn (ssn varchar(11))
RETURNS varchar(11)
/* substitui 020-12-9876 por xxx-xx-9876 */
AS
$$ SELECT overlay (ssn placing 'xxx-xx' FROM 1) ;$$
LANGUAGE SQL SECURITY DEFINER;
```

Figura 3.36. Função de redação ou mascaramento dos dados de ssn [Linster 2018]

Em seguida, é necessário criar uma visão da tabela de empregado (Figura 3.37) que chame a função definida anteriormente para a coluna `ssn` [Linster 2018]. Existem funções similares para as colunas de telefone (`phone`) e data de nascimento (`birthday`).

```
CREATE OR REPLACE VIEW redacteddata.employees
AS
SELECT
id,
name,
redact_ssn(ssn) ssn,
redact_phone(phone) phone,
redact_date(birthday) birthday
FROM employeedata.employees;
```

Figura 3.37. Exemplo de visão que realiza a chamada para as funções de mascaramento [Linster 2018]

Posteriormente, os usuários comuns obtêm acesso à visão criada e os usuários privilegiados obtêm acesso à tabela original. Além disso, o esquema padrão do papel do usuário comum passa a ser o esquema da visão (`redacteddata`) e o esquema padrão do papel do usuário privilegiado passa a ser o esquema dos dados originais (`employeedata`). A indicação dos esquemas é mostrada na Figura 3.38. Finalmente, os dados podem ser consultados pelos usuários comuns (Figura 3.39) e pelos usuários privilegiados (Figura 3.40).

```
ALTER ROLE redacteduser IN DATABASE mycompany SET search_path TO "$user", public, redacteddata;
ALTER ROLE privilegeduser IN DATABASE mycompany SET search path TO "$user", public, employeedata;
```

Figura 3.38. Alteração de esquema padrão para as papéis [Linster 2018]

```
SELECT * FROM employees;
id | name          | ssn         | phone      | birthday
---+---------------+-------------+------------+---------------------
 1 | Sally Sample  | xxx-xx-9345 | 5081234567 | 02-FEB-02 00:00:00
 2 | Jane Doe      | xxx-xx-9345 | 6171234567 | 14-FEB-02 00:00:00
 3 | Bill Foo      | xxx-xx-9345 | 9781234567 | 14-FEB-02 00:00:00
(3 rows)
```

Figura 3.39. Dados consultados por usuários comuns [Linster 2018]

```
SELECT * FROM employees;
id | name         | ssn         | phone      | birthday
---+--------------+-------------+------------+---------------------
 1 | Sally Sample | 020-78-9345 | 5081234567 | 02-FEB-61 00:00:00
 2 | Jane Doe     | 123-33-9345 | 6171234567 | 14-FEB-63 00:00:00
 3 | Bill Foo     | 123-89-9345 | 9781234567 | 14-FEB-63 00:00:00
(3 rows)
```

Figura 3.40. Dados consultados por usuários privilegiados [Linster 2018]

Outra forma de redação de dados confidenciais pode ser vista na extensão chamada *postgresql anonymizer* [Clochard 2018]. A Figura 3.41 mostra os dados originais. A Figura 3.42 mostra como a extensão pode ser criada e ativada no SGBD. Em seguida, na Figura 3.43, tem-se a criação de um papel para o usuário que verá os dados mascarados. Na Figura 3.44 apresenta-se a declaração das regras de mascaramento, onde o nome terá seus caracteres substituídos de forma randômica e o telefone só terá os 2 primeiros e os 2 últimos caracteres apresentados de forma real. Finalmente, na Figura 3.45, os dados são apresentados de forma mascarada.

```
SELECT * FROM people;
id   | name           | phone
-----+----------------+------------
T800 | Schwarzenegger | 0609110911
(1 row)
```

Figura 3.41. Dados originais [Clochard 2018]

```
CREATE EXTENSION IF NOT EXISTS anon CASCADE;
SELECT anon.mask_init();
```

Figura 3.42. Criação e ativação da extensão PostgreSQL Anonymizer [Clochard 2018]

```
CREATE ROLE skynet;
COMMENT ON ROLE skynet IS 'MASKED';
```

Figura 3.43. Criação do papel do usuário que verá os dados mascarados [Clochard 2018]

```
COMMENT ON COLUMN people.name IS 'MASKED WITH FUNCTION anon.random_last_name()';
COMMENT ON COLUMN people.phone IS 'MASKED WITH FUNCTION anon.partial(phone, 2, $$******$$, 2)';
```

Figura 3.44. Descrição do mascaramento das colunas de nome e telefone [Clochard 2018]

```
psql test -U skynet -c 'SELECT * FROM people;'
id   | name           | phone
-----+----------------+------------
T800 | Nunziata       | 06******11
(1 row)
```

Figura 3.45. Exibição de dados mascarados [Clochard 2018]

A criptografia de dados pode ser obtida através da aquisição da camada TDE proposta pela ***FUJITSU Enterprise Postgres*** [Downey 2019]. Essa camada não demanda alterações no SGBD e é disponibilizada de forma gratuita pela empresa. A Figura 3.46 mostra um exemplo da arquitetura com essa camada. Como pode ser visto na figura, os dados podem ser armazenados no banco de dados e no arquivo de backup de forma criptografada e somente os usuários autorizados que conseguem visualizar os dados originais. A criptografia pode ocorrer em nível de *tablespaces*, dados de backup, *log* de registro prévio de escrita (WAL), arquivos temporários e replicação de *streaming*.

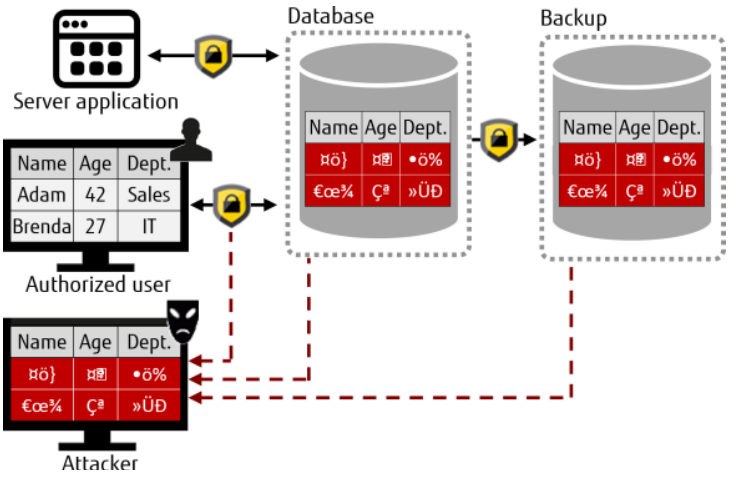

Figura 3.46. Arquitetura da camada TDE do FUJITSU Enterprise Postgres [Downey 2019]

Além disso, o PostgreSQL 10 possui algumas opções de criptografia, na própria camada do banco de dados, através do módulo *pgcrypto*, para: senha, colunas específicas, *storage* (nível de sistema e de bloco), dos dados durante o tráfego de rede, autenticação do host (cliente e servidor) por certificado e do lado do cliente [PostgreSQL 2019].

A Figura 3.47 mostra um exemplo de inserção de senha, usando as funções `crypt()` e `gen_salt()` para senhas *hashing*. A primeira função faz o *hashing* e a segunda prepara os parâmetros do algoritmo para ela.

```
INSERT INTO "login" (login, password, employee_id)
VALUES ('email', crypt('password', gen_salt('bf')));
```

Figura 3.47. Exemplo de criptografia de senha no banco de dados [stackoverflow 2013]

3.4.4.2. PostgreSQL - Monitoramento/Detecção

A auditoria de dados pode ser implementada através do módulo *pgAudit* [Riggs et al 2019]. Esse módulo é gratuito e provê um *log* detalhado de auditoria de objeto e/ou de sessão. Por exemplo, caso o(a) auditor(a) queira verificar qual a tabela que foi criada no período de janela de manutenção, ele(a) pode verificar o *log*, que informará os detalhes conforme demonstrado na Figura 3.48. É importante reparar que o usuário que criou a tabela tentou mascarar a execução do comando, ao incluí-lo em uma transação, com o uso do comando `EXECUTE`.

```
AUDIT: SESSION, 33, 1,FUNCTION, DO,,,"DO $$
BEGIN
    EXECUTE 'CREATE TABLE import' || 'ant_table (id INT)';
END $$."
AUDIT: SESSION, 33, 2,DDL, CREATE TABLE, TABLE, public.important table, CREATE TABLE important table(id INT)
```

Figura 3.48. Exemplo de registro no log de auditoria [Riggs et al 2019]

Na Figura 3.49 tem-se um exemplo de como o *log* de auditoria pode ser habilitado para todos os comandos DDL e de escrita (DML), descrevendo também as relações envolvidas nos comandos DML.

```
SET pgaudit.log = 'write, ddl';
SET pgaudit.log_relation = ON;
```

Figura 3.49. Comandos para habilitar o log de auditoria [Riggs et al 2019]

3.5. Conclusão

Este minicurso apresenta uma discussão sobre a LGPD e seus princípios, bem como os recursos disponíveis em alguns dos principais ambientes de bancos de dados existentes que podem auxiliar os controladores de dados sensíveis a manter o ambiente de banco de dados em conformidade com esta lei. Através das análises apresentadas no presente capítulo, percebe-se que os recursos para controle de acesso aos dados encontrados nos ambientes de Bancos de Dados analisados não são suficientes para contemplar os princípios da LGPD em relação a todas as questões de privacidade. Neste sentido é preciso também se preocupar com o sigilo, a integridade, o tempo de vida, o anonimato, o escopo de uso pela aplicação e a separação de funções dos dados.

Como já descrito por Elmasri e Navathe (2019), o avanço rápido do uso da tecnologia da informação (TI) na indústria, no governo e no meio acadêmico gera questões e problemas desafiadores com relação à proteção e ao uso de informações pessoais. Questões como quem e quais direitos à informação sobre indivíduos, para quais finalidades, tornam-se cada vez mais importantes à medida que seguimos para um mundo em que é tecnicamente possível conhecer quase tudo sobre qualquer um. Decidir como projetar considerações de privacidade na tecnologia para o futuro inclui dimensões filosóficas, legais e práticas.

Dessa forma, as organizações ainda se questionam qual o setor responsável em sua estrutura pela implantação e manutenção da LGPD, pela multidisciplinaridade inerente ao tema. Embora não tenha sido o foco do presente minicurso, essa questão ainda fica em aberto para futuras discussões.

Referências

Adithe, S., Singh, S. (2018) “Comprehensive Identity and Access Management in the Cloud”, SAPinsider, Volume 19, Issue 2, disponível em: https://sapinsider.wispubs.com/Assets/Articles/2018/May/Comprehensive-Identity-and-Access-Management-in-the-Cloud, acessado em outubro de 2019.

Amazon Web Services (AWS). (2019a) “Detalhes do Amazon Macie”, disponível em: https://aws.amazon.com/pt/macie/details/, acessado em outubro de 2019.

Amazon Web Services (AWS). (2019b) “Amazon Inspector - Guia do usuário”, disponível em: https://docs.aws.amazon.com/pt_br/inspector/latest/userguide/inspector_introduction.html, acessado em outubro de 2019.

Amazon Web Services (AWS). (2019c) “Conceitos básicos do AWS IAM”, disponível em: https://aws.amazon.com/pt/iam/getting-started/, acessado em outubro de 2019.

Amazon Web Services (AWS). (2019d) “Amazon GuardDuty”, disponível em: https://aws.amazon.com/pt/guardduty/, acessado em outubro de 2019.

Amazon Web Services (AWS). (2019e) “AWS Config – Guia do desenvolvedor”, disponível em: https://docs.aws.amazon.com/pt_br/config/latest/developerguide/viewing-the-aws-config-dashboard.html, acessado em outubro de 2019.

Amazon Web Services (AWS). (2019f) “Recursos do AWS Key Management Service”, disponível em: https://aws.amazon.com/pt/kms/features/, acessado em outubro de 2019.

Amazon Web Services (AWS). (2015) “2015 Webinar Series – AWS Inspector”, disponível em: https://www.youtube.com/watch?v=ddz0JmCTTsU, acessado em outubro de 2019.

Barr, J. (2017) “Amazon GuardDuty – Continuous Security Monitoring & Threat Detection”, disponível em: https://aws.amazon.com/pt/blogs/aws/amazon-guardduty-continuous-security-monitoring-threat-detection/, acessado em outubro de 2019.

Clochard, D. (2018) “Introducing PostgreSQL Anonymizer”, disponível em: https://blog.taadeem.net/english/2018/10/29/Introducing-PostgreSQL-Anonymizer, acessado em outubro de 2019.

Dean, B. (2017) “Privacy vs. Security”, disponível em: https://www.secureworks.com/blog/privacy-vs-security, acessado em outubro de 2019.

Downey, P. (2019) “Providing maximum data security with minimal impact to your business using transparent data encryption”, disponível em: https://www.postgresql.fastware.com/blog/transparent-data-encryption-tde, acessado em outubro de 2019.

Elmasri, R., Navathe, S. B. (2019) “Sistemas de Banco de Dados”, 7ª edição (versão traduzida), editora Pearson Universidades.

Feinberg, D., Heudecker, N., Adrian, M. (2018) “Magic Quadrant for Operational Database Management Systems”, In: Gartner Research, ID: G00346575, disponível em: https://www.gartner.com/en/documents/3891967, acessado em outubro de 2019.

Granet, E. (2018) “Static Data Masking for Azure SQL Database and SQL Server”, disponível em: https://azure.microsoft.com/pt-br/blog/static-data-masking-preview/, acessado em outubro de 2019.

Linster, M. (2018) “Creating a Data Redaction Capability to Meet GDPR Requirements Using EDB Postgres”, disponível em: https://www.enterprisedb.com/blog/creating-data-redaction-capability-meet-gdpr-requirements-using-edb-postgres, acessado em outubro de 2019.

Mahajan, G. (2019) “SQL Server Static Data Masking Example”, disponível em : https://www.mssqltips.com/sqlservertip/5939/sql-server-static-data-masking-example/, acessado em outubro de 2019.

Microsoft. (2019a) “Descoberta e classificação de dados SQL”, disponível em: https://docs.microsoft.com/pt-br/sql/relational-databases/security/sql-data-discovery-and-classification?view=sql-server-2017, acessado em outubro de 2019.

Microsoft. (2019b) “Segurança em nível de linha’, disponível em: https://docs.microsoft.com/pt-br/sql/relational-databases/security/row-level-security?view=sql-server-2017, acessado em outubro de 2019.

Microsoft. (2019c) “Mascaramento de dados dinâmicos”, disponível em: https://docs.microsoft.com/pt-br/sql/relational-databases/security/dynamic-data-masking?view=sql-server-ver15, acessado em outubro de 2019.

Microsoft (2019d) “Habilitar conexões criptografadas com o Mecanismo de Banco de Dados”, disponível em: https://docs.microsoft.com/pt-br/sql/database-

engine/configure-windows/enable-encrypted-connections-to-the-database-engine?view=sql-server-ver15, acessado em outubro de 2019.

Microsoft (2019e) "Criptografia de Dados Transparente (TDE)", disponível em: https://docs.microsoft.com/pt-br/sql/relational-databases/security/encryption/transparent-data-encryption?view=sql-server-ver15, acessado em outubro de 2019.

Microsoft. (2018a) "SQL Server and Azure SQL Database GDPR Guidance", disponível em: https://azurepartnerportal.blob.core.windows.net/media/Resources/SQL%20Server%20GDPR%20Guidance%20Paper.pdf, acessado em outubro de 2019.

Microsoft. (2018b) "Autenticação no SQL Server", disponível em: https://docs.microsoft.com/pt-br/dotnet/framework/data/adonet/sql/authentication-in-sql-server, acessado em outubro de 2019.

Microsoft. (2017a) "Server and Database Roles in SQL Server", disponível em: https://docs.microsoft.com/en-us/dotnet/framework/data/adonet/sql/server-and-database-roles-in-sql-server, acessado em outubro de 2019.

Microsoft. (2017b) "Create Database Audit Specification (Transact-SQL)", disponível em: https://docs.microsoft.com/pt-br/sql/t-sql/statements/create-database-audit-specification-transact-sql?view=sql-server-ver15, acessado em outubro de 2019.

Microsoft. (2017c) "Sempre criptografados (mecanismo de banco de dados)", disponível em: https://docs.microsoft.com/pt-br/sql/relational-databases/security/encryption/always-encrypted-database-engine?view=sql-server-ver15, acessado em outubro de 2019.

Microsoft. (2017d) "Avaliação de Vulnerabilidades SQL", disponível em: https://docs.microsoft.com/pt-br/sql/relational-databases/security/sql-vulnerability-assessment?view=sql-server-ver15, acessado em outubro de 2019.

Microsoft. (2016a) "Auditoria do SQL Server (Mecanismo de Banco de Dados)", disponível em: https://docs.microsoft.com/pt-br/sql/relational-databases/security/auditing/sql-server-audit-database-engine?view=sql-server-ver15, acessado em outubro de 2019.

Microsoft. (2016b) "Tabelas temporais", disponível em: https://docs.microsoft.com/pt-br/sql/relational-databases/tables/temporal-tables?view=sql-server-ver15, acessado em outubro de 2019.

Microsoft (2014) "SQL Server – Object Level Permissions details", disponível em: https://gallery.technet.microsoft.com/scriptcenter/SQL-Server-Object-Level-fc2f1cb6, acessado em outubro de 2019.

Oracle. (2019a) "Secure Critical Data with Oracle Data Safe – Improve the Security of Cloud Databases with a Unified Control Center for Managing Sensitive Data", White Paper disponível em: https://www.oracle.com/a/tech/docs/dbsec/data-safe/wp-security-data-safe.pdf, acessado em outubro de 2019.

Oracle. (2019b) "Encryption and Redaction with Oracle Advanced Security", White Paper disponível em: https://www.oracle.com/a/tech/docs/dbsec/aso/advanced-security-wp-19c.pdf, acessado em outubro de 2019.

Oracle. (2019c) “Oracle Database Vault”, White Paper disponível em: https://www.oracle.com/a/tech/docs/dbsec/dbv/wp-dv-19c.pdf, acessado em outubro de 2019.

Oracle. (2018a) “Managing Oracle Database Encryption Keys in Oracle Cloud Infrastructure with Oracle Key Vault”, White Paper disponível em: https://docs.cloud.oracle.com/iaas/Content/Resources/Assets/whitepapers/manage-encryption-keys-oci-okv.pdf, acessado em outubro de 2019.

Oracle (2018b) “Oracle Label Security”, White Paper disponível em: https://www.oracle.com/technetwork/wp-dbsec-ols-201702-3634252.pdf, acessado em outubro de 2019.

Oracle. (2017) “Data Masking and Subsetting Guide”, disponível em: https://docs.oracle.com/en/database/oracle/oracle-database/12.2/dmksb/index.html, acessado em outubro de 2019.

PostgreSQL. (2019) “PostgreSQL 10 - Encryption Options”, disponível em: https://www.postgresql.org/docs/10/encryption-options.html, acessado em outubro de 2019.

Rajasekharan, D. (2017) “Accelerate Your Response to the EU General Data Protection Regulation (GDPR)”, Oracle White Paper, disponível em: https://www.oracle.com/technetwork/database/security/wp-security-dbsec-gdpr-3073228.pdf, acessado em outubro de 2019.

Riggs, S., Menon-Sem, A., Barwick, I. (2019) “pgAudit Open Source PostgreSQL Audit Logging”, disponível em: https://github.com/pgaudit/pgaudit/blob/master/README.md, acessado em outubro de 2019.

Stackoverflow. (2018) “Key Management Services”, disponível em: https://stackoverflow.com/questions/47904805/key-management-services, acessado em outubro de 2019.

Stackoverflow. (2013) “How do I encrypt passwords with PostgreSQL?”, disponível em: https://stackoverflow.com/questions/18656528/how-do-i-encrypt-passwords-with-postgresql, acessado em outubro de 2019.

Teixeira, B., Schwabe, D., Santoro, F., Baião, F., Luiza Campos, M., Verona, L., Laufer, C., Barbosa, S., Lifschitz, S., Costa, R. (2019). Privacy and Transparency within the 4IR: Two faces of the same coin. In Companion Proceedings of The 2019 World Wide Web Conference (pp. 581-593). ACM.

Tutorialspoint. (2019) “SAP GRC Tutorial – SAP GRC - Overview”, disponível em: https://www.tutorialspoint.com/sap_grc/sap_grc_overview.htm, acessado em outubro de 2019.

Autores

Ana Carolina Brito de Almeida (Instrutora) é professora adjunta da Universidade do Estado do Rio de Janeiro (UERJ), alocada no Departamento de Informática e analista judiciário com especialidade em Informática no Tribunal Regional Federal da 2ª Região (TRF2), atuando como DBA. Realizou pós-doutorado com ênfase em *Big Data* na Universidade Federal do Rio de Janeiro (UFRJ) (Out/2014 a Abr/2015). Doutora em Informática com especialização em *Tuning* de BD pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) (2013). Mestre em Sistemas e Computação pelo Instituto Militar de Engenharia (IME/RJ) (2006). Pesquisadora na área de BD com ênfase em: (i) sistemas de (auto) sintonia de BD e (ii) ontologias. Já foi consultora em BD e ministrou cursos na Universidade Petrobras. Detalhes em http://lattes.cnpq.br/8306729029606464.

Letícia Dias Verona (Instrutora) possui mestrado em Informática pela UFRJ (2018) e graduação em Ciência da Computação pela UFRJ. Possui experiência com desenvolvimento de sistemas web, gestão de projetos e equipes. Pesquisadora em: integração de informações heterogêneas, análise de redes políticas e econômicas, metadados, ontologias, modelagem conceitual, BD e web semântica. Atualmente participa da coordenação do grupo de pesquisa de Dados Abertos, vinculado ao grupo GRECO (PPGI/UFRJ), cursando o doutorado em Gestão de Sistemas Complexos.
Detalhes em http://lattes.cnpq.br/2165808131875029.

Maria Luiza Machado Campos é professora no Departamento de Ciência da Computação da UFRJ, e uma das coordenadoras do grupo de pesquisas GRECO, atuando como pesquisadora e orientadora de mestrado e doutorado no PPGI da mesma universidade. Possui graduação em Engenharia Civil pela Universidade Federal do Rio Grande do Sul (1978), mestrado em Engenharia de Sistemas e Computação pela COPPE/UFRJ (1984) e doutorado em *Information Systems - University Of East Anglia* (1993), Inglaterra. Em 2015, realizou Pós-doutorado no *Laboratory of Applied Ontology*, CNRS, Italia. Foi coordenadora do Bacharelado em Ciência da Computação e do Programa de Pós-graduação em Informática, assim como Diretora Adjunta de Extensão do Instituto de Matemática da UFRJ. Participou de diversos projetos de desenvolvimento, pesquisa e extensão, assim como de numerosas orientações de trabalhos de conclusão de curso, dissertações de mestrado e de doutorado ao longo de mais de 30 anos de carreira. Seus principais temas de pesquisa estão associados à integração de informações heterogêneas, abordando principalmente os seguintes temas: metadados, ontologias, modelagem conceitual, banco de dados, *data warehousing* e web semântica. Detalhes em http://lattes.cnpq.br/0659658820912418.

Fernanda Araujo Baião é Professora no Departamento de Engenharia Industrial da PUC-Rio. Seus temas de pesquisa são nas áreas de Ciência de Dados, Modelagem Conceitual e Ontologias, Gestão de Processos de Negócio (BPM) e Integração de Dados Distribuídos através de Alinhamento de Ontologias, particularmente investigando a interação entre Ciências Cognitivas e Processamento de Linguagem Natural com BPM e Gestão de Dados, assim como o desenvolvimento de sistemas de Alinhamento de Ontologias para apoio à Integração de Dados sobre ambientes distribuídos de larga escala. De outubro de 2018 a Agosto de 2019 atuou como Cientista de Dados chefe em uma iniciativa na Caixa Econômica Federal para Prevenção de Fraudes no Programa Seguro desemprego. De 2004 a 2018 foi Professora da Universidade Federal do Estado do Rio de Janeiro. De 2001 a 2004 atuou como post-doc na COPPE/UFRJ, onde obteve seu título de Doutorado (2001) e de Mestrado (1997) em Engenharia de Sistemas e Computação. No ano de 2000 foi pesquisadora visitante na *University of Wisconsin-Madison* (USA). É autora de mais de 150 publicações com revisões por pares, muitas em colaboração com pesquisadores renomados na comunidade científica internacional resultantes da sua participação em projetos de pesquisa nacionais e internacionais, como o *Brazilian Institute of WebScience Research*, financiado pelo CNPq, e o Rise-BPM (rise-bpm.eu), financiado pela União Europeia. Coordena ou já coordenou projetos de pesquisa financiados pelo CNPq e FAPERJ. Participa e coordena diversos comitês de programa e e de editoração de conferências e periódicos nacionais e internacionais, como a Applied Ontology Journal, ER, BPM, dentre outras. Desenvolveu expertise valioso em projetos de transferência de conhecimento entre a Academia e a Indústria, e atuou como líder técnico em projetos de P&D sobre Data Science, BPM, Arquitetura Empresarial, Gestão de Dados e Segurança da Informação, em domínios de Exploração e Produção de Óleo e Gás, Seguros, Gestão de Serviços de TI e Predição de Fraudes. Detalhes em `http://lattes.cnpq.br/5068302552861597`.