LEI GERAL DE PROTEÇÃO DE DADOS E A APLICABILIDADE NA ANONIMIZAÇÃO

Fabíola Gonçalves Barreto[1], Fabricio Gustavo Henrique[1]

[1]Faculdade de Tecnologia de FATEC Ribeirão Preto (FATEC)

fabiola.barreto3@fatec.sp.gov.br,
fabricio.henrique@fatec.sp.gov.br

Resumo. *Diante do crescimento cada vez mais acelerado das tecnologias, surgem os questionamentos sobre a segurança das informações corporativas e de seus clientes. Criada em 2018 e em vigor desde 2020, ainda se discute muito sobre a adaptabilidade da tecnologia à Lei Geral de Proteção de Dados. O presente estudo de pesquisa bibliográfica tem como objetivo, reunir através de pesquisas, informações que englobam sobre a Lei Geral de Proteção de Dados e a necessidade da aplicação de técnicas de proteção de dados, como a de anonimização, que está isenta da incidência normativa.*

Abstract*. Given the increasingly accelerated growth of technologies, questions arise about the security of corporate information and its customers. Created in 2018 and in force since 2020, there is still much debate about the adaptability of technology to the General Data Protection Law. This bibliographical research study aims, to gather, through research, information that encompasses the General Data Protection Law and the need to apply data protection techniques, such as anonymization, which is exempt from normative incidence.*

1 INTRODUÇÃO

Com o advento da internet nos últimos anos é possível acessar redes sociais, acompanhar notícias, fazer compras, ver vídeos, fazer pagamentos, entre outras atividades, que tornam as pessoas cada vez mais dependentes dela. Porém essas tarefas têm um preço que muitas vezes passa despercebido pelos usuários: a coleta dos seus dados pessoais.

Dados pessoais são qualquer tipo de informação que permite identificar o indivíduo, por exemplo: nome, CPF, gênero, preferências de lazer, gosto musical, dentre outros.

> Art. 5º Para os fins desta Lei, considera-se:
> I - dado pessoal: informação relacionada a pessoa natural identificada ou identificável;

Os dados pessoais se tornaram muito valiosos, no mundo dos negócios e é comparado ao petróleo, a diferença é que os dados são fontes rentáveis infinitas. Um estudo feito pela Oracle e o Instituto de Tecnologia de Massachusetts (MIT), indicou que *startups* estão em voga atualmente, porque muitas delas tratam a informação do usuário como um produto, vendendo para outras empresas e em consequência reduzindo seus custos de operação.

Em 2013 houve um escândalo envolvendo o governo dos Estados Unidos, Edward Snowden, revelou documentos ultrassecretos que evidenciaram um amplo projeto de

espionagem pelos Estados Unidos, o monitoramento envolvia até países da Europa e América latina, incluindo o Brasil. Em reação a isso, a União Europeia lançou uma lei a proteção de dados, o Regulamento Geral de Proteção de Dados (GPDR[1], 2016). Essa nova lei que entrou em vigor em maio de 2018, influenciaria as relações com a União Europeia.

Países que não tivessem uma legislação aplicada à proteção de dados, seriam considerados menos confiáveis em negociações comerciais e de respeito aos direitos e garantias fundamentais. Publicada em 2018 no Brasil, a Lei Geral de Proteção de Dados entrou em vigor somente em agosto de 2021 – com multas aplicáveis. A norma tem como objetivo garantir ao titular de dados, o controle sobre quais dados estão sendo tratados pelas empresas e como estão sendo utilizados.

Em um artigo publicado por Weverton Guedes em agosto de 2020, ele menciona que apesar das possíveis sansões ao descumprimento da lei, existe um lado positivo, as novas regulações forçam as organizações a criarem ordem sobre os dados que elas possuem, ao passo que esta ordem habilita as organizações a obterem *insights* que antes estavam escondidos em informações desorganizadas. Por consequência, além de garantir a proteção de direitos e liberdades fundamentais, se bem-encarada, gera fomento ao desenvolvimento econômico e tecnológico.

1.1. A Lei Geral de Proteção de Dados e a demanda profissional

Para que o sistema de proteção de dados do titular funcione dentro da empresa é necessário a atuação do controlador, operador e o encarregado, cargos esses criados pela Lei Geral de Proteção de dados (LGPD, 2018).

O controlador fica responsável pela tomada de decisão em relação aos dados pessoais, além das diretrizes a serem seguidas pelo operador, este com responsabilidade em supervisionar e fiscalizar o cumprimento das regras e atuação técnica no tratamento de dados. O encarregado agirá como canal de comunicação entre a empresa e o titular dos dados, além da Autoridade Nacional de Proteção de Dados [2](ANPD).

Por ter facilidade em interpretar as leis, os profissionais de advocacia vêm ganhando mercado para o cargo de encarregado, onde há a expectativa de gerar 50.000 de vagas para advogados orientados a Tecnologia da informação, segundo a OAB/SP.[3]

Em uma pesquisa da Robert Half foi mostrada uma expectativa de crescimento para analista de negócios, encarregado de proteção de dados, gestor de projetos e profissional de conformidade (*compliance*).[4]

2 METODOLOGIA

No presente estudo foi utilizado o método de pesquisa bibliográfica com a finalidade de definir e analisar a Lei Geral de Proteção de Dados (LGPD, 2018), apresentando seus principais pontos. Adiante foi exposto o conceito e aplicabilidade dos tipos de

[1] GPDR abreviação da nomenclatura da lei em inglês *General Data Protection Regulation*
[2] A ANPD é o órgão responsável pela fiscalização e aplicação da Lei Geral de Proteção de Dados
[3] Disponível em:<https://www.migalhas.com.br/quentes/348927/lgpd-pode-gerar-abertura-de-novos-postos-de-trabalho-no-mundo-juridico> Acesso em: 10 out, 2021.
[4] Disponível em: < https://www.dgabc.com.br/Noticia/3293041/lgpd-x-demanda-por-profissionais> Acesso em: 10 out, 2021

anonimização, a fim de entender de que maneiras satisfazem os requisitos da Lei Geral de Proteção de Dados (LGPD, 2018), tal como não ser possível a identificação do usuário. Para o estudo foram selecionados artigos de conceituados autores da área, tendo como base sites confiáveis acerca do assunto, além de outros meios de pesquisa julgados necessários.

3 CONCEITO E PRINCÍPIOS DA LEI GERAL DE PROTEÇÃO DE DADOS

A Lei Geral de Proteção de Dados Pessoais (LGPD, 2018) no seu Art. 1º, dispõe sobre o tratamento de dados pessoais, seja em hospitais, bancos, escolas ou meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural. A lei teve como base a Regulação Geral de Base de Dados da União Europeia, considerada uma das maiores referência em lei de proteção de dados[5]. É regulamentada e fiscalizada pela Autoridade Nacional de Proteção dos Dados (ANPD).

Independentemente da localização da sede da organização ou centro de dados, seja no Brasil ou no exterior, se houver o processamento de conteúdo de pessoas que estão no território brasileiro, a lei deve ser cumprida.

É possível observar, que a lei segue o escopo PII (*Personally Identifiable Information*)[6], termo usado em segurança da informação, que se refere as informações que podem ser usadas para identificar, contatar ou localizar uma pessoa, não se limitando a nome, idade, endereço residencial ou e-mail, por exemplo, pode incluir dados acadêmicos, *cookies*, histórico de compras ou médico, entre outros.

Para fins da lei, no Art. 5º considera dado pessoal qualquer informação relacionada à pessoa natural identificada ou identificável e dado pessoal sensível - este exige uma atenção maior, qualquer dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou à 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.

O tratamento dos dados sensíveis deve ocorrer quando houver o consentimento explícito da pessoa e para um fim definido ou no caso sem o consentimento do titular, nas hipóteses que este for indispensável em situações como: a uma obrigação legal; a políticas públicas; a estudos via órgão de pesquisa; a um direito, em contrato ou processo; à preservação da vida e da integridade física de uma pessoa; à tutela de procedimentos feitos por profissionais das áreas da saúde ou sanitária; à prevenção de fraudes contra o titular.

Algo interessante de comentar, seria o uso da nomenclatura “pessoa natural” - o ser humano capaz de direitos e obrigações na esfera civil - em vez de “pessoa física” - sujeito do direito, ente único, do qual e para o qual decorrem normas - na lei, com o avanço da tecnologia, cada vez mais robôs com inteligência artificial são criados, a tendência é em

[5] Disponível em: <https://www.compugraf.com.br/diferencas-entre-lgpd-e-gdpr/>. Acesso em: 05 jun, 2021

[6] Disponível em: < https://digitalks.com.br/artigos/lgpd-e-marketing-guia-pratico-para-o-profissional-de-marketing-digital/ >. Acesso em: 06 jun, 2021

breve ter "pessoas digitais", como a Sophia, primeiro robô a ter cidadania oficial em um país, inclusive possui mais direitos que muitas pessoas na Arábia Saudita, país onde a sua cidadania foi concedida[7].

3.1 Casos que a LGPD não se aplica

Existem casos que a LGPD não se aplica, conforme descrito no Art. 4º, quando realizado por pessoa para fins exclusivamente não econômicos e particulares, e realizados para fins exclusivamente jornalísticos, artísticos ou acadêmicos, onde se aplicam à hipótese os Arts. 7º e 11º desta lei. Não se aplica também para fins de segurança pública ou do Estado, a defesa nacional, como investigações, repressão de infrações penais, ou dados provenientes de fora do território nacional e que não sejam objetos de comunicação, uso compartilhado dos dados entre agentes de tratamentos brasileiros e internacionais, desde que o país de origem utilize o grau de proteção adequado ao previsto por lei.

Os dados jurídicos, como CNPJ, telefones, endereço, desde que não seja identificada uma pessoa também não é abrangido pela Lei Geral de Proteção de Dados.

3.2 Princípios da LGPD

No Art 6º a Lei Geral de Proteção de Dados é determinada por dez princípios que as empresas devem considerar ao tratar dados pessoais, lembrando que tratamentos de dados é qualquer tipo de atividade realizada com os dados pessoais.

Os princípios a serem observados são explicados abaixo.

3.2.1 Finalidade

Sobre este princípio a LGPD (2018) diz: "realização do tratamento para propósitos legítimos, específicos, explícitos e informados ao titular, sem possibilidade de tratamento posterior de forma incompatível com essas finalidades", ou seja, as empresas precisam deixar claras as suas intenções para o titular dos dados, justificando e apontando o uso dos dados pessoais.

Por exemplo, se um endereço de e-mail é coletado com a finalidade exclusiva de enviar um boleto bancário para o cliente, a empresa não pode utilizar este e-mail para enviar ofertas e promoções;

3.2.2 Adequação

A LGPD (2018) descreve o princípio da adequação como "compatibilidade do tratamento com as finalidades informadas ao titular, de acordo com o contexto do tratamento", dessa forma as empresas precisam justificar e garantir que os dados coletados tenham valor e sejam condizentes com o modelo de negócio da organização.

Exemplificando, em um *site* de compras *on-line* de uma farmácia, é solicitado no cadastro informações sobre a orientação sexual, a coleta deste tipo de dado não é compatível com o tipo de negócio do *site*, sendo passível de punição e multa;

---

[7] LGPD: Entendendo o básico da lei. Carlos Eduardo Pelosi. 2021, 60 min. Disponível em: <https://www.udemy.com/course/lgpd-entendendo-o-basico-da-lei/learn/lecture/25744878#overview> Acesso em:10 jun, 2021

3.2.3 Necessidade

O princípio da necessidade leva em consideração a responsabilidade das empresas acerca dos dados tratados. Ou seja, quanto mais dados pessoais tratados, maior é a responsabilidade e, por consequência, maior é a cobrança e as multas em casos de erros e falhas;

A LGPD (2018) afirma que o princípio da necessidade envolve "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";

3.2.4 Livre acesso

O livre acesso é a "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", como está descrito na lei.

As empresas devem criar mecanismos para que o titular dos dados tenha o direito de consultar os seus próprios dados e informações de forma gratuita, deixar evidente os seus objetivos e o período que os dados serão utilizados;

3.2.5 Qualidade dos dados

Refere-se à "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";

3.2.6 Transparência

O princípio da transparência, de acordo com a lei, é a "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";

3.2.7 Segurança

O princípio da segurança envolve a adoção de procedimentos, tecnologias e soluções que garantam maior proteção dos dados pessoais em casos de acessos não autorizados, como em ataques hackers, por exemplo.

A lei diz: "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";

3.2.8 Prevenção

O princípio da prevenção determina a "adoção de medidas para prevenir a ocorrência de danos em virtude do tratamento de dados pessoais";

3.2.9 Não discriminação

De acordo com a LGPD (2018), refere-se à "impossibilidade de realização do tratamento para fins discriminatórios ilícitos ou abusivos", sendo assim, o tratamento de dados pessoais, como os sensíveis, jamais podem ser utilizados com objetivos de discriminar ou de promover abusos contra os seus titulares;

3.2.10 Responsabilização e Prestação de Contas

Sobre o princípio de segurança, a LGPD (2018) diz: “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”.

3.3 Consequências de não cumprimentos

No Art. 52 da legislação é estabelecido as sanções administrativas aplicáveis aos agentes de tratamento de dados caso descumpra a lei que só serão aplicadas após análise do órgão responsável, possibilitando oportunidade de defesa:

a) Advertência com prazo indicado para adoção de medidas corretivas (Inciso I);
b) Multa de até 2% do faturamento da empresa, podendo ser diário e se limitando a cinquenta milhões de reais por infração (Incisos II e III);
c) Mostrar publicamente o problema que causou (Inciso IV), este as vezes poderá custar mais que a aplicação da multa, uma vez que os usuários sentirão menos confiança na empresa;
d) Bloqueio dos dados ou eliminação a que se refere a infração (Incisos V e VI);
e) Suspensão parcial do funcionamento do banco de dados pelo período máximo de seis meses ou mais caso não regularizado (Inciso X);
f) Suspensão ou proibição (parcial ou total) do exercício de atividades relacionadas aos tratamentos de dados, no caso da suspensão, pode ser no período de até seis meses, podendo ser prorrogável (Inciso XI e XII).

4 LGPD E SEGURANÇA DA INFORMAÇÃO

As vezes usados como sinônimos até por profissionais da tecnologia, Segurança da Informação e Proteção de Dados apresentam conceitos diferentes. Segurança da Informação é a proteção contra diversos tipos de ameaças, protegendo qualquer forma de dados (físicos ou eletrônicos) e qualquer tipo de dados (pessoais ou não pessoais), enquanto Proteção de Dados é o direito à privacidade ao processamento de dados pessoais.

Na perspectiva de Hintzbergen et al. (2018) Segurança da Informação possui três princípios: confidencialidade, integridade e disponibilidade que são os pilares para a preservação da informação.

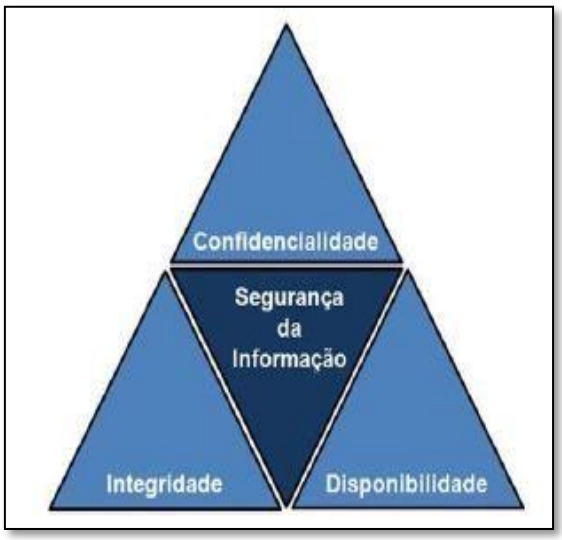

Figura 1 - Segurança da Informação: princípios básicos
Fonte: Lopes, 2017

Além desses pilares, existem outros que podem ser incluídos como: autenticidade, responsabilidade, não repúdio e confiabilidade. Na Figura1 estão ilustrados os principais pilares.

Não é possível garantir a conformidade de proteção de dados sem considerar a Segurança da Informação, existe Segurança sem privacidade, mas não existe privacidade sem segurança.

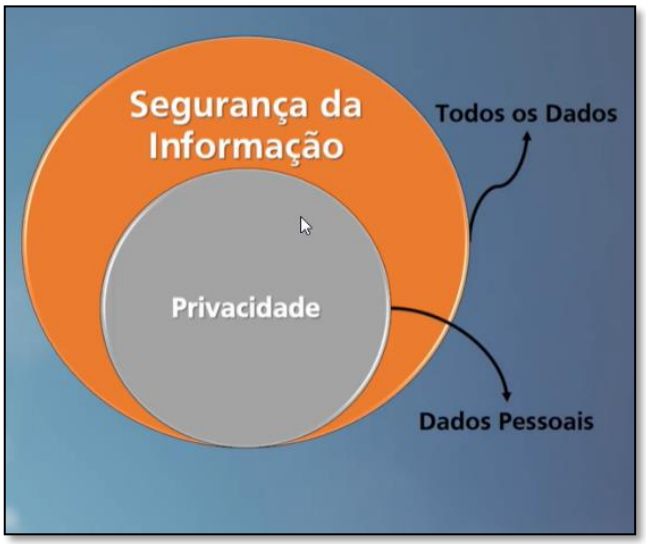

Figura 2 - Segurança da Informação e Privacidade
Fonte: Fernando Martins Silva, 2021

5 ANONIMIZAÇÃO DE DADOS

5.1 Banco de dados

Antigamente as empresas armazenavam dados em papéis, que por sua vez, eram organizados em arquivos físicos, como as pastas. Com o surgimento e evolução dos computadores, essa forma de extrair informações e mantê-los organizados, foram ficando para trás.

Na década de 60 os computadores se tornaram uma parte importante nas empresas, junto a capacidade de armazenar dados, os arquivos digitais começaram a ficar complexos e manipulá-los através de arquivos do sistema operacional não era mais adequado, já que existia situações que relacionar entidades era algo trivial. Surge então na empresa IBM os fundamentos de banco de dados relacionais, atualmente o mais utilizado no mercado. Junto com a evolução do armazenamento de dados, surgiu o termo Sistemas Gerenciadores de Banco de Dados (SGBD), que é uma coleção de programas que permitem o usuário definir, construir e manipular bases de dados de acordo com a necessidade.

Atualmente os principais SGBDs relacionais do mercado mundial são MySql, Oracle, SQL Server e PostgreSQL[8]. Todos compartilham da linguagem SQL[9].

5.2 Conceito de anonimização

Anonimização (Também conhecida como data masking ou data sanitization)[10] é uma técnica de processamento de dados que remove ou altera informações onde seja possível identificar uma pessoa, resultando em dados que não podem ser associados a nenhum indivíduo específico, logo sem incidência na LGPD (2018).

Doutorando em Direito Comercial e um dos pesquisadores [11]que participou do processo de elaboração da Lei Geral de Proteção de Dados (Lei nº 13.709/2018), Bruno Bioni descreve anonimização da seguinte forma:

> Diante do próprio significado do termo, anônimo seria aquele que não tem nome nem rosto. Essa inaptidão pode ser fruto de um processo pelo qual é quebrado o vínculo entre o(s) dado(s) e seu(s) respectivo(s) titular(es), o que é chamado de anonimização. Esse processo pode se valer de diferentes técnicas que buscam eliminar tais elementos identificadores de uma base de dados.
> (BIONI, 2020)

Na anonimização não existe a possibilidade de reversão, se existir alguma identificação, então não é considerado um dado anonimizado – apenas um dado pseudonimizado,[12] estando sujeito a LGPD (2018).

> Art. 5º Para os fins desta Lei, considera-se:
> III - dado anonimizado: dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento;
> XI - anonimização: utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo;

[8] Diponivel em: < https://www.opservices.com.br/banco-de-dados/>. Acesso em: 03 abr, 2021
[9] SQL: *Structured Query Language*
[10] Disponível em: < https://www.totvs.com/blog/negocios/anonimizacao/ >. Acesso em: 10 jun, 2021.
[11] Disponível em: < https://teletime.com.br/11/03/2021/data-privacy-brasil-indica-o-pesquisador-bruno-bioni-para-conselho-consultivo-da-anpd/>. Acesso em: 16 nov, 2021.
[12] Disponível em: <https://www.serpro.gov.br/lgpd/menu/protecao-de-dados/dados-anonimizados-lgpd>. Acesso em: 10 jun, 2021.

5.3 Tipos de anonimização

A anonimização de dados é comum para fins de estudos, como pesquisas com análise de dados estatísticos feito por organizações, por exemplo.

Através de técnicas, um conjunto de dados original é transformado em um novo conjunto, por meio de alterações. As principais são: supressão, generalização e randomização, que serão exemplificadas com a tabela a seguir.

Nome	Data Nascimento	CEP	Sexo	Fumante
Elvis	15/01/1980	14160-010	Masculino	Sim
Freddie	22/12/1996	14150-000	Masculino	Nao
Tina	01/02/1992	14160-025	Feminino	Sim

Tabela 1: Representação de dados simulados
Fonte: autoria própria

5.3.1 Supressão

Nesta técnica os valores em um conjunto de dados [13]são removidos.

Sua aplicabilidade é indicada quando o atributo não puder ser anonimizado com outra técnica ou os atributos não são necessários, deve ser aplicada no início do processo de anonimização, diminuindo a identificação do titular.[14]

Sexo	Fumante
Masculino	Sim
Masculino	Nao
Feminino	Sim

Tabela 2: Representação de dados simulados suprimidos
Fonte: autoria própria

No exemplo acima, foi removido as colunas Nome, Data Nascimento e CEP, sendo possível uma estática de fumantes entre homens e mulheres.

Para Ciriani et al. (2007), a supressão pode ocorrer de três maneiras:

- No registro: onde, um registro é removido inteiramente do conjunto de dados;
- No atributo: todos os valores de um atributo em uma coluna são suprimidos;
- Em células individuais: apenas células de um determinado registro são suprimidas, caracterizando uma supressão local.

[13] No banco de dados é referido como coluna.

[14] Guia Básico de Técnicas de Anonimização da Comissão de Proteção de Dados da Singapura. Disponível em: < https://iapp.org/media/pdf/resource_center/Guide_to_Anonymisation.pdf>. Acesso em: 17 jul, 2021.

5.3.2 Generalização

Na anonimização por generalização, os dados são substituídos por categorias mais amplas e genéricas, tendo o devido cuidado de não categorizar de forma estreita e consequentemente facilite a identificação do titular, são mais bem aproveitadas para dados estatísticos que possam ser agrupados.

Nome	Data Nascimento	CEP	Sexo	Fumante
Elvis	15/01/1980	14160	Masculino	Sim
Freddie	22/12/1996	14150	Masculino	Nao
Tina	01/02/1992	14160	Feminino	Sim

Tabela 3: Representação de dados simulados generalizados

Fonte: autoria própria

Foram mantidos os 5 (cinco) primeiros dígitos, que são capazes de identificar cidade, região, bairro e subsetor e retirados os 3 (três) últimos que identifica a rua.
Para CIRIANI et al. (2007) a generalização pode ser aplicada em relação a:

a) Atributo: a generalização ocorre nos valores da coluna;
b) Células: a generalização é realizada em células individuais, como resultado, a tabela pode conter em uma coluna específica, valores distintos nos atributos.

A partir da generalização é possível alcançar o k-anonimato, l-diversidade e t-proximidade:

a) k-anonimato

Técnica criada por Latanya Sweeney e Pierangela Samarati, refere-se ao processo no qual diversos registros são agrupados em um único, onde pelo menos $k - 1$ registros compartilharão os mesmos valores para todos os seus atributos, garantindo a privacidade deste registro (indivíduo), k representa o tamanho do grupo.

Nome	Idade	CEP	Sexo	Fumante
Elvis	33	14060-010	Masculino	Sim
Freddie	25	14150-000	Masculino	Nao
Tina	29	14160-025	Feminino	Sim
Aretha	40	14070-210	Feminino	Sim
Ritchie	45	14010-120	Masculino	Sim
Marie	33	14012-123	Feminino	Nao
Per	25	14125-000	Masculino	Sim
Axl	28	14102-252	Masculino	Nao

Idade	CEP	Fumante
21-30	141**-***	Sim
21-30	141**-***	Nao
21-30	141**-***	Sim
21-30	14**-***	Nao
31-40	140**-***	Sim
31-40	140**-***	Sim
31-40	140**-***	Não
31-40	140**-***	Sim

Tabela 4: Dados generalizados com aplicação de k-anonimato, conjunto de 4 anônimos

Fonte: autoria própria

Apesar de não haver uma definição clara sobre o nível aceitável de k, vários artigos argumentam que um nível de k = 5 ou k = 10 é o preferido. A maioria dos indivíduos na esfera acadêmica parece concordar que k = 2 é insuficiente e k = 3 é o mínimo necessário para preservar a privacidade.[15]

b) *l*-diversidade e *t*-proximidade

As técnicas de generalização por *l*-diversidade são definidas por Daniel Versoza em sua monografia, como importantes para os casos em que a recorrência de registros, permitam a inferência de determinada situação, mesmo que anonimizados com garantia do *k*-anonimato.
Exemplificando[16]:

> Imagine que um grupo de pessoas tenha pesquisado o mesmo tópico de saúde (por exemplo, sintomas da gripe), todas ao mesmo tempo. Se analisarmos esse conjunto de dados, não conseguiremos dizer quem pesquisou o tópico, graças ao k-anonimato. No entanto, ainda poderá haver alguma preocupação em relação à privacidade, uma vez que todos compartilham do mesmo atributo de confidencialidade (ou seja, o tópico da pesquisa). Com a l-diversidade, o conjunto de dados anonimizados não incluiria apenas pesquisas sobre a gripe, mas poderia incluir também outras pesquisas para proteger ainda mais a privacidade do usuário.

A *t*-proximidade surge a partir da *l*-diversidade, que consiste na criação de classes equivalentes de registros que garantam a distribuição de valores de forma próxima à distribuição da base de dados original. [17]

5.3.3 Mascaramento de caracteres

É uma técnica de mascaramento onde é feito alteração dos caracteres de um valor de dados por um símbolo de constante – por exemplo "*" (asterisco) ou "x" (xis). Normalmente é aplicado apenas em alguns caracteres no atributo.

Por meio de algum ruído ao dado, os valores originais são substituídos por outros fictícios, sem afetar análises estatísticas e impossibilitando a identificação do titular.

Nome	Data Nascimento	CEP	Sexo	Fumante
Elvis	15/01/1980	14160-xxx	Masculino	Sim
Freddie	22/12/1996	14150-xxx	Masculino	Nao
Tina	01/02/1992	14160-xxx	Feminino	Sim

Tabela 5: Dados com a aplicação de ruídos
Fonte: autoria própria

[15] Disponível em: < https://ichi.pro/pt/privacidade-de-dados-na-era-do-big-data-213872290125064 >. Acesso em: 16 nov, 2021.
[16] Disponível em: <https://policies.google.com/technologies/anonymization?hl=pt-BR>. Acesso em: 16 nov, 2021.
[17] Disponível em: <https://revista.internetlab.org.br/dados-nao-pessoais-a-retorica-da-anonimizacao-no-enfrentamento-a-covid-19-e-o-privacywashing/>. Acesso em 16 nov, 2021.

5.4 A complexidade da anonimização

Há quem defende que apenas a mera possibilidade de um dado ser atrelado a uma pessoa, não seja o suficiente para classificá-lo como identificável, como defendido por Bruno Bioni (2019, p. 76), "se para a correlação entre um dado e uma pessoa demanda-se um esforço fora do razoável, não há que se falar em dados pessoais. Nessa situação o dado é considerado como anônimo".

Por outro lado, existem os que sustentam que essa possibilidade por si só o torna identificado e/ou identificável.

Cada vez mais são publicados estudos que demonstram o processo de anonimização como algo falível, tornando um mito a garantia de 100% (cem por cento) de eficiência do anonimato das pessoas.

Dentre esses estudos está o da pesquisadora Latanya Sweeney, feito em 2000 nos Estados Unidos, mostrando que a anonimização é um processo complexo e que nem sempre garante o anonimato do usuário. No seu experimento, ela cruzou uma base de registros médicos com uma de dados eleitorais, e com essa junção conseguiu identificar vários usuários, concluiu-se que 87% (oitenta e sete por cento) dos americanos tem características relatadas que os tornam únicos somente com um código *ZIP*[18], gênero e data de nascimento.[19]

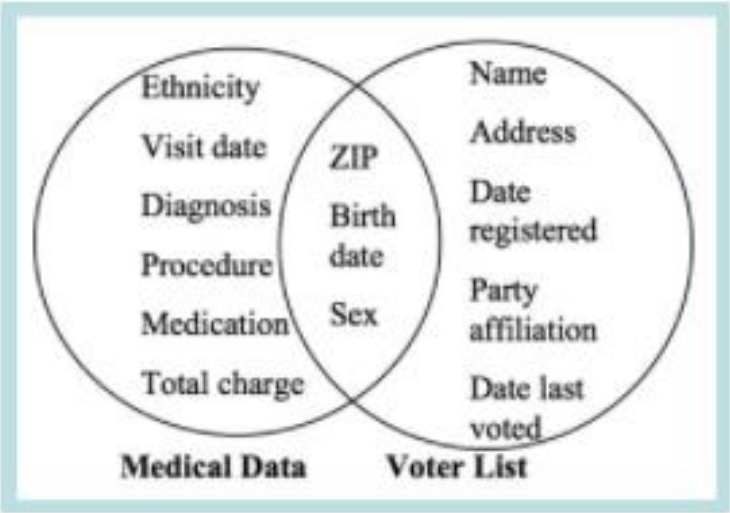

Figura 3 - Segurança da Informação e Privacidade
Fonte: Latanya Sweeney, Ph.D., 2001

Moisés Simões, Mestre em Engenharia de Software destaca:

> Hoje em dia, com o volume, dinamismo, complexidade e velocidade de *Big Datas* cada vez mais expandidos em progressão crescente, basta uma ferramenta analítica mais sofisticada e de alto desempenho para pescar e interpretar qualquer tipo de dado, criando correlações a partir das quais é possível identificar pessoas com uma precisão cada vez mais eficiente.[20]

[18] *ZIP* é um código de endereçamento equivalente ao CEP brasileiro.
[19] Disponível em: < http://latanyasweeney.org/work/identifiability.html >. Acesso em: 17 jul, 2021.
[20] Disponível em:< https://www.serpro.gov.br/lgpd/noticias/2019/anonimizacao-pseudonimizacao-dados-suficientes-adequar-lgpd>. Acesso em: 16 nov, 2021

Paul Ohm reforça o apresentado por Latanya Sweeney: "A partir do momento em que um adversário linkou dois bancos de dados anonimizados, ele pode adicionar os novos dados linkados para sua coleção de informações externas e usá-los para destravar outros bancos de dados anonimizados".[21]

Por se tratar de uma lei estratégica (Lei 13.709/2018)[22], não é indicado como dever ser feito a anonimização, fazendo ressalvas de trabalhar com um filtro de razoabilidade, conforme descrito no artigo 12:

> Art. 12. Os dados anonimizados não serão considerados dados pessoais para os fins desta Lei, salvo quando o processo de anonimização ao qual foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser revertido.
> § 1º. A determinação do que seja razoável deve levar em consideração fatores objetivos, tais como custo e tempo necessários para reverter o processo de anonimização, de acordo com as tecnologias disponíveis, e a utilização exclusiva de meios próprios.
> § 2º. Poderão ser igualmente considerados como dados pessoais, para os fins desta Lei, aqueles utilizados para formação do perfil comportamental de determinada pessoa natural, se identificada.

A dificuldade está na aferição do que é ou não um "esforço razoável" nessa definição da potencialidade de que um dado seja vinculado a uma pessoa, podendo ser estabelecido regras mais específicas para caracterizar a anonimização pela Autoridade Nacional de Proteção de Dados – Como base no parágrafo § 3ºdo Art. 12:

> § 3º. A autoridade nacional poderá dispor sobre padrões e técnicas utilizados em processos de anonimização e realizar verificações acerca de sua segurança, ouvido o Conselho Nacional de Proteção de Dados Pessoais.

6 CONCLUSÃO

Com a Lei Geral de Proteção de Dados, várias áreas e processos de uma empresa sofrerão impactos e as organizações deverão se adequar a esta nova Lei. É necessário que o controlador e operador tenham ciência da importância do seu papel dentro da organização, mantendo a segurança das informações, dados de colaboradores, clientes e demais dados pessoais sensíveis coletadas pela empresa.

O presente trabalho buscou construir uma análise a respeito dos principais conceitos da Lei Geral de Proteção de Dados, além disso, apresentando os principais pontos sobre anonimização, que apresenta maior garantia de privacidade aos titulares dos dados. Por conseguinte, a apresentação das técnicas de anonimização e exemplos de sua aplicação, que podem ser das mais radicais, como a supressão, até as técnicas de aleatorização, como a de mascaramento de dados. Dividindo opiniões quando sua aplicabilidade é válida ou não, a anonimização é atualmente a técnica mais indicada, sempre considerando o escopo, contexto e fins de processamento, uma vez que a forma que a anonimização for aplicada, terá influência de forma direta na probabilidade de riscos de re-identificação.

---

[21] OHM, 2010 apud MOURA Jose Luiz de. MAGALHÃES Guilherme. **PROTEÇÃO DE DADOS E ANONIMIZAÇÃO: PERSPECTIVAS À LUZ DA LEI Nº 13.709/2018**. Documento eletrônico. Disponível em: <https://www.estudosinstitucionais.com/REI/article/download/476/681>. Acesso em 16 nov, 2021.

[22] ANONIMIZAÇÃO & LGPD, Prof. Davis Alves, Ph.D. 2021. 180min.Transmitido ao vivo em 15 de setembro pela Ead TI Exames. Link não disponível.

7 REFERÊNCIAS

ANPD e o desafio de regulamentar a LGPD. 24 de janeiro de 2021. Disponível em: https://www.lickslegal.com/post/anpd-e-o-desafio-de-regulamentar-a-lgpd. Acesso em: 18/08/2021.

ALVES, Daniel Versoza. 2021. **Técnicas de anonimização de dados pessoais e a lei n. 13.709/2018**. Disponível em: https://acervodigital.ufpr.br/handle/1884/71173. Acesso em: 01/11/2021.

ARAKAKI, Ingrid Luize Bonadiman. LGPD e a anonimização de dados pessoais. 2020. Disponível em: https://www.migalhas.com.br/depeso/337227/lgpd-e-a-anonimizacao-de-dados-pessoais. Acesso em: 10/06/2021.

BIONI, Bruno. **Dados "anônimos" como antítese de dados pessoais: o filtro da razoabilidade**. 2019. Disponível em: http://genjuridico.com.br/2019/10/11/dados-anonimos-antitese-dados-pessoais/ Acesso em: 10/06/2021.

BRASIL. Lei nº 13.709, DE 14 DE AGOSTO DE 2018. **Diário Oficial da União**, Poder Executivo, Brasília, DF, 15 ago 2021. Seção 1, p. 59.
Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 07/03/2021.

FERREIRA, Tamires. LGPD: qual a diferença entre dados pessoais, sensíveis e anonimizados?. Disponível em: https://olhardigital.com.br/2021/08/17/tira-duvidas/lgpd-qual-a-diferenca-entre-dados-pessoais-sensiveis-e-anonimizados/. 02 de dezembro de 2021. Acesso em: 18/08/2021.

HINTZBERGEN, Jule et al. **Fundamentos de Segurança da Informação com base na ISO 27001 e na ISO 27002.** 1 ed. Rio de Janeiro: Brasport, 2018.

HORMAZABAL, Renato. SEGURANÇA da Informação x Privacidade dos Dados x Proteção de Dados. 12 de julho de 2021. Disponível em: https://hypeflame.blog/2020/12/07/seguranca-da-informacao-x-privacidade-dos-dados-x-protecao-de-dados/. 2020. Acesso em: 10/06/2021.

LOPES, Petter Vilela. PSI – Política de Segurança da Informação. 25 de janeiro de 2017. Disponível em:
https://periciacomputacional.com/psi-politica-de-seguranca-da-informacao/. Acesso em: 07/03/2021.

RAMEZ ELMASRI E SHAMKANT B. NAVATHE. **Sistemas de banco de dados** / revisor técnico Luis Ricardo de Figueiredo. Ed. São Paulo : Pearson Addison Wesley, 2005.

RODRIGUES, Juciana. **Anonimização como forma de proteção de dados**. 31 de janeiro de 2020. Disponível em: https://abracd.org/anonimizacao-como-forma-de-protecao-de-dados/. Acesso em: 18/08/2021.