A LEI GERAL DE PROTEÇÃO DE DADOS BRASILEIRA (Lei n.º 13.709/2018) NOÇÕES LEGAIS E TECNOLÓGICAS PARA ANONIMIZAÇÃO, PSEUDONIMIZAÇÃO E ELIMINAÇÃO DE DADOS DIGITAIS

ANDERSON CAMPOS DA COSTA

Dissertação para obtenção do Grau de Mestre em

Segurança de Informação e Direito no Ciberespaço

Orientação:

Professor Doutor Carlos Manuel Costa Lourenço Caleiro e

Professora Doutora Raquel Alexandra Jesus G. M. Brízida Castro

Júri:

Presidente: Professor Doutor Paulo Alexandre Carreira Mateus

Vogal: Professora Doutora Ana Isabel Barceló Caldeira Fouto

Vogal: Professora Doutora Raquel Alexandra Jesus G. M. Brízida Castro

Lisboa, julho de 2021.

AGRADECIMENTOS

Agradeço aos meus orientadores Professor Doutor Carlos Caleiro e a Professora Doutora Raquel Castro pelos conselhos e partilha de conhecimento.

Muito agradeço meus pais Francisco e Neiva, e a minha filha Maria Clara pelo incentivo para o retorno à academia.

Também agradeço aos meus amigos e sócios, Daisy Noroefé e Sigisfredo Hoepers pelo apoio pessoal e profissional nesta empreitada.

Aos meus colegas advogados Jodelismarko Mamoré e Cecilia Retamoso, pela disponibilidade e opiniões, registro meu agradecimento.

Por fim, agradeço e dedico este trabalho a minha querida esposa Marta Falcão, pois, sem seu apoio incondicional, paciência e carinho não seria possível esta conquista.

RESUMO

As leis devem ter a capacidade de transmitir a seus destinatários, de forma clara e precisa, os seus direitos e deveres. Nem sempre isso sucede e não raras as vezes encontramos pessoas perdidas na complexidade da teia legislativa[1].

A Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018) estabelece a necessidade de processos de anonimização, pseudonimização e eliminação de dados como indispensáveis para proteção e privacidade dos dados pessoais. Entretanto, não indica quais técnicas podem ser utilizadas para este fim.

Por esse motivo, a presente dissertação tem o propósito de realizar uma análise descomplicada da origem, fundamentos e relacionamentos legais, bem como, tratar de forma clara e objetiva, métodos e técnicas de anonimização, pseudonimização e eliminação de dados digitais alinhados à Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018).

Nessa perspectiva, inicialmente, investiga-se as origens históricas da Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018), bem como seu inter-relacionamento direto com o ordenamento jurídico pátrio, em seguida aborda-se a questão do Direito à Privacidade e Direito à Autodeterminação Informacional, para, na sequência, discorrer sobre o arcabouço legal vinculado aos processos de anonimização, pseudonimização e eliminação de dados.

Sob a ótica tecnológica investiga-se a aplicabilidade e funcionalidade de métodos e técnicas de anonimização e pseudonimização, largamente difundidos na Europa e Estados Unidos, e plenamente recebidos na comunidade brasileira.

Por fim, este trabalho propõem um modelo para a execução do procedimento de eliminação de dados digitais através da aplicação de criptografia homomórfica, perfeitamente alinhada à Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018).

[1] MATIAS MAGALHÃES, Filipa; LEITÃO PEREIRA, Maria. Regulamento Geral de Proteção de Dados, Manual Prático (2.ª Edição revista e ampliada), Porto: Vida Económica, 2018. p.78

PALAVRAS-CHAVE: Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018); Direito à Privacidade e Direito à Autodeterminação Informacional; anonimização; pseudonimização; eliminação de dados; criptografia e homomórfica.

ABSTRACT

Laws must be able to convey to their recipients, clearly and precisely, their rights and duties. This does not always happen and it is uncommon to find people lost in the complexity of the legislative web.

The Brazilian General Data Protection Law (Law No. 13.709/2018) establishes the need for anonymization, pseudonymization and data deletion processes as essential for the protection and privacy of personal data. However, it does not indicate which techniques can been used for this purpose.

For this reason, this dissertation aims to carry out an uncomplicated analysis of the origin, fundamentals and legal relationships, as well as to deal in a clear and objective way, methods and techniques of anonymization, pseudonymization and elimination of digital data in line with the General Law of Brazilian Data Protection (Law No. 13.709/2018).

From this perspective, initially, the historical origins of the Brazilian General Data Protection Law (Law No. 13.709/2018) is investigated, as well as its direct interrelationship with the national legal system, then the issue of Right to Privacy and Right to Informational Self-Determination, to, subsequently, discuss the legal framework linked to the processes of anonymization, pseudonymization and deletion of data established in the General Law for the Protection of Brazilian Data (Law No. 13.709/2018).

From a technological perspective, the applicability and functionality of anonymization and pseudonymization methods and techniques, widely disseminated in Europe and the United States, and fully received by the Brazilian community, is been investigated.

Finally, this work proposes a model for the execution of the procedure for deleting digital data through the application of homomorphic cryptography, perfectly aligned with the General Law for the Protection of Brazilian Data (Law No. 13.709/2018).

.

KEYWORDS: Brazilian General Data Protection Law (Law No. 13.709/2018); Right to Privacy and Right to Informational Self-Determination; anonymization; pseudonymization; data erasure; cryptography and homomorphic.

.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

ANPD – Autoridade Nacional de Proteção de Dados
BACEN – Banco Central
BCR – *Binding Corporate Rules*
CC – Código Civil
CCPA – *California Consumer Privacy Act*
CDC – Código de Defesa do Consumidor
CEE – Comunidade Econômica Europeia
CEP – Código de Endereçamento Postal
CF – Constituição Federal
DIN – Instituto Alemão de Normalização
DSS - *Data Security Standard*
EMD - *Earth Mover's Distance*
EPD – Encarregado de Proteção de Dados
FIPE – Fundação Instituto de Pesquisas Econômicas
GDPR – *General Data Protection Regulation*
IBGE – Instituto Brasileiro de Geografia e Estatística
INPI – Instituto Nacional da Propriedade Industrial
ISO – *International Organization for Standardization*
LAI – Lei de Acesso à Informação
LGPD – Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018)
MCI – Marco Civil da Internet
PCI – *Payment Card Industry*
RGPD – Regulamento Geral de Proteção de Dados
R$ - Real (Moeda da República Federativa do Brasil)

SUMÁRIO

1 - INTRODUÇÃO

A globalização modificou a forma das pessoas lidarem com seus dados pessoais. Desde o advento da internet, das redes sociais, das compras on-line e de outros canais, a divulgação massiva de dados pessoais é regra.

Tudo o que fazemos em nossas vidas deixa (ou em breve deixará) um traço digital, que pode ser analisado. Os recentes avanços na captura e análise de big data nos ajudam a melhorar o congestionamento do tráfego, prever com precisão o comportamento e as necessidades humanas em várias situações e muito mais. No entanto, essa coleta em massa de dados também pode ser usada contra pessoas. Exemplos simples disso seriam cobrar dos indivíduos maiores prêmios de seguro de automóvel ou recusar hipotecas e empregos com base no perfil de um indivíduo, conforme apresentado pelos dados coletados. Na pior das hipóteses, essa riqueza de informações poderia ser usada por governos totalitários para perseguir seus cidadãos anos após a coleta dos dados. Dessa forma, a vasta coleta de dados pessoais tem o potencial de apresentar uma grave infração à liberdade pessoal. Os indivíduos podem enfrentar perpétua ou periodicamente a estigmatização como consequência de uma ação passada específica, mesmo que já tenha sido adequadamente penalizada. Isso, por sua vez, ameaça à democracia como um todo, pois pode forçar os indivíduos a autocensurar opiniões e ações pessoais por medo de retaliações posteriores[2].

Isso alterou a conduta da sociedade com reflexos diretos na economia, na medida que as empresas precisaram, rapidamente, aprender a lidar com altos volumes dos dados pessoais, muitas vezes sem as devidas cautelas de proteção e privacidade dos seus titulares.

Assim, evolução socioeconômica fomentou a necessidade de criar diretrizes legais sobre a privacidade e a proteção de dados, cada vez mais específicas e globais.

Neste contexto, em agosto de 2018 o estado brasileiro entrou em vigor a Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018), ou somente LGPD como

[2] Sanjam Garg, Shafi Goldwasser, Prashant Nalini Vasudevan. "Formalizando a Exclusão de Dados no Contexto do Direito de Ser Esquecido." EUROCRYPT (2) 2020: 373-402.

será referida neste trabalho. A LGPD impôs um amplo rol de direitos para os titulares de dados e de obrigações para empresas controladoras dos dados pessoais.

Para que o propósito da LGPD seja atendido, determinou-se a execução de processos de anonimização, pseudonimização e eliminação dos dados. Contudo o texto legal não abordou a forma de execução dos procedimentos, ou seja, deixou a cargo do controlador dos dados a decisão de quais métodos e técnicas adotar para atendimento da LGPD.

Nessa perspectiva, o presente trabalho abordará os fundamentos e objetivos legais, em seguida, sob a ótica tecnológica, serão tratadas as funcionalidades e aplicabilidades dos métodos e técnicas de anonimização, pseudonimização com foco especial na eliminação segura de dados com a geração de evidências auditáveis.

Quanto ao conhecimento sobre as origens históricas do surgimento da LGPD brasileira, inicialmente discorre-se sobre o direito à privacidade que nasce em paralelo com a própria história da civilização, elenca-se as principais leis Europeias e Norte Americanas e, em seguida, as leis brasileiras que precederam a LGPD expondo com elas se relacionam.

O Direito à Privacidade e o Direito à Autodeterminação Informacional são objeto de investigação, desde as origens doutrinárias até sua influência na LGPD.

A LGPD propriamente dita, tem a abordagem direta dos envolvidos na proteção e privacidade de dados na empresa, ao apresentar os fundamentos, conceitos, direitos e obrigações, bem como, as sanções cíveis, administrativas em vigor e a existência do Anteprojeto de Lei que trata dos crimes contra a proteção de dados pessoais.

Os métodos e técnicas de pseudonimização, anonimização são tratados de forma bastante explicativa, com a exposição de diagramas, gráficos e tabelas por meio dos quais a dinâmica do processamento de dados fica claramente demonstrada, possibilitando a realização de periódica de testes de eficiência.

A geração de evidências é o que torna sustentável o processo de eliminação segura de dados, pois prova inequivocamente o cumprimento da lei. Neste sentido, o

uso de técnicas criptografia em especial a geração de chave aleatória e de propriedades homomórficas são abordadas, pois asseguram a irreversibilidade do processo, bem como a geração de evidência, atendendo o disposto nos Art. 5º, XIV; Art. 18, IV; e Art.19 da LGPD, abaixo:

> Art. 5º Para os fins desta Lei, considera-se:
>
> XIV - eliminação: exclusão de dado ou de conjunto de dados armazenados em banco de dados, independentemente do procedimento empregado;
>
> Art. 18. O titular dos dados pessoais tem direito a obter do controlador, em relação aos dados do titular por ele
>
> tratados, a qualquer momento e mediante requisição:
>
> IV - anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou tratados em desconformidade com o disposto nesta Lei;
>
> Art. 19. A confirmação de existência ou o acesso a dados pessoais serão providenciados, mediante requisição do titular:
>
> I - em formato simplificado, imediatamente; ou
>
> II - por meio de declaração clara e completa, que indique a origem dos dados, a inexistência de registro, os
>
> critérios utilizados e a finalidade do tratamento, observados os segredos comercial e industrial, fornecida no prazo de até 15 (quinze) dias, contado da data do requerimento do titular.
>
> § 1º Os dados pessoais serão armazenados em formato que favoreça o exercício do direito de acesso.
>
> § 2º As informações e os dados poderão ser fornecidos, a critério do titular:
>
> I - por meio eletrônico, seguro e idôneo para esse fim; ou
>
> II - sob forma impressa.
>
> § 3º Quando o tratamento tiver origem no consentimento do titular ou em contrato, o titular poderá solicitar cópia eletrônica integral de seus dados pessoais, observados os segredos comercial e industrial, nos termos deregulamentação da autoridade nacional, em formato que permita a sua utilização subsequente, inclusive em outras operações de tratamento.
>
> § 4º A autoridade nacional poderá dispor de forma diferenciada acerca dos prazos previstos nos incisos I e II do caput deste artigo para os setores específicos.

Para a redação dessa dissertação, adotou-se o uso da língua portuguesa conforme o Acordo Ortográfico, em vigor desde 2009, do qual Portugal e Brasil são signatários.

2 - HISTÓRICO DO SURGIMENTO DA LGPD NO BRASIL

A Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018) em vigor desde agosto de 2020, trouxe consigo novos encargos para empresas e pessoas físicas que lidam com dados pessoais, salvo no tocante às punições que só poderão ser aplicadas a partir de agosto de 2021[3]. Todavia, antes de submergir no mérito do assunto, faz-se necessário estabelecer o contexto histórico do surgimento da legislação, visando compreender suas origens e o objeto de que a Lei se ocupa.

O direito à privacidade tem surgimento em paralelo com a própria história da civilização, pois é algo relativo ao íntimo, ao privado. Nesse aspecto, o site da *Global Internet Liberty Campaign*[4] apoia entendimento de que:

> Privacidade tem raízes profundas na história. A Bíblia tem numerosas referências à privacidade. Também havia uma proteção real da privacidade nos primórdios das culturas hebraica, grega clássica e chinesa. A maioria destas proteções focavam no direito a ser deixado só.

Conforme entendimento de Mikhail Vieira de Lorenzi Cancelier[5], na antiguidade clássica a sociedade grega já diferenciava a vida pública da privada, por meio do isolamento entre a vida familiar e a política. Sendo que, ao adentrar na vida profissional, o indivíduo não mais se relacionava com aquilo que lhe era particular, como sua casa e família, mas, sim, com o que lhe era comum à coletividade.

Já na Idade Média, os sujeitos mais opulentos passaram a estimar formas de afastamento, período em que as famílias nobres adquiriram o hábito de praticar determinadas atitudes em ambientes privados, como por exemplo atos sexuais e das necessidades fisiológicas[6].

O progresso disto foi a valorização da esfera privada em que os lares e o ambiente familiar passaram a ser vistos como “centros de representação do poder

---

[3] BRASIL, **Senado Federal**. Disponível em: <https://www12.senado.leg.br/assessoria-de-imprensa/notas/nota-de-esclarecimento-vigencia-da-lgpd>. Acesso em: 7 set. 2020.

[4] GLOBAL INTERNET LIBERTY CAMPAIGN. PRIVACY AND HUMAN RIGHTS: ***An International Survey of Privacy Laws and Practice***. Disponível em: <http://gilc.org/privacy/survey/intro.html>. Acesso em: 20 jul. 2020.

[5] CANCELIER, Mikhail Vieira de Lorenzi. **O Direito à Privacidade hoje**: perspectiva histórica e o cenário brasileiro. Disponível em: <http://www.scielo.br/pdf/seq/n76/2177-7055-seq-76-00213.pdf>. Acesso em: 20 jul. 2020.

[6] Ibidem.

político[7]". Assim, como o público ofereceu status, a possibilidade de viver com privacidade também passa a ser um estilo dos mais afortunados, tornando-se essa conduta comum às famílias nobres[8].

Ainda segundo Cancelier, com o crescimento da classe abastada, potencializa-se a procura pela individualidade como expressão da própria personalidade dos indivíduos[9].

Destarte, essa busca pelo direito à privacidade, vez que existia uma necessidade da confirmação de um local, íntimo e particular, para essa nova classe, que não estava sob as asas das monarquias ou do clero, acabou representando parte dos atos que, por fim, culminaram na destruição do absolutismo.

Isso porque, assim como o direito à privacidade, o direito ao esquecimento não é, também, novidade chegada com século XXI. Desde o século XVIII, e com mais destaque no início do século XIX, há notícias de aperfeiçoamentos muito empoderados em prol desse mote[10].

Assim, tendo em vista as modificações socioeconômicas trazidas pela revolução industrial, surgiu a obrigação de tutelar o direito à privacidade contra ingerências alheias, cujo marco se deu por meio da Declaração Universal dos Direitos Humanos da ONU em 1948, que trouxe no seu artigo XII[11] a previsão do direito em comento.

A partir disso, estabeleceu-se a necessidade de criar diretrizes legais sobre a proteção de dados cada vez mais específicas e globais. Na visão de Bauman, o cenário

---

[7] Ibidem.

[8] DONEDA, Danilo. **Considerações iniciais sobre os bancos de dados informatizados e o direito à privacidade.** Disponível em: <http://www.edtig.ipbeja.pt/Consideracoes.pdf>. Acesso em: 22 jul. 2020.

[9] CANCELIER, Mikhail Vieira de Lorenzi. **O Direito à Privacidade hoje**: perspectiva histórica e o cenário brasileiro. Disponível em: <http://www.scielo.br/pdf/seq/n76/2177-7055-seq-76-00213.pdf>. Acesso em: 20 jul. 2020.

[10] MACHADO, Ulysses. **Direitos ao esquecimento, à privacidade e à informação: como eles se relacionam?** Disponível em: < serpro.gov.br/menu/noticias/noticias-2020/direito-esquecimento-privacidade-lgpd>. Acesso em: 7 jul. 2020.

[11] Artigo XII - Ninguém será sujeito a interferências na sua vida privada, na sua família, no seu lar ou na sua correspondência, nem a ataques à sua honra e reputação. Toda pessoa tem direito à proteção da lei contra tais interferências ou ataques.

moderno atual exige essa “vigilância em massa”, o do direito sobre todos os indivíduos. Conforme o autor:

> No cenário atual da modernidade, a separação entre política e direito permanece com um ponto de reflexão e delineamento do debate. Enquanto o poder se tornou um fenômeno global e extraterritorial, com a proliferação dos centros de poder para além das relações estatais, a política tem se reduzido drasticamente a um fenômeno local – com expressiva repercussão nas vilas globais -, incapazes de reagir, tanto na esfera pública quanto privada, e fornecer mecanismos de controle político às incertezas e avanços tecnológicos da modernidade. Num cenário de vilas globais com redes regulatórias próprias, mais do que nunca o papel da política tornou-se essencial para proporcionar o acoplamento estrutural entre atuação estatal e atividade privada para evitar a falta de interoperabilidade entre regulações sobre privacidade e proteção de dados[12].

Pelo que se expôs, um dos mais inquietantes aspectos na análise da proteção de dados e de privacidade, envolve a delimitação do conceito adequado para proteger a esfera individual.

Para compreender como o Estado está agindo neste sentido, faz-se necessário conhecer o histórico das principais leis relacionadas à privacidade na internet e aos direitos dos indivíduos sobre seus dados pessoais, para assim entender os pormenores que antecederam a criação da LGPD no Brasil.

2.1 - Leis europeias e nos Estados Unidos nas décadas de 60 e 70

Nos anos 1960 apareceram projetos de processamento de dados em larga escala e de formato centralizado, liderados pelos Estados Unidos e por alguns países europeus.

Essa inquietação se justificava pelos efeitos do uso do poder de processamento tecnológico e procedeu no bloqueio de sua estruturação, assim como na elaboração de normas focadas nos dados de proteção ao crédito nos Estados Unidos e na normatização de atividades de bancos de dados eletrônicos na Europa.

Ao mesmo tempo que a sociedade foi se tornando cada dia mais complexa e os fluxos de dados relevantes, os interesses do Estado nesses dados foram ficando cada

[12] BAUMAN, Zygmunt. **Vigilância líquida:** diálogos com David Lyon. Trad. Carlos Alberto Medeiros. Rio de Janeiro: Zahar, 2013. p. 13.

vez maiores, com vistas a exercer mais controle. Surgiram nos Estados Unidos, na década de 70, por exemplo, os censos de pesquisas, que passaram a coletar dados sobre as habitações privadas de toda sociedade. Com isso, sobreveio a busca de informações sobre as minúcias íntimas e sobre a vida privada dos cidadãos, que resultou na citação:

> A hipótese que explica o porquê desta crescente forma de invasão é o fato de que simplesmente tornou-se factível, para a tecnologia da época, processar estas informações e delas extrair alguma utilidade – e o que era novo não era a utilidade, mas o fato de sua obtenção ter sido tornada possível[13].

A concepção da Comunidade Econômica Europeia – CEE – impulsionou a indigência de uma legislação una para tratamento de dados. Contudo os processos de regulação da União Europeia – adepta à denominação graças ao Tratado de Maastricht em 1992 – levam em conta a dessemelhança jurídica de seus componentes[14].

Desta forma, apenas em 1995 surgiu a primeira legislação europeia unificada: Diretiva 95/46/EC, a qual dispunha a proteção dos indivíduos quanto ao processamento de seus dados pessoais e a circulação deles no âmbito da União Europeia. Em que pese especialistas confiarem que a ansiedade com dados pessoais brotou nos Estados Unidos anos 60, a primeira lei oficialmente direcionada ao tema foi concebida na década de 70 em Hessen, na Alemanha[15].

Nesse interim, o progresso da cibernética e da indústria nos países mais adiantados impulsionou o estado alemão a criar cláusulas para regular a privacidade no país, sendo também essa a primeira vez que o conceito de proteção de dados foi inserido no cenário jurídico da Alemanha.

Conquanto o conceito tenha sido criado no início da década de 70, a legislação só foi concluída e praticada em 1978. No mesmo ano, países como França, Noruega, Suécia e Áustria também estabeleceram suas próprias normas sobre como os dados de seus cidadãos poderiam ser utilizados. Já em 1981, uma convenção reunindo os

---

[13] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**. Rio de Janeiro: Renovar, 2006, p. 9-10.
[14] VIDOR, Daniel MARTINS, **LGPD: origem e implicações**. Disponível em: <http://mercurylbc.com/lgpd-origem-e-implicacoes/>. Acesso em: 8 set. 2020.
[15] Ibidem.

países membros do Conselho da Europa acudiu a unificar e expandir as normas para o tratamento automatizado dos dados pessoais[16].

O modelo dos Estados Unidos, onde há grande recurso à autorregularão e onde o estado busca interferir, criando normas, apenas em áreas consideradas sensíveis, como saúde, finanças e menores de idade. Nesse modelo, apesar de haver alguma tendência à judicialização, o principal foco é a liberdade de contratação e a possibilidade de o mercado definir, em sua interação com os consumidores, os comportamentos aceitáveis[17].

A União Europeia e sua abordagem à proteção dos dados pessoais. O modelo eclético é assim nomeado pois resulta de um grande ajuntamento de estratégias, incluindo o recurso a normas estatais e mercadológicas, mecanismos alternativos de resolução de controvérsias, soluções tecnológicas e o engajamento ativo, tanto do titular dos dados quanto do responsável pelo tratamento de dados, no cumprimento e fiscalização da lei. Esse último ponto talvez seja o mais interessante, por trabalhar não pela repressão de condutas (onde cumprir a lei é um modo de não perder dinheiro), mas pela indução, onde o respeito à privacidade é interessante para todos os envolvidos, pelos benefícios daí advindos[18].

Da leitura desse breve histórico, se percebe que a evolução da proteção de dados pessoais como um dos direitos da personalidade. Isso ocorreu sobretudo no modelo europeu, que se apresenta diametralmente oposto ao modelo norte-americano. Esses dois são os exemplos mais consagrados em termos de proteção de dados pessoais e sensíveis em nível mundial[19].

2.2 - Constituição Federal Brasileira de 1988

Em 1988 o Brasil abraçou uma nova Carta Magna, que trouxe no seu bojo alguns pontos sobre proteção de dados. O inciso X do artigo 5º, referente aos direitos e

---

[16] Ibidem.

[17] MODELOS REGULATÓRIOS PARA PROTEÇÃO DE DADOS PESSOAIS; Guilherme Berti de Campos Guidi, Disponível em: https://itsrio.org/wp-content/uploads/2017/03/Guilherme-Guidi-V-revisado.pdf. Acesso em: 17 jul. 2021.

[18] Ibidem.

[19] RAMINELLI, Francieli Puntel; RODEGHERI, Letícia Bodanese. **A Proteção de Dados Pessoais na Internet no Brasil: Análise de decisões proferidas pelo Supremo Tribunal Federal.** Cadernos do Programa de Pós-graduação em Direito. Disponível em: <https://seer.ufrgs.br/ppgdir/article/view/61960/39936>. Acesso em: 7 jul. 2020.

deveres dos cidadãos trata, de forma geral, da privacidade dos brasileiros: “são invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito à indenização pelo dano material ou moral decorrente de sua violação”[20].

Isto é, a Constituição Federal antevê o direito à privacidade, incluindo-se a inviolabilidade do sigilo de comunicações, de dados e comunicações telefônicas (artigo 5º, inciso XII), além da garantia de acesso a informações pessoais e de retificação de dados, constantes de bancos de dados públicos, por meio do *Habeas Data* (artigo 5º, inciso LXXII) e, posteriormente, regulado pela Lei n.º 9.507 de 1997[21].

O artigo 7º da Lei n.º 9.507/97 trata da concessão do *Habeas Data*:

> Art. 7° Conceder-se-á *habeas data*:
>
> I - para assegurar o conhecimento de informações relativas à pessoa do impetrante, constantes de registro ou banco de dados de entidades governamentais ou de caráter público;
>
> II - para a retificação de dados, quando não se prefira fazê-lo por processo sigiloso, judicial ou administrativo;
>
> III - para a anotação nos assentamentos do interessado, de contestação ou explicação sobre dado verdadeiro mas justificável e que esteja sob pendência judicial ou amigável.

Esse colírio constitucional permite que o impetrante tenha acesso aos registros ou aos bancos de dados de entidades governamentais ou de caráter público, a fim de ter informação de seus dados ou de retificá-los. A abundância desse instrumento ocorreu especialmente em países saídos de regimes ditatoriais, como por exemplo o Brasil, que também foi marcado pelo “uso autoritário da informação”[22].

Embora de forma restrita, o aludido remédio constitucional não deixa de ter relação com a proteção da privacidade e dos dados pessoais, desde que incluídos em bancos de dados governamentais ou de caráter público. Isto é, a garantia seria um desdobramento do princípio insculpido no mencionado inciso X, do artigo 5º, da Constituição Federal[23]. Portanto os elementos constitucionais que se aproximam do

---

[20] BRASIL. **Constituição Federal de 1988**. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 7 set. 2020.

[21] BRASIL. **Lei nº 9.507, de 12 de novembro de 1997**. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/L9507.htm. Acesso em: 7 set. 2020.

[22] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**. Rio de Janeiro: Renovar, 2006, p. 347.

[23] X - são invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou moral decorrente de sua violação.

assunto se concentram em alguns incisos do artigo 5º, o que lhes confere caráter de direito fundamental[24].

Contudo, importa inferir que a interpretação conferida ao artigo 5º, inciso XII[25] da Constituição Federal, extraída da Lei n.º 9.296 de 1996 não abrange os dados estáticos[26], pois somente o fluxo de dados estaria protegido pelo mencionado dispositivo constitucional, de modo que a proteção aos dados pessoais estaria abrangida pela previsão genérica, constante do inciso X da Constituição Federal[27].

Destarte, o que se nota é que não há menção expressa, no texto constitucional, ao direito à privacidade. Apesar disso, entende-se que as espécies intimidade e vida privada estão dentro do gênero privacidade. Da previsão Constitucional se extraem diversos subsídios que levam à tutela da privacidade, com o benefício formal de possuírem caráter pétreo e fundamental.

2.3 - Lei n.º 8.078 de 1990 (Código de Defesa do Consumidor)

Primordial destacar partes do Código de Defesa do Consumidor brasileiro, promulgado em 1990 que, não obstante seja uma lei anterior ao Código Civil do Brasil, é relevante quando se trata da temática abordada pela presente pesquisa acadêmica.

Isso tudo porque a redação dos artigos 43 e 44[28] da referida norma garante que o consumidor tenha acesso a informações pessoais registradas pelos provedores e

---

[24] TAVARES, Letícia Antunes; ALVAREZ, Bruna Acosta**. Da proteção dos dados pessoais: uma análise comparada dos modelos de regulação da Europa, dos Estados Unidos da América e do Brasil.** Disponível em: <http://www.tjsp.jus.br/download/EPM/Publicacoes/ObrasJuridicas/ii%204.pdf?d=636680444556135606>. Acesso em: 7 set. 2020.

[25] XII - é inviolável o sigilo da correspondência e das comunicações telegráficas, de dados e das comunicações telefônicas, salvo, no último caso, por ordem judicial, nas hipóteses e na forma que a lei estabelecer para fins de investigação criminal ou instrução processual penal.

[26] CANOTILHO, J. J. Gomes et al. (Coord.). **Comentários à Constituição do Brasil.** São Paulo: Saraiva, 2013. p. 293.

[27] TAVARES, Letícia Antunes; ALVAREZ, Bruna Acosta**. Da proteção dos dados pessoais: uma análise comparada dos modelos de regulação da Europa, dos Estados Unidos da América e do Brasil.** Disponível em: <http://www.tjsp.jus.br/download/EPM/Publicacoes/ObrasJuridicas/ii%204.pdf?d=636680444556135606>. Acesso em: 7 set. 2020.

[28] Redação completa do dispositivo dos artigos 43 e 44 e seus parágrafos: "Art. 43. O consumidor, sem prejuízo do disposto no art. 86, terá acesso às informações existentes em cadastros, fichas, registros e dados pessoais e de consumo arquivados sobre ele, bem como sobre as suas respectivas fontes.

§ 1° Os cadastros e dados de consumidores devem ser objetivos, claros, verdadeiros e em linguagem de fácil compreensão, não podendo conter informações negativas referentes a período superior a cinco anos.

§ 2° A abertura de cadastro, ficha, registro e dados pessoais e de consumo deverá ser comunicada por escrito ao consumidor, quando não solicitada por ele.

possa solicitar sua retificação, estabelecendo, inclusive, que o consumidor deverá ser avisado a respeito, quando seus dados forem incluídos em um cadastro e que eles serão armazenados pelo prazo máximo de 5 anos[29].

Com isso, o Código de Defesa do Consumidor progrediu na procura por tutelar as informações das pessoas, contendo, outrossim, seção destinada especificamente a cadastros e banco de dados, defendendo o direito do consumidor acessar os dados que um empreendimento privado tem sobre ele e solicitar sua correção, em caso de algum dado estar incorreto[30].

Por fim, cabe mencionar que há artigos que afiançam a privacidade e responsabilizam as empresas sobre a segurança dos dados, como o artigo 11º, capítulo 3 do CDC: “Os dados pessoais do consumidor serão preservados, mantidos em sigilo e utilizados exclusivamente para os fins do atendimento”[31].

Destarte, é possível aferir que, desde a criação da lei em questão, a sociedade já ambicionava maior proteção de dados pessoais e essa ansiedade alcançou os legisladores, fazendo com que se ocupassem com a questão.

---

§ 3° O consumidor, sempre que encontrar inexatidão nos seus dados e cadastros, poderá exigir sua imediata correção, devendo o arquivista, no prazo de cinco dias úteis, comunicar a alteração aos eventuais destinatários das informações incorretas.
§ 4° Os bancos de dados e cadastros relativos a consumidores, os serviços de proteção ao crédito e congêneres são considerados entidades de caráter público.
§ 5° Consumada a prescrição relativa à cobrança de débitos do consumidor, não serão fornecidas, pelos respectivos Sistemas de Proteção ao Crédito, quaisquer informações que possam impedir ou dificultar novo acesso ao crédito junto aos fornecedores.
§ 6º Todas as informações de que trata o caput deste artigo devem ser disponibilizadas em formatos acessíveis, inclusive para a pessoa com deficiência, mediante solicitação do consumidor
Art. 44. Os órgãos públicos de defesa do consumidor manterão cadastros atualizados de reclamações fundamentadas contra fornecedores de produtos e serviços, devendo divulgá-lo pública e anualmente. A divulgação indicará se a reclamação foi atendida ou não pelo fornecedor.
§ 1° É facultado o acesso às informações lá constantes para orientação e consulta por qualquer interessado.
§ 2° Aplicam-se a este artigo, no que couber, as mesmas regras enunciadas no artigo anterior e as do parágrafo único do art. 22 deste código.

[29] Ibidem.

[30] Ibidem.

[31] BRASIL. **Código de Defesa do Consumidor**, Lei 8.078 de 1990. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm>. Acesso em: 10 jun. 2020.

2.4 - Lei n.º 9.279/1996 (Lei de Propriedade Industrial)

A propriedade industrial no Brasil está consubstanciada no ordenamento jurídico, mormente no artigo 5º, inciso XXIX, da Constituição Federal:

> Art. 5º - Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade, nos termos seguintes:
> XXIX - a lei assegurará aos autores de inventos industriais privilégio temporário para sua utilização, bem como proteção às criações industriais, à propriedade das marcas, aos nomes de empresas e a outros signos distintivos, tendo em vista o interesse social e o desenvolvimento tecnológico e econômico do País[32].

Por certo, a propriedade industrial encontra maior respaldo na Lei n.º 9.279, promulgada em 14 de maio de 1996. Tal norma veio para regular direitos e obrigações relativos à propriedade industrial, determinando os direitos de uso exclusivo das respectivas propriedades industriais: patentes e registros.

O Ministério do Desenvolvimento, Indústria e Comércio Exterior é responsável pelo aprimoramento, dispersão e gestão do sistema brasileiro e por garantir direitos de propriedade intelectual à indústria. Subordinando-se a este Ministério, a autarquia federal do Instituto Nacional da Propriedade Industrial – INPI –, que tem a seu encargo a análise e concessão da patente (carta-patente) e do registro (certificado)[33].

Ou seja, como é possível perceber, no ano de 1996, o Brasil ocupou-se em ter uma forma legal e regulamentada de proteger marcas, patentes, desenhos industriais e invenções. Porém tal legislação não tratava diretamente da proteção do direito de proteção aos dados das marcas, patentes e direitos autorais, nem mesmo atendia às modernidades de aplicativos e programas de informática.

---

[32] BRASIL. **Constituição Federal de 1988.** Disponível em: <planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 10 de jun. 2020.

[33] BRASIL. **Lei da Propriedade Industrial, Lei 9.727 de 1996.** Disponível em: <http://www.planalto.gov.br/ccivil_03/Leis/L9279.htm>. Acesso em: 10 de jun. 2020.

2.5 - Lei n.º 9.296/1996 (Lei de Interceptação Telefônica)

Adiante, mas ainda no ano de 1996, promulgou-se a Lei 9.296, que dispunha essencialmente acerca da validade jurídica de interceptações telefônicas postuladas e deferidas legalmente.

Isso porque, antes disso, as supremas cortes do país entendiam que essa violação da privacidade – ainda que para fins de cessação de crimes – era considerada afronta ao artigo 5º, inciso XII, que estabelece que "é inviolável o sigilo da correspondência e das comunicações telegráficas, de dados e das comunicações telefônicas, salvo, no último caso, por ordem judicial, nas hipóteses e na forma que a lei estabelecer para fins de investigação criminal ou instrução processual penal"[34].

Embora se tenha entendido que a interceptação constituía uma excelente fonte de prova, percebeu-se a necessidade de regulamentá-la para que não ocorressem abusos. Com isso, a edição da Lei de Interceptações Telefônicas foi salutar, como mecanismo capaz de positivar o procedimento para esses acessos no campo pessoal dos indivíduos que porventura estivessem envolvidos em algum tipo de delito[35].

Mais uma vez percebe-se que o Brasil vinha em uma escalada constante e necessária para desenvolver processos legais para relativizar o direito à privacidade constituído legalmente na Carta Magna pátria, em razão da necessidade de relativização do direito em alguns aspectos.

2.6 - Lei n.º 9.610/1998 (Lei de Direitos Autorais)

A Lei 9.610, de 19 de fevereiro de 1998, entrou em vigor no dia 19 de junho de 1998, alterando, modernizando e concretizando a legislação sobre os direitos autorais. Outra vez foi possível aferir a necessidade do brasileiro em proteger seus dados e publicações, assim como ocorreu com o advento da Lei de Propriedade Industrial.

---

[34] BRASIL. **Constituição Federal de 1988.** Disponível em: <planalto.gov.br/ccivil_03/ constituição /constituicao.htm>. Acesso em: 10 de jun. 2020.

[35] GRECO FILHO, Rogério. **Interceptação telefônica**: considerações sobre a Lei n.º 9.296, de 24 de julho de 1996. 2ªed. São Paulo: Saraiva, 2009. p. 20.

Em seu artigo 5º conceitua publicação, transmissão ou emissão, retransmissão, distribuição, comunicação ao público, reprodução, contratação, obra (em coautoria, anônima, pseudônima, inédita, póstuma, originária, derivada, coletiva, audiovisual), fonograma, editor, produtor, radiodifusão, artistas intérpretes ou executantes[36].

Note-se que o artigo 6º, diz que "não serão de domínio da União, dos Estados, do Distrito Federal ou dos municípios as obras por eles simplesmente subvencionadas"[37]. Esse artigo vem explicar a problemática que vinha gerando muita discussão à época.

Com isso, nota-se que o Brasil seguia no rumo da normatização de tudo aquilo que carecia de proteção com relação aos direitos pessoais e ainda aos direitos profissionais dos indivíduos, fazendo com que leis como esta emergissem em meio à necessidade da época. Entretanto, ainda pouco se falava sobre a proteção e posterior eliminação de dados desses registros.

2.7 - Lei Complementar nº 105/2001 (Lei do Sigilo Bancário)

Ainda no âmbito das relações da Lei Geral de Proteção de Dados do Brasil com outros atos normativos no país, que também falam da proteção de dados, passa-se à análise da Lei do Sigilo Bancário.

Com efeito, as instituições financeiras estão sujeitas, desde 2001, à lei específica no tratamento de dados de seus clientes, cujo teor obriga à manutenção do sigilo das relações financeiras confiadas a cada uma das instituições.

A Lei do Sigilo Bancário (Lei Complementar nº 105/2001) trouxe uma vantagem para os Bancos no cenário LGPD, já que a nova lei brasileira também traz em seu texto exigências quanto a quesitos de segurança mínimos, características que os Bancos já têm atendido, não apenas por conta da Lei Complementar, mas também pelas regulamentações do Banco Central do Brasil, órgão regulador do tema no país.

---

[36] BRASIL. **Direitos Autorais, Lei 9.610 de 1998.** Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9610.htm>. Acesso em: 10 de jun. 2020.
[37] Ibidem.

Com isso, importa dizer que a LGPD dialoga com a Lei do Sigilo Bancário não só nas relações dos clientes com as instituições, mas também nas afinidades jurídicas das instituições financeiras entre si e com terceiros.

Como dito, a Lei Complementar nº 105/2001 estabelece o sigilo das transações realizadas pelas instituições financeiras (listadas em seu art. 1º, §§ 1º e 2º[38]) e as circunstâncias em que esse segredo pode ser relativizado.

Como a própria denominação indica, a Lei do Sigilo Bancário trata do sigilo como regra geral nas atividades bancárias, em razão da natureza jurídica dos dados e em respeito à privacidade da vida financeira das pessoas. Entretanto, a Lei do Sigilo Bancário também traz normas para prevenir que o sigilo seja utilizado como meio para dar azo às práticas criminosas.

Em suma, as normas de proteção de dados na Lei do Sigilo Bancário são as seguintes: (i) Dever de sigilo (artigo 1º, *caput*): o direito à privacidade previsto na Constituição é tratado como regra nas atividades das instituições financeiras, com arrimo na natureza jurídica dos dados dos clientes. Além disso, o § 3º do art. 1º da Lei complementar descreve as atividades que podem ser realizadas sem a violação do dever de sigilo, enquanto o seu § 4º prevê as hipóteses de quebra do sigilo. Desta forma, nota-se a preocupação legislativa com amparar o sigilo dos dados pessoais, em detrimento da tutela da vida privada, o que se repete na Lei Geral de Proteção de Dados, que tem entre os seus fundamentos a inviolabilidade da intimidade, da honra e

[38] Art. 1o As instituições financeiras conservarão sigilo em suas operações ativas e passivas e serviços prestados.
§ 1o São consideradas instituições financeiras, para os efeitos desta Lei Complementar:
I – os bancos de qualquer espécie;
II – distribuidoras de valores mobiliários;
III – corretoras de câmbio e de valores mobiliários;
IV – sociedades de crédito, financiamento e investimentos;
V – sociedades de crédito imobiliário;
VI – administradoras de cartões de crédito;
VII – sociedades de arrendamento mercantil;
VIII – administradoras de mercado de balcão organizado;
IX – cooperativas de crédito;
X – associações de poupança e empréstimo;
XI – bolsas de valores e de mercadorias e futuros;
XII – entidades de liquidação e compensação;
XIII – outras sociedades que, em razão da natureza de suas operações, assim venham a ser consideradas pelo Conselho Monetário Nacional.
§ 2o As empresas de fomento comercial ou factoring, para os efeitos desta Lei Complementar, obedecerão às normas aplicáveis às instituições financeiras previstas no § 1º.

da imagem (artigo 2º, IV, da LGPD[39]) e os direitos do titular têm sua base principal nos direitos fundamentais de liberdade, de intimidade e de privacidade (artigos 1º e 17 da LGPD[40]); (ii) Compartilhamento de dados cadastrais entre instituições financeiras e centrais de risco (artigo 1º, § 3º, I, da Lei do Sigilo Bancário[41]): conforme a primeira hipótese de ausência de violação de sigilo, as instituições financeiras podem trocar informações entre si, para verificar a veracidade/autenticidade dos dados, bem como para noticiar a ocorrência de eventuais riscos nas contratações.

Essa troca de dados deve mirar sempre os princípios da finalidade e da necessidade (artigo 6º, I e III, da LGPD[42]), razão pela qual os dados compartilhados devem ser apenas os necessários para o objetivo pretendido e devem ser utilizados para a finalidade específica. (iii) Fornecimento de informações em cadastros de emitentes de cheques sem provisão de fundos e de inadimplentes a entidades de proteção de crédito (artigo 1º, § 3º, II, da Lei do Sigilo Bancário[43]): na segunda hipótese de ausência de violação de sigilo, as instituições financeiras não precisam examinar previamente o devedor sobre a comunicação do fato aos órgãos de proteção de crédito. Apesar disso, ele deve ser notificado previamente da inscrição, o que lhe dá, por exemplo, o direito de acesso e de correção de dados eventualmente incorretos (artigo 18, II e III, da LGPD[44]). Ainda, tal fornecimento e tratamento dos dados enquadra-se na

[39] Art. 2º A disciplina da proteção de dados pessoais tem como fundamentos:
IV - a inviolabilidade da intimidade, da honra e da imagem;
[40] Art. 17. Toda pessoa natural tem assegurada a titularidade de seus dados pessoais e garantidos os direitos fundamentais de liberdade, de intimidade e de privacidade, nos termos desta Lei.
[41] Art. 1o As instituições financeiras conservarão sigilo em suas operações ativas e passivas e serviços prestados.
§ 3o Não constitui violação do dever de sigilo:
I – a troca de informações entre instituições financeiras, para fins cadastrais, inclusive por intermédio de centrais de risco, observadas as normas baixadas pelo Conselho Monetário Nacional e pelo Banco Central do Brasil;
[42] Art. 6º As atividades de tratamento de dados pessoais deverão observar a boa-fé e os seguintes princípios:
II - adequação: compatibilidade do tratamento com as finalidades informadas ao titular, de acordo com o contexto do tratamento;
III - 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;
[43] Art. 1o As instituições financeiras conservarão sigilo em suas operações ativas e passivas e serviços prestados.
§ 3o Não constitui violação do dever de sigilo:
II - o fornecimento de informações constantes de cadastro de emitentes de cheques sem provisão de fundos e de devedores inadimplentes, a entidades de proteção ao crédito, observadas as normas baixadas pelo Conselho Monetário Nacional e pelo Banco Central do Brasil;
[44] Art. 18. O titular dos dados pessoais tem direito a obter do controlador, em relação aos dados do titular por ele tratados, a qualquer momento e mediante requisição:
II - acesso aos dados;
III - correção de dados incompletos, inexatos ou desatualizados;

hipótese de proteção do crédito prevista no artigo 7º, X, da LGPD[45], e o titular tem o direito de obter informações sobre o compartilhamento dos dados e da sua finalidade (artigo 9º, V, da LGPD[46]); (iv) Fornecimento de dados sobre contribuições tributárias que incidam em operações financeiras (artigo 1º, § 3º, III, da Lei do Sigilo Bancário[47]): a terceira hipótese de ausência de violação de sigilo trata do fornecimento de dados que eram prestados pelas instituições financeiras à Receita Federal, sobre a retenção e o recolhimento da contribuição provisória sobre movimentação ou transmissão de valores e de créditos e direitos de natureza financeira, com a identificação dos contribuintes e suas operações. Esse tratamento dos dados pelas instituições financeiras e, de forma compartilhada, pela Receita Federal, enquadra-se na hipótese de cumprimento de dever legal prevista no artigo 7º, II, da LGPD[48]. Assim, o titular tem o direito de obter informações sobre o compartilhamento de seus dados e da sua finalidade (art. 9º, V, da LGPD); (v) Comunicação da prática de ilícitos penais ou administrativos (artigo 1º, § 3º, IV, da Lei do Sigilo Bancário[49]): de acordo com a quarta hipótese de exceção à violação, o sigilo bancário não abrange a prática de crimes, o que compreende também as operações posteriormente realizadas com os recursos oriundos das práticas criminosas (a lavagem de dinheiro); (vi) Fornecimento consentido de dados pelos interessados (artigo 1º, § 3º, V, da Lei do Sigilo Bancário[50]): na quinta hipótese flexibilização do violação de sigilo, tem-se a possibilidade do consentimento do titular como um requisito para o tratamento dos dados, ou seja, é lícita a revelação dos dados bancários com o consentimento expresso do titular. Na Lei Geral de Proteção de Dados, o consentimento deve ser livre, expresso, inequívoco, por escrito, revogável, de finalidade específica e limitada (artigo 5º, II, da LGPD[51]). (vii) Prestação

[45] Art. 7º O tratamento de dados pessoais somente poderá ser realizado nas seguintes hipóteses:
X - para a proteção do crédito, inclusive quanto ao disposto na legislação pertinente.
[46] Art. 9º O titular tem direito ao acesso facilitado às informações sobre o tratamento de seus dados, que deverão ser disponibilizadas de forma clara, adequada e ostensiva acerca de, entre outras características previstas em regulamentação para o atendimento do princípio do livre acesso:
V - informações acerca do uso compartilhado de dados pelo controlador e a finalidade;
[47] III – o fornecimento das informações de que trata o § 2o do art. 11 da Lei no 9.311, de 24 de outubro de 1996;
[48] Art. 7º O tratamento de dados pessoais somente poderá ser realizado nas seguintes hipóteses:
II - para o cumprimento de obrigação legal ou regulatória pelo controlador;
[49] Art. 1o As instituições financeiras conservarão sigilo em suas operações ativas e passivas e serviços prestados.
§ 3o Não constitui violação do dever de sigilo:
VI – a prestação de informações nos termos e condições estabelecidos nos artigos 2o, 3o, 4o, 5o, 6o, 7o e 9 desta Lei Complementar.
[50] V – a revelação de informações sigilosas com o consentimento expresso dos interessados;
[51] Art. 5º Para os fins desta Lei, considera-se:

de informações no cumprimento de dever legal (artigo 1º, §3º, VI, da Lei do Sigilo Bancário[52]): a sexta hipótese prevê que devem ser fornecidas informações pelas instituições financeiras nos termos e condições dos casos previstos nos artigos 2º a 7º e 9º da lei, que compreendem, por exemplo, o envio de dados bancários para cumprimento de decisão judicial ou para a instrução de processo administrativo fiscal, por exemplo. Com isso, o titular tem o direito de obter informações sobre o compartilhamento de seus dados e da sua finalidade (artigo 9º, V, da LGPD[53]); (viii) Dados de pagamentos e operações de créditos para cadastros positivos de crédito (artigo 1º, § 3º, VII, da Lei do Sigilo Bancário[54]): ao cabo, conforme a sétima hipótese de exceção à regra do sigilo, os dados podem ser compartilhados não apenas para os bancos de dados de cadastros de inadimplentes, assim como para apuração de nota/cálculo de pontuação em registros positivos de crédito, em benefício do titular dos dados. Ou seja, para caso de cadastros positivos, o titular tem o direito de obter informações sobre o compartilhamento dos dados e da sua finalidade (artigo 9º, V, da LGPD) e de correção de dados eventualmente incorretos (artigo 18, II e III, da LGPD[55]).

Por fim, em dezembro de 2019 o Supremo Tribunal Federal firmou posição sobre a constitucionalidade do compartilhamento de dados bancários com a Receita Federal e com o Ministério Público para fins penais, independentemente de autorização prévia em processo judicial, e fixou as seguintes tese no Tema nº 990 da Repercussão Geral:

> "1. É constitucional o compartilhamento dos relatórios de inteligência financeira da UIF e da íntegra do procedimento fiscalizatório da Receita Federal do Brasil, que define o lançamento do tributo, com os órgãos de persecução penal para fins criminais, sem a obrigatoriedade de prévia autorização judicial, devendo ser resguardado o sigilo das informações em procedimentos formalmente instaurados e sujeitos a posterior controle jurisdicional.

II - dado pessoal sensível: 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;

[52] VI – descontos de duplicatas, notas promissórias e outros títulos de crédito;

[53] Art. 9º O titular tem direito ao acesso facilitado às informações sobre o tratamento de seus dados, que deverão ser disponibilizadas de forma clara, adequada e ostensiva acerca de, entre outras características previstas em regulamentação para o atendimento do princípio do livre acesso:
V - informações acerca do uso compartilhado de dados pelo controlador e a finalidade;

[54] VII - o fornecimento de dados financeiros e de pagamentos, relativos a operações de crédito e obrigações de pagamento adimplidas ou em andamento de pessoas naturais ou jurídicas, a gestores de bancos de dados, para formação de histórico de crédito, nos termos de lei específica.

[55] Art. 18. O titular dos dados pessoais tem direito a obter do controlador, em relação aos dados do titular por ele tratados, a qualquer momento e mediante requisição:
II - acesso aos dados;
III - correção de dados incompletos, inexatos ou desatualizados;

> 2. O compartilhamento pela UIF e pela RFB, referente ao item anterior, deve ser feito unicamente por meio de comunicações formais, com garantia de sigilo, certificação do destinatário e estabelecimento de instrumentos efetivos de apuração e correção de eventuais desvios".

2.8 - Lei nº 10.741/2003 (Estatuto do Idoso)

A designação de Estatuto do Idoso refere-se a Lei Federal nº 10.471/2003[56], que se destina a regular os direitos e garantias das pessoas com idade a partir dos 60 anos, conforme conceito da Organização Mundial de Saúde.

O objetivo deste estatuto é ampliar e garantir à pessoa idosa, o acesso aos direitos fundamentais inerentes à pessoa humana.

A LGPD em seu texto não dedica qualquer ponto exclusivamente aplicado à pessoa idosa. Entretanto, seu Artigo 2º, VII, estabelece como um de seus fundamentos os direitos humanos, o livre desenvolvimento da personalidade, a dignidade e o exercício da cidadania pelas pessoas naturais, ou seja, a LGPD ampara a aplicação de seus termos, para os indivíduos amparados pelo Estatuto do Idoso.

Neste sentido, recorremos ao Decreto 10.474[57] de 26/08/2020, em seu Anexo I, Capitulo I, Artigo 2º, que estabelece as competências da Autoridade Nacional de Proteção de Dados, dentre elas destacamos o Inciso XIX, que diz: "garantir que o tratamento de dados de idosos seja efetuado de maneira simples, clara, acessível e adequada ao seu entendimento, nos termos da LGPD e do Estatuto do Idoso".

Assim sendo, constamos que Brasil inovou no cenário internacional, na medida garantiu o tratamento diferenciado dos dados pessoais de indivíduos contemplados pelo Estatuto do Idoso.

---

[56] http://www.planalto.gov.br/ccivil_03/leis/2003/l10.741.htm, acessado em 10/02/2021.

[56] https://www.in.gov.br/web/dou/-/decreto-n-10.474-de-26-de-agosto-de-2020-274389226, acessado em 10/02/2021.

2.9 - Lei n.º 12.527/2011 (Lei de Acesso à Informação)

As instituições, com o fito de ampliar sua legitimidade, visando democratizar as informações junto à sociedade organizada e reforçar cidadania, compreenderam que o acesso aos dados de forma transparente trata-se de um instrumento fundamental para uma república federativa.

Desta maneira, a Lei de Acesso à Informação[58] granjeou proeminência no ordenamento jurídico brasileiro, na medida que essa garantia de acesso transparente constituiu direito fundamental. Indicando-se como diretrizes básicas, aventadas pelo próprio Senado Federal em seu preâmbulo legal, como sendo o princípio a publicidade, o sigilo como exceção, a divulgação de informações de interesse público independentemente de solicitação, a cultura da transparência e controle social do que é feito na administração pública[59].

Não obstante, a LAI constitui direito adquirido importante para o povo brasileiro, vez que assegura aos cidadãos o direito se receber dos órgãos públicos informações do seu interesse pessoal ou coletivo geral. Prestigiando, em última análise, a publicidade como princípio basilar, vedando práticas sigilosas, que agora são exceção à regra.

2.10 - Lei n.º 12.846/2013 (Lei Anticorrupção)

Outro marco importantíssimo na história legal da proteção de dados no Brasil foi a Lei Anticorrupção. Isso porque, trouxe em seu bojo a possibilidade de que empresas possam ser responsabilizadas administrativa e civilmente por seus atos contra a administração pública, nacional ou estrangeira. Isto é, a LAC inovou no cenário brasileiro ao prever a responsabilização objetiva de pessoas jurídicas.

Muito embora existissem iniciativas de combate à corrupção no âmbito internacional e nacional, percebeu-se a ocorrência de atos corruptos promovidos por

---

[58] BRASIL. **Direitos de Acesso à Informação, Lei 12.527 de 2011.** Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em 11 de jun. 2020.

[59] BRASIL. Senado Federal. **Lei de Acesso à Informação no Brasil:** O que você precisa saber. Disponível em: <https://www12.senado.leg.br/transparencia/arquivos/sobre/cartilha-lai/>. Acesso em: 10 jun. 2020.

países, de maneira especial os desenvolvidos, em detrimento de economias e democracias mais fragilizadas[60].

Como se pode perceber, a Lei em comento foi de suma importância para que as empresas privadas reforçassem a ideia da aplicabilidade plena do compliance digital, criando programas de conduta para zelar por sua integridade com reflexo direto na privacidade e proteção dos dados.

2.11 - Lei n.º 12.850/2013 (Organizações Criminosas)

Em que pese o escopo principal da legislação seja o desmembramento e identificação/responsabilização de membros de organizações criminosas, a norma trouxe importante inovação no que se refere às provas eletrônicas. Isso porque, o artigo 3º do mencionado dispositivo prevê que serão admitidos como meios de provas, por exemplo:

a) os registros de ligações telefônicas e telemáticas a dados cadastrais constantes em banco de dados públicos ou privados e a informações eleitorais ou comerciais;

b) a interceptação de comunicações telefônicas e telemáticas, nos termos da legislação específica; e

c) o afastamento dos sigilos financeiro, bancário e fiscal, nos termos da legislação específica[61].

Portanto, a Lei em tela concedeu aos órgãos de investigação novas exceções à garantia constitucional da privacidade e, para aqueles casos em que os dados aclarados sirvam para descobrir delitos ou desmantelar organizações criminosas.

2.12 - Lei n.º 12.965/2014 (Marco Civil da Internet)

---

[60] SIMON, Pedro. **A impunidade veste colarinho branco.** Brasília: Senado Federal, 2010. p. 10.

[61] BRASIL. **Direitos das Organizações Criminosas, Lei 12.850 de 2013.** Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/lei/l12850.htm>. Acesso em: 11 de jun. 2020.

O Marco Civil da Internet foi a primeira lei que regulou a aplicação do direito digital no Brasil. Pode-se dizer que foi o pontapé inicial para que a justiça brasileira começasse a entender que os acontecimentos na internet repercutem diretamente no mundo real.

O estudioso Ronaldo Lemos elucida que o processo legislativo que deu origem ao Marco Civil foi impulsionado, principalmente, por dois acontecimentos de relevância nacional: o primeiro consistiu no escândalo provocado pela descoberta de que o governo brasileiro sofreu espionagem americana, fazendo com que as autoridades governamentais passassem a tratar com urgência e relevância a proteção de dados e a criação de regulamentação específica sobre a Internet; e o segundo consubstanciado na existência de um projeto de lei (Lei Azeredo[62]), que tinha como propósito a criação de legislação criminal específica para a Internet, a qual pretendia a criminalizar de um extenso rol de condutas praticadas na rede[63].

Em razão da importância do tema, o MCI arrolou, entre os princípios basilares para disciplinar o uso e funcionamento da Internet no Brasil, a "proteção da privacidade" e a "proteção dos dados pessoais" nos termos dispostos nos incisos II e III do artigo 3º[64], que são os nortes debruçados neste estudo.

Esses atributos foram relevantes para que se atingisse a meta de confeccionar uma lei comprometida em garantir liberdades civis, dentro das quais se encontra a possibilidade de conter a circulação quase irrestrita de dados pessoais e a proteção da privacidade. Além de a internet ter potencializado as possibilidades do fluxo de informações, "tudo o que ocorre na rede ocorre de acordo com seus protocolos e pode ser acompanhado e ter seus passos meticulosamente registrados".

Portanto, uma das principais papeis do MCI é impedir que "um prestador de serviços em determinado contexto extrapole sua função, obtenha e valha-se de dados

---

[62] O projeto de Lei 84 de 1999, dispõe sobre os crimes cometidos na área de informática, suas penalidades e dá outras providências.

[63] LEMOS, Ronaldo. **O marco civil como símbolo do desejo por inovação no Brasil**. In: LEITE, George Salomão; LEMOS, Ronaldo (coord.). Marco Civil da Internet. São Paulo, Atlas, 2014, p. 04.

[64] Redação completa do artigo 3º e dos incisos II e III: Art. 3º A disciplina do uso da internet no Brasil tem os seguintes princípios: II - proteção da privacidade; III - proteção dos dados pessoais, na forma da lei.

que nada têm a ver com a transação específica que ele executa*".* Sobre esse formato inovador de consulta pública, Fabro Steibel afirma:

> Consultas públicas *online* são uma forma de *e-rulemaking* (criação de políticas públicas online) na qual há um desafio permanente de traduzir princípios de governança em formas de governança (ou seja, de traduzir princípios de administração pública em processos, como no caso do design de software e portais online)[65].

Certos predicados dessa legislação compõem o que se compreende como "Sociedade em Rede", termo utilizado e difundido por Manuel Castells[66] a partir da década de 1990, como sendo "o conceito 'sociedade em rede' enfatiza a forma, o intercâmbio e a organização do processamento de informação. Uma infraestrutura das redes sociais e da mídia se encarrega disso". Assim, o marco regulatório da internet deve possuir uma afinidade com as construções para a proteção do fluxo de informações e seu controle, pois em grande medida é a internet que dá suporte a esse modelo de sociedade.

Na lei consta, ainda, uma seção específica sobre a Proteção aos Registros, aos Dados Pessoais e às Comunicações Privadas. Do artigo 10 ao artigo 12[67] estabelece-

---

[65] STEIBEL, Fabro. **O portal da consulta pública do Marco Civil da Internet**, p. 18-28. In. LEITE, G. L, LEMOS, R. **Marco Civil da Internet.** São Paulo: Altas, 2014. p. 19.

[66] MOLINARO, Carlos Alberto; SARLET, Ingo Wolfgang. **Breves notas acerca das relações entre a sociedade em rede, a internet e o assim chamado estado de vigilância**, p. 29-49. In. LEITE, G. L, LEMOS, R. **Marco Civil da Internet.** São Paulo: Altas, 2014. p. 31.

[67] Redação completa dos artigos 10 ao 12 da Lei 12.965 de 2014: Art. 10. A guarda e a disponibilização dos registros de conexão e de acesso a aplicações de internet de que trata esta Lei, bem como de dados pessoais e do conteúdo de comunicações privadas, devem atender à preservação da intimidade, da vida privada, da honra e da imagem das partes direta ou indiretamente envolvidas.

§ 1º O provedor responsável pela guarda somente será obrigado a disponibilizar os registros mencionados no *caput*, de forma autônoma ou associados a dados pessoais ou a outras informações que possam contribuir para a identificação do usuário ou do terminal, mediante ordem judicial, na forma do disposto na Seção IV deste Capítulo, respeitado o disposto no art. 7º.

§ 2º O conteúdo das comunicações privadas somente poderá ser disponibilizado mediante ordem judicial, nas hipóteses e na forma que a lei estabelecer, respeitado o disposto nos incisos II e III do art. 7º.

§ 3º O disposto no *caput* não impede o acesso aos dados cadastrais que informem qualificação pessoal, filiação e endereço, na forma da lei, pelas autoridades administrativas que detenham competência legal para a sua requisição.

§ 4º As medidas e os procedimentos de segurança e de sigilo devem ser informados pelo responsável pela provisão de serviços de forma clara e atender a padrões definidos em regulamento, respeitado seu direito de confidencialidade quanto a segredos empresariais.

Art. 11. Em qualquer operação de coleta, armazenamento, guarda e tratamento de registros, de dados pessoais ou de comunicações por provedores de conexão e de aplicações de internet em que pelo menos um desses atos ocorra em território nacional, deverão ser obrigatoriamente respeitados a legislação brasileira e os direitos à privacidade, à proteção dos dados pessoais e ao sigilo das comunicações privadas e dos registros.

§ 1º O disposto no *caput* aplica-se aos dados coletados em território nacional e ao conteúdo das comunicações, desde que pelo menos um dos terminais esteja localizado no Brasil.

se, resumidamente, que a guarda e a disponibilização dos dados pessoais e de comunicações privadas devem atender à preservação da intimidade e da vida privada, bem como prevê que a disponibilização de tais informações deverá ocorrer mediante autorização judicial, mas simultaneamente, permite o acesso aos dados cadastrais (qualificação pessoal, filiação, endereço) pelas autoridades administrativas competentes[68].

Também estão expressamente legislados as garantias e os princípios da proteção dos dados pessoais com fundamento na preservação da privacidade, o que faz concluir pela sua enorme relevância legal, especialmente ao se considerar que é no âmbito da internet que ocorre a circulação em massa de informações pessoais. O MCI, portanto, é instrumento jurídico relevante para se discutir situações que envolvam a utilização indevida de dados pessoais.

Neste compasso, entende-se que de toda a legislação infraconstitucional sob análise, o MCI é, sem dúvidas, a lei mais inovadora que abordou o tema. As suas previsões, além de atenderem a preceitos constitucionais, fornecem maior segurança jurídica aos indivíduos e apresentam novos recursos técnicos para se proteger a privacidade e a circulação dos dados pessoais.

Assim, seguindo os mesmos vetores do Marco Civil da Internet, a Lei Geral de Proteção de Dados surge com o fito de garantir o direito digital do cidadão no que se

---

§ 2º O disposto no *caput* aplica-se mesmo que as atividades sejam realizadas por pessoa jurídica sediada no exterior, desde que oferte serviço ao público brasileiro ou pelo menos uma integrante do mesmo grupo econômico possua estabelecimento no Brasil.
§ 3º Os provedores de conexão e de aplicações de internet deverão prestar, na forma da regulamentação, informações que permitam a verificação quanto ao cumprimento da legislação brasileira referente à coleta, à guarda, ao armazenamento ou ao tratamento de dados, bem como quanto ao respeito à privacidade e ao sigilo de comunicações.
§ 4º Decreto regulamentará o procedimento para apuração de infrações ao disposto neste artigo.
Art. 12. Sem prejuízo das demais sanções cíveis, criminais ou administrativas, as infrações às normas previstas nos arts. 10 e 11 ficam sujeitas, conforme o caso, às seguintes sanções, aplicadas de forma isolada ou cumulativa:
I - advertência, com indicação de prazo para adoção de medidas corretivas;
II - multa de até 10% (dez por cento) do faturamento do grupo econômico no Brasil no seu último exercício, excluídos os tributos, considerados os condição econômica do infrator e o princípio da proporcionalidade entre a gravidade da falta e a intensidade da sanção;
III - suspensão temporária das atividades que envolvam os atos previstos no art. 11; ou
IV - proibição de exercício das atividades que envolvam os atos previstos no art. 11.
Parágrafo único. Tratando-se de empresa estrangeira, responde solidariamente pelo pagamento da multa de que trata o caput sua filial, sucursal, escritório ou estabelecimento situado no País.
[68] BRASIL. **Marco Civil da Internet, Lei 12.965 de 2014.** Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm>. Acesso em 11 de jun. 2020.

refere à coleta, manipulação e transferência de seus dados pessoais (digitais, biológicos, laboratoriais, associativos, financeiros, biométricos, parentais, médicos, saúde, judiciais, entre outros).

3 - DIREITO A PRIVACIDADE E A AUTODETERMINAÇÃO INFORMACIONAL

O direito à privacidade, em sua formulação inicial, relacionava-se com a proteção à vida íntima, familiar e pessoal dos indivíduos. Isto é, o conceito de privacidade mostra-se intimamente ligado à definição de liberdade, eis que o exercício do direito à privacidade representa o exercício do direito à liberdade, tanto àquela de se expor ou àquela do poder de decisão de em que medida pretende o titular mostrar sua intimidade e vida privada para o mundo exterior[69].

No mesmo aspecto ensina Gilberto Haddad Jabur:

> O direito à privacidade decorre do direito à liberdade, na medida em que o primeiro abriga o direito à quietude, à paz interior, à solidão e ao isolamento contra a curiosidade pública, em relação a tudo o quanto possa interessar à pessoa, impedindo que se desnude sua vida particular; enquanto o segundo resguarda o direito a uma livre escolha daquilo que o indivíduo pretende ou não expor para terceiros, protegendo o seu círculo restrito da forma como lhe aprouver.

Com isso, pode-se entender como direito à privacidade:

> "... a faculdade que tem cada indivíduo de obstar a intromissão de estranhos em sua vida privada e familiar, assim como de impedir-lhes o acesso a informações sobre a privacidade de cada um, e também impedir que sejam divulgadas informações sobre esta área de manifestação existencial do ser humano"[70].

O doutrinador Danilo Doneda afirma que o direito à privacidade compreende algo mais complexo do que uma simples concepção de vida íntima *"zero-relationship*", devendo abranger também o direito de se relacionar com outros indivíduos[71]. Eis que

---

[69] VIEIRA, Tatiana Malta. **O direito à privacidade na sociedade da informação: efetividade desse direito fundamental diante dos avanços da tecnologia da informação.** Brasília, 2007. p. 296. ps. Dissertação (Mestrado em Direito, Estado e Sociedade). Universidade de Brasília, Brasília, 2007. Pág. 22.

[70] BASTOS, Celso Ribeiro; MARTINS, Ives Gandra. **Comentários à Constituição do Brasil.** São Paulo: Saraiva, 1989, vol. 2, p. 63.

[71] DONEDA, Danilo. **A proteção dos dados pessoais como um direito fundamental**. Espaço Jurídico Journal of Law. vol. 12, n. 2, p. 91-108.

a pessoa pertence a um certo grupo social, e, assim, a constituição da sua identidade não surge de forma disjunta.

Entender a afirmação de que “socializar o uso da informação privada” é indispensável para colocar em um plano de similitude todos aqueles que estão interessados em contribuir, por meio da livre discussão, é importante para a determinação da política do próprio país[72].

Ainda é necessário, que também se reconheça a existência de um núcleo duro do direito à privacidade, considerando-se mecanismos de proteção a serem criados pelo Estado devem ser diferenciados para cada grupo de informações. De modo que medidas legais devem ser tomadas para manutenção do sigilo necessário àquelas informações cuja circulação pode trazer riscos irreparáveis a seus titulares.

O segundo conceito deste tópico é o direito à Autodeterminação Informativa, que se entende como o “direito que cabe a cada indivíduo de controlar e de proteger os próprios dados pessoais, tendo em vista a moderna tecnologia e processamento de informação[73]”. Também é denominado de direito à privacidade informacional e decisional por alguns doutrinadores, em razão de ser uma subespécie do gênero direito à privacidade.

Em suma, a boa doutrina constitucional considera como sendo a existência de um "direito geral à autodeterminação informativa que se traduz, fundamentalmente, na faculdade de o particular determinar e controlar a utilização de seus dados[74]".

Vale dizer, que tal direito foi trazido à baila pela primeira vez pelo Tribunal Constitucional Alemão, no julgamento do caso da Lei Alemã do Censo de 1982. Segundo esta Corte, o direito à autodeterminação informativa “pressupõe que, mesmo sob as condições da moderna tecnologia de processamento de informações (...), que o

[72] RODOTÀ, Stefano. **A vida na sociedade da vigilância** (coord. Maria Celina Bodin de Moraes). Rio de Janeiro: Renovar, 2008. p. 33.
[73] MARTINS, Leonardo (Org.). **Cinquenta anos de jurisprudência do Tribunal Constitucional Federal Alemão. Montevidéu**: Fundação Konrad Adenauer, 2005, p. 233-235.
[74] CANOTILHO, JJ Gomes. **Direito Constitucional e Teoria da Constituição**. Coimbra. Ed. Almedina, 2003. p. 20.

indivíduo exerça sua liberdade de decisão sobre as ações a serem precedidas ou omitidas em relação a seus dados[75]".

"Não existem mais dados insignificantes", essa foi mais uma constatação cirúrgica do Tribunal Constitucional Alemão, neste julgamento que marcou um novo paradigma na tutela jurídica dos dados em todo o mundo.

Vale ainda ressaltar os princípios que tutelam o direito à autodeterminação informativa, os quais podem ser extraídos de dois diplomas legais como a Recomendação da OCDE de 23/09/1980 e a Convenção do Conselho da Europa:

> a. Princípio da correção: deve ser facultado ao cidadão o direito de retificar quaisquer dados coletados a seu respeito, a qualquer tempo sem quaisquer ônus;
>
> b. Princípio da exatidão: os dados coletados devem guardar pertinência exata com os dados fornecidos pelo cidadão àquele destinatário, vedando-se o uso de meios suplementares não autorizados de coletas de dados;
>
> c. Princípio da finalidade: deve haver uma relação direta de pertinência entre as finalidades da ação executada pelo coletor das informações e os dados que podem ser legitimamente coletados. Este princípio veda a licença indiscriminada, genérica e ampla para coletar dados pessoais em quaisquer formulários[15].
>
> d. Princípio da publicidade dos bancos de dados: existência de um registro público prévio, com amplo acesso, dos bancos de dados;
>
> e. Princípio do acesso individual: deve o indivíduo conhecer quais são as informações coletadas sobre si próprio, obter cópias e correção das informações, a integração das incompletas e a eliminação daquelas coletadas ilegitimamente;
>
> f. Princípio da segurança física e lógica: os bancos de dados devem ser mantidos sob estruturas seguras o suficiente para impedir o acesso não autorizados dos dados por terceiros.

Por fim, em 2016, a União Europeia, instituiu o Regulamento Geral sobre a Proteção de Dados (RGPD) (UE) 2016/679, que uniformizou o quadro regulamentar europeu sobre direito à privacidade e proteção de dados pessoais dos cidadãos e residentes na União Europeia e Espaço Económico Europeu[76].

---

[75] VIEIRA, Tatiana Malta. **O direito à privacidade na sociedade da informação: efetividade desse direito fundamental diante dos avanços da tecnologia da informação.** Brasília, 2007. 296 ps. Dissertação (Mestrado em Direito, Estado e Sociedade). Universidade de Brasília, Brasília, 2007. p. 88.

[76] Presidency of the Council: "Compromise text. Several partial general approaches have been instrumental in converging views in Council on the proposal for a General Data Protection Regulation in its entirety. The text on

Assim, baseado nos princípios abordados, para que se promova a proteção aos dados pessoais frente ao progresso tecnológico, deve-se admitir que o cidadão exerça um real poder sobre a exatidão das informações coletadas, os destinatários e a finalidade. Sendo esse o conceito do “direito à autodeterminação informativa”, cujo objetivo é a garantia de que suas informações ante as múltiplas possibilidades de coleta de dados oferecidas pela tecnologia estejam seguras.

Portanto, o direito à autodeterminação informativa surge como um desmembramento do direito à privacidade, com o fito de proteger, de forma efetiva, o conjunto de dados considerados pessoais dos cidadãos, garantindo-lhes o controle sobre eles.

4 - A LEI GERAL DE PROTEÇÃO DE DADOS BRASILEIRA

Neste cenário surgiu a Lei Geral de Proteção de Dados (Lei n.º 13.709/2018), que entrou vigor em agosto de 2020, trazendo consigo novos encargos para empresas e pessoas físicas que lidam com dados pessoais, salvo no tocante às punições que só poderão ser aplicadas a partir de agosto de 2021[77].

Influenciada pelos princípios do RGPD europeu, a LGPD, em seu artigo 3º, define que as pessoas naturais ou jurídicas que coletam ou realizam tratamento de dados pessoais no território brasileiro, independente do meio, do país de sua sede ou da localização dos dados, estão obrigadas em seus termos.

Os principais pontos da LGPD são: direito para o titular acessar, editar ou solicitar a exclusão de seus dados, recolhimento autorizado (com exceção em casos específicos), maior cuidado com dados sensíveis, portabilidade de dados e sanções administrativas se houver descumprimento.

Isto é, a LGPD se situa como meio de efetivação dos direitos da personalidade, por meio de princípios e regras, estabelecendo regulamento nacional sobre o

---

the Regulation which the Presidency submits for approval as a General Approach appears in annex," 201 pages, 11 June 2015, PDF, http://data.consilium.europa.eu/doc/document/ST-9565-2015-INIT/en/pd.

[77] BRASIL, **Senado Federal**. Disponível em: <https://www12.senado.leg.br/assessoria-de-imprensa/notas/nota-de-esclarecimento-vigencia-da-lgpd>. Acesso em: 7 set. 2020.

tratamento de dados, a fim de evitar que ocorram distorções no tratamento de informações consideradas dados pessoais[78].

A aplicação da LGPD objetiva a pessoas físicas e naturais, cujos dados pessoais sejam artefato de tratamento por pessoa jurídica ou natural, pública ou privada, conforme estabelecido em seu artigo gênese:

> Art. 1º Esta Lei dispõe sobre o tratamento de dados pessoais, inclusive nos 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[79].

A LGPD não é aplicada a dados relacionados às pessoas jurídicas, os quais já estão tutelados no campo da propriedade intelectual. Além disso, consoante ao que dispõe o artigo 4º[80], a referida lei também não se aplica ao tratamento de dados pessoais realizados por pessoa natural para fins particulares e não econômicos; ou exclusivamente para fins jornalístico, artístico ou acadêmico; bem como para fins de segurança pública, defesa nacional, segurança do Estado ou atividades de investigação e repressão de infrações penais.

Por meio da implementação, a autonomia do titular ganha relevância no 3º Capítulo da LGPD, que trata sobre os direitos do titular dos dados é essa seção que aborda também as questões da anonimização, bloqueio e direito à eliminação dos dados pessoais.

A LGPD, portanto, traz a previsão de que os titulares dos dados têm o direito de vê-los eliminados de forma segura. Entretanto, não traduz a maneira pela qual esses

---

[78] COELHO, A. C. B. **A Lei Geral de Proteção de Dados Pessoais Brasileira como meio de efetivação dos direitos da personalidade.** João Pessoa: [s.n.], 2019. p. 55.

[79] BRASIL. **Lei Geral de Proteção de Dados, Lei 13.709 de 2018.** Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 11 de jun. 2020.

[80] Art. 4º Esta Lei não se aplica ao tratamento de dados pessoais:
I - realizado por pessoa natural para fins exclusivamente particulares e não econômicos;
II - realizado para fins exclusivamente:
a) jornalístico e artísticos; ou
b) acadêmicos, aplicando-se a esta hipótese os arts. 7º e 11 desta Lei;
III - realizado para fins exclusivos de:
a) segurança pública;
b) defesa nacional;
c) segurança do Estado; ou
d) atividades de investigação e repressão de infrações penais; [...]

dados devem ser descartados. Motivo pelo qual o objetivo dessa dissertação é apresentar possível solução ao mote.

4.1 - Conceitos relativos à proteção de dados

Para melhor esclarecimento do tema proposto no presente estudo, é necessário que se faça breve explicação acerca dos conceitos-chave referente à proteção de dados, nomeadamente de dados pessoais, de tratamento de dados, de titular dos dados, conjuntos especiais de dados e de violação de dados pessoais.

Anteriormente à promulgação da LGPD, a definição de tais conceitos vinha descrita no bojo da correlata legislação europeia. Logo no nº. 1 do artigo 4.º do RGPD, definiu-se como “dados pessoais” a informação relativa a uma pessoa singular identificada ou identificável do “titular dos dados”, sendo considerada uma pessoa singular identificável aquela que possa ser identificada, direta ou indiretamente, em especial por referência a um identificador, como por exemplo um nome, um número de identificação, dados de localização, identificadores por via eletrônica ou a um ou mais elementos da identidade física, fisiológica, genética, mental, económica, cultural ou social dessa pessoa singular[81].

Na mesma norma encontra-se a definição de ”tratamento” como sendo a operação ou conjunto de operações efetuadas sobre dados pessoais ou sobre conjuntos de dados pessoais por meios automatizados ou não automatizados, tais como a recolha, o registo, a organização, a estruturação, a conservação, a adaptação ou alteração, a recuperação, a consulta, a utilização, a divulgação por transmissão, difusão ou qualquer outra forma de disponibilização, a comparação ou interconexão, a limitação, o apagamento ou a destruição[82].

Com isso, espelhando-se no RGPD da UE, a LGPD, também faz uma separação entre dados pessoais e dados pessoais sensíveis. De igual sorte, a LGPD define os conceitos acima no artigo 5º e incluí mais algumas especificidades:

---

[81] **RGPD**. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679#d1e2479-1-1>. Acesso em 5 fev 2021.

[82] Ibidem.

Art. 5º Para os fins desta Lei, considera-se:

I - dado pessoal: informação relacionada a pessoa natural identificada ou identificável;

II - dado pessoal sensível: 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;

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;

IV - banco de dados: conjunto estruturado de dados pessoais, estabelecido em um ou em vários locais, em suporte eletrônico ou físico;

V - titular: pessoa natural a quem se referem os dados pessoais que são objeto de tratamento;

VI - controlador: pessoa natural ou jurídica, de direito público ou privado, a quem competem as decisões referentes ao tratamento de dados pessoais;

VII - operador: pessoa natural ou jurídica, de direito público ou privado, que realiza o tratamento de dados pessoais em nome do controlador;

VIII - encarregado: pessoa natural, indicada pelo controlador, que atua como canal de comunicação entre o controlador e os titulares e a autoridade nacional;

IX - agentes de tratamento: o controlador e o operador;

X - tratamento: toda operação realizada com dados pessoais, como as que se referem a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração;

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;

XII - consentimento: manifestação livre, informada e inequívoca pela qual o titular concorda com o tratamento de seus dados pessoais para uma finalidade determinada;

XIII - bloqueio: suspensão temporária de qualquer operação de tratamento, mediante guarda do dado pessoal ou do banco de dados;

XIV - eliminação: exclusão de dado ou de conjunto de dados armazenados em banco de dados, independentemente do procedimento empregado;

XV - transferência internacional de dados: transferência de dados pessoais para país estrangeiro ou organismo internacional do qual o país seja membro;

XVI - uso compartilhado de dados: comunicação, difusão, transferência internacional, interconexão de dados pessoais ou tratamento compartilhado de bancos de dados pessoais por órgãos e entidades públicos no cumprimento de suas competências legais, ou entre esses e entes privados, reciprocamente, com autorização específica, para uma ou mais modalidades de tratamento permitidas por esses entes públicos, ou entre entes privados;

XVII - relatório de impacto à proteção de dados pessoais: documentação do controlador que contém a descrição dos processos de tratamento de dados pessoais que podem gerar riscos às liberdades civis e aos

> direitos fundamentais, bem como medidas, salvaguardas e mecanismos de mitigação de risco;
>
> XVIII - órgão de pesquisa: órgão ou entidade da administração pública direta ou indireta ou pessoa jurídica de direito privado sem fins lucrativos legalmente constituída sob as leis brasileiras, com sede e foro no País, que inclua em sua missão institucional ou em seu objetivo social ou estatutário a pesquisa básica ou aplicada de caráter histórico, científico, tecnológico ou estatístico; e
>
> XIX - autoridade nacional: órgão da administração pública responsável por zelar, implementar e fiscalizar o cumprimento desta Lei em todo o território nacional.

Portanto, a LGPD sofre clara influência do RGPD, pois ambos os diplomas são bastante semelhantes a respeito dos conceitos relativos à proteção de dados pessoais. Sendo que, como se viu, ambas as leis trazem em seu bojo os aspectos gerais dos direitos do titular sobre suas informações pessoais, como as empresas podem ou não recolher e processar dados e quais as punições cabíveis caso descumpram as regras

A LGPD e o RGPD também se equiparam no conceito de dado pessoal sensível como sendo a informação íntima que precisa de tratamento especial, porque sua divulgação ou utilização inadequada pode causar prejuízos para o titular.

4.2 - Fundamentos e objetivos da LGPD Brasileira

Estudar o tema da proteção de dados sem associá-lo às questões do ciberespaço e internet, atualmente, é absolutamente inconcebível.

A grande circulação se dá pelas redes e pelos meios instantâneos de comunicação. Ao lado de inúmeras facilidades, o processo trouxe consigo a falta de conhecimento acerca da forma como esses dados são administrados, como são capturados, como são guardados, quem os armazena, para quais fins são utilizados, por qual período permanecem acessíveis, dentre várias outras nuances sobre seu conceito[83].

[83] LUZ, Pedro Henrique Machado da; LOUREIRO, Maria Fernanda Battaglin**. Privacidade e proteção de dados pessoais:** os novos desafios na sociedade em rede. Disponível em: <http://www.fumec.br/revistas/meritum/article/view/5811>. Acesso em: 8 jul. 2020.

A Lei Geral de Proteção de Dados Pessoais, portanto, visa regulamentar a permissão e a utilização de dados na internet, sendo que sua aprovação veio para inserir o Brasil entre os países que contam com instrumentos para a abrigo desse importante aspecto do direito fundamental à privacidade[84].

A LGPD é aplicável ao tratamento de dados tanto por pessoas jurídicas de direito público como de direito privado. Assim, o regime geral da tutela dos dados pessoais incide sobre a regulação da Identidade Civil Nacional.

A Lei arrola princípios que devem reger as atividades de tratamento de dados, como boa-fé (objetiva), finalidade, adequação, transparência, entre outros. Exige-se, em suma, que o tratamento de dados seja realizado a partir de propósitos compatíveis com a ordem jurídica, que os dados coletados sejam empregados unicamente nestas finalidades e que o tratamento se dê de modo acautelado e cristalino, fornecendo amplo abrigo à pessoa humana[85].

O que salta aos olhos ao ler a norma é a preocupação com os chamados dados sensíveis, que são aquelas informações pessoais 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[86]".

Para o doutrinador Flávio Tartuce, ao dispor sobre o tratamento dos dados sensíveis, exige consentimento do titular ou de seu responsável legal "de forma específica e destacada, para finalidades específicas", salvo nos casos em que o dado em questão for indispensável para proteção da vida ou da incolumidade física do titular ou de terceiros, cumprimento de obrigação legal ou regulatória pelo controlador, ou outros fins previstos na lei[87].

---

[84] SCHREIBER, Anderson. **Manual de direito civil:** 3. ed. São Paulo: Saraiva Educação, 2020. p. 207.
[85] Ibidem.
[86] Inteligência do artigo 5º, II, da LGPD: Art. 5º Para os fins desta Lei, considera-se: II - dado pessoal sensível: 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.
[87] TARTUCE, Flávio. **Manual de direito civil.** 10. ed. Rio de Janeiro: Forense; São Paulo: Método, 2020. p. 204.

Com isso, o uso de dados sensíveis sem o consentimento do titular somente é possível nas hipóteses em que for indispensável para cumprimento de obrigação legal pelo controlador; tratamento compartilhado de dados necessários à execução, de políticas órgão de pesquisa, pela administração pública, garantida, sempre que possível, a anonimização dos dados pessoais sensíveis; exercício regular de direitos, até mesmo em contrato e em processo judicial, administrativo e arbitral; proteção da vida do titular ou de terceiros; tutela da saúde, em procedimento realizado por profissionais da área da saúde ou por entidades sanitárias; ou garantia da prevenção à fraude e à segurança do titular, nos processos de identificação e autenticação de cadastro em sistemas eletrônicos[88].

Nas observações de Marcos Ehrhardt Junior, entende-se necessário compatibilizar a necessária proteção dos dados pessoais sensíveis, tais como informações relativas ao estado de saúde das pessoas, com a premissa do interesse público. Contudo, sob a perspectiva da coexistência de direitos o importante é proteger o interesse coletivo e não excluir a necessária proteção da pessoa natural[89].

Cumpre ainda destacar, no que se refere às diretrizes e objetivos da Lei, que ela fora aprovada com vetos, importando dizer que o mais relevante deles é aquele aposto à criação de agência reguladora com a atribuição precípua de zelar pela proteção dos dados pessoais. Como razão do veto, foi alegado vício de iniciativa. No ano seguinte, a Lei n.º 13.853/2019 (oriunda da conversão da Medida Provisória n.º 869/2018[90]) veio inserir na LGPD os artigos 55-A e seguintes, criando a Autoridade Nacional de Proteção de Dados (ANPD). Neste aspecto, os parágrafos 1º e 2º do artigo 55-A[91] esclarecem, porém, que:

---

[88] BRASIL. **Lei Geral de Proteção de Dados, Lei 13.709 de 2018.** Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em 11 de jun. 2020.

[89] JÚNIOR, Marcos Ehrhardt. **Privacidade e proteção de dados pessoais durante a pandemia da COVID-19**. Disponível em: <https://direitocivilbrasileiro.jusbrasil.com.br/artigos/824478175/privacidade-e-protecao-de-dados-pessoais-durante-a-pandemia-da-covid-19>. Acesso em: 5 mai. 2020.

[90] Alterava a Lei nº 13.709, de 14 de agosto de 2018, para dispor sobre a proteção de dados pessoais e para criar a Autoridade Nacional de Proteção de Dados, e dava outras providências.

[91] Art. 55-A. Fica criada, sem aumento de despesa, a Autoridade Nacional de Proteção de Dados (ANPD), órgão da administração pública federal, integrante da Presidência da República.

§ 1º A natureza jurídica da ANPD é transitória e poderá ser transformada pelo Poder Executivo em entidade da administração pública federal indireta, submetida a regime autárquico especial e vinculada à Presidência da República.

> [...] tal natureza afigura-se transitória, podendo a Autoridade ser transformada pelo Poder Executivo em entidade da administração pública federal indireta, submetida a regime autárquico especial e vinculada à Presidência da República, devendo tal transformação ser avaliada em até dois anos da data da entrada em vigor da estrutura regimental da ANPD[92].

O apontamento faz-se necessário na medida em que a experiência de outros países mostra que a total independência da Autoridade Fiscalizadora em relação ao Poder Executivo torna-se indispensável, porque o Poder Público, invariavelmente, é um violador da privacidade dos indivíduos.

Outra diretriz estabelecida na Lei e que norteia o presente estudo, é a necessidade de eliminação dos dados. Ou seja, é o direito do cidadão da exclusão de seus dados ou de conjunto de dados armazenados em banco, independentemente do procedimento empregado. Com isso, a LGPD consagra o poder do titular em requerer, junto ao controlador, a eliminação de seus dados pessoais tratados sob a base do consentimento (artigo 18, VI[93]).

Ocorre que, para os casos em que o titular tenha consentido ao controlador o compartilhamento de seus dados com outro agente de tratamento, ao requerer a eliminação junto ao controlador, espera-se que o terceiro envolvido, que recebeu os dados compartilhados, também os elimine.

Nesta esteira, o artigo 18, §6º[94], prevê expressamente o dever do responsável pelo tratamento de informar, de maneira imediata, o pedido de eliminação dos dados do titular a outros agentes na cadeia de tratamento, com os quais tenha compartilhado os mesmos.

Neste aspecto, há de se reconhecer a dificuldade a ser enfrentada pelo controlador para cumprir o requerimento do titular para exclusão integral dos dados.

---

§ 2º A avaliação quanto à transformação de que dispõe o § 1º deste artigo deverá ocorrer em até 2 (dois) anos da data da entrada em vigor da estrutura regimental da ANPD. [...]

[92] SCHREIBER, Anderson. **Manual de direito civil:** 3. ed. São Paulo: Saraiva Educação, 2020. p. 209.

[93] Art. 18. O titular dos dados pessoais tem direito a obter do controlador, em relação aos dados do titular por ele tratados, a qualquer momento e mediante requisição: VI - eliminação dos dados pessoais tratados com o consentimento do titular, exceto nas hipóteses previstas no art. 16 desta Lei.

[94] § 6º O responsável deverá informar, de maneira imediata, aos agentes de tratamento com os quais tenha realizado uso compartilhado de dados a correção, a eliminação, a anonimização ou o bloqueio dos dados, para que repitam idêntico procedimento, exceto nos casos em que esta comunicação seja comprovadamente impossível ou implique esforço desproporcional.

Isso, pois, nos atuais padrões de negócios as dinâmicas de compartilhamento de dados muitas vezes não comportam obrigações contratuais garantindo auditorias e conferições *in loco.*

Além disso, existe uma lacuna referente à limitação técnica do controlador para gerenciar o cumprimento da determinação de eliminação de dados face à empresa terceira, pois sobre ela não possui ingerência. Motivo pelo qual, os tópicos a seguir debruçar-se-ão em técnicas seguras para garantir o cumprimento a Lei.

O objetivo da LGPD é, portanto, evidente: a proteção do cidadão quanto ao uso indiscriminado e imoderado de seus dados pessoais, além institucionalizar a cultura de transparência na guarda dessas informações.

Finalmente, importante dizer que a implantação das diretrizes da LGPD implicará em mudanças nas empresas, cuja não adequação pode gerar a aplicação de advertências até a imposição de multa com base no faturamento da pessoa jurídica, com o alcance de cifras milionárias, sem prejuízos de perdas e danos causados ao titular dos dados.

4.3 - Da Responsabilidade Civil e das Sansões Administrativas

O Capítulo IV, Seção III da LGPD[95], estabelece pontos de responsabilidade civil e reparação de danos, com o objetivo de assegurar ao titular dos dados a efetiva indenização pelos danos causados pelo controlador dos dados.

---

[95] Art. 42. O controlador ou o operador que, em razão do exercício de atividade de tratamento de dados pessoais, causar a outrem dano patrimonial, moral, individual ou coletivo, em violação à legislação de proteção de dados pessoais, é obrigado a repará-lo.
§ 1º A fim de assegurar a efetiva indenização ao titular dos dados:
I - o operador responde solidariamente pelos danos causados pelo tratamento quando descumprir as obrigações da legislação de proteção de dados ou quando não tiver seguido as instruções lícitas do controlador, hipótese em que o operador equipara-se ao controlador, salvo nos casos de exclusão previstos no art. 43 desta Lei;
II - os controladores que estiverem diretamente envolvidos no tratamento do qual decorreram danos ao titular dos dados respondem solidariamente, salvo nos casos de exclusão previstos no art. 43 desta Lei.
§ 2º O juiz, no processo civil, poderá inverter o ônus da prova a favor do titular dos dados quando, a seu juízo, for verossímil a alegação, houver hipossuficiência para fins de produção de prova ou quando a produção de prova pelo titular resultar-lhe excessivamente onerosa.
§ 3º As ações de reparação por danos coletivos que tenham por objeto a responsabilização nos termos do caput deste artigo podem ser exercidas coletivamente em juízo, observado o disposto na legislação pertinente.
§ 4º Aquele que reparar o dano ao titular tem direito de regresso contra os demais responsáveis, na medida de sua participação no evento danoso.

As sanções administrativas aplicáveis do controlador dos dados, estão previstas no Capitulo VIII, Seção I da LGPD[96], sendo a imputação sob a responsabilidade da

Art. 43. Os agentes de tratamento só não serão responsabilizados quando provarem:
I - que não realizaram o tratamento de dados pessoais que lhes é atribuído;
II - que, embora tenham realizado o tratamento de dados pessoais que lhes é atribuído, não houve violação à legislação de proteção de dados; ou
III - que o dano é decorrente de culpa exclusiva do titular dos dados ou de terceiro.
Art. 44. O tratamento de dados pessoais será irregular quando deixar de observar a legislação ou quando não fornecer a segurança que o titular dele pode esperar, consideradas as circunstâncias relevantes, entre as quais:
I - o modo pelo qual é realizado;
II - o resultado e os riscos que razoavelmente dele se esperam;
III - as técnicas de tratamento de dados pessoais disponíveis à época em que foi realizado.
Parágrafo único. Responde pelos danos decorrentes da violação da segurança dos dados o controlador ou o operador que, ao deixar de adotar as medidas de segurança previstas no art. 46 desta Lei, der causa ao dano.
Art. 45. As hipóteses de violação do direito do titular no âmbito das relações de consumo permanecem sujeitas às regras de responsabilidade previstas na legislação pertinente.

[96] Art. 52. Os agentes de tratamento de dados, em razão das infrações cometidas às normas previstas nesta Lei, ficam sujeitos às seguintes sanções administrativas aplicáveis pela autoridade nacional:
I - advertência, com indicação de prazo para adoção de medidas corretivas;
II - multa simples, de até 2% (dois por cento) do faturamento da pessoa jurídica de direito privado, grupo ou conglomerado no Brasil no seu último exercício, excluídos os tributos, limitada, no total, a R$ 50.000.000,00 (cinquenta milhões de reais) por infração;
III - multa diária, observado o limite total a que se refere o inciso II;
IV - publicização da infração após devidamente apurada e confirmada a sua ocorrência;
V - bloqueio dos dados pessoais a que se refere a infração até a sua regularização;
VI - eliminação dos dados pessoais a que se refere a infração;
§ 1º As sanções serão aplicadas após procedimento administrativo que possibilite a oportunidade da ampla defesa, de forma gradativa, isolada ou cumulativa, de acordo com as peculiaridades do caso concreto e considerados os seguintes parâmetros e critérios:
I - a gravidade e a natureza das infrações e dos direitos pessoais afetados;
II - a boa-fé do infrator;
III - a vantagem auferida ou pretendida pelo infrator;
IV - a condição econômica do infrator;
V - a reincidência;
VI - o grau do dano;
VII - a cooperação do infrator;
VIII - a adoção reiterada e demonstrada de mecanismos e procedimentos internos capazes de minimizar o dano, voltados ao tratamento seguro e adequado de dados, em consonância com o disposto no inciso II do § 2º do art. 48 desta Lei;
IX - a adoção de política de boas práticas e governança;
X - a pronta adoção de medidas corretivas; e
XI - a proporcionalidade entre a gravidade da falta e a intensidade da sanção.
§ 2º O disposto neste artigo não substitui a aplicação de sanções administrativas, civis ou penais definidas em legislação específica.
§ 2º O disposto neste artigo não substitui a aplicação de sanções administrativas, civis ou penais definidas na Lei nº 8.078, de 11 de setembro de 1990, e em legislação específica. (Redação dada pela Lei nº 13.853, de 2019)
§ 3º O disposto nos incisos I, IV, V, VI, VII, VIII e IX do caput deste artigo poderá ser aplicado às entidades e aos órgãos públicos, sem prejuízo do disposto na Lei nº 8.112, de 11 de dezembro de 1990 (Estatuto do Servidor Público Federal) , na Lei nº 8.429, de 2 de junho de 1992 (Lei de Improbidade Administrativa) , e na Lei nº 12.527, de 18 de novembro de 2011 (Lei de Acesso à Informação) .
§ 4º No cálculo do valor da multa de que trata o inciso II do caput deste artigo, a autoridade nacional poderá considerar o faturamento total da empresa ou grupo de empresas, quando não dispuser do valor do faturamento no ramo de atividade empresarial em que ocorreu a infração, definido pela autoridade nacional, ou quando o valor for apresentado de forma incompleta ou não for demonstrado de forma inequívoca e idônea.

Autoridade Nacional de Proteção de Dados, e vão de uma simples advertência até a aplicação de multas pecuniárias de R$ 50.000.000,00 (cinquenta milhões de reais) por infração.

Verificamos que, muito além da reparação pecuniária dos titulares de dados alvo de danos causados pelos controladores, a LGPD no que tange às sansões administrativas, busca um efeito pedagógico imediato no tratamento de dados realizado pelos controladores.

4.4 - Crimes contra proteção de dados pessoais – Anteprojeto de lei

Apesar da grande influência do RGPD, na LGPD o legislador brasileiro não cuidou dos aspectos criminais relacionados às más práticas que envolvam dados pessoais.

Por isso, em 26 de novembro de 2019, foi instituída por Ato do Presidente da Câmara dos Deputados, uma Comissão de Juristas liderada pelo ministro Nefi Cordeiro, do Superior Tribunal de Justiça com objetivo de elaborar o Anteprojeto de Lei de Proteção de Dados para segurança pública e investigação criminal.

---

§ 5º O produto da arrecadação das multas aplicadas pela ANPD, inscritas ou não em dívida ativa, será destinado ao Fundo de Defesa de Direitos Difusos de que tratam o art. 13 da Lei nº 7.347, de 24 de julho de 1985, e a Lei nº 9.008, de 21 de março de 1995. (Incluído pela Lei nº 13.853, de 2019)
§ 6º (VETADO). (Incluído pela Lei nº 13.853, de 2019)
§ 7º Os vazamentos individuais ou os acessos não autorizados de que trata o caput do art. 46 desta Lei poderão ser objeto de conciliação direta entre controlador e titular e, caso não haja acordo, o controlador estará sujeito à aplicação das penalidades de que trata este artigo. (Incluído pela Lei nº 13.853, de 2019).
Art. 53. A autoridade nacional definirá, por meio de regulamento próprio sobre sanções administrativas a infrações a esta Lei, que deverá ser objeto de consulta pública, as metodologias que orientarão o cálculo do valor-base das sanções de multa.
§ 1º As metodologias a que se refere o caput deste artigo devem ser previamente publicadas, para ciência dos agentes de tratamento, e devem apresentar objetivamente as formas e dosimetrias para o cálculo do valor-base das sanções de multa, que deverão conter fundamentação detalhada de todos os seus elementos, demonstrando a observância dos critérios previstos nesta Lei.
§ 2º O regulamento de sanções e metodologias correspondentes deve estabelecer as circunstâncias e as condições para a adoção de multa simples ou diária.
Art. 54. O valor da sanção de multa diária aplicável às infrações a esta Lei deve observar a gravidade da falta e a extensão do dano ou prejuízo causado e ser fundamentado pela autoridade nacional.
Parágrafo único. A intimação da sanção de multa diária deverá conter, no mínimo, a descrição da obrigação imposta, o prazo razoável e estipulado pelo órgão para o seu cumprimento e o valor da multa diária a ser aplicada pelo seu descumprimento.

Em novembro de 2020 foi entregue ao Presidente da Câmara dos Deputados a Minuta do Anteprojeto de Lei de Proteção de Dados para segurança pública e investigação criminal[97], que define no 1º artigo seus objetivos:

> Art. 1º Esta Lei dispõe sobre o tratamento de dados pessoais realizado por autoridades competentes para atividades de segurança pública e persecução penal, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural.

O Art. 62[98] o Anteprojeto propõe a alteração do Decreto-Lei n. 2.848, de 7 de dezembro de 1940 (Código Penal), visando a criação de novos tipos penais denominados “Crimes contra a proteção de dados pessoais”.

> **Invasão de dispositivo informático**
>
> Art. 154-A. Acessar indevidamente ou invadir dispositivo informático alheio, conectado ou não à rede de computadores, ou nele instalar vulnerabilidade:
>
> Pena – reclusão, de 1 (um) a 4 (quatro) anos, e multa.
>
> **Petrechos para invasão**
>
> §1º Produzir, oferecer ou difundir dispositivo, programa de computador, técnica ou vulnerabilidade com o intuito de permitir a prática da conduta definida no caput:
>
> Pena – detenção, de 6 (seis) meses a 2 (anos), e multa. (NR)
>
> § 2º A pena será de reclusão, de 2 (dois) a 5 (cinco) anos, se o fato não constitui crime mais grave, se resultar:
>
> I - na obtenção, modificação ou eliminação de:
>
> a) dados de comunicações eletrônicas privadas;
>
> b) segredos profissionais, comerciais ou industriais;
>
> c) informações sigilosas, assim definidas em lei ou decisão judicial;
>
> d) dados pessoais sensíveis.
>
> II - na modificação, alteração ou interrupção do funcionamento de sistema informático ou no impedimento de seu restabelecimento:
>
> § 3º Aumenta-se a pena de um a dois terços se:
>
> I - o crime for praticado por funcionário público em razão do exercício de suas funções;
>
> II – o sistema informático pertencer à Administração Pública, nacional ou estrangeira, ou a organização internacional;
>
> II – o crime for praticado com o fim de obtenção de vantagem indevida;
>
> IV - resultar em prejuízo econômico a outrem;
>
> V - o agente obtiver o controle remoto não autorizado do dispositivo;
>
> VI - o agente obtiver informações classificadas como reservadas, secretas ou ultrassecretas, conforme a lei;
>
> VII - houver divulgação, comercialização ou transmissão a terceiro, a qualquer título, dos dados ou informações obtidas;

---

[97] BRASIL, Minuta do Anteprojeto de Lei de Proteção de Dados para Segurança Pública e Investigação Criminal. Disponível em: https://www.conjur.com.br/dl/anteprojeto-lei-disciplina-protecao.pdf. Acesso em 20 de dezembro de 2020.

[98] Ibidem.

III - houver divulgação, comercialização ou transmissão a terceiro, a qualquer título, dos dados ou informações obtidas.

§4º Considera-se dispositivo informático todo equipamento tecnológico com capacidade computacional, fixo ou móvel.

Transmissão ilegal de dados pessoais

Art. 154-C. Transmitir, distribuir, usar de forma compartilhada, transferir, comunicar, difundir dados pessoais ou interconectar bancos de dados pessoais sem autorização legal:

Pena - reclusão, de 1 (um) a 4 (quatro), anos e multa.

Parágrafo único. Aumenta-se a pena de um a dois terços se:

I - os dados pessoais forem sensíveis ou sigilosos;

II – praticado por funcionário público em razão do exercício de suas funções;

III – praticado com o fim de obtenção de vantagem indevida;

IV – a conduta causar dano ao titular dos dados ou a terceiros a ele relacionados.

Interrupção ou perturbação de serviço telegráfico, telefônico, informático, telemático ou de informação de utilidade pública

Art. 266 - Interromper ou perturbar serviço telegráfico, radiotelegráfico ou telefônico, impedir ou dificultar-lhe o restabelecimento:

Pena - detenção, de um a três anos, e multa.

Parágrafo único. Aplicam-se as penas em dobro se o crime é cometido por ocasião de calamidade pública.

Inserção de dados falsos em sistema informático

Art. 313-A. Inserir ou facilitar a inserção de dados falsos, alterar ou excluir indevidamente dados corretos no sistema informático ou em bancos de dados da Administração Pública com o fim de obter vantagem indevida para si ou para outrem ou para causar dano:

Pena – reclusão, de 3 (três) a 8 (oito) anos, e multa.

Modificação ou alteração não autorizada de sistema informático

Art. 313-B. Modificar ou alterar o funcionamento de sistema informático sem autorização ou solicitação de autoridade competente:

Pena – detenção, de 3 (três) meses a 2 (dois) anos, e multa.

Parágrafo único. As penas são aumentadas de um terço até a metade se da modificação ou alteração resulta prejuízo para a Administração Pública ou para o administrado.

Desobediência

Art. 330…….

Pena - reclusão, de 1 (um) a 3 (três) anos, e multa.”

Art. 63. O artigo 1º, parágrafo único, da Lei 8.137, de 27 de dezembro de 1990, passa a vigorar com a seguinte redação:

“Art. 1º.....

Parágrafo único. A falta de atendimento da exigência da autoridade, no prazo de 10 (dez) dias, que poderá ser convertido em horas em razão da maior ou menor complexidade da matéria ou da dificuldade quanto ao atendimento da exigência, sujeitará o autor à pena de reclusão, de 1 (um) a 3 (anos), e multa.”

Art. 64. O disposto no artigo 69 da Lei n. 9.605, de 12 de fevereiro de 1998, passa a vigorar com a seguinte redação:

“Art. 69.....

Pena – reclusão, de um a três anos, e multa”

Ainda é prematuro dizer como ou quando o Anteprojeto evoluirá para um Projeto de Lei a ser submetido aos debates, eventuais vetos e substitutivos nas diversas comissões do Congresso Nacional, com início na famosa CCJ – Comissão de Constituição e Justiça, que se ocupará de certificar seu alinhamento ao texto constitucional.

Certo é que foi dado o primeiro passo para a criminalização das condutas contra a proteção de dados pessoais e, considerando a velocidade dos acontecimentos na internet com impactos cada vez maiores na esfera jurídica e socioeconômica, não seria pessimista acreditar que sua promulgação não tardará.

4.5. - Aspectos gerais da LGPD Brasileira

4.5.1. - Conceito de proteção de dados e privacidade

O conceito de *dados* pode ser definido como "um conjunto de registros sobre fatos, passíveis de serem ordenados, analisados e estudados para se alcançar conclusões". Estes dados, quando "organizados e ordenados de forma coerente e significativa para fins de compreensão e análise", são denominados de Informação.[99] E, quando se adiciona a palavra "pessoais" ao termo "dados", há uma personalização do conceito, de modo que os "dados pessoais" seriam um conjunto de registros referentes a um indivíduo[100].

Elegeu-se por utilizar o termo privacidade, eis que, segundo conceituação do estudioso Danilo Doneda:

> O termo é específico o suficiente para distinguir-se de outras locuções com as quais eventualmente deve medir-se, como a imagem, honra ou a identidade pessoal; e também é claro o bastante para especificar seu conteúdo, um efeito da sua atualidade. Mas esta escolha não é consequência somente das fragilidades das demais opções: ao contrário, ela revela-se por si só a mais adequada, justamente por unificar os valores expressos pelos termos intimidade e vida privada[101]

---

[99] LACOMBE, Francisco José Masset et al. **Administração – princípios e tendências**. São Paulo: Saraiva, 2003. p. 490.

[100] TAVARES, Letícia Antunes; ALVAREZ, Bruna Acosta**. Da proteção dos dados pessoais: uma análise comparada dos modelos de regulação da Europa, dos Estados Unidos da América e do Brasil.** Disponível em: <http://www.tjsp.jus.br/download/EPM/Publicacoes/ObrasJuridicas/ii%204.pdf?d=636680444556135606>. Acesso em: 7 set. 2020.

[101]DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**. Rio de Janeiro: Renovar, 2006, p. 347.

A ocupação dos dados pessoais justifica-se pela eventual necessidade de conhecer as pessoas, oferecer e fiscalizar serviços (públicos ou privados) e afiançar a segurança do Estado e do coletivo.

Ocorre que, a legislação não aponta, especificamente, a forma como essa proteção de dados deve ser feita, sendo essa a principal ocupação do presente trabalho, como ver-se-á nos tópicos a seguir.

4.5.2. - A Anonimização

No artigo 5º da LGPD, inciso III, consta uma definição de "dado anonimizado", sendo qualquer "dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento"[102].

Ainda, no mesmo artigo, no inciso XI[103], tem-se a acepção do processo que transforma uma informação em um dado anonimizado: a anonimização. Sendo está a "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".

Dados anônimos são aqueles pertinentes a um titular que não pode ser identificado pelo controlador ou outra pessoa, tendo em vista todos os meios e tempo razoavelmente necessários. Esses dados anonimizados não se submetem à aplicação dos marcos regulatórios se puderem ser objeto de reidentificação[104].

Ainda segundo entendimento de Thiago Luis dos Santos Sombra, para se alcançar a anonimização, é necessário que os dados pessoais tenham sido coletados e tratados em conformidade com as normas da categoria, vez que se o processo tiver

---

[102] Redação completa do artigo 5º, inciso III, da LGPD: 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.

[103] 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.

[104] SOMBRA, Thiago Luis dos Santos. **Fundamentos da regulação da privacidade e proteção de dados pessoais**: pluralismo jurídico e transparência em perspectiva. São Paulo: Thomson Reuters Brasil, 2019. p. 170.

algum vício, o controlador responderá pelo período em que os dados eram passíveis de serem associados a uma pessoa identificável[105].

Desta maneira, é admissível interpretar da leitura da nova regulamentação, que será forçoso os entes os envolvidos usarem de técnicas para tornar possível a anonimização. Entretanto a lei não informa como isso deve ser feito.

Isso porque, para ter a exata dimensão da efetividade das técnicas de anonimização empregadas, conforme conclusão do doutrinador Thiago Luis dos Santos Sombra, três perguntas devem ser feitas: 1) Ainda é possível identificar um indivíduo? 2) Ainda é possível vincular dados a um indivíduo? 3) É possível inferir alguma informação referente a um indivíduo[106]?

Se todas as respostas às questões apresentadas forem negativas, o processo de anonimização obteve êxito, pois a reversão não é mais provável entre as premissas de tempo e custo razoáveis. Por isso, para que se tenha a precisa dissociação exigida pela resposta das perguntas, é necessário observar se o dado anonimizado perdeu a capacidade de identificação potencial do indivíduo. Ou seja, o dado deve ter perdido a possibilidade de vinculação de dados a pessoas e inferência de informações[107].

4.5.3. - A Pseudonimização

No Brasil, a definição de pseudonimização está descrita no Artigo 13, § 4º da Lei Geral de Proteção de Dados:

> Para os efeitos deste artigo, a pseudonimização é o tratamento por meio do qual um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo, senão pelo uso de informação adicional mantida separadamente pelo controlador em ambiente controlado e seguro. [108]

---

[105] Ibidem. p. 171.
[106] Ibidem.
[107] Ibidem.
[108] BRASIL. **Lei Geral de Proteção de Dados**, Lei 13.709 de 2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 8 set. 2020.

Quanto ao uso de técnicas, a aplicação da pseudonimização para proteção de dados pessoais, na União Europeia, ocorre de forma ampla, enquanto no Brasil é prevista apenas no tratamento de dados pessoais sensíveis[109].

[109] Redação completa da **Seção II, Do Tratamento de Dados Pessoais Sensíveis**: Art. 11. O tratamento de dados pessoais sensíveis somente poderá ocorrer nas seguintes hipóteses: I - quando o titular ou seu responsável legal consentir, de forma específica e destacada, para finalidades específicas; II - sem fornecimento de consentimento do titular, nas hipóteses em que for indispensável para: a) cumprimento de obrigação legal ou regulatória pelo controlador; b) tratamento compartilhado de dados necessários à execução, pela administração pública, de políticas públicas previstas em leis ou regulamentos; c) realização de estudos por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais sensíveis; d) exercício regular de direitos, inclusive em contrato e em processo judicial, administrativo e arbitral, este último nos termos da Lei nº 9.307, de 23 de setembro de 1996 (Lei de Arbitragem);e) proteção da vida ou da incolumidade física do titular ou de terceiro; f) tutela da saúde, em procedimento realizado por profissionais da área da saúde ou por entidades sanitárias; ou f) tutela da saúde, exclusivamente, em procedimento realizado por profissionais de saúde, serviços de saúde ou autoridade sanitária; ou g) garantia da prevenção à fraude e à segurança do titular, nos processos de identificação e autenticação de cadastro em sistemas eletrônicos, resguardados os direitos mencionados no art. 9º desta Lei e exceto no caso de prevalecerem direitos e liberdades fundamentais do titular que exijam a proteção dos dados pessoais.§ 1º Aplica-se o disposto neste artigo a qualquer tratamento de dados pessoais que revele dados pessoais sensíveis e que possa causar dano ao titular, ressalvado o disposto em legislação específica.§ 2º Nos casos de aplicação do disposto nas alíneas “a” e “b” do inciso II do caput deste artigo pelos órgãos e pelas entidades públicas, será dada publicidade à referida dispensa de consentimento, nos termos do inciso I do *caput* do art. 23 desta Lei.§ 3º A comunicação ou o uso compartilhado de dados pessoais sensíveis entre controladores com objetivo de obter vantagem econômica poderá ser objeto de vedação ou de regulamentação por parte da autoridade nacional, ouvidos os órgãos setoriais do Poder Público, no âmbito de suas competências.§ 4º É vedada a comunicação ou o uso compartilhado entre controladores de dados pessoais sensíveis referentes à saúde com objetivo de obter vantagem econômica, exceto nas hipóteses relativas a prestação de serviços de saúde, de assistência farmacêutica e de assistência à saúde, desde que observado o § 5º deste artigo, incluídos os serviços auxiliares de diagnose e terapia, em benefício dos interesses dos titulares de dados, e para permitir: I - a portabilidade de dados quando solicitada pelo titular; ou II - as transações financeiras e administrativas resultantes do uso e da prestação dos serviços de que trata este parágrafo. § 5º É vedado às operadoras de planos privados de assistência à saúde o tratamento de dados de saúde para a prática de seleção de riscos na contratação de qualquer modalidade, assim como na contratação e exclusão de beneficiários. 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. § 3º A autoridade nacional poderá dispor sobre padrões e técnicas utilizadas em processos de anonimização e realizar verificações acerca de sua segurança, ouvido o Conselho Nacional de Proteção de Dados Pessoais. Art. 13. Na realização de estudos em saúde pública, os órgãos de pesquisa poderão ter acesso a bases de dados pessoais, que serão tratados exclusivamente dentro do órgão e estritamente para a finalidade de realização de estudos e pesquisas e mantidos em ambiente controlado e seguro, conforme práticas de segurança previstas em regulamento específico e que incluam, sempre que possível, a anonimização ou pseudonimização dos dados, bem como considerem os devidos padrões éticos relacionados a estudos e pesquisas.§ 1º A divulgação dos resultados ou de qualquer excerto do estudo ou da pesquisa de que trata o caput deste artigo em nenhuma hipótese poderá revelar dados pessoais.§ 2º O órgão de pesquisa será o responsável pela segurança da informação prevista no caput deste artigo, não permitida, em circunstância alguma, a transferência dos dados a terceiro.§ 3º O acesso aos dados de que trata este artigo será objeto de regulamentação por parte da autoridade nacional e das autoridades da área de saúde e sanitárias, no âmbito de suas competências.§ 4º Para os efeitos deste artigo, a pseudonimização é o tratamento por meio do qual um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo, senão pelo uso de informação adicional mantida separadamente pelo controlador em ambiente

A *contrario sensu* do que diz respeito aos dados anonimizados, que repelem a aplicação direta das leis de proteção de dados, para os dados pseudonimizados não há previsão legal específica sobre quais benefícios ou desobrigações o controlador pode auferir com seu uso.

Do ponto de vista da segurança da informação, a pseudonimização coopera para garantir segurança dos dados, mitigando danos causados por eventuais vazamentos, se os dados afetados forem somente aqueles não identificáveis, sem o acesso aos dados complementares mantidos em apartado.

A interpretação da pseudonimização como medida extra de segurança foi adotada pelo *Working Party 29 na Opinion on Anonymisation Techniques*[110], quando a *Data Protection Directive* 95/46 estava em vigor. A *RGPD* menciona no artigo 25[111] que a pseudonimização é uma forma de se efetivar o princípio de *privacy by design*.

Portanto, a pseudonimização é um processo voltado a disfarçar a identificação de um titular de dados pessoais, de forma a buscar maior nível de segurança, mediante substituição de um atributo exclusivo do titular por outro tipo de registro.

Na ótica do estudioso Thiago Sombra sobre o tema:

> Ao contrário do que muitos sustentam, dados pseudonimização são considerados dados pessoais, ou seja, não envolvem um processo de

controlado e seguro. BRASIL. Lei Geral de Proteção de Dados, Lei 13.709 de 2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 8 set. 2020.

[110] REINO UNIDO. ***Information Commissioner's Office. Anonymisation:*** *Managing Data Protection Risk Code of Practice.* 2012. p. 22. Disponível em: <https://bit.ly/2qwK1xy>. Acesso em 13 set. 2020.

[111] CAPÍTULO IV: Responsável pelo tratamento e subcontratante. Artigo 25. Proteção de dados desde a conceção e por defeito 1. Tendo em conta as técnicas mais avançadas, os custos da sua aplicação, e a natureza, o âmbito, o contexto e as finalidades do tratamento dos dados, bem como os riscos decorrentes do tratamento para os direitos e liberdades das pessoas singulares, cuja probabilidade e gravidade podem ser variáveis, o responsável pelo tratamento aplica, tanto no momento de definição dos meios de tratamento como no momento do próprio tratamento, as medidas técnicas e organizativas adequadas, como a pseudonimização, destinadas a aplicar com eficácia os princípios da proteção de dados, tais como a minimização, e a incluir as garantias necessárias no tratamento, de uma forma que este cumpra os requisitos do presente regulamento e proteja os direitos dos titulares dos dados. 2. O responsável pelo tratamento aplica medidas técnicas e organizativas para assegurar que, por defeito, só sejam tratados os dados pessoais que forem necessários para cada finalidade específica do tratamento. Essa obrigação aplica-se à quantidade de dados pessoais recolhidos, à extensão do seu tratamento, ao seu prazo de conservação e à sua acessibilidade. Em especial, essas medidas asseguram que, por defeito, os dados pessoais não sejam disponibilizados sem intervenção humana a um número indeterminado de pessoas singulares. 3. Pode ser utilizado como elemento para demonstrar o cumprimento das obrigações estabelecidas nos n. os 1 e 2 do presente artigo, um procedimento de certificação aprovado nos termos do artigo 42.

> anonimização, na medida em que o controlador ainda tem condições de identificar o titular. Na pseudonimização, o controlador tem informações adicionais capazes de refazer toda a cadeia de identificação até chegar novamente no titular dos dados.
>
> O objetivo da pseudonimização é coletar os dados adicionais relacionados ao mesmo indivíduo sem a necessidade de saber sua identidade e, consequentemente, reduzir a possibilidade de identificação. Trata-se, portanto, de uma técnica de segurança particularmente relevante para pesquisas estatísticas[112].

Em última análise, na pseudonimização não são estabelecidas vantagens jurídicas específicas na LGPD, para além de ser considerada uma técnica de aperfeiçoamento da segurança da informação.

Com isso, a pseudonimização não tem origem jurídica específica e exaustiva, sendo que o papel dos desenvolvedores e agentes do mercado pode ser considerado indispensável para fins de cumprimento da Lei.

4.5.4. - A Eliminação de dados

O direito à eliminação de dados aparece no artigo 18 da Lei Geral de Proteção de Dados, que estabelece que o titular pode solicitar ao controlador, a qualquer momento e mediante requisição, entre outros: acesso aos dados tratados pelo controlador; correção de dados incompletos, inexatos ou desatualizados; anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou em desconformidade com a LGPD; portabilidade dos dados a outro fornecedor de serviços ou produtos; revogação do consentimento; eliminação dos dados pessoais tratados com o consentimento do titular[113].

O item, entretanto, tem uma advertência: o artigo 16 elenca uma série de hipóteses em que a eliminação dos dados pessoais é dispensada, são elas: o cumprimento de obrigação legal ou regulatória pelo controlador; estudo por órgão de pesquisa (garantida, sempre que possível, a anonimização dos dados pessoais); uso

---

[112] SOMBRA, Thiago Luis dos Santos. **Fundamentos da regulação da privacidade e proteção de dados pessoais**: pluralismo jurídico e transparência em perspectiva. São Paulo: Thomson Reuters Brasil, 2019. p. 159.

[113] BRASIL. **Lei Geral de Proteção de Dados, Lei 13.709 de 2018.** Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 8 set. 2020.

exclusivo do controlador, vedado seu acesso por terceiro, e desde que anonimizados os dados[114].

Isto é, o controlador deve demonstrar que não irá vincular o nome do indivíduo aos dados armazenados, e que não serão utilizados/passados para quaisquer outras organizações, o que geraria uma exceção quanto ao pedido de eliminação dos dados.

Como exposto, o direito a eliminação dos dados compreende-se pelo fato de os indivíduos não carecerem de permanecer vinculados a informações sobre eles que não são completas/corretas, ou que não são mais estimadas como sendo relevantes. Assim, resta constatada a crescente valorização da privacidade, desde os primórdios, como visto nos tópicos anteriores, bem como conectado ao controle dos dados por parte dos titulares.

Tudo isso está incluído um contexto maior, que levou à implementação da LGPD e que reforça a necessidade do controlador agir sempre de acordo, e, com o consentimento do titular.

4.5.5. - O direito a eliminação de dados

A partir de toda explanação legal que abarca da LGPD, é necessário que se faça uma abordagem acerca do que a legislação afirma ser o direito a eliminação de dados.

Com o fito de proporcionar mais controle por parte dos indivíduos sobre as suas próprias informações, a LGPD regulamentou um conceito que já havia sido previsto no Marco Civil da Internet: o direito à eliminação de dados.

Como se viu, o titular dos dados tem o direito de requerer ao responsável pelo tratamento, a eliminação dos seus dados, no sentido de que o responsável cessa a utilização dos mesmos, com fundamento na inexatidão, alteração das finalidades para que forem recolhidos ou tratados posteriormente[115], e ainda, por razões poderosas e

[114] Ibidem.
[115] CASTRO, Catarina Sarmento e. **Direito da Informática, Privacidade e Dados Pessoais**. Coimbra: Almedina, 2005. p. 229-237.

legítimas relacionadas com a sua situação particular dos dados conservados para além do prazo legalmente previsto.

A eliminação dos dados impede o tratamento destes, salvo para apuramento de eventuais responsabilidades e colocados à disposição, por exemplo, dos órgãos administrativos ou judiciais durante o prazo de prescrição de qualquer ilícito contra rede nacional ou criminal[116].

A manipulação indiscriminada de dados pessoais incorporados em um arquivo representa uma violação às normas previstas na LGPD, com enorme impacto no direito ao acesso de dados pessoais. Isso porque, o direito de eliminação dos dados não é uma questão nova na jurisprudência[117].

Com isso, a regulação do direito de eliminação dos dados resulta das características da internet, à comunicação e à universalização da informação.

Por conseguinte, a LGPD dispõe, no direito do titular dos dados, o poder de requerer ao controlador dos dados que elimine essas informações de seus sistemas, sendo essa a conceituação do direito à eliminação dos dados. Todavia, o legislador não informa de que maneira (técnica) isso deve ser feito, cingindo-se a determinar de forma genérica a obrigação do apagamento.

Por fim, repisa-se que, adiante, este trabalho apresentará métodos e técnicas que, se implementadas corretamente, atendem a ordem legal.

4.5.6. - A obrigatoriedade de ter o consentimento do titular

Os artigos 7ºe 8º da LGPD regulam a obrigatoriedade de consentimento do titular dos dados, para que o controlador realize quaisquer movimentações de dos dados

[116] Ibidem.

[117] Recorde-se que o Tribunal do Estado da Califórnia considerou que todos tem direito à felicidade e a garantia de reparar os erros do passado e por isso não se pode atacar a reputação das pessoas, conforme decidido em Melvin vs. Reid, em 1931, devido à produção de um filme - O Kimono Vermelho – em que retratava a vida de ex-prostituta acusada de homicídio e absolvida e o Tribunal de Paris, considerou que as reminiscências da vida privada de um indivíduo fazem parte de seu capital moral e que ninguém tem o direito, mesmo de boa-fé, de publicá-las sem a autorização clara e explícita da pessoa cuja vida está sendo relatada como se fosse a própria, mas ganhou uma relevância sem precedentes no contexto virtual.

pessoais e/ou alteração da finalidade. Por isso, a LGPD exige a apresentação de informações claras e acessíveis sobre quais dados estão sendo coletados.

Desta forma, é indispensável para o controlador ter a certeza de consentimento do titular, não cabendo ao titular buscar informações relacionadas ao uso de seus dados pessoais. Isso segue o espírito do *Privacy by Design*, que determina que qualquer produto ou serviço deve ter a privacidade inerente desde a sua concepção e, também, como arquétipo de configuração.

Vale registrar que, o primeiro descumprimento da GDPR foi identificado em janeiro de 2019 na França, quando o Google foi multado em $57.000.000,00 (cinquenta e sete milhões de dólares) pela Comissão Nacional de Informática e Liberdade (CNIL), um órgão independente francês que regulamenta questões de privacidade, vez que a empresa não esclareceu como a informação pessoal de seus usuários é coletada, o fez com elas e pela não solicitação de consentimento para a exibição de anúncios personalizados[118].

Destarte, a grande relevância do tema, visto que mais de 8 meses depois da entrada em vigor da RGPD, até mesmo empresas como o Google ainda não estavam preparadas o suficiente para garantir esse consentimento de seus usuários e titulares dos dados, bem como não estabeleceram formas seguras para eliminar os dados.

4.5.7. - O término do tratamento de dados

Tratar, armazenar e compartilhar dados não são tarefas simples. No entanto, no contexto da Lei Geral de Proteção de Dados, é necessário redobrar esses acurados e implementar aparelhamentos técnicos. O artigo 15[119] da lei enfatiza esses contextos, falando principalmente sobre o término do tratamento dos dados.

---

[118]PACETE, Gustavo Luiz. **Multa aplicada ao Google é emblemática para a GDPR.** Disponível em: <https://www.meioemensagem.com.br/home/midia/2019/01/22/multa-aplicada-ao-google-e-divisor-de-aguas-para-a-gdpr.html>. Acesso em 10 set. 2020.

[119] Redação completa do artigo 15 e parágrafos: O término do tratamento de dados pessoais ocorrerá nas seguintes hipóteses: I - verificação de que a finalidade foi alcançada ou de que os dados deixaram de ser necessários ou pertinentes ao alcance da finalidade específica almejada; II - fim do período de tratamento; III - comunicação do titular, inclusive no exercício de seu direito de revogação do consentimento conforme disposto no § 5º do art. 8º

A LGPD ainda aborda todos os quesitos legais para o armazenamento e traz recomendações sobre governança de dados, regras para transferência internacional de dados, treinamento, responsabilidade dos controladores, entre outras questões.

O término do tratamento dos dados é o evento que determina o início do prazo para eliminação dos dados. Motivo pelo qual o presente trabalho é tão relevante.

As hipóteses de término do tratamento são abordadas nos incisos I a IV do artigo 15[120]. A regra geral é que os dados devem ser eliminados pelos controladores. Porém essa regra admite quatro exceções, reguladas no art.16.

A primeira é próprio cumprimento legal ou regulatório, como é o caso dos dados de cidadãos reunidos em censo nacional, por exemplo, que devem ser mantidos.

A segunda é a exceção dirigida aos órgãos de pesquisa, como o FIPE, IBGE etc., e determina que os dados podem ser retidos, desde que se garanta a anonimização.

A terceira exceção diz respeito à transferência dos dados a terceiros, desde que os requisitos do tratamento de dados (consentimento, sigilo, segurança, boa-fé, finalidade etc.) tenham sido cumpridos.

E, por último, a quarta exceção trazida no art. 15, é aquela que permite manter os dados para uso exclusivo do controlador, vedado qualquer acesso de terceiro e somente se os dados forem anonimizados.

4.5.8. - Eventos que determinam o término do tratamento de dados

O encerramento pelo atingimento da finalidade proposta, ou quando os dados deixem de ser necessários para o atingimento da finalidade consentida, é a primeira hipótese de término de tratamento.

desta Lei, resguardado o interesse público; ou IV - determinação da autoridade nacional, quando houver violação ao disposto nesta Lei.

[120] Ibidem.

O segundo caso é o fim do período de tratamento. Se uma coleta de dados é realizada durante determinado prazo, para verificar dados específicos daquele lapso temporal, com o final do interim há o término do tratamento de dados.

Além desses, temos o terceiro caso, que é a comunicação do titular dos dados, que deseja revogar seu consentimento, ou solicitar a exclusão de seus dados, nos termos do §5º do artigo 8º e do artigo 18, LGPD[121].

O quarto caso se dá por determinação da Autoridade Nacional, quando houver violação ao disposto na LGPD. Os direitos do titular estão dispostos do artigo 17 ao artigo 22[122], Capítulo III, da LGPD. O artigo 17 define que qualquer pessoa natural é titular de dados pessoais – e assim podem exercer diretamente seus direitos – e tem garantidos os direitos de privacidade, liberdade e intimidade.

Ao avançar para o artigo 18[123], a LGPD materializa essa titularidade. O *caput* do artigo determina o período de exercício do direito e de qual forma processual, além de limitar o titular a exercer direito sobre seus dados e só sobre eles.

---

[121] Ibidem.

[122] Art. 17. Toda pessoa natural tem assegurada a titularidade de seus dados pessoais e garantidos os direitos fundamentais de liberdade, de intimidade e de privacidade, nos termos desta Lei.

[123] Art. 18. O titular dos dados pessoais tem direito a obter do controlador, em relação aos dados do titular por ele tratados, a qualquer momento e mediante requisição: I - confirmação da existência de tratamento; II - acesso aos dados; III - correção de dados incompletos, inexatos ou desatualizados; IV - anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou tratados em desconformidade com o disposto nesta Lei; V - portabilidade dos dados a outro fornecedor de serviço ou produto, mediante requisição expressa, de acordo com a regulamentação da autoridade nacional, observados os segredos comercial e industrial; VI - eliminação dos dados pessoais tratados com o consentimento do titular, exceto nas hipóteses previstas no art. 16 desta Lei; VII - informação das entidades públicas e privadas com as quais o controlador realizou uso compartilhado de dados; VIII - informação sobre a possibilidade de não fornecer consentimento e sobre as consequências da negativa; IX - revogação do consentimento, nos termos do § 5º do art. 8º desta Lei. § 1º O titular dos dados pessoais tem o direito de peticionar em relação aos seus dados contra o controlador perante a autoridade nacional. § 2º O titular pode opor-se a tratamento realizado com fundamento em uma das hipóteses de dispensa de consentimento, em caso de descumprimento ao disposto nesta Lei. § 3º Os direitos previstos neste artigo serão exercidos mediante requerimento expresso do titular ou de representante legalmente constituído, a agente de tratamento. § 4º Em caso de impossibilidade de adoção imediata da providência de que trata o § 3º deste artigo, o controlador enviará ao titular resposta em que poderá: I - comunicar que não é agente de tratamento dos dados e indicar, sempre que possível, o agente; ou II - indicar as razões de fato ou de direito que impedem a adoção imediata da providência. § 5º O requerimento referido no § 3º deste artigo será atendido sem custos para o titular, nos prazos e nos termos previstos em regulamento. § 6º O responsável deverá informar, de maneira imediata, aos agentes de tratamento com os quais tenha realizado uso compartilhado de dados a correção, a eliminação, a anonimização ou o bloqueio dos dados, para que repitam idêntico procedimento, exceto nos casos em que esta comunicação seja comprovadamente impossível ou implique esforço desproporcional. § 7º A portabilidade dos dados pessoais a que se refere o inciso V do *caput* deste artigo não inclui dados que já tenham sido anonimizados pelo controlador. § 8º

Com isso, fecha-se a questão legal acerca das normativas trazidas pela LGPD com relação à proteção, privacidade dos dados e o direito à eliminação de dados pessoais.

5 – PREMISSAS PARA PROTEÇÃO E PRIVACIDADE DE DADOS

A premissa fundamental para que privacidade e proteção de dados sejam efetivas, considera que controlador de dados na figura do EPD seja absolutamente honesto e fiel as suas obrigações, ou seja, somente os profissionais interessados em atender plenamente a legislação estão aptos a beneficiar-se do modelo sugerido, pois neste cenário, os principais adversários podem ser os próprios tutores.

Desta forma, é indispensável que o EPD siga as recomendações e atue com diligência e profissionalismo com suas obrigações legais, pois atos de negligência, imprudência ou imperícia caracterizam grave infração, com reflexos legais diretos, em especial a reparação dos danos causados ao titular dos dados, bem como, intervenções regulatórias.

Por fim, as premissas de honestidade e profissionalismo devem permear todos processos de segurança da informação, fundamentalmente os relacionados à proteção de dados pessoais estabelecidos pela LGPD.

5.1 As Normas

Existem muitas normas que abordam a privacidade dos dados, as principais no âmbito da segurança da informação são a RGPD, a LGPD e a CCPA.

Além dessas, é importante destacar as normas que determinam o padrão de segurança de dados da Indústria de Cartões de Pagamento (PCI DSS), a resolução do Banco Central do Brasil 4.658, além das normas ISO27001:2013, ISO27002:2013, ISO27003:2017 e ISO27701:2019.

O direito a que se refere o § 1º deste artigo também poderá ser exercido perante os organismos de defesa do consumidor.

São essenciais para garantir a privacidade e a segurança da informação em uma organização as seguintes ações: manter estrutura de governança; manter inventário de dados pessoais; manter a política de privacidade e proteção de dados; incorporar privacidade e proteção de dados nas operações; gerenciar riscos de segurança da informação; gerenciar riscos de terceiros; atualizar os avisos; responder a solicitações e reclamações dos titulares de dados; monitorar novas práticas operacionais; manter o programa de gerenciamento de violação de privacidade e proteção de dados; monitorar o tratamento de dados e rastrear critérios externos.

A organização deve, ainda, realizar uma avaliação de risco de privacidade e proteção de dados da empresa, exigir que os funcionários reconheçam e concordem em aderir às políticas de privacidade e proteção de dados e manter programa de treinamento e conscientização.

Além disso, para alcançar esse objetivo é recomendável que a organização designe um profissional capacitado para função de EPD, desenvolva os níveis de privacidade e proteção de dados e convoque a participação de todos os envolvidos na organização em questões de privacidade e proteção de dados.

Por fim, a organização deve manter em lugar seguro o inventário e o mapa de dados pessoais. Desta forma, reduzirá risco de acesso indevido dos dados pessoais sob sua responsabilidade.

6 - A PSEUDONIMIZAÇÃO

Conforme já abordado neste trabalho, a definição de pseudonimização está descrita no Artigo 13, § 4º da LGPD[124]:

> Para os efeitos deste artigo, a pseudonimização é o tratamento por meio do qual um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo, senão pelo uso de informação adicional mantida separadamente pelo controlador em ambiente controlado e seguro.

[124] BRASIL. **Lei Geral de Proteção de Dados**, Lei 13.709 de 2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 8 set. 2020.

A inspiração na lei europeia fica patente, quando o RGPD define a pseudonimização:

> Tratamento de dados pessoais de forma a que deixem de poder ser atribuídos a um titular de dados específico sem recorrer a informações suplementares, desde que essas informações suplementares sejam mantidas separadamente e sujeitas a medidas técnicas e organizativas para assegurar que os dados pessoais não possam ser atribuídos a uma pessoa singular identificada ou identificável[125].

A pseudonimização para proteção de dados pessoais ocorre na União Europeia de forma ampla, ao passo que no Brasil a técnica é prevista apenas no tratamento de dados pessoais sensíveis[126].

A RGPD salienta que a técnica de pseudonimização aplicada aos dados pessoais pode reduzir os riscos para os titulares de dados em questão e auxiliar as organizações, os responsáveis pelo tratamento de dados e os seus subcontratantes a cumprir as suas obrigações de proteção de dados[127].

Esse processo consiste na substituição de dados identificadores por dados inventados, esta técnica também pode ser designada como codificação. Em alguns casos[128] os dados originais devem ser preservados, sendo possível a divulgação apenas dos pseudônimos.

Em muitos casos, a aplicação do processo de pseudonimização é suficiente para proteger os dados do titular, mas não deve ser considerada segura se aplicada isoladamente.

---

[125] UNIÃO EUROPEIA (UE). Regulamento Geral sobre a Proteção de Dados (RGPD). Regulamento (UE) 2016/679 do Parlamento Europeu e do conselho de 27 de abril de 2016 relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). Disponível em: https://www.privacy-regulation.eu/pt/4.htm. Acesso em: 109 ago. 2020

[126] BRASIL. Lei Geral de Proteção de Dados, Lei 13.709 de 2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 8 set. 2020.

[127] UNIÃO EUROPEIA (UE). Regulamento Geral sobre a Proteção de Dados (RGPD). Regulamento (UE) 2016/679 do Parlamento Europeu e do conselho de 27 de abril de 2016 relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). Disponível em: https://www.privacy-regulation.eu/pt/4.htm. Acesso em: 109 ago. 2020.

[128] Por exemplo, na divulgação de casos de investigação criminal os dados originais devem ser mantidos em segurança, mas os pseudónimos podem ser divulgados.

Em regra, a pseudonimização deve ser aplicada concomitantemente a outras técnicas de anonimização, pois quando usada isoladamente, a pseudonimização não resultará em um conjunto de dados anônimos[129], ou seja, o titular do dado provavelmente será identificado indiretamente, então, o processo pode ser revertido.

A aplicação concomitante de técnicas reduz a probabilidade de correlação de um conjunto de dados com a identidade original de um titular determinado, ou seja, é uma medida de segurança útil, mas quando aplicada individualmente não garante a segurança do dado.

O resultado do processo de pseudonimização pode ser independente ou derivado do valor original, no primeiro caso, um número aleatório é gerado pelo controlador ou qualquer identificador único é criado. No segundo caso, é utilizada uma função *hash* ou esquema de criptografia para gerar os identificadores[130].

Os dados originais devem ser mantidos de forma segura pela organização, mas podem ser obtidos e relacionados novamente ao pseudônimo, quando houver necessidade. Neste contexto, os dados originais de identificação não podem ser compartilhados, esses devem ser mantidos em segurança e apenas podem ser utilizados pela organização para resolver quaisquer questões específicas determinadas na LGPD.

Os pseudônimos podem ser gerados aleatoriamente ou de forma determinista[131]. Caso um mesmo pseudônimo seja utilizado para identificar a mesma informação em base de dados diferentes, esse pseudônimo é persistente, ou seja, permite correlação de informações quando múltiplos conjuntos de dados são analisados.

---

[129] "Opinion 05/2014 on Anonymisation Techniques". Article 29 Data Protection Working Party (European Commission), em 10 de Abril de 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf. Acesso em: 10 ago. 2020.

[130] Ibidem.

[131] P. D. P. C. Singapore, Guide To Basic Data Anonymization Techniques, 2018. Disponível em: https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Other-Guides/Guide-to-Anonymisation_v1-(250118).pdf. Acesso em: 10 ago. 2020.

De outro lado, os pseudônimos podem ser do tipo transitório. Neste caso, pseudônimos diferentes são utilizados para apresentar a mesma informação em conjuntos de dados diferentes.

Os pseudônimos persistentes tipicamente fornecem melhor utilidade ao manter a integridade referencial entre conjuntos de dados[132], já os pseudônimos transitórios impossibilitam a referência indireta em conjuntos de dados diferentes.

O processo de pseudonimização mapeia identificadores para pseudônimos de acordo com a definição das regras deste processo. Existem algumas informações adicionais que permitem a associação dos pseudônimos com os identificadores originais; este é o segredo da pseudonimização, ou seja, o segredo de pseudonimização é a tabela de mapeamento entre identificadores e seus pseudônimos[133].

Ocasionalmente, os pseudônimos devem seguir uma estrutura, ou possuir um tipo de dado relacionado ao valor original. Uma possibilidade é gerar os pseudônimos a partir de um processo de criptografia. Nestes casos, geradores especiais de pseudônimos podem ser necessários para criar bases de dados sintéticas.

Na geração de pseudônimos por meio do processo de criptografia, a chave de encriptação utilizada no processo não deve ser partilhada, com vistas a sua proteção contra acessos não autorizados.

O nível de segurança desse processo está ligado à capacidade da organização em garantir que a chave de encriptação não sofra acesso não autorizado, pois tal evento resultaria em grave violação de segurança, que permitiria a descriptografia dos dados.

---

[132] Ibidem.

[133] ENISA. Pseudonymisation techniques and best practices em novembro de 2019. Disponível em: https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices/at_download/fullReport. Acesso em: 10 ago. 2020.

Exemplificando o processo de pseudonimização, a tabela a seguir mostra os dados originais fictícios.

Neste exemplo, os registros identificadores, ou seja, os nomes das pessoas foram substituídos por pseudônimos, assim, esse é um exemplo de pseudonimização que pode ser revertido se for necessário[134]. Antes do processo de pseudonimização os dados originais são:

Nome	**Sexo**	**Nacionalidade**	**Idade**	**Cep**
Aline Campos	Feminino	Brasileira	30	17116-125
Ana Maria	Feminino	Brasileira	32	12778-190
Beatriz da Silva	Feminino	Brasileira	33	17116-125
Bernardo Castro	Masculino	Brasileira	25	19136-019
Ingrid Xavier	Feminino	Brasileira	29	79415-011
José Ferreira	Masculino	Brasileira	20	05176-030
Lucas Batista	Masculino	Brasileira	28	13186-032
Mauricio de Prado	Masculino	Brasileira	36	10170-171
Miguel Barros	Masculino	Brasileira	34	10170-171
Ricardo Alcantara	Masculino	Brasileira	27	05176-030
Tais Almeida	Feminino	Brasileira	31	12778-190
Thiago de Melo	Masculino	Brasileira	26	13186-032
Vanessa Amorim	Masculino	Brasileira	34	19136-019

Tabela 1 - Dados antes da pseudonimização

Após o processo de pseudonimização:

Identificador Único	**Sexo**	**Nacionalidade**	**Idade**	**Cep**
344556	Feminino	Brasileira	30	17116-125
677889	Feminino	Brasileira	32	12778-190
124567	Feminino	Brasileira	33	17116-125
907856	Masculino	Brasileira	25	19136-019
237890	Feminino	Brasileira	29	79415-011
347809	Masculino	Brasileira	20	05176-030

[134] P. D. P. C. Singapore, Guide To Basic Data Anonymization Techniques, 2018. Disponível em: https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Other-Guides/Guide-to-Anonymisation_v1-(250118).pdf. Acesso em: 10 ago. 2020.

446773	Masculino	Brasileira	28	13186-032
908756	Masculino	Brasileira	36	10170-171
743927	Masculino	Brasileira	34	10170-171
374610	Masculino	Brasileira	27	05176-030
769312	Feminino	Brasileira	31	12778-190
111037	Masculino	Brasileira	26	13186-032
297462	Masculino	Brasileira	34	19136-019

Tabela 2 - Dados Após a pseudonimização

O dado que possibilita relacionar o nome da pessoa ao seu pseudônimo, aqui definido como “Identificador Único”, é a base de identidade e deve ser mantida em segurança pela organização.

Caso sobrevenha uma razão legítima para identificar indivíduos, essa base poderá ser consultada. Portanto, os controles de segurança também devem ser usados para proteger esses dados:

Identificador Único	**Nome**
344556	Aline Campos
677889	Ana Maria
124567	Beatriz da Silva
907856	Bernardo Castro
237890	Ingrid Xavier
347809	José Ferreira
446773	Lucas Batista
908756	Mauricio de Prado
743927	Miguel Barros
374610	Ricardo Alcantara
769312	Tais Almeida
111037	Thiago de Melo
297462	Vanessa Amorim

Tabela 3 – Base de Identidade

6.1 - Os riscos à pseudonimização

Se os dados forem anonimizados corretamente, não há risco de reidentificação do titular dos dados. No entanto, no processo de pseudonimização, mesmo que tenha

assegurado os mais altos níveis de segurança, sempre haverá o risco de reidentificação do titular dos dados. Portanto, contrariamente aos dados anonimizados, os dados pseudonimizados ainda estão no âmbito de aplicação do RGPD, já que existe risco substancial de reidentificação dos titulares dos dados[135]. Sendo certo que, a afirmação também se aplica a LGPD.

Na pseudonimização, ainda que o responsável pelo tratamento de dados aplique todas a orientações relativas à segurança da informação e proceda com as devidas precauções, ele não pode garantir a total eliminação dos riscos de reidentificação dos titulares dos dados pessoais.

O indivíduo não precisa de estar identificado para os dados estarem sobre a égide da legislação de proteção de dados, a mera possibilidade de identificação do titular dos dados garante a proteção legal. Assim sendo, é possível definir o grau de identificabilidade dos dados de acordo com as técnicas de segurança aplicadas.

Os incidentes de violação de dados pessoais também são um atentado às regras da segurança da informação, tanto no vazamento de dados pessoais quanto no descumprimento de regras de segurança há comprometimento da confidencialidade dos dados. No entanto, atualmente, a proteção aos dados pessoais tem um alcance diferente das infrações de segurança da informação. O primeiro é protegido por lei em sentido formal enquanto o segundo é regido por normas de segurança, essas normas não são leis criadas pelo estado, elas são recomendações fomentadas pela iniciativa privada.

Consequentemente, há incidentes de segurança da informação quando não ocorre a violação de dados pessoais. Por exemplo, em um cenário de acesso não autorizado a dados estatísticos ou dados de recenseamento que não contenha qualquer informação específica de um determinado indivíduo. Trata-se de uma transgressão às regras segurança da informação, mas não uma violação de dados

[135] SEQUEIRA, C. S. A. “Identidade Digital – o Espectro desde a Anonimização à Identificação,” Master’s thesis, Instituto Superior Técnico, 2019.

pessoais. Já o inverso não se aplica, porquanto o vazamento de dados pessoais sempre implica um incidente de segurança da informação[136].

A grande quantidade de dados pessoais gerada na internet pelo próprio titular dos dados, pode facilitar o ataque. Ainda que de modo isolado, esses dados podem parecer anônimos, ao combinar esses dados com outras fontes, pode-se eventualmente re-identificar o titular dos dados.

Um conjunto de dados pode sofrer vários tipos de ataque de reidentificação. Atualmente há muitas pesquisas que se propõem a demonstrar como um conjunto de dados pode ser re-identificado, por exemplo os dados públicos inseridos em redes sociais[137], divulgação de dados de estudos demográficos[138], dados de estudos de sequências do genoma[139], informações de padrões de mobilidade[140], avaliações de filmes[141] e até mesmo o estilo de escrita[142] das pessoas podem ser utilizados para ataques de reidentificação.

---

[136] COSTA, G. M. N. P. "Notificações de violações de dados: a mudança de paradigma com o regulamento geral de proteção de dados," Master's thesis, Instituto Superior Técnico, 2018.

[137] D. J. Crandall, L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, and J. Kleinberg, "Inferring social ties from geographic coincidences", Proceedings of the National Academy of Sciences, vol. 107, no. 52, pp.22 436–22 441, 2010.

[138] P. Golle, "Revisiting the uniqueness of simple demographics in the us population," in Proceedings of the 5th ACM Workshop on Privacy in Electronic Society, 2006, pp. 77–80.

[139] GYMREK M [et al.] - Identifying personal genomes by surname inference. In: Science.

HOMER, Nils [et al.] - Resolving individuals contributing trace amounts of DNA to highly complex mixtures using highdensity SNP genotyping microarrays. In: PLoS Genet.

[140] DE MONTJOYE [et al.] -Unique in the Crowd: The privacy bounds of human mobility. In. Sci Rep.

GAMBS, Sébastien; KILLIJIAN, Marc-Olivier; NÚÑEZ DEL PRADO CORTEZ, Miguel - De-anonymization attack on geolocated data. In: Journal of Computer and System Science.

[141] NARAYANAN, Arvind; SHMATIKOV, Vitaly - Robust De-anonymization of Large Datasets. In: SP '08 Proceedings of the 2008 IEEE Symposium on Security and Privacy.

[142] NARAYANAN, Arvind [et al.] - On the feasibility of internet-scale author identification. In: Proceedings of 33rd IEEE Symposium on Security and Privacy.

MARKOV, Ilia; BAPTISTA, Jorge; PICHARDO-LAGUNAS, Obdulia - Authorship Attribution in Portuguese Using Character N-grams. In: Acta Polytechnica Hungarica.

Muitos desses ataques utilizam dados divulgados por seu próprio titular em suas redes sociais como: a morada[143]; o gênero[144]; atribuição de autoria[145], dentre outros. Em um conjunto de dados pode-se encontrar vários níveis do risco de reidentificação, pois quanto mais informação o usuário divulga em suas redes sociais, maior o risco de reidentificação. Desta forma, quanto maior a eficácia da aplicação das técnicas de anonimização, menor o risco de reidentificação do titular dos dados.

No entanto o aumento da segurança tem reflexo negativo na utilidade prática dos dados, pois quanto maior a anonimização, menor inidentificabilidade[146].

7 - A ANONIMIZAÇÃO DE DADOS

No processo de anonimização, os dados que possibilitam a identificação do titular são removidos ou modificados. Em outras palavras, os dados submetidos ao processo de anonimização não podem ser relacionados a qualquer indivíduo específico.

Quanto aos tipos, a anonimização, essa será simples quando executada com a aplicação de uma única técnica, o que, aliás, limita bastante sua eficácia. A anonimização será composta quando processada por meio da aplicação em conjunto de mais de uma técnica, ampliando consideravelmente sua eficácia.

Existem várias técnicas de anonimização, algumas podem modificar os dados de modo significativo, outras modificam apenas partes dos dados. Há técnica que substitui o valor de um atributo em vários registos, existem outras que substituem o

[143] ELMONGUI, Hicham G.; MORSY, Hader; MANSOUR, Riham - Inference models for Twitter user's home location prediction. 2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA).
[144] DEITRICK, William [et al.] - Gender Identification on Twitter Using the Modified Balanced Winnow, In: Communications and Network;
BURGER, John [et al.] - Discriminating gender on Twitter. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP '11).
[145] YAN, Jinny; MATTHEWS, Suzanne - Applying clustering algorithms to determine authorship of chinese twitter messages, In: IEEE MIT Undergraduate Research Technology Conference (URTC).
ALMISHARI Mishari [et al.] - Stylometric linkability of Tweets. In: Proceedings of 13th Workshop on Privacy in the Electronic Society.
[146] De inidentificável, segundo o dicionário é aquilo que não se consegue identificar, in Dicionário Online de Português, https://www.dicio.com.br/inidentificavel/ [consultado em 16-02-2021].

atributo com informação não relacionada, mas consistente. Além disso, algumas técnicas removem algum atributo totalmente[147].

Como já dito, não é recomendável a aplicação de técnicas de anonimização isoladamente, elas podem e devem ser combinadas entre si, visando a maior segurança no resultado. Por exemplo, remover alguma informação após a aplicação da técnica de generalização. A seguir está o modelo de processo de anonimização:

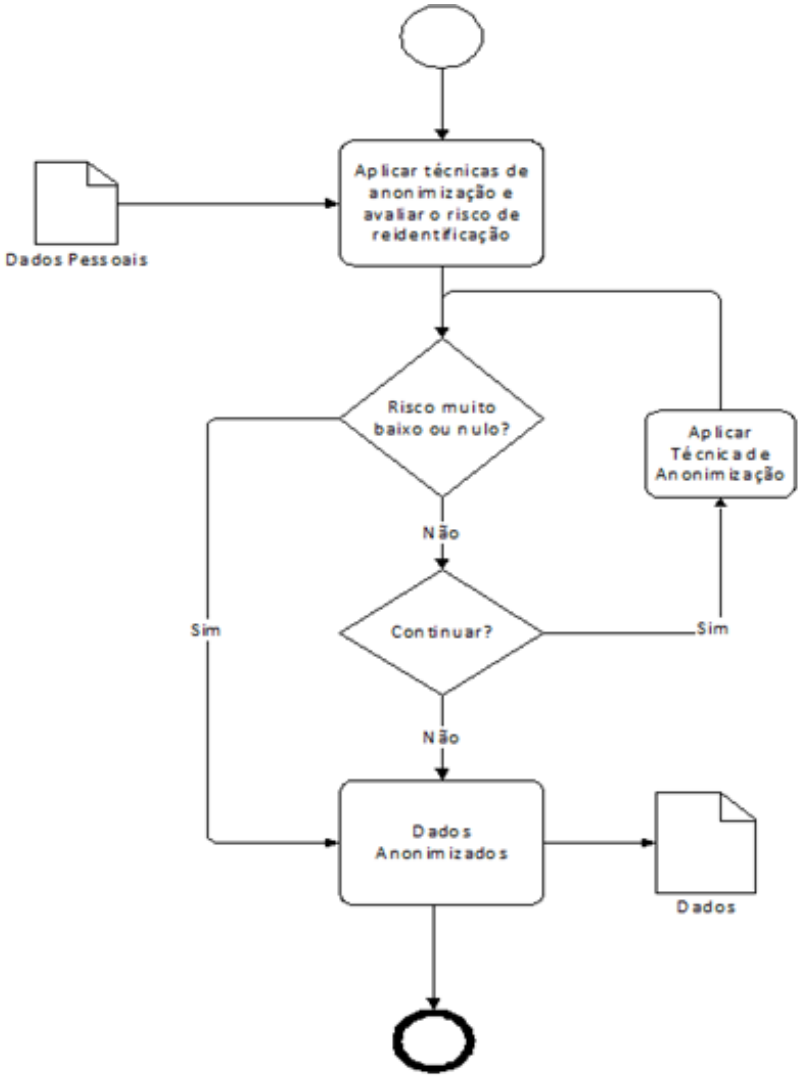

Figura 1 Fluxograma Aplicação de Técnicas de Anonimização
Fonte: Compilação do Autor[148].

[147] "Opinion 05/2014 on Anonymisation Techniques". Article 29 Data Protection Working Party (European Commission), em 10 de Abril de 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf. Acesso em: 10 ago. 2020.

[148] Imagem Elaborada pelo autor a partir de imagens coletadas no Guia para Técnicas Básicas de Anonimização de Dados, Governo da Região Administrativa Especial de Macau Gabinete para a Protecção de Dados Pessoais, 2019.

7.1 - Técnicas de Anonimização

Há diferentes técnicas de anonimização, cada uma com características e aplicabilidades específicas. Antes da utilização, recomenda-se a análise e o alinhamento da técnica em relação ao cenário existente e ao objetivo desejado.

Para anonimizar dados com identificadores diretos, a técnica de encobrimento de caracteres é a mais recomendada, para a anonimização de informações com identificadores indiretos, a técnica de agregação é a mais apropriada[149].

As técnicas de generalização são mais pertinentes para os dados que podem ser agrupados em um conjunto com características semelhantes. Por exemplo, a classificação de animais, que podem ser agrupados como mamíferos, aves, répteis, peixes, anfíbios etc.

Em alguns casos, não é conveniente ou existe a impossibilidade legal de anonimização completa dos dados, por exemplo, estudos de intervenção, dados médicos, dados de investigação policial.

Nestes casos, então, deve-se procurar uma anonimização parcial, ou seja, os dados serem codificados e a chave de leitura desse código, que permite ligar os dados ao titular, fique à guarda de quem se responsabilize pela proteção dos dados, conforme demonstrado no processo de pseudonimização.

Se os dados puderem ser completamente anonimizados deve-se optar por essa solução e, assim, garantir que se cumpra integralmente a LGPD. Desde que se garanta a completa anonimização dos dados, ou seja, a impossibilidade de identificar o titular dos dados, o processo estará alinhado a LGPD. Vale ressaltar que todo o procedimento deve ser auditável e permanecer à disposição dos entes fiscalizadores.

---

[149] P. D. P. C. Singapore, Guide To Basic Data Anonymization Techniques, 2018. Disponível em: https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Other-Guides/Guide-to-Anonymisation_v1-(250118).pdf. Acesso em: 10 ago. 2020.

A partir de agora, passa-se a analisar as principais técnicas de anonimização de dados[150], são elas:

a) Anatomização – baseada na estruturação de todos os atributos semi-identificadores e identificadores em duas tabelas distintas;

b) Generalização – baseada na substituição dos valores dos atributos semi-identificadores por valores semanticamente semelhantes, mas menos específicos;

c) Supressão – baseada na remição ou substituição de um ou mais atributos em um conjunto de dados por algum valor especial;

d) Permutação – baseada na alteração de atributos entre os dados do conjunto;

e) Perturbação – baseada na inserção de dados externo ao conjunto de dados original;

f) Dados Sintéticos – baseada na substituição dos dados reais por dados fictícios;

g) Agregação de dados – baseada no agrupamento de dados semelhantes em um conjunto de dados menos específico.

7.1.1 - A Anatomização

Esta técnica de anonimização dissocia a relação entre os atributos semi-identificadores e os atributos sensíveis, e produz duas tabelas distintas com atributos não sobrepostos[151].

---

[150] Idem

[151] Susan, V.S.; Christopher, T. Anatomisation with slicing: A new privacy preservation approach for multiple sensitive attributes. SpringerPlus 2016, 5, 964.

A anatomização, em contraste com a generalização e a supressão, não faz modificações nos atributos semi-identificadores ou nos atributos sensíveis, mas interrompe o elo entre os dois elementos[152].

Atributos semi-identificadores são aqueles atributos que podem ser associados a dados externos e, assim, expor o indivíduo. Ou ainda, podem reduzir a incerteza sobre a identidade do titular dos dados, por exemplo: data do nascimento, cargo, função, tipo sanguíneo.

Os atributos sensíveis são os que contêm dados sensíveis sobre os indivíduos, por exemplo: doenças, salário, exames médicos, lançamentos do cartão de crédito[153]. Estes atributos também são chamados de atributos confidenciais.

Portanto esta técnica separa os dados semi-identificadores e os dados sensíveis em duas tabelas distintas. Uma tabela com os semi-identificadores, e outra contendo os dados sensíveis, sendo que as duas tabelas têm um atributo comum, que é o grupo a que pertencem.

Esse Identificador do grupo é o elemento que indica a segurança do processo. Quanto maior o número de identificadores distintos, mais improvável é a descoberta de atributos sensíveis dos titulares dos dados.

O maior benefício da anatomização está na manutenção dos dados originais tanto na tabela com os semi-identificadores quanto na tabela que contém os dados sensíveis. Esta abordagem minimiza a perda de informações e auxilia na preservação da correlação e utilidade dos dados.

Além disso, pode-se aplicar às tabelas anatomizadas outros métodos de anonimização, como por exemplo a permutação.

---

[152] Xiao X, Tao Y (2006) Anatomy: simple and effective privacy preservation. In: Proceedings of international conference on very large data bases (VLDB), 1 September 2006.

[153] VIMERCATI, Sabrina de Capitani. FORESTI, Sara; LIVRAGA, Giovanni.; SAMARATI, Pierangela. Data Privacy: Definicitions and Techiniques. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems Vol. 20, No. 6 (2012) 793–817 World Scientific Publishing Company, 2012.

O conjunto de dados disponibilizado para este exemplo consiste em dados de funcionários de uma organização e seus registros médicos, os atributos são: Sexo, Data de Nascimento, Cargo, Salário, Enfermidade, Data Diagnóstico e o Médico do funcionário.

Código	Sexo	Data Nascimento	Cargo	Salário	Enfermidade	Data diagnóstico	Médico
1931	Masculino	06/01/1979	Analista	R$ 4 800,00	LER	01/04/2019	Carlos
1932	Feminino	07/03/1980	Técnico	R$ 2 800,00	DORT	22/05/2018	Beatriz
1933	Feminino	01/04/1980	Advogado	R$ 5 200,00	Ansiedade	10/02/2020	Maria
1934	Masculino	10/02/1981	Engenheiro	R$ 5 000,00	Estresse	04/12/2019	Thiago
1935	Masculino	09/04/1981	Contador	R$ 4 900,00	Estresse	05/05/2019	Thiago
1936	Feminino	10/06/1982	Técnico	R$ 2 800,00	Depressão	06/01/2020	Felipe
1937	Feminino	13/04/1982	Analista	R$ 4 800,00	DORT	07/03/2020	Carlos
1938	Masculino	22/05/1982	Técnico	R$ 2 800,00	LER	08/04/2019	Beatriz
1939	Masculino	11/04/1983	Técnico	R$ 2 800,00	LER	09/03/2019	Beatriz
1940	Masculino	05/05/1983	Advogado	R$ 5 200,00	Estresse	10/02/2020	Thiago
1941	Feminino	12/04/1984	Analista	R$ 4 800,00	Depressão	11/01/2020	Felipe
1942	Masculino	04/12/1985	Engenheiro	R$ 5 000,00	DORT	12/04/2019	Carlos
1943	Feminino	08/02/1986	Contador	R$ 4 900,00	Ansiedade	13/05/2020	Thiago
1944	Masculino	14/03/1986	Técnico	R$ 2 800,00	DORT	14/11/2018	Carlos
1945	Feminino	15/05/1988	Engenheiro	R$ 5 000,00	Ansiedade	15/10/2019	Maria
1946	Feminino	16/06/1989	Técnico	R$ 2 800,00	LER	16/09/2018	Carlos

Tabela 4 - Dados antes da Anatomização

Sexo	Data Nascimento	Cargo	Salário
Masculino	06/01/1979	Analista	R$ 4 800,00
Feminino	07/03/1980	Técnico	R$ 2 800,00
Feminino	01/04/1980	Advogado	R$ 5 200,00
Masculino	10/02/1981	Engenheiro	R$ 5 000,00
Masculino	09/04/1981	Contador	R$ 4 900,00
Feminino	10/06/1982	Técnico	R$ 2 800,00
Feminino	13/04/1982	Analista	R$ 4 800,00
Masculino	22/05/1982	Técnico	R$ 2 800,00
Masculino	11/04/1983	Técnico	R$ 2 800,00
Masculino	05/05/1983	Advogado	R$ 5 200,00
Feminino	12/04/1984	Analista	R$ 4 800,00
Masculino	04/12/1985	Engenheiro	R$ 5 000,00
Feminino	08/02/1986	Contador	R$ 4 900,00
Masculino	14/03/1986	Técnico	R$ 2 800,00
Feminino	15/05/1988	Engenheiro	R$ 5 000,00
Feminino	16/06/1989	Técnico	R$ 2 800,00

Enfermidade	Data diagnóstico	Médico
LER	01/04/2019	Carlos
DORT	22/05/2018	Beatriz
Ansiedade	10/02/2020	Maria
Estresse	04/12/2019	Thiago
Estresse	05/05/2019	Thiago
Depressão	06/01/2020	Felipe
DORT	07/03/2020	Carlos
LER	08/04/2019	Beatriz
LER	09/03/2019	Beatriz
Estresse	10/02/2020	Thiago
Depressão	11/01/2020	Felipe
DORT	12/04/2019	Carlos
Ansiedade	13/05/2020	Thiago
DORT	14/11/2018	Carlos
Ansiedade	15/10/2019	Maria
LER	16/09/2018	Carlos

Tabela 5 - Semi-identificadores e dados Sensíveis

O primeiro passo deste método de anonimização é separar o conjunto de dados em duas tabelas, a primeira tabela contém os dados semi-identificadores, e a segunda contém os dados sensíveis.

Subsequentemente, agrupa-se o conjunto de dados, em grupos distintos e permutam-se os valores confidenciais dentro de cada grupo.

Código	Enfermidade	Médico	Grupo
1932	DORT	Beatriz	1
1938	LER		
1939			
1931	LER	Carlos	2
1946			
1937	DORT		
1942			
1944			
1936	Depressão	Felipe	3
1941			
1933	Ansiedade	Maria	4
1945			
1934	Estresse	Thiago	5
1935			
1940			
1943	Ansiedade		

Código	Sexo	Cargo	Salário	Grupo
1933	Feminino	Advogado	R$ 5 200,00	1
1940	Masculino			
1937	Feminino	Analista	R$ 4 800,00	2
1941				
1931	Masculino			
1943	Feminino	Contador	R$ 4 900,00	3
1935	Masculino			
1945	Feminino	Engenheiro	R$ 5 000,00	4
1934	Masculino			
1942				
1932	Feminino	Técnico	R$ 2 800,00	5
1936				
1946				
1938	Masculino			
1939				
1944				

Código	Data Diagnóstico	Grupo
1932	22/05/2018	1
1946	16/09/2018	
1944	14/11/2018	
1939	09/03/2019	2
1931	01/04/2019	
1938	08/04/2019	
1942	12/04/2019	
1935	05/05/2019	
1945	15/10/2019	
1934	04/12/2019	
1936	06/01/2020	3
1941	11/01/2020	
1933	10/02/2020	
1940	10/02/2020	
1937	07/03/2020	
1943	13/05/2020	

Tabela 6 - Grupos

O produto deste processo é uma tabela que contém todas as informações que podem ser divulgadas, ou seja, este método de anonimização possibilita a publicação de todos os dados disponíveis em uma única tabela e, ainda assim, a privacidade é preservada[154].

Nascimento	(sexo, cargo, salário)	(enfermidade, médico)	Data diagnóstico
01/1979	2	2	2
03/1980	5	1	1
04/1980	1	4	3
02/1981	4	5	2
04/1981	3	5	2
06/1982	5	3	3
04/1982	2	2	3
05/1982	5	1	2
04/1983	5	1	2
05/1983	1	5	3
04/1984	2	3	3
12/1985	4	2	2
02/1986	3	5	3

[154] Xiao X, Tao Y (2006) Anatomy: simple and effective privacy preservation. In: Proceedings of international conference on very large data bases (VLDB), 1 September 2006.

03/1986	5	2	1
05/1988	4	4	2
06/1989	5	2	1

Tabela 7 - Danos Anonimizados por Anatomização

Desta forma, a técnica garante que os dados divulgados não podem ser revertidos e associado a qualquer titular especificamente e, ainda, minimiza a perda de informações por meio da liberação direta dos atributos semi-identificadores.

7.1.2 - A Generalização

Esta abordagem consiste em generalizar, ou seja, diminuir a especificidade dos dados identificadores. Esta redução na precisão dos dados é possível modificando a escala ou a sua ordem de magnitude[155].

Esta técnica pode ser uma alternativa quando se busca o equilíbrio entre utilidade e privacidade dos dados. No Brasil, um exemplo de generalização é o código de endereçamento postal (CEP). Os números representam o escopo geográfico, ou seja, quanto mais à esquerda o número, maior o seu alcance[156].

Outro exemplo de generalização é a conversão a idade de uma pessoa em uma faixa etária, ou de um ponto preciso para um ponto menos preciso. Esta técnica também é denominada de recodificação.

É importante que os valores substituídos mantenham a finalidade dos dados, pois valores muito genéricos podem inviabilizar a utilização dos dados. Desta forma, a aplicação da técnica deve promover a generalização dos dados e, ainda assim, preservar sua utilidade para o objetivo pretendido.

[155] "Opinion 05/2014 on Anonymisation Techniques". Article 29 Data Protection Working Party (European Commission), em 10 de Abril de 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf. Acesso em: 10 ago. 2020.

[156] CORREIOS, B. Estrutura CEP. [S.l.: s.n.], 2018. https://www.correios.com.br/enviar-e-receber/ferramentas/cep/estrutura-do-cep. Acesso em: 13 ago. 2020.

Por exemplo, faixas etárias muito grandes podem implicar que os dados sejam muito “modificados”, enquanto as faixas muito pequenas podem significar que os dados sejam pouco alterados e por isso fáceis de reidentificar.

Os autores Nergiz e Gök[157] salientam que uma das vantagens da generalização é a preservação da veracidade dos dados e, De Capitani di Vimercati[158] relata que a combinação de generalização e de supressão pode disponibilizar informações mais verdadeiras, embora, não completas, isso devido à redução dos detalhes dos dados que serão divulgados, preservando assim a utilidade dos dados.

A tabela a seguir foi usada como a base de dados original. Neste exemplo, o conjunto de dados contém um número de sequência, um identificador único para cada registro, a idade e o endereço:

Número	Id. Único	Idade	Endereço
1	340928	25	Rua das Peras, 58 – Planalto, Cuiabá/MT
2	230956	33	Rua Presidente Torres, 104 – Jardins, Sorriso/MT
3	450967	42	Rua da Saudade, 2910 – Verdão, Boa Vista/RR
4	102938	45	Rua Primavera, 382 – Morumbi, Santarém/PA
5	382910	22	Rua Portugal, 010 – Cipoal, Campinas/SP
6	467843	75	Rua Brasília, 4927 – Itaipava, Diadema/SP
7	012935	29	Rua das Pedras, 1936 – Caçari, Osasco/SP
8	926530	17	Rua Primeira, 1932 – Abial, Tefé/AM
9	173029	31	Rua das Flores, 1599 – Periperi, Salvador/BA
10	294710	36	Rua Alegrete, 1023 – Ibiapaba, Cariacica/ES
11	392185	20	Rua Amazonas, 2847 – Setiba, Serra/ES
12	402915	24	Rua Apa, 2201 – Prado, Belo Horizonte/MG
13	229577	40	Rua Camilo Rama, 2011 – Caravelas, Ipatinga/MG

Tabela 8 - Dados antes da Generalização

Uma abordagem possível para a generalização da idade é substituir essa informação por uma faixa etária definida anteriormente.

< 20
21-30
31-40

[157] NERGIZ, M. E.; GÖK, M. Z. Hybrid k-Anonymity. Computers & Security, v. 44, p. 51-63, 2014.

[158] DE CAPITANI DI VIMERCATI, S. et al. Data privacy: Definitions and techniques. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, v. 20, n. 06, p. 793-817, 2012. Disponível em: http://spdp.di.unimi.it/papers/ijufks2012.pdf. Acesso em: 17 ago. 2020.

41-50
51-60
> 60

Tabela 9 - Generalização por faixa etária

Além disso, quanto à generalização do endereço, uma opção é remover o número da casa e deixar apenas o nome da rua:

Numero	Id. Único	Idade	Endereço
1	357703	21-30	Cuiabá/MT
2	233121	31-40	Sorriso/MT
3	938637	41-50	Boa Vista/RR
4	591493	21-30	Santarém/PA
5	202626	21-30	Campinas/SP
6	888948	>60	Diadema/SP
7	175878	21-30	Osasco/SP
8	312304	41-50	Tefé/AM
9	214025	21-30	Salvador/BA
10	271714	31-40	Cariacica/ES
11	341338	21-30	Serra/ES
12	529057	21-30	Belo Horizonte/MG
13	390438	31-40	Ipatinga/MG

Tabela 10 - Dados após a Generalização

7.1.3 - A Supressão

A supressão pode ser a remoção de um atributo ou o de um registro. A supressão de um atributo refere-se à remoção de uma seção inteira do conjunto de dados, já a supressão de um registro refere-se à remoção de um registo inteiro do conjunto de dados.

Esta técnica de anonimização exclui alguns valores de atributos que possibilitam a identificação do titular dos dados. É muito utilizada no contexto de bancos de dados estatísticos, os quais disponibilizam apenas resumos estatísticos dos dados da tabela, em vez de dados originais[159].

Quando um atributo ou um registro não é necessário no conjunto de dados anonimizados, ou seja, a falta dessas informações não inviabiliza a utilidade dos dados, esta técnica é a mais utilizada.

[159] Samarati, P. (2001). Protecting respondents identities in microdata release. Knowledge and Data Engineering, IEEE Transactions on, 13(6):1010–1027.

Além disso, em muitos casos, o conjunto de dados não pode ser anonimizado, adequadamente, com outra técnica. Neste cenário, a supressão é a única técnica que possibilita uma anonimização apropriada.

É necessária a remoção dos atributos ou registros, ou seja, os dados devem ser efetivamente eliminados da base de dados. A supressão é, de fato, uma remoção permanente de dados, e não apenas “ocultamento da coluna ou do registro”.

A remoção de um registro, em contraste com a maioria das técnicas de anonimização, afeta vários atributos ao mesmo tempo. Nesse caso, é recomendável sua aplicação quando nenhuma outra técnica de anonimização for suficiente para gerar dados completamente anonimizados.

Apenas ocultar alguns dados, não é o suficiente para garantir a anonimização dos dados ou inviabilizar sua utilidade. Assim sendo, pode-se criar um “atributo derivado” que preserva a utilidade, mas ainda é menos sensível que os atributos originais que podem, então, ser suprimidos.

Por exemplo, criar um atributo “tempo de permanência”, a partir dos atributos “hora de entrada” e “hora de saída”.

Neste exemplo, o conjunto de dados consiste em pontuações obtidas pelos estudantes em avaliações disciplinares. O objetivo é analisar a pontuação obtida pelos estudantes em relação ao sexo e disciplinas, assim, a supressão desse atributo não afetará o objetivo do conjunto de dados:

Estudante	Sexo	Professor	Disciplina	Pontuação
Thiago Maia	Masculino	Felipe	Matemática	82
Cristiane Oliveira	Feminino	Thiago	Português	76
Tereza Barroso	Feminino	Felipe	Matemática	63
Carlos Rodrigues	Masculino	Helena	História	70
Edinaldo da Costa	Masculino	Rodrigo	Geografia	82
Ana Beatriz	Feminino	Felipe	Matemática	90
Maria João	Feminino	Carlos	Física	85
Pedro Santos	Masculino	Thiago	Português	69
Eduardo da Silva	Masculino	Carlos	Física	63
Antonio	Masculino	Bianca	Química	71
Rafaela de Lima	Feminino	Rodrigo	Geografia	89

Marcelo Brito	Masculino	Bianca	Química	85
Vanessa Martins	Feminino	Helena	História	90
Robinho Ferreira	Masculino	Thiago	Português	81
Vânia Lopes	Feminino	Rodrigo	Geografia	87
Vanessa Carvalho	Feminino	Bianca	Química	91

Tabela 11 - Dados antes da Supressão

Esta técnica de anonimização é irreversível, por isso, é considerada bastante eficaz, pois não há forma de reverter o processo, ou seja, recuperar os dados removidos. A seguir, encontra-se a tabela resultante do processo de supressão:

Sexo	**Disciplina**	**Pontuação**
Feminino	Física	85
Masculino	Física	63
Feminino	Geografia	87
Feminino	Geografia	89
Masculino	Geografia	82
Feminino	História	90
Masculino	História	70
Feminino	Matemática	63
Feminino	Matemática	90
Masculino	Matemática	82
Feminino	Português	76
Masculino	Português	69
Masculino	Português	81
Feminino	Química	91
Masculino	Química	71
Masculino	Química	85

Tabela 12 - Dados após a Supressão

7.1.4 - A Permutação

A finalidade desta técnica é reorganizar os dados em um determinado conjunto, de tal forma que os valores dos atributos individuais ainda estejam disponíveis, mas em geral não correspondam ao registo original. É, também, denominada embaralhamento ou troca.

Tal técnica é mais utilizada quando é necessário fazer uma análise apenas dos dados no agregado, ou seja, não é necessário manter as relações entre os atributos em cada registro.

No exemplo a seguir, a tabela apresenta um conjunto de dados com o registro de clientes que fizeram *login* em uma determinada aplicação:

Sessão	Sexo	Profissão	Idade	Categoria	Login mês
200436	Masculino	Desenvolvedor	30	Prata	0
200968	Feminino	Arquiteto	32	Platina	5
200473	Masculino	Advogado	30	Ouro	2
200333	Feminino	Médico	29	Prata	1
200194	Feminino	Enfermeira	31	Prata	2
200579	Masculino	Advogado	31	Ouro	4

Tabela 13 - Dados antes da Permutação

Após o processo, os valores dos atributos foram trocados, ou seja, o resultado produzido não mantém a correlação entre os atributos.

Sessão	Sexo	Profissão	Idade	Categoria	Login mês
200968	Feminino	Médico	32	Ouro	0
200579	Feminino	Enfermeira	31	Ouro	1
200473	Feminino	Desenvolvedor	31	Platina	2
200436	Masculino	Arquiteto	30	Prata	2
200333	Masculino	Advogado	30	Prata	4
200194	Masculino	Advogado	29	Prata	5

Tabela 14 - Dados após a Permutação

Existem vários algoritmos para a realização da permutação[160] dos dados como o algoritmo de Heap[161], o algoritmo Trotter-Johnson[162] e o algoritmo de Dijkstra[163] para geração de permutações em ordem lexicográfica e sem repetições. Para o exemplo acima, escolhe-se uma das permutações possíveis para cada um dos atributos individualmente, ou seja, escolhe uma das permutações possíveis para o atributo sessão, depois para o atributo sexo, para o atributo profissão e assim sucessivamente.

A quantidade de permutações possíveis é dada pela função P = n!, portanto, no exemplo supracitado existem 720 combinações possíveis para cada atributo da tabela. Com a utilização deste método, todas as relações entre os atributos são rompidos, ou seja, este método só é indicado para cenários em que o conjunto de dados, de uma forma geral, é utilizado posteriormente.

[160] Robert Sedgewick. 1977. Permutation Generation Methods. ACM Comput. Surv. 9, 2 (June 1977), 137–164. DOI:https://doi.org/10.1145/356689.356692

[161] Otto Seppälä, Lauri Malmi & Ari Korhonen (2006) Observations on student misconceptions—A case study of the Build – Heap Algorithm, Computer Science Education, 16:3, 241-255, DOI: 10.1080/08993400600913523

[162] Phillip J. Chase. 1970. Algorithm 383: permutations of a set with repetitions [G6]. Commun. ACM 13, 6 (June 1970), 368–369. DOI:https://doi.org/10.1145/362384.362503

[163] GERSTING, Judith L. Fundamentos matemáticos para a ciência da computação. LTC, 2001.

Por exemplo, se o objetivo for analisar o conjunto de dados para verificar quantos *logins* foram realizados durante o mês, ou qual é a idade das pessoas que mais utilizam o aplicativo, ou ainda, quantas pessoas do sexo masculino ou do sexo feminino fizeram o login na aplicação. Neste caso, o que importa é o conjunto de dados e não cada elemento individualmente.

7.1.5 - Perturbação

Esta técnica modifica ligeiramente os valores do conjunto de dados, sendo recomendada para semi-identificadores, que podem ser potencialmente identificadores quando combinados com outras fontes de dados, e ligeiras mudanças nos valores sejam aceitáveis.

Para exemplificar essa técnica, utilizamos o conjunto de dados a seguir. Os dados serão utilizados para investigar a possível ligação entre o tempo de serviço, profissão, salário do colaborador com as enfermidades laborais. Para tanto, definimos os seguintes arredondamentos:

Atributo	**Técnica de anonimização**
Idade (em anos)	O número escolhido para o arredondamento da idade é o número 5 por ser proporcional aos valores próprios de Idade.
Salário (em R$)	O número 100 é escolhido para o arredondamento do atributo salário, normalmente os valores salariais ficam entre 1000 e 5000 reais.
Tempo de Serviço (em anos)	No tempo de serviço a base utilizada é a 4, a maioria dos funcionários da organização possuem entre 4 e 40 anos de trabalho.

Tabela 15 – Arredondamento de Dados

Conjunto de dados antes da anonimização:

Funcionário	Idade (anos)	Tempo de Serviço	Salário (R$)	Fumante	Tratamento
367190	47	21	3654,58	Não	Não
392018	42	11	3135,98	Não	Não
342190	52	31	4234,99	Sim	Sim
687032	55	35	4695,29	Não	Não
665731	58	29	4038,23	Sim	Sim

Tabela 16 - Dados antes a Perturbação

Conjunto de dados após o processo de anonimização:

Idade (anos)	Tempo de Serviço (anos)	Salário (R$)	Fumante	Tratamento Médico
45	20	3700,00	Não	Não
40	12	3100,00	Não	Não
50	32	4200,00	Sim	Sim
55	36	4700,00	Não	Não
60	28	4000,00	Sim	Sim

Tabela 17 - Dados após a Perturbação

7.1.6 - Dados sintéticos

Esta técnica é usada para gerar conjuntos de dados sintéticos e separá-los dos dados originais, em vez de modificar o conjunto original.

Os dados sintéticos podem ser considerados dados “falsos”, criados a partir de dados “reais”. A vantagem é que a base é uma base de dados com dados reais e informações reais, o que a torna quase indistinguível dos dados originais[164].

Como os dados reais não podem ser divulgados por estarem protegidos pela legislação, divulgam-se os dados sintéticos. Além disso, muitas vezes, sem os atributos identificadores os dados podem não ser tão úteis quanto deveriam.

Normalmente, este método é utilizado quando há um grande volume de dados, no entanto os dados reais não podem ser utilizados e, ainda assim, os dados devem ser verdadeiros em certos aspectos, por exemplo: tipo de dados, relação entre atributos.

Dados para teste de uma determinada aplicação é um cenário próprio para aplicação deste método de anonimização. Neste contexto é necessário que os dados de teste contenham informações semelhantes aos dados originais, mas não necessariamente sejam dados reais. Deste modo, os padrões do conjunto original precisam ser aplicados ao conjunto de dados anonimizados.

[164] Bellovin, Steven M. and Dutta, Preetam K. and Reitinger, Nathan, Privacy and Synthetic Datasets (August 20, 2018). Stanford Technology Law Review, Forthcoming. Available at SSRN: https://ssrn.com/abstract=3255766.

A vantagem desta técnica é que o volume de dados gerado para teste pode ser de qualquer tamanho, sem comprometer a anonimização dos dados. A base de dados anonimizada é de pouca utilidade para atividade que não seja o teste da aplicação porque os dados não são reais.

O balanceamento de carga em aplicações de gerenciamento de banco de dados que a quantidade de dados seja muito grande, mas não precisem ser reais pode ser um ótimo cenário para a aplicação deste método de anonimização.

Neste exemplo o objetivo é criar um conjunto de dados sintéticos para realizar teste de simulações em um aplicativo de disponibilidade de vagas de estacionamento em uma organização. Alguns atributos podem ser substituídos por valores médios ou medianos sem reduzir a utilidade dos dados. Este é o conjunto de dados original de clientes que utilizam o estacionamento durante uma semana:

Cliente	Dia da Semana	Hora de Entrada	Tempo de Permanência
Pedro Botelho	segunda-feira	10:27	04:37
Daniel da Silva	segunda-feira	11:20	06:50
Flavio Fernandes	terça-feira	12:45	04:32
Indira Barros	terça-feira	13:55	05:59
Daniela Batista	quarta-feira	13:18	05:30
Luzia Carvalho	quarta-feira	13:02	07:00
Naiara Costa	quinta-feira	14:25	05:58
Leila Martins	quinta-feira	10:00	07:31
José Vieira	sexta-feira	15:04	05:35
Leandro Vani	sexta-feira	16:17	04:54

Tabela 18 - Dados originais (Dados Sintéticos)

Estatísticas obtidas do conjunto de dados original:

Dia da Semana	N.º Médio de Utilizadores	Permanência
segunda-feira	1	04:37
terça-feira	2	05:41
quarta-feira	4	06:06
quinta-feira	1	07:31
sexta-feira	2	05:14

Tabela 19 - Estatísticas dos dados

O conjunto de dados sintético para uma semana são criados com base em estatísticas derivadas do conjunto original, a média é o elemento mais utilizado.

Código	Data	Hora de Entrada	Tempo de Permanência
100001	segunda-feira	10:00	04:37
100002	segunda-feira	11:00	04:37
100003	terça-feira	12:00	05:41
100004	terça-feira	13:00	05:41
100005	quarta-feira	13:00	06:06
100006	quarta-feira	13:00	06:06
100007	quinta-feira	14:00	07:31
100008	quinta-feira	10:00	07:31
100009	sexta-feira	15:00	05:14
100010	sexta-feira	16:00	05:14

Tabela 20 - Dados Sintéticos

6.1.7 - Agregação de dados

A agregação de dados transforma um conjunto de dados originais em conjunto de dados derivados com as informações, de certa forma, agrupadas. Ou seja, um resumo com as informações que constam no conjunto de dados originais.

Esta técnica é mais utilizada quando os registros individuais não são necessários e os dados agregados são suficientes para o objetivo da análise. Por esse motivo, é necessário conhecer bem qual o objetivo da análise dos dados, assim, pode-se encontrar um equilíbrio entre o nível de agregação e a utilidade das informações.

Aplica-se, portanto, o processo de generalização dos atributos no conjunto de dados originais. Os valores são reunidos de modo que cada registro partilhe o mesmo valor em determinado atributo. Por exemplo, ao reduzir a granularidade de um local de uma cidade para um país é incluído maior número de pessoas[165].

As datas de nascimento individuais podem ser generalizadas em um intervalo de datas ou agrupadas por mês ou ano. Outros atributos numéricos como por exemplo: salários, peso e altura podem ser generalizados por intervalos de valores como faixa salarial, faixa etária[166].

[165] "Opinion 05/2014 on Anonymisation Techniques". Article 29 Data Protection Working Party (European Commission), em 10 de Abril de 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf. Acesso em: 10 ago. 2020.
[166] Ibidem.

Este método deve ser aplicado com cautela para grupos que tenham poucos registos após realizar a agregação, pois se os dados agregados incluem um único registro em cada uma das categorias, pode ser fácil para alguém com algum conhecimento adicional identificar o titular dos dados.

Em determinados casos, será necessária a aplicação de outros métodos de anonimização, combinando a agregação. Alguns atributos podem ser suprimidos, porque contém informações que não podem ser agregadas, e outros atributos podem ser modificados ou permutados.

Para exemplificar a técnica de agregação, utilizamos os dados fictícios relacionados a uma organização de caridade que tem o registo das doações, bem como alguns dados sobre os doadores. A organização de caridade avaliou que os dados agregados são suficientes para um consultor externo realizar a análise de dados, e por isso realiza a agregação no conjunto de dados originais:

Nome	Salário mensal	Quantia doada
Ana Beatriz	R$ 4 000,00	R$ 210,00
Antonio Fernandes	R$ 4 900,00	R$ 420,00
Carlos Rodrigues	R$ 2 200,00	R$ 150,00
Cristiane Oliveira	R$ 4 200,00	R$ 110,00
Daniel da Silva	R$ 5 500,00	R$ 260,00
Daniela Batista	R$ 2 600,00	R$ 40,00
Edinaldo da Costa	R$ 3 300,00	R$ 130,00
Eduardo da Silva	R$ 5 500,00	R$ 210,00
Flavio Fernandes	R$ 1 600,00	R$ 380,00
Indira Barros	R$ 3 200,00	R$ 80,00
José Vieira	R$ 2 000,00	R$ 440,00
Leandro Vani	R$ 5 800,00	R$ 400,00
Leila Martins	R$ 4 600,00	R$ 390,00
Luzia Carvalho	R$ 1 900,00	R$ 480,00
Marcelo Brito	R$ 1 700,00	R$ 320,00
Maria João	R$ 2 400,00	R$ 330,00
Naiara Costa	R$ 4 300,00	R$ 390,00
Pedro Botelho	R$ 2 300,00	R$ 260,00
Pedro Santos	R$ 3 500,00	R$ 80,00
Rafaela de Lima	R$ 1 700,00	R$ 290,00

Tabela 21 - Dados originais (Agregação)

A seguir é apresentado o resultado do processo de anonimização aplicando a técnica de agregação aos dados originais:

Faixa Salarial	Nº de doações	Quantia doada
R$ 1000- R$ 1999	4	R$ 1470,00
R$ 2000- R$ 2999	5	R$ 1220,00
R$ 3000- R$ 3999	3	R$ 290,00
R$ 4000- R$ 4999	5	R$ 1520,00
R$ 5000- R$ 6000	3	R$ 870,00
Total	**20**	R$ 5 370,00

Tabela 22 - Dados Anonimizados (Agregação)

7.2 - Metodologias de Anonimização

A anonimização visa ocultar a identidade ou dados confidenciais das pessoas, de forma que, ainda que a informação seja divulgada e útil para análise por parte de receptores, a privacidade individual seja preservada[167].

Existem vários modelos de anonimização com a finalidade de proteger a privacidade do titular do dado, ao se disponibilizar estes dados publicamente. Os principais modelos de anonimização de dados são k-anonimato, l-diversidade, LKC-Privacidade, t-proximidade, b-semelhante e δ-presença.

Os seguintes passos são esperados de uma metodologia sugerida para a realização de anonimização: determinar o modelo de divulgação; determinar o nível de risco de reidentificação aceitável, bem como a utilidade esperada e o nível de risco pretendido ou aceitável; classificar os atributos dos dados; remover atributos de dados que não são utilizados; anonimizar identificadores diretos e indiretos; determinar o risco real e comparar o mesmo com o limite; realizar mais anonimização, se for necessário; avaliar a solução; determinar os controles necessários; documentar o processo de anonimização.

A metodologia de anonimização deve determinar o modelo de divulgação dos dados, que pode ser *público* ou *não público*.

O *público* refere-se a disponibilizar os dados para qualquer pessoa, enquanto o *não-público* refere-se a uma divulgação controlada para receptores limitados.

[167] QUEIROZ, M. J.; LINO, N. C.; MOTTA, G. Uma ontologia de domínio para preservação de privacidade em dados publicados pelo governo brasileiro. XII Simpósio Brasileiro de Sistemas de Informação. Florianópolis, SC, Brasil, 2016.

Obviamente, o modelo de divulgação *público* demanda mais cautela na aplicação das técnicas de anonimização.

É possível, assim, determinar o nível de risco de reidentificação aceitável, bem como a utilidade esperada e o nível de risco pretendido ou necessário. Classificar os atributos no conjunto de dados como identificadores, semi-identificadores ou não-identificadores afeta como estes atributos serão subsequentemente processados.

Principalmente, espera-se deste método a anonimização dos identificadores e dos semi-identificadores, ou seja, a aplicação das técnicas de anonimização, em conjunto ou separadamente, até que os dados estejam anonimizados. Além disso, deve-se remover qualquer atributo que não seja claramente necessário no conjunto de dados anonimizados.

O método deve, ainda, determinar o risco real e comparar estes riscos com o limite. Caso seja necessário, o método deve realizar mais processos de anonimização, ou seja, aplicar novamente uma técnica de anonimização se o risco de reidentificação ainda estiver fora dos limites determinados.

É importante fazer uma avaliação da solução proposta, isto inclui examinar o conjunto de dados anonimizados, para avaliar se sua utilidade vai ao encontro do objetivo determinado. Se a utilidade for insuficiente, o processo de anonimização pode ter de ser redesenhado, ou pode ser necessário considerar se a anonimização é exequível para este conjunto de dados.

Ademais, os controles necessários devem ser determinados e documentados. Por fim, mas não menos importante, o método deve documentar o todo o processo de anonimização. Os detalhes do processo de anonimização, os parâmetros e controles usados devem ser claramente registados para referência futura. Esta documentação facilita a revisão, manutenção, aperfeiçoamento e auditoria. É importante salientar que tal documentação deve ser mantida sob estrita segurança, pois a divulgação dos parâmetros pode facilitar a reidentificação.

O anonimato pode ser computacional, ou seja, deverá ser informaticamente difícil, mesmo para o responsável pelo tratamento de dados em colaboração com

terceiros, identificar direta ou indiretamente um dos titulares de dados. Pode ser, também, um anonimato perfeito, ou seja, deverá ser impossível, mesmo para o responsável pelo tratamento de dados em colaboração com terceiros, identificar direta ou indiretamente um dos titulares de dados[168].

7.2.1 - k-anonimato

O modelo k-anonimato requer que qualquer combinação de atributos semi-identificadores seja compartilhada por, pelo menos, mais de um registro em um banco de dados anonimizado[169]. Logo, o principal objetivo do modelo é transformar um conjunto de dados de forma que ninguém possa fazer associações entre os registros na tabela e as entidades correspondentes, mesmo que os atacantes tenham acesso a informações externas.

Um conjunto de dados anonimizado por esse método pode ter níveis diferentes de anonimização. A variável K no nome do método indica qual é o nível aceitável de registros que compartilham os mesmos semi-identificadores, o valor de k, onde k é um inteiro positivo definido pelo proprietário dos dados, possivelmente como resultado de negociações com outras partes interessadas[170].

A utilização deste método assegura que os atributos identificadores ou semi-identificadores são compartilhados por pelo menos k registros. Esta é uma característica fundamental do método, pois é ela que impede os ataques de ligação. Ou seja, isto garante a impossibilidade de ligar ou isolar o registo de um indivíduo porque os registos k são idênticos nos atributos identificadores ou semi-identificadores.

Geralmente, quanto mais alto o valor de k, mais difícil é para os titulares dos dados serem identificados, no entanto a utilidade dos dados pode-se tornar muito pequena à medida que o k aumenta e mais registos são suprimidos.

---

[168] "Opinion 05/2014 on Anonymisation Techniques". Article 29 Data Protection Working Party (European Commission), em 10 de Abril de 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf. Acesso em: 10 ago. 2020

[169] SAMARATI, P. Protecting respondents identities in microdata release. Knowledge and Data Engineering, IEEE Transactions on, v. 13, n. 6, p. 1010–1027, 2001.

[170] Ibidem.

Portanto possuir um elevado valor de k em conjunto de dados anonimizados indica que o risco de reidentificação caso os dados sejam divulgados é baixo. Embora este método não proteja o conjunto de dados contra a divulgação dos dados, a probabilidade de reidentificar um registro é de 1/k[171].

Vale ressaltar que, não obstante o atacante não ter a capacidade de reidentificar o registro, ele pode ter acesso a atributos sensíveis no conjunto de dados anonimizado.

Neste exemplo, o k é igual a 2, ou seja, cada conjunto de semi-identificadores deve ser partilhado por dois registros, após a anonimização. As seguintes técnicas de anonimização são utilizadas em combinação e o nível de resolução é um exemplo que permite atingir o nível K necessário.

Atributo	Técnica de anonimização
Idade	Generalização em faixa etária
Ocupação	Generalização -Tanto "Analista de Sistemas" como "Desenvolvedor de Software" são generalizados para "Tecnologia da Informação"
Número Médio de Viagens por semana	Perturbação – o número médio de viagens foi substituído pela média de viagens para cada ocupação
Supressão de Registo	Supressão – nas extremidades, as técnicas de generalização e perturbação não são suficientes para impedir a reidentificação. Neste caso aplica-se a técnica de supressão.

Tabela 23 - Técnicas aplicadas na anonimização dos dados

Conjunto de dados original:

Idade	Gênero	Ocupação	Média de viagens por semana
21	Feminino	Administrador	14
38	Masculino	Contador	2
25	Feminino	Advogado	8
44	Feminino	Analista de Sistemas	3
25	Feminino	Analista Administrativo	1
31	Masculino	Contador	6
42	Feminino	Desenvolvedor de	3
22	Feminino	Analista Administrativo	3
30	Feminino	Administrador	2

[171] BASSO, T.; MATSUNAGA, R.; MORAES, R.; ANTUNES, N. Challenges on anonymity, privacy, and big data. In: IEEE. Dependable Computing (LADC), 2016 Seventh Latin-American Symposium on. [S.l.], 2016. p. 164–171.

Tabela 24 - Conjunto de Dados Original

Conjunto de dados anonimizados pela técnica de generalização da idade, ocupação e o número médio de viagens foi substituído pela média de viagens para cada ocupação. Na extremidade é aplicada a técnica de supressão:

Idade	Gênero	Ocupação	Média semanal de viagens
21 a 30	Feminino	Administrador	8
31 a 40	Masculino	Contador	4
41 a 50	Feminino	Tecnologia da Informação	3
21 a 30	Feminino	Analista Administrativo	2
21 a 30	Masculino	Advogado	8
21 a 30	Feminino	Administrador	8

Tabela 25 - Conjunto de Dados Anonimizado

Neste exemplo, para que o K seja igual a 2, cada atributo deve aparecer em, no mínimo, dois registros. Para atingir este fim, foi aplicado a técnica de supressão ao registro que possui a ocupação “Advogado”. Além disso, as ocupações “Analista de Sistemas” e “Desenvolvedor de Software” foram substituídas por “Tecnologia da Informação”. Outra mudança, ocorreu na idade, o atributo idade foi substituído pela faixa etária em cada registro, deste modo, k=2 porque cada atributo está apresentado em pelo menos 2 vezes na tabela.

7.2.2 - l-diversidade

O modelo l-diversidade[172] tem o objetivo de impedir o ataque de ligação de atributos sensíveis por inferência. Por exemplo, casos em que o atacante pode inferir informações sensíveis sobre registros mesmo sem identificá-los.

Este método teve origem como uma maneira de sanar as limitações da metodologia k-anonimato. Para evitar a identificação do titular dos dados por inferência, o modelo exige que cada classe de equivalência possua l valores distintos para cada atributo sensível.

A proposta desta metodologia é permitir maior diversidade de registros com os mesmos atributos sensíveis. Assim sendo, um conjunto é considerado l-diversificado

[172] Machanavajjhala, A., Kifer, D., Gehrke, J., and Venkitasubramaniam, M. (2007). l-diversity: Privacy beyond k-anonymity. In ACM Transactions on Knowledge Discovery from Data (TKDD).

se possuir pelo menos dois registros distintos para cada atributo semi-identificador do conjunto.

Isto garante que um atacante, mesmo com conhecimento prévio que lhe permita descobrir a classe de equivalência de um indivíduo, não consiga inferir o atributo sensível do titular dos dados com probabilidade maior que $\frac{1}{l}$[173].

A tabela a seguir mostra que, apesar de o conjunto de dados atender aos requisitos da metodologia k-anonimato, ainda assim, pode se presumir qual é a enfermidade do indivíduo, caso se conheça algumas informações.

ID	Sexo	Idade	CEP	Enfermidade
1	Masculino	>40	79415000	Sinusite
2	Masculino	>40	79415000	Gripe
3	Feminino	>40	79415000	Sinusite
4	Feminino	>40	79415000	Gripe
5	Masculino	<40	78005-000	Bronquite
6	Masculino	<40	78005-000	Bronquite
7	Feminino	<40	78005-000	Bronquite
8	Feminino	<40	78005-000	Bronquite

Tabela 26 - K-anonimato passível ataque de inferência

Se um atacante possuir acesso à tabela anterior e posterior, tendo conhecimento que os dados são de uma clínica médica, e que os titulares figuram em ambas as tabelas, poderia inferir que a Rafaela Santo e a Vanessa Carmo possuem a enfermidade Bronquite.

ID	Nome	Ano de Nascimento	Sexo
1	Luzia Brito	1971	Feminino
2	Cristiane de Barros	1969	Feminino
3	Jessica Martins	1974	Feminino
4	Rafaela Santos	1997	Feminino
5	Maria Ferreira	1973	Feminino
6	Ana Maria	1970	Feminino
7	Vanessa Carmo	1998	Feminino
8	Taís Campos	1977	Feminino

Tabela 27 - Ataque de Inferência

[173] Machanavajjhala, A., Kifer, D., Gehrke, J., and Venkitasubramaniam, M. (2007). l-diversity: Privacy beyond k-anonymity. In ACM Transactions on Knowledge Discovery from Data (TKDD).

Por esse motivo, a metodologia l-diversidade requer que o número de valores distintos de atributos sensíveis em cada classe de equivalência seja no mínimo l.

A definição de l-diversidade é a seguinte: um grupo de semi-identificadores é l-diverso, se contiver pelo menos l valores bem representados para os atributos sensíveis. Uma tabela é l-diversa se cada grupo de semi-identificadores for l-diverso[174].

O modelo garante privacidade, mesmo quando não são conhecidas quais as informações que o atacante possui, pois garante a existência de, pelo menos, l valores de atributos sensíveis em cada grupo de atributos semi-identificador[175].

A tabela abaixo é um exemplo da aplicação da técnica l-diversidade cujo ataque de inferência não é possível, visto que, se o atacante conhecer uma pessoa dessa tabela, que seja do sexo masculino, que tenha menos de 40 anos e que mora no CEP 78005-000, ainda assim, não será possível inferir qual é a enfermidade dessa pessoa. Neste caso, pode ser Bronquite ou Diabetes. Portanto, a tabela 29 é a implementação desta técnica com a resolução dos problemas encontrados na tabela 26.

ID	Sexo	Idade	CEP	Enfermidade
1	Masculino	>40	79415000	Sinusite
2	Masculino	>40	79415000	Gripe
3	Feminino	>40	79415000	Sinusite
4	Feminino	>40	79415000	Gripe
5	Masculino	<40	78005-000	Bronquite
6	Masculino	<40	78005-000	Diabetes
7	Feminino	<40	78005-000	Bronquite
8	Feminino	<40	78005-000	Diabetes

Tabela 28 - l-diversidade

Neste cenário, o l é igual a dois, ou seja, P = $\frac{1}{2}$. Portanto, a probabilidade de inferir qual é a enfermidade de uma determinada pessoa é de 50%. O valor de l é inversamente proporcional a possibilidade do atacante inferir informações de um determinado registro.

[174] LI, N.; LI, T.; VENKATASUBRAMANIAN, S. t-closeness: Privacy beyond k-anonymity and l-diversity. In: In Proc. of IEEE 23rd Int'l Conf. on Data Engineering (ICDE'07). [S.l.: s.n.], 2007. p. 106–115.

[175] ZHANG, X.; LIU, Q.; LIU, D.; XIE, W. A survey on anonymity for privacy preserving data mining. In: CRC PRESS. Information Science and Electronic Engineering: Proceedings of the 3rd International Conference of Electronic Engineering and Information Science (ICEEIS 2016), 4-5 January, 2016, Harbin, China. [S.l.], 2016. p. 343–346.

É importante destacar que a metodologia l-diversidade não trata todas as falhas do modelo anterior. Assim, este modelo apresenta algumas fragilidades e a sua utilização pode impactar a utilidade dos dados em determinados cenários. Por exemplo, quando o conjunto de dados possuir grande número de repetições para o valor de um atributo sensível e pouco ou nenhum de outros valores, pois, neste caso, é necessário introduzir grande distorção ou supressão nos dados[176].

Além disso, essa metodologia pode sofrer o ataque de simetria que ocorre quando o atacante tem conhecimento prévio e descobre tanto a classe de equivalência de um indivíduo quanto a distribuição dos atributos sensíveis, que pode ser obtida apenas analisando a tabela publicada.

Em um conjunto de dados que contêm dez mil registros sobre um vírus, que afeta apenas um por cento da população, para um subconjunto desses dados que contenha apenas casos negativos, medidas fortes de privacidade provavelmente não são necessárias, pois as pessoas não têm a doença e não se importam se sua identidade for exposta.

No entanto, um subconjunto desses mesmo dados em que metade dos registros seja positiva e a outra metade seja negativa, ou seja, contém um número igual de registros positivos e negativos. Isso dá a todos, neste conjunto de dados, uma chance de cinquenta por cento de ter o vírus, que é muito maior do que a distribuição global[177].

Ainda pode sofrer o ataque de similaridade quando, mesmo que os atributos sensíveis sejam distintos, qualquer um deles forneça uma informação sensível ao atacante, por exemplo, no caso apresentado no modelo i-diversidade disponível na tabela 28, e os valores para a classe de equivalência de determinado indivíduo são descobertos pelo atacante, a partir de conhecimento prévio das doenças (sinusite e

[176] "Opinion 05/2014 on Anonymisation Techniques". Article 29 Data Protection Working Party (European Commission), em 10 de Abril de 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf. Acesso em: 10 ago. 2020.

[177] Li, N., Li, T., and Venkatasubramanian, S. (2007). t-closeness: Privacy beyond k-anonymity and l-diversity. In 23th ICDE International Conference on Data Engineering (ICDE), pages 106–115.

bronquite), por similaridade conseguirá inferir que o indivíduo tem uma doença respiratória[178].

7.2.3 - t-proximidade

A metodologia k-anonimato não impede totalmente a divulgação de atributos, mas é eficiente na proteção contra a sua divulgação. Por esse motivo, diversas melhorias foram implementadas na metodologia, visando a maior proteção de informações sensíveis e outros riscos de divulgação, dentre as quais a metodologia t-proximidade tem destaque.

A metodologia t-proximidade[179] propõe corrigir algumas limitações do l-diversidade no que diz respeito à proteção contra o ataque de simetria, onde o adversário pode inferir informações sobre atributos sensíveis a partir do conhecimento da frequência de ocorrência dos atributos no conjunto de dados definido.

Esta metodologia é um refinamento da l-diversidade, na medida em que visa criar classes equivalentes, que se assemelhem à distribuição inicial de atributos na tabela. Esta técnica é útil quando é importante manter os dados tão próximos quanto possível do original. Para esse efeito, é colocada uma nova restrição na classe de equivalência, designadamente, que devem existir não só pelo menos “l” valores diferentes em cada classe de equivalência, mas também que cada valor seja representado as vezes que forem necessárias para refletir a distribuição inicial de cada atributo[180].

Vale salientar que, como uma melhoria para o k-anonimato, é necessário que, para um conjunto de dados ser t-proximidade, esse conjunto também precise satisfazer as condições do k-anonimato, para um k escolhido. O t-proximidade propõe-se a garantir a distribuição dos dados de um atributo sensível em cada classe de equivalência e que essa seja próxima a sua distribuição global.

---

[178] Idem.

[179] Idem.

[180] “Opinion 05/2014 on Anonymisation Techniques”. Article 29 Data Protection Working Party (European Commission), em 10 de Abril de 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf. Acesso em: 10 ago. 2020.

Uma classe de equivalência satisfaz a metodologia t-proximidade quando a distância entre a distribuição de um atributo sensível na classe e a distribuição do atributo na tabela não for maior que o valor atribuído a t. A tabela será t-próxima se todas as classes de equivalência forem t-próximas[181].

O modelo t-proximidade utiliza o conceito de “conhecimento global de retaguarda", que pressupõe que o adversário pode inferir informações sobre atributos sensíveis, a partir do conhecimento da frequência de ocorrência destes atributos na tabela. Como os dados anonimizados disponibilizados devem conter a maior parte ou todos os registros da tabela original, é possível para o atacante calcular a medida da distribuição do atributo sensível em relação ao total de registros da tabela[182].

No exemplo a seguir, um atacante pode ter uma informação sobre dados estatísticos acerca dos colaboradores de uma empresa, com essas informações é possível inferir que 80% dos funcionários com idade superior a 41 anos têm Hipertensão.

Semi-identificadores			**Dados Sensíveis**	
Faixa Etária	**Sexo**	**CEP**	**Enfermidade**	**Identificador**
41 - 50	Masculino	75349*	Hipertensão	012457
41 - 50	Masculino	75349*	Câncer	239865
41 - 50	Masculino	75349*	Hipertensão	659043
41 - 50	Masculino	75349*	Hipertensão	102938
41 - 50	Masculino	75349*	Hipertensão	896523

Tabela 29 – fragilidade l-diversidade

Portanto, embora os dados estejam anonimizados na tabela acima e cumpram os requisitos da metodologia l-diversidade, é insuficiente para garantir a privacidade dos dados.

[181] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In International Conference on Data Engineering, volume 23, pages 106 – 115, April 2007.

[182] JUNIOR, Eliseu Castelo Branco, Uma Estrategia Para Assegurar a Confidencialidade De Dados Armazenados Em Nuvem. Disponível em: < http://repositorio.ufc.br/bitstream/riufc/23917/1/2017_tese _ ecastelobrancojunior.pdf>. Acesso em: 05 set. 2020.

Caso um atacante tenha acesso aos dados de saúde dos funcionários, o ataque de inferência tem uma probabilidade muito alta de sucesso em razão de 80% dos funcionários possuírem a enfermidade Hipertensão.

A privacidade é considerada violada quando a distribuição do atributo sensível em uma classe de equivalência não está próxima da distribuição em toda a tabela.

A fim de superar a vulnerabilidade apresentada na l-diversidade, a metodologia t-proximidade propõe adicionar mais restrições na classe de equivalência, ou seja, além de existir pelo menos “l” valores diferentes em cada classe de equivalência, cada valor de atributo deverá ser representado tantas vezes quantas forem necessárias para refletir a distribuição inicial de cada atributo.

Para este exemplo, se considera na tabela abaixo os dados de saúde, originais não anonimizados dos funcionários de uma determinada instituição.

ID	CEP	Idade	Sexo	Enfermidade
1	75349-721	32	Masculino	Úlcera gástrica
2	79311-220	23	Masculino	Gastrite
3	75349-802	27	Masculino	Câncer de estômago
4	79311-911	31	Masculino	Pneumonia
5	75349-599	38	Masculino	Bronquite
6	73821-122	42	Masculino	Câncer de estômago
7	73821-505	44	Masculino	Gastrite
8	73821-900	50	Masculino	Gripe
9	73821-673	48	Masculino	Bronquite
10	79311-562	28	Masculino	Bronquite
11	75349-377	25	Masculino	Pneumonia
12	79311-730	33	Masculino	Câncer de estômago

Tabela 30 - t-proximidade Dados Originais

A partir da tabela original, se obtém os dados anonimizados da tabela a seguir.

ID	CEP	Idade	Sexo	Enfermidade
1	79311*	<40	Masculino	Gastrite
2	79311*	<40	Masculino	Bronquite
3	79311*	<40	Masculino	Pneumonia
4	79311*	<40	Masculino	Câncer de estômago
5	75349*	<40	Masculino	Úlcera gástrica
6	75349*	<40	Masculino	Câncer de estômago
7	75349*	<40	Masculino	Bronquite
8	75349*	<40	Masculino	Pneumonia

9	73821*	>40	Masculino	Gastrite
10	73821*	>40	Masculino	Gripe
11	73821*	>40	Masculino	Câncer de estômago
12	73821*	>40	Masculino	Bronquite

Tabela 31 - t-proximidade Dados Anonimizados

Considerando o seguinte cenário: o atacante tem acesso a tabela com os dados anonimizados e sabe as seguintes informações de um dos funcionários: uma pessoa do sexo masculino que tem 27 anos de idade e reside no CEP 75349-190.

Desta forma, mesmo que o atacante conheça os dados acima citados, não é possível ao atacante inferir com relevante grau de certeza, qual é a enfermidade dessa pessoa.

Apesar do esforço da metodologia t-proximidade, ainda há algumas limitações[183], como a falta de flexibilidade para especificação de diferentes níveis de privacidade para cada atributo sensível; a função EMD não é adequada para ataques de ligação ao atributo quando estes são numéricos e, por fim, para garantir a mesma distribuição em todos as classes de equivalência, poderá acarretar uma baixa utilidade dos dados.

Observamos que, a proximidade **T** protege contra a divulgação de atributos, mas não trata da divulgação de identidade. Assim, pode ser desejável usar tanto t-proximidade quanto k-anonimato ao mesmo tempo.

7.2.4 - δ-presença

O δ-presença foi proposto por Nergiz, M. E., Atzori, M., and Clifton[184], e tem como objetivo evitar a vinculação de um indivíduo a uma tabela publicada, ou seja, busca proteger a privacidade de indivíduos contra o ataque de ligação à tabela. Este modelo define o limite δ = (δmax,δmin) para a probabilidade de um adversário inferir a presença de um indivíduo em um conjunto de dados tabulados.

[183] Fung, B. C., Wang, K., Fu, A. W.-C., and Yu, P. S. (2010). Introduction to Privacy Preserving Data Publishing: Concepts and Techniques. Chapman & Hall/CRC, 1st edition. ISBN 978-1-4200-9148-9.

[184] Nergiz, M. E., Atzori, M., and Clifton, C. (2007). Hiding the presence of individuals from shared databases. In Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data, SIGMOD '07, pages 665–676, New York, NY, USA. ACM.

Para exemplificar um ataque de ligação, vamos utilizar o seguinte exemplo: Suponha que um adversário tenha acesso a dados públicos da tabela a seguir.

Nome	Profissão	Gênero	Idade
Sandra	Professor	Feminino	[25-30)
Valentina	Professor	Feminino	[25-30)
Amanda	Professor	Feminino	[25-30)
Aline	Professor	Feminino	[25-30)
Sthefany	Professor	Feminino	[25-30)
Ricardo	Advogado	Masculino	[30-35)
Carlos	Advogado	Masculino	[30-35)
Mateus	Advogado	Masculino	[30-35)
Francisco	Advogado	Masculino	[30-35)

Tabela 32 - Tabela externa no formato 4-anonimato

Suponha ainda que, o atacante tenha acesso às informações anonimizadas da seguinte tabela, e tenha conhecimento de que todos os indivíduos da tabela 33 estão na tabela 32.

Profissão	Gênero	Idade	Enfermidade
Advogado	Masculino	[30-35)	Sinusite
Advogado	Masculino	[30-35)	Câncer
Professor	Feminino	[25-30)	Sinusite
Professor	Feminino	[25-30)	Bronquite
Professor	Feminino	[25-30)	Hipertensão
Professor	Feminino	[25-30)	Câncer

Tabela 33 - Tabela de pacientes no formato 3-anonimimato

Neste contexto, pode-se observar que a tabela 32 possui duas classes de equivalência: a classe de equivalência E1 (Professor, Feminino, [25,30]) que possui cinco indivíduos e a classe de equivalência E2 (Advogado, Masculino, [30-35]) que possui quatro indivíduos.

Assim, será possível deduzir que a probabilidade de Sandra estar presente na tabela 33 é de 80%, pois há 4 registros nessa tabela, e 5 registros na tabela 32 com a mesma classe de equivalência indicado pelos valores: “Professor, Feminino, [25−30)”. Pelo mesmo raciocínio, a probabilidade de Carlos estar presente na tabela 33 é de 50%.

7.2.5 - β -semelhante

Esta metodologia revela-se como uma proposta de solução ao problema da quebra de privacidade de valores de atributos sensíveis que ocorrem com pouca frequência.

A metodologia b-semelhante[185] garante que a estimativa de um atacante sobre o valor de um atributo sensível não aumentará em termos relativos, mais que um limite b pré-estabelecido, após o atacante tomar conhecimento dos dados anonimizados publicados.

Esta metodologia difere das anteriores em relação ao uso da uma função de distância para estabelecer o limite de distância máximo, em vez da distância cumulativa entre os atributos sensíveis das metodologias anteriores.

A definição básica de b-semelhante é de que dada uma tabela qualquer que contenha atributos sensíveis, seja V = (v_1, v_2, v_3, ..., v_m,) o domínio de atributos sensíveis e P = (p_1, p_2, p_3, ..., p_m) a distribuição global de atributos sensíveis na tabela[186].

Uma classe de equivalência G com distribuição de atributos sensíveis Q = (q_1, q_2, q_3, ..., q_m) satisfaz um limite básico b-semelhante, se e somente se $\forall_{q_i}$, $\mathrm{D}(p_i, q_i) = \frac{q_i - p_i}{p_i} \leq \min\{\beta, lnp_i\}$, onde β > 0 é um limite e ln p_i é o logaritmo natural de p_i[187].

A distância D deve ser grande o suficiente para proteger os dados de ataques de assimetria e de similaridade e é calculada pela fórmula: D(p_i , q_i) = $\frac{q_i - p_i}{2}$.

A restrição imposta à função D(p_i , q_i) de ser menor ou igual ao limite b, tem como consequência a criação de um limite superior para a frequência de v_i ∈ V em

[185] CAO, J.; KARRAS, P. Publishing microdata with a robust privacy guarantee. Proc. VLDB Endow., v. 5, n. 11, p. 1388–1399, 2012.
[186] Ibidem.
[187] Ibidem.

qualquer classe de equivalência G, de acordo com a seguinte expressão: $\frac{(q_i - p_i)}{p_i} \leq \beta \rightarrow q_i \leq p_i \times (1 + \beta)$.

Se estas restrições não forem atendidas pela anonimização dos dados é possível ao atacante saber que o registro da vítima está presente na classe de equivalência. Neste caso, a inferência do valor atributo sensível pode ser de cem por cento de certeza.

7.2.6 - LKC-privacidade

O modelo LKC-Privacidade foi criado como uma tentativa de resolver a perda de utilidade dos dados causada pela técnica de supressão em grande quantidade de dados semi-identificadores[188]. Esta metodologia parte do pressuposto que o atacante não possui todos os dados do seu alvo, no que tange aos atributos semi-identificadores.

O k-anonimato utiliza-se da técnica de generalização de dados em grupos de atributos sensíveis para evitar o ataque de ligação. Assim, uma vez que cada grupo contenha k registros com os mesmos valores de semi-identificadores e diversificação, isso dificulta ao atacante correlacionar atributos da vítima que ele conhece.

O problema de aplicar esta técnica, quando a quantidade de atributos semi-identificadores é muito grande, é que a maior parte dos atributos tem de ser suprimida para se obter k-anonimização, o que diminui a utilidade desses dados anonimizados[189]. Este é conhecido como problema da alta dimensionalidade dos dados em k-anonimização[190].

Portanto, com o propósito de superar essas anomalias a metodologia LKC-Privacidade busca garantir que, em um conjunto S de valores de atributos semi-identificadores, cada combinação de valores de tamanho máximo L em uma tabela T

[188] MOHAMMED, N.; FUNG, B. C.; HUNG, P. C.; LEE, C.-k. Anonymizing healthcare data: A case study on the blood transfusion service. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: ACM, 2009. (KDD '09), p. 1285–1294. ISBN 978-1-60558-495-9. Disponível em: < https://dl.acm.org/doi/10.1145/1557019.1557157>

[189] FUNG, B. C.; WANG, K.; FU, A. W.-C.; YU, P. S. Introduction to Privacy-Preserving Data Publishing: Concepts and Techniques. [S.l.]: Chapman-Hall, 2010. 376 p

[190] AGGARWAL, C. C. On k-anonymity and the curse of dimensionality. In: Proceedings of the 31st international conference on Very large data bases. [S.l.]: VLDB Endowment, 2005. p. 901–909.

seja compartilhada por pelo menos K registros, e a confiança da inferência de qualquer valor sensível em S não seja maior do que um valor probabilístico C que limita o espaço amostral de possibilidade da descoberta do proprietário do atributo, onde L, K e C são valores limites definidos[191].

Esta metodologia limita a probabilidade de sucesso na identificação do registro da vítima a ser menor ou igual a 1/K e a probabilidade de sucesso no ataque de ligação de atributo a ser menor ou igual a C, considerando que o conhecimento prévio do adversário não excede o valor de L.

O modelo LKC-Privacidade é adequado para anonimização de dados com alta dimensionalidade por possuir as seguintes propriedades[192]: requer que apenas um subconjunto de atributos semi-identificadores seja compartilhado por k registros; generaliza vários modelos tradicionais como por exemplo o k-anonimato; é flexível para ajustar o dilema entre privacidade de dados e utilidade de dados. Além disso, ele é um modelo de privacidade geral que evita ataques de ligação de registro e ligação de atributos[193].

Por exemplo[194], uma clínica publica um conjunto de dados de trajetória de pacientes com seus respectivos diagnósticos de saúde, os dados estão anonimizados como mostrado na Tabela 34. Cada registro contém apenas um atributo sensível - enfermidade, e que uma trajetória é representada por uma sequência de localizações L = (a,b, c,d, e, f,g,h).

Se um valor sensível aparece frequentemente em alguma sequência de localizações, informações sensíveis podem ser inferidas a partir de tal sequência.

---

[191] JUNIOR, Eliseu Castelo Branco, Uma Estrategia Para Assegurar a Confidencialidade De Dados Armazenados Em Nuvem. Disponível em: < http://repositorio.ufc.br/bitstream/riufc/23917/1/2017_tese_ecastelobrancojunior.pdf>. Acesso em: 05 set. 2020.

[192] Idem

[193] Mohammed, Noman & Fung, Benjamin & Hung, Patrick & Lee, Cheuk-kwong. (2009). Anonymizing healthcare data: A case study on the blood transfusion service. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 1285-1294. 10.1145/1557019.1557157.

[194] BRITO, Felipe Timbó. Uma abordagem distribuída para preservação de privacidade na publicação de dados de trajetória. 2016.

Suponha que um atacante tenha o conhecimento de que uma das pessoas esteve nas localizações *b* e *f*.

	Classe				Atributos Sensíveis
ID	**Trajetória**	**Sexo**	**Idade**	**Doador**	**Enfermidade**
1	b → d → c → f → h	Masculino	<20	Não	Hipertensão
2	f → h → e	Masculino	>20<22	Sim	Câncer
3	d → c → f → e	Masculino	>22<25	Não	Pneumonia
4	b → g → h → e	Masculino	>25<28	Sim	Dengue
5	d → f → e → b	Masculino	>30	Sim	Hipertensão
6	g → f → e	Masculino	>22<25	Não	Diabetes
7	b → f → h → e	Masculino	>25<28	Sim	Diabetes
8	b → g → f → h	Masculino	>20<22	Sim	Hipertensão
9	e → b → c	Masculino	<20	Sim	Pneumonia
10	c → h → f	Masculino	>30	Não	Arritmias cardíacas

Tabela 34 –Trajetórias de pacientes e seus diagnósticos de saúde

No conjunto de dados apresentado, dos quatro registros que possuem as localizações *b* e *f*, nomeadamente, os registros com IDs 1,5,7,8 apresentam o valor de atributo sensível Hipertensão. Assim, o atacante pode inferir que a pessoa em questão foi diagnosticada com Hipertensão com 75% de certeza.

	Semi-identificadores				Atributos Sensíveis
ID	**Trajetória**	**Sexo**	**Idade**	**Doador**	**Enfermidade**
1	d → f → h	Masculino	>30	Não	Hipetensão
2	f → h → e	Masculino	>20<22	Sim	Dengue
3	d → f → e	Masculino	>22<25	Não	Pneumonia
4	g → h → e	Masculino	>25<28	Sim	Dengue
5	d → h → e	Masculino	>30	Não	Doença cardiovascular
6	g → f → e	Masculino	>22<25	Não	Diabetes
7	f → h → e	Masculino	>25<28	Sim	Diabetes
8	g → f → h	Masculino	>20<22	Sim	Hipertensão
9	d →g →h	Masculino	<20	Sim	Pneumonia
10	f →d → g	Masculino	>30	Não	Doença cardiovascular

Tabela 35 – Versão anonimizada dos dados com L = 2, K = 2, C = 0.5

Por outro lado, tal inferência não seria possível com valor acima de 50% caso o conjunto de dados da Tabela 35 fosse publicado, já que foi anonimizado utilizando o modelo de privacidade LKC-privacidade com o valores L = 2, K = 2, ou seja, cada trajetória com duas localizações deve aparecer em pelo menos dois registros, consequentemente, C = 50%.

Para garantir que um conjunto de dados de trajetória atenda a esse modelo, itentifica-se as sequências de violação (IDs 5,10) e aplicam a generalização do atributo semi-identificador nos pares selecionados de violações.

8 - A CRIPTOGRAFIA

Para que o problema da eliminação segura de dados, seja tratado de forma eficiente, o conhecimento acerca da utilização de técnicas de criptografia é indispensável.

Heródoto[195] conta a história de um grego que precisava transmitir uma mensagem secretamente. Ele, então, raspa o cabelo do mensageiro, tatua a mensagem na cabeça e espera que o cabelo cresça.

Ao chegar ao destinatário, o mensageiro raspa a cabeça, revelando a mensagem. A principal deficiência deste tipo de técnica é que caso a mensagem seja descoberta, ela está aberta, podendo ser lida por qualquer um[196].

A criptografia utiliza o conceito de modificar a mensagem de forma que somente o destinatário possa entendê-la. Para que isso aconteça, a mensagem é embaralhada, usando alguma técnica combinada entre o emissor e o receptor, de forma que o segundo, e apenas ele, saiba arrumar, e recompor a mensagem original que o primeiro embaralhou[197].

É verdade que a utilização da criptografia tem aumentado muito nas últimas décadas, tanto em volume de dados criptografados quanto em quantidade de sistemas que utilizam a criptografia de dados.

Não obstante o já tradicional uso da técnica em sistemas criptográficos autônomos, tais como encriptação ou decriptação e sistemas de assinatura digital ou verificação digital, há diversas situações novas que utilizam a criptografia. A criptografia

[195] O historiador grego Heródoto (484 a.C. - 426 a.C) é considerado o pai da História. Viajou por vários países e sua obra principal é a “História”, dividida em nove livros.
[196] COSTA, Celso. FIGUEIREDO, Luiz Manoel. Introdução à Criptografia. Disponível em: < https://canal.cecierj.edu.br/recurso/4687>. Acesso em: 05 set. 2020.
[197] Ibidem.

homomórfica é uma técnica de criptografia que surgiu para atender a essas novas demandas do mercado.

8.1 Criptografia Homomórfica

Criptografia homomórfica são estruturas que permitem executar operações com segurança, em dados criptografados sem decodificá-los e sem fornecer a chave com que foram encriptados.

Desta forma, não havendo necessidade descriptografar os dados para utilizá-los, impõe-se um controle muito mais rigoroso sobre a disponibilidade dos dados, garantindo a sua integridade e confidencialidade.

A criptografia homomórfica está longe de ser uma descoberta recente, os especialistas em criptografia estão cientes de sua promessa desde a segunda metade da década de 70.

O conceito foi inicialmente apresentado em 1978 por Ronald Rivest, Leonard Adleman e M. L. Dertouzos[198] com a expressão homomorfismos privados.

O principal objetivo da criptografia homomórfica é poder realizar operações diretamente sobre os dados criptografados. Esse assunto reuniu profissionais de matemática e computação para viabilizar uma possível aplicabilidade dessa técnica de criptografia.

Os esquemas de criptografia homomórficos permitem delegar com segurança o processamento de dados a terceiros, que tem um significado importante em muitos aplicativos cliente-servidor. A segurança desses esquemas de criptografia homomórfica depende da dureza de alguns problemas matemáticos.

Essa criptografia permite realizar um conjunto de procedimentos computacionais diretamente sobre um texto cifrado sem a necessidade de decifrá-lo, impedindo assim

[198] R L Rivest, L Adleman, and M L Dertouzos. On data banks and privacy homomorphisms. Foundations of Secure Computation. Academic Press, 1978.

que a mensagem original seja de conhecimento do servidor ou de qualquer usuário que acompanhe a realização de tais procedimentos.

Neste sentido, é possível garantir a privacidade dos dados nos processos de comunicação e de armazenamento externo, sendo possível, por exemplo, delegar a realização cálculos ou o armazenamento dos dados.

Usando um esquema de criptografia homomórfica, o controlador criptografa os dados e os envia para realização dos cálculos sem decodificá-los.

Deste modo, a criptografia homomórfica surge como uma das principais ferramentas para sanar diversos problemas da criptografia moderna. A técnica faz parte da chamada criptografia pós-quântica: ramo da criptografia que estuda as classes de algoritmos criptográficos resistentes à criptoanálise baseada na computação quântica[199].

Os sistemas criptográficos homomórficos podem ser classificados em parcialmente ou completamente homomórficos. Os sistemas criptográficos completamente homomórficos suportam qualquer tipo de operações sobre os dados criptografados, ou seja, é possível realizar operações de adição e multiplicação nos dados. Já os sistemas criptográficos parcialmente homomórficos suportam apenas operações de adição.

8.1.1 Criptografia homomórfica completa

Ao longo dos anos 80 e 90, os pesquisadores criaram vários sistemas criptográficos interessantes. Muitos apoiaram algum tipo de homomorfismo, geralmente multiplicação ou adição. No entanto, nenhum parecia capaz de lidar com ambas as operações simultaneamente, pelo menos não sem sérias limitações.

[199] Daniel J. Bernstein. Introduction to post-quantum cryptography. In Post-Quantum Cryptography. 2009.

Esquemas de criptografia totalmente homomórficos são mais úteis na prática, mas, consequentemente, muito difíceis de conseguir, pois esse esquema é mais caro e trabalhoso.

Craig Gentry apresentou uma solução em sua tese de doutorado usando um esquema de criptografia baseado em rede[200]. Ele foi o primeiro a demonstrar que a criptografia homomórfica completa é teoricamente possível.

O esquema apresentado por ele suporta as operações lógicas básicas em textos cifrados, a partir dos quais é possível construir circuitos para realizar computação arbitrária.

Para construir um esquema de criptografia homomórfica, ele propôs um esquema dividido em três partes: primeiro, construir um esquema de criptografia que seja parcialmente homomórfico, ou seja, que possa avaliar funções de complexidade limitada; em segundo lugar, simplificar a função de descriptografia desse esquema o máximo possível; em terceiro lugar, avaliar esta função de descriptografia simplificada homomorficamente[201].

Teoricamente, todas as operações são possíveis de serem aplicadas em um texto cifrado, no entanto, verificou-se que a realização das operações de adição e multiplicação são suficientes para suportar todas as operações sobre um texto criptografado.

Portanto, se um esquema de criptografia homomórfica suporta tanto adição quanto multiplicação, ele também pode avaliar qualquer circuito aritmético em dados criptografados[202] e, portanto, pode-se dizer que é um esquema de criptografia completamente homomórfica.

Não obstante as vantagens apresentadas por este sistema criptográfico, o grande problema dos métodos propostos para os sistemas completamente

[200] Gentry, C. “Fully homomorphic encryption using ideal lattices”. In: Proceedings of the 41st annual ACM Symposium on Theory of computing, pp. 169–178. ACM, 2009.
[201] Idem.
[202] Ibidem.

homomórficos é que os seus tempos de execução, bem como o tamanho dos parâmetros utilizados, especialmente os das chaves públicas, crescem ciclicamente a cada iteração em complexidade na ordem de $O(\lambda^{10})$[203]. Onde λ é o parâmetro de segurança de todo o regime, e que define o tamanho, em bits de comprimento, das chaves geradas[204].

Deste modo, o grande desafio para tornar este sistema viável na prática consiste em resolver o elevado custo computacional exigido pelo sistema, além da necessidade de uma estrutura de armazenamento robusta em razão do excessivo tamanho das chaves públicas criadas por esses métodos.

8.1.2 Criptografia parcialmente homomórfica

É importante salientar que, em muitos dos algoritmos que implementam a criptografia homomórfica, a propriedade do homomorfismo foi observada antes mesmo de sua definição formal, ou seja, já era conhecida essa característica de permitir a realização de operações matemáticas em dados criptografados.

Esquemas de criptografia homomórficos são difíceis de implementar, e as construções conhecidas, podem não ser completamente suficientes para as necessidades de aplicações práticas do cotidiano.

Alguns esquemas de criptografia homomórfica multiplicativa e/ou esquemas de criptografia homomórfica aditiva podem alcançar o sigilo perfeito, ou seja, sem a chave secreta é impossível revelar o dado criptografado.

Para um esquema de criptografia homomórfica ser funcional, ele deve pelo menos permitir a avaliação de funções matemáticas úteis como adição, multiplicação e polinomiais.

---

[203] Coron, J., Mandal, A., Naccache, D., et al. “Fully homomorphic encryption over theintegers with shorter public keys”, Advances in Cryptology-, pp. 487–504, 2011.

[204] Gavinho Filho, J., Oliveira, J., Miceli, C., & da Silva, G. P. Aplicação do Algoritmo de Colônia de Formigas para Redução do Tamanho de Chaves Públicas em Criptografia Completamente Homomórfica.

Durante a busca pela identificação dos potenciais candidatos a algoritmo padrão, muitos algoritmos apresentaram vulnerabilidades críticas e logo foram descartados pelos estudiosos de criptografia. Embora a criptografia parcialmente homomórfica não tenha muita aplicabilidade no mundo real, os estudos destes sistemas contribuíram para o desenvolvimento de outros criptosistemas aceitos pelo meio científico atualmente.

No decorrer destes estudos, os pesquisadores criaram muitos criptosistemas que, geralmente, suportavam apenas uma operação: ou a operação de multiplicação, ou a de adição. Então, esses sistemas de criptografia são conhecidos como criptografia parcialmente homomórfico.

Dessa forma, o sistema de criptografia de chave pública RSA[205] é o início do estudo sobre criptografia homomórfica, pois foi a partir dele que as propriedades sobre homomorfismo foram observadas e, pela primeira vez, descritas em um trabalho científico[206].

8.1.3 Pesquisa de palavras sobre texto criptografado

A literatura apresenta alguns esquemas de criptografia homomórfica pesquisáveis, por exemplo, o esquema apresentado por Ning Cao ; Cong Wang ; Ming Li ; Kui Ren ; Wenjing Lou[207]. Outro esquema de criptografia que possui estas características é o esquema apresentado por Sky Faber, Stanislaw Jarecki, Hugo Krawczyk, Quan Nguyen, Marcel Rosu e Michael Steiner[208]. Outra pesquisa muito explorada nessa área é a procura em imagem[209].

---

[205] RIVEST, R. L.; SHAMIR, A.; ADLEMAN, L. A method for obtaining digital signatures and public-key cryptosystems. CACM, l21, p. 120–126, 1978.

[206] ADLEMAN, R.; ADLEMAN, L.; DERTOUZOS, M. On Data Banks and Privacy Homomorphisms. In: Foundations of Secure Computation. [S.l.]: New York: Academic Press, 1978. p. 169–180.

[207] N. Cao, C. Wang, M. Li, K. Ren and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data", in *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 222-233, Jan. 2014, doi: 10.1109/TPDS.2013.45.

208 S. Faber, S. Jarecki, H. Krawczyk, Q. Nguyen, M. Rosu, and M. Steiner. Rich queries on encrypted data: Beyond exact matches. In Computer Security - ESORICS 2015 - 20th European Symposium on Research in Computer Security, pages 123–145, 2015.

[209] B. Ferreira, J. Rodrigues, J. Leitao, and H. Domingos. Privacy-preserving content-based image retrieval in the cloud. CoRR, abs/1411.4862, 2014.

Realizar pesquisas de palavras-chave em um texto criptografado sem a necessidade de reverter a criptografia, é uma questão tratada neste trabalho, pois se aplicado um esquema de criptografia determinística do texto, palavra por palavra, esse funcionaria. Contudo, revelaria informações desnecessárias sobre o texto.

Portanto a solução está em criptografar um conjunto auxiliar de dados, que será usado para a pesquisa, em vez de criptografar o conjunto completo. O método para criar esse conjunto auxiliar de dados é[210]:

- dividir o conjunto original em palavras usando um determinado conjunto de separadores. Por exemplo; espaços, mudanças de linha, vírgulas;
- extrair todas as palavras distintas, sem repetir palavras;
- reorganizar o conjunto de palavras aleatoriamente;
- criptografar cada palavra com um esquema de criptografia determinístico;
- concatenar as palavras criptografadas usando um separador;
- concatenar tudo isso, com o texto real, depois de criptografá-lo com um forte esquema de criptografia com chave aleatória.

O atacante saberá quantas palavras distintas o conjunto simples possui, mas não a ordem ou a frequência em que aparecem nele, embora, com esse método, seja possível fazer toda a pesquisa de palavras-chave.

[210] SILVA, Eugénio Alves da (2016). Practical use of Partially Homomorphic Cryptography. Monografia do curso de Mestrado Bolonha em Segurança de Informação e Direito no Ciberespaço, Universidade de Lisboa em parceria com o Instituto Superior Técnico e Escola Naval, Lisboa - PT, 67p.

9 - ELIMINAÇÃO SEGURA DE DADOS

O acesso à internet é essencial ao exercício da cidadania e, ao usuário, é assegurado o direito da exclusão definitiva dos dados pessoais que tiver fornecido a determinada aplicação de internet, a seu requerimento, ao término da relação entre as partes, ressalvadas as hipóteses de guarda obrigatória de registros previstas na Lei que estabelece princípios, garantias, direitos e deveres para o uso da internet no Brasil [211] e na LGPD[212].

A previsão legal de que os titulares dos dados têm o direito de vê-los eliminados de forma segura foi consagrada na LGPD[213]. Contudo o legislador foi omisso quanto a maneira pela qual os dados devem ser eliminados. Motivo pelo qual este trabalho pretende propor uma solução.

Neste sentido, duas grandes questões se apresentam: a 1ª é como o controlador pode comprovar que os dados foram eliminados; e a derradeira, como os órgãos de fiscalização podem certificar a efetividade do processo de eliminação de dados.

A importância da resolução dessas questões é resultante da possível responsabilização civil e administrativa do controlador e dos responsáveis pelo tratamento dos dados. Além disso, é função dos órgãos de fiscalização certificar que os dados foram efetivamente eliminados.

Portanto, é fundamental e urgente o estabelecer processos estruturados com a geração de evidências físicas e lógicas, para eliminação de dados pessoais alinhados a LGPD[214].

Este procedimento servirá como alicerce para inequívoca certificação eliminação dos dados pelos órgãos de fiscalização.

---

[211] BRASIL, 2014, Lei n. 12.965, de 23 de abr. de 2014. Lei Brasileira que estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm; acesso em: 10 set 2020.

[212] BRASIL, 2018, Lei n. 13.709, de 14 de ago. de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm; acesso em: 10 set 2020.

[213] Ibidem.

[214] Ibidem.

9.1 Eliminação de dados por Criptografia

Atualmente muito se discute sobre a eliminação segura de dados pessoais, mas os principais métodos que se propõe aplicar não utilizam criptografia.

Em geral, a forma como os modernos sistemas de informações realiza o armazenamento, impede a eliminação definitiva dos dados digitais, ou seja, o processo de apagamento não garante a eliminação definitiva dos dados, pois apenas a referência aos dados é excluída. Desta forma, os dados continuam acessíveis e úteis.

Desta maneira, a remoção de arquivos em disco é um problema quando se deseja a eliminação definitiva dos dados[215].

Os cientistas Joel Reardon, Srdjan Capkun e David Basin do departamento de ciência da computação do instituto federal de tecnologia de Zurique propuseram o Data Node Encrypted File System (DNEFS)[216], que usa o esquema de criptografia para eliminar dados de forma eficiente e segura em sistemas de memória *flash*.

Os pesquisadores Jaeheung Lee, Junyoung Heo, Yookun Cho, Jiman Hong, e Sung Y. Shin3 no artigo "Secure Deletion for NAND Flash File System"[217] também apresentaram uma maneira de eliminar dados definitivamente de uma memória *flash*.

O mesmo assunto é abordado no artigo "An Efficient Secure Deletion Scheme for Flash File Systems" escrito por Zachary N. J. Peterson, Randal Burns, Joe Herring, Adam Stubblefield e Aviel D. Rubin[218], que apresentaram algoritmos e uma arquitetura para a eliminação segura de arquivo.

---

[215] Joel, R.; Srdjan, C., Basin, D. "Data Node Encrypted File System: Efficient Secure Deletion for Flash Memory." In: Proceedings of the 21st USENIX Conference on Security Symposium, USENIX Association, 2012, pp. 17–17.

[216] Joel, R.; Srdjan, C., Basin, D. "Data Node Encrypted File System: Efficient Secure Deletion for Flash Memory." In: Proceedings of the 21st USENIX Conference on Security Symposium, USENIX Association, 2012, pp. 17–17.

[217] Lee, J.; Heo, J.; Cho, Y.; Hong, J.; Shin, S. "Secure deletion for NAND flash file system". In Proceedings of ACM Symposium on Applied Computing, 2008, pp. 1710-1714.

[218] Peterson. Z.N.J.; Burns. R.; Herring. J.; Stubblefield. A.; Rubin. A.D. "Secure Deletion for a Versioning File System". In: Proceedings of the 4th Conference on USENIX Conference on File and Storage Technologies (FAST), 2005, vol. 4, pp. 143-154.

Os autores Feng Hao, Dylan Clarke e Avelino Francisco Zorzo no artigo “Deleting Secret Data with Public Verifiability”[219] descreveram uma solução criptográfica com o objetivo tornar o processo de eliminação de dados mais transparente e verificável.

Há na literatura muitos estudos publicados sobre esse tema, pois nas duas últimas décadas, este assunto foi amplamente tratado por pesquisadores na academia e na iniciativa privada. A preocupação em eliminar definitivamente os dados pairou sobre os estudiosos de tecnologia da informação. Algoritmos que usam criptografia para cifrar blocos, antes de escrevê-los na memória, foram bem recebidos pela comunidade.

Este método de criptografia faz com que cada bloco escrito no disco seja criptografado por uma chave distinta, e quando um bloco deve ser eliminado simplesmente se elimina a chave. Portanto, este sistema mostrou-se seguro tanto para dados existentes na memória quanto para dados a serem removidos.

No entanto o problema de eliminação segura de um arquivo é substituído pelo problema de eliminação segura de uma chave criptográfica. Utilizando esta técnica não é mais necessário eliminar grandes quantidades de dados, mas sim eliminar uma chave de 128 a 256 bits[220].

Visto que o principal objetivo é realizar uma eliminação segura, as chaves são geradas automaticamente, por meio de números randômicos, pelo sistema de proteção, e gerenciadas pelo próprio sistema, de forma que o controlador não precisa gerenciar chaves.

Em mídias de armazenamento tradicionais, a exclusão segura de dados é implementada substituindo o conteúdo de um arquivo por outras informações, por exemplo, escrevendo zeros ou uns sobre as informações que deverão ser apagadas[221],

[219] Hao, F.; Clarke, D.; Zorzo, A. F. “Deleting Secret Data with Public Verifiability.” IEEE Transactions on Dependable and Secure Computing, vol. 13 (6), 2016, pp. 617–29.
[220] Hao, F.; Clarke, D.; Zorzo, A. F. “Deleting Secret Data with Public Verifiability.” IEEE Transactions on Dependable and Secure Computing, vol. 13 (6), 2016, pp. 617–29.
[221] Plumb, C. “shred(1) - Linux man page”. Página do comando shred. Capturado em: https://linux.die.net/man/1/shred. Dezembro 2015.

ou modificando o sistema de arquivos para, assim, substituir automaticamente qualquer setor utilizado para armazenamento de dados[222].

Os métodos tradicionais de sobrescrita utilizados em meios magnéticos não são efetivos em sistemas de armazenamento de dados mais modernos. Por isso é preciso garantir uma maneira de eliminar os dados destes sistemas de forma segura, impossibilitando que eles sejam recuperados.

De volta a utilização de criptografia para eliminação de dados, se a matemática por trás da criptografia for boa, a reversibilidade do processo é impossível e improvável.

Por exemplo, é improvável que uma chave simétrica de 256 bits (com matemática sólida) seja quebrada tão cedo. Como essa afirmação ainda é verdadeira no mundo atual de milhões de computadores baseados em nuvem e unidades de processador gráfico mais rápidas (GPUs)? Os atacantes não podem tentar adivinhar todas as combinações possíveis resultantes de 256 0 e 1 diferentes? A resposta é não, devido ao grande número de chaves possíveis que podem ser formuladas a partir de todas as combinações das chaves de 256 bits[223].

Todos os computadores do mundo não têm memória ou espaço suficiente no disco rígido para armazenar tudo o que seria necessário para quebrar uma chave de criptografia decente. Se você pudesse, de alguma forma, criar e conectar um bilhão de computadores, cada um fazendo 2 bilhões de palpites por segundo, ainda levaria mais tempo do que a idade estimada dos computadores existentes no universo (por exemplo, 14 bilhões de anos). E não há energia suficiente no universo para realmente realizar o intervalo[224].

[222] Bauer, S.; Priyantha, N. B. “Secure Data Deletion for Linux File Systems”. In: Proceedings of the Usenix Security Symposium, 2001, pp. 153–164.
[223] “O que é apagamento por criptografia? Disponível em: https://www.blancco.com/resources/article-what-is-cryptograhic-erasure. Acesso em 03/08/2020.
[224] Ibidem.

9.1.1 Método de criptografia (método de J. Lee)

Este método, proposto por Jaeheung Lee e seus colegas Junyoung Heo, Yookun Cho, Jiman Hong, Sung Y. Shin[225], propõe utilizar criptografia para realizar uma eliminação segura de dados.

As chaves de criptografia são geradas aleatoriamente, e os dados são criptografados usando essas chaves de criptografia.

As chaves de criptografia de um arquivo específico são armazenadas em uma unidade de apagamento, como pode ser visto a seguir.

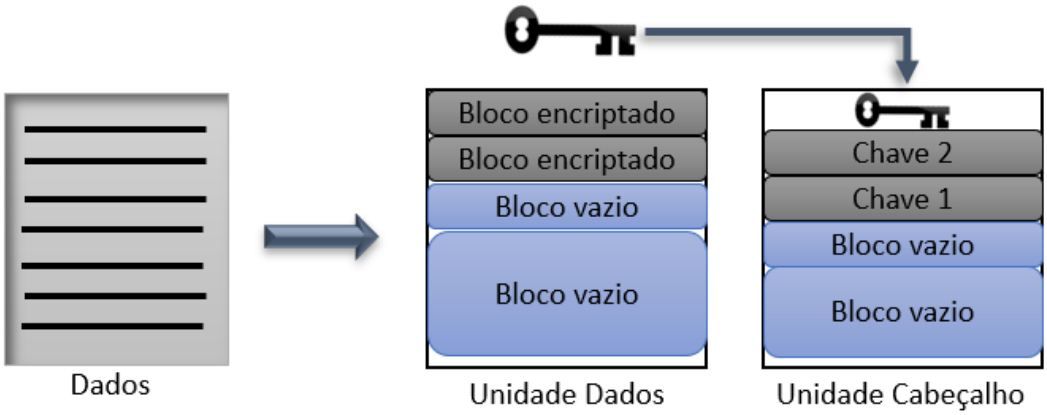

Figura 2 - Processo de armazenamento de dados com criptografia.
Fonte: J. Lee, et al, tradução nossa.

Portanto, a chave que é usada para criptografar os dados é armazenada em um lugar diferente dos dados e, quando for necessário eliminar os dados, é suficiente que se apague a chave com a qual foi criptografado.

[225] J. Lee, J. Heo, Y. Cho, J. Hong, and S. Shin, "Secure deletion for NAND flash file system," in Proceedings of ACM Symposium on Applied Computing, 2008, pp. 1710- 1714.

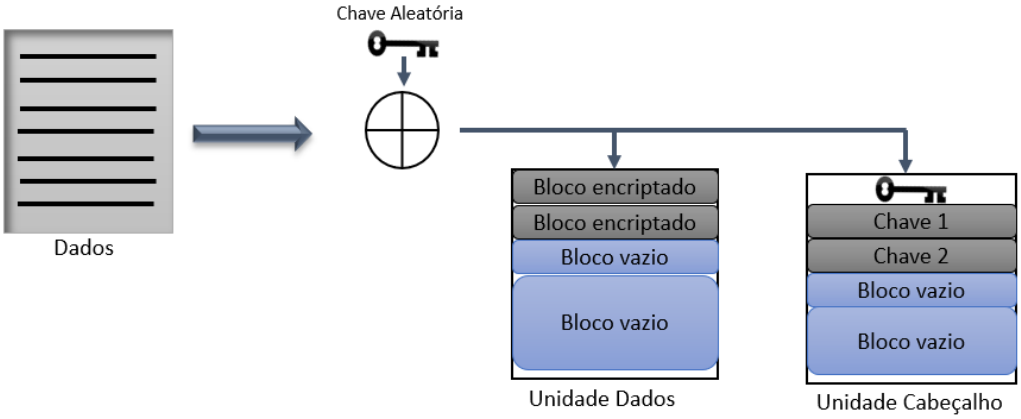

Figura 3 - Armazenamento de dados e chaves geradas aleatoriamente.
Fonte: J. Lee, et al, tradução nossa.

Durante o processo de eliminação, a chave de um arquivo específico é apagada. Embora os dados eliminados ainda permaneçam na memória *flash*, estes permanecem criptografados, e as chaves de decodificação já eliminadas. Desta forma, considera-se seguro o processo de eliminação de dados.

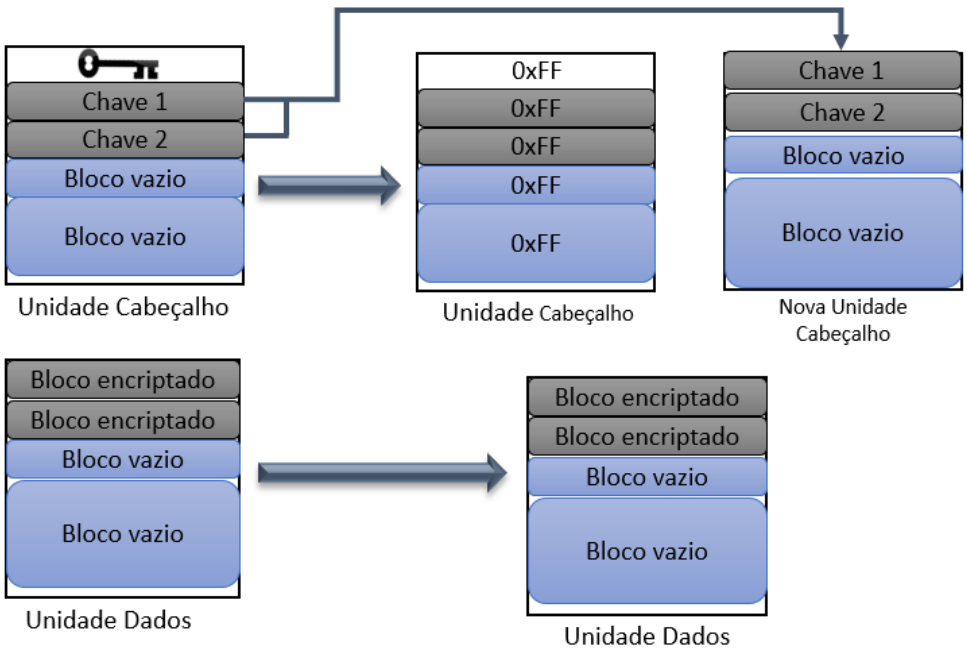

Figura 4 - Processo de remoção de dados no método cifrado
Fonte: J. Lee, et al, tradução nossa.

O método de Lee[226] estabelece que, para realizar a eliminação de dados, inicialmente é necessário pesquisar uma unidade de cabeçalho que armazena as chaves de criptografia de dados apagados.

Depois da pesquisa realizada, deve-se verificar se esta unidade contém outras chaves de criptografia válidas, além das chaves que devem ser excluídas. Se chaves de criptografia válidas estão armazenadas, deve ser eliminada. Por fim, deve-se apagar a unidade de cabeçalho que armazena as chaves pesquisadas no início do processo.

Portanto, somente uma operação de apagamento é necessária para eliminar dados de forma segura. Além disso, os metadados desse arquivo que está na mesma unidade, também são eliminados.

9.1.2 Método de criptografia com apagamento (método de B. Lee)

Este método foi proposto por Byunghee Lee, Kyungho Son, Dongho Won e Seungjoo Kim[227] no artigo “Secure Data Deletion for USB Flash Memory” e utiliza apagamento criptográfico por meio de duas passagens.

Na primeira passagem, os blocos do arquivo são sobrescritos com zero. Na segunda, realiza-se a eliminação da unidade de apagamento que contém estes blocos, após copiar os blocos válidos para uma nova posição.

Para maior segurança na remoção da chave criptográfica, o método de B. Lee realiza uma sobrescrita da chave com zeros.

---

[226] J. Lee, J. Heo, Y. Cho, J. Hong, and S. Shin, “Secure deletion for NAND flash file system,” in Proceedings of ACM Symposium on Applied Computing, 2008, pp. 1710- 1714.

[227] Lee, B.; Son, K.; Won, D.; Kim, S. “Secure Data Deletion for USB Flash Memory”. Journal of Information Science and Engineering, vol. 27, 2011, pp. 933-952.

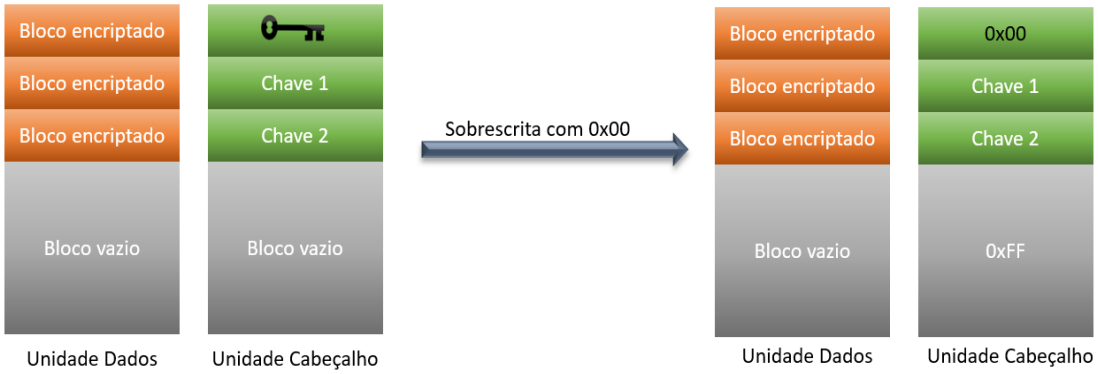

Figura 5 - Processo de sobrescrita da chave no método de B. Lee
Fonte: Byunghee Lee, et al, tradução nossa.

Após a realização da etapa anterior, efetua-se o apagamento da chave, como mostrado na figura a seguir:

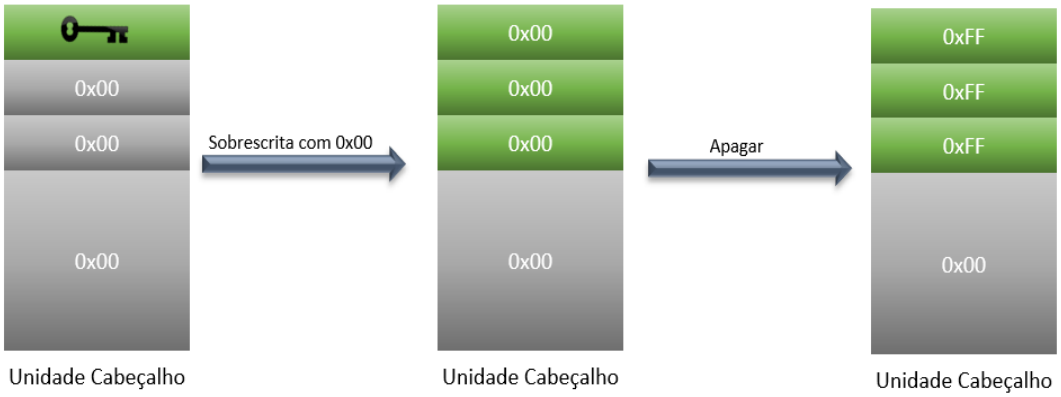

Figura 6 - Processo de apagamento no método de B. Lee
Fonte: Byunghee Lee, et al, tradução nossa.

Basicamente, deve-se procurar por uma unidade de cabeçalho que contenha a chave criptográfica do arquivo a ser eliminado, ao encontrar a chave deve-se sobrescreve está chave com um padrão de 0x00. Depois, verifica-se a existência de outras chaves ainda válidas na unidade de cabeçalho.

Se existirem outras chaves validas na unidade de cabeçalho, o processo de eliminação termina. Caso não exista mais qualquer chave válida na unidade de cabeçalho, então se realiza o apagamento da unidade de cabeçalho[228].

[228] Lee, B.; Son, K.; Won, D.; Kim, S. “Secure Data Deletion for USB Flash Memory”. Journal of Information Science and Engineering, vol. 27, 2011, pp. 933-952.

9.2. Proposta de processo para eliminação definitiva dados

Idealmente, o sistema projetado para eliminação de dados deve: permitir a seleção de um padrão específico, com base em necessidades únicas; remover os dados em todo o dispositivo; averiguar se o método de remoção foi feito corretamente; e suportar o processo de auditoria quando necessário.

A eliminação definitiva de dados vai além dos comandos básicos de exclusão de arquivos, que só removem ponteiros diretos para os setores de disco de dados e tornam a recuperação de dados possível com ferramentas de software comuns.

Geralmente, o processo de eliminação de dados não se destina a apagar todos os dados da mídia. Muitas vezes sequer é possível o apagamento de todos os dados no dispositivo, como por exemplo, quando os dados estão armazenados na nuvem.

A exclusão de arquivos na nuvem pode sugerir que o usuário não está deixando rastros do arquivo, mas é bem o contrário. Quando alguém opta por excluir um arquivo da nuvem, na maioria dos casos, ainda restam cópias desse registro em servidores de backup e armazenamentos na redundância.

Ou seja, mesmo depois que os arquivos são excluídos, na prática uma cópia permanece no servidor. Na maioria das vezes, há backups de seus dados excluídos permanentemente no servidor por um determinado período. A boa notícia sobre esses backups é que eles não estão prontamente acessíveis, pois é necessário a interferência de um técnico do serviço de nuvem para viabilizar a restauração, uma vez que estão sob várias camadas de protocolos de segurança. Diante disso, o vazamento desses dados é bastante improvável.

Portanto, no cenário apresentado e, em muitos outros casos, não é viável a destruição física das mídias de armazenamento, pois com esse procedimento as mídias tornam inutilizáveis. Ao contrário, a eliminação segura dos dados deixa o dispositivo de armazenamento completamente funcional.

Existem algumas diferenças entre a eliminação de dados e outros métodos de sobrescrever dados, que podem deixar os dados intactos e aumentar o risco de violação de dados, roubo de identidade ou falha em alcançar a conformidade normativa.

Inúmeros sistemas para erradicação de dados também fornecem vários tipos de sobrescritas para que eles usem padrões reconhecidos pelo governo e do setor privados, embora uma sobrescrita de passe único seja amplamente considerada suficiente para os discos rígidos modernos.

As ferramentas de eliminação de dados também podem direcionar dados específicos em um disco para eliminação de rotina, fornecendo um método de proteção de *hackers*, mais rápido que a criptografia de software.

Estabelecer um procedimento para a realização desta atividade é imprescindível para mitigar os riscos de responsabilização civil ou criminal, tanto da organização, quanto para o encarregado pela proteção de dados.

O procedimento a seguir baseia-se na utilização de criptografia e deve ser utilizado como um suporte para orientar, organizar, oficializar e padronizar as atividades de eliminação definitiva de dados. Além disso, contempla a utilização racional dos recursos, a uniformidade das tarefas, irreversibilidade e geração de evidências para o suporte aos processos de auditoria e fiscalização.

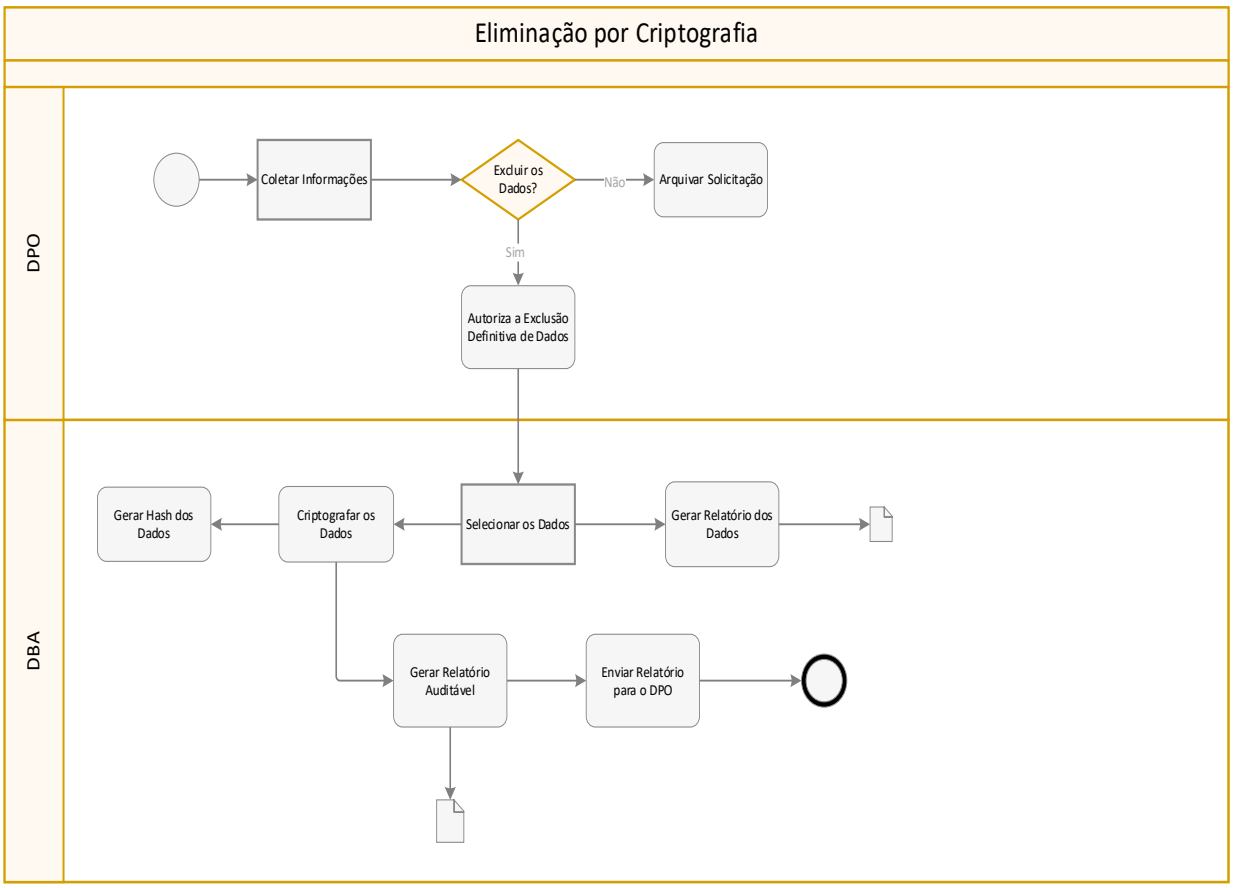

Figura 7 Procedimento Eliminação definitiva de dados por Criptografia
Fonte: Elaborado pelo Autor.

Para execução do procedimento, são imprescindíveis a segregação das funções executivas e a dupla checagem.

Sobre o EPD recai a integral responsabilidade de adotar as providências relacionadas ao tratamento de dados pessoais. Por isso, o EPD deve ter profundos conhecimentos em segurança da informação e proteção de dados, dispor de recursos adequados, reporte direto à alta gestão e fundamentalmente atuar com absoluta independência.

A eliminação dos dados será processada conforme descrito na figura apresentada. Inicialmente será gerado o *hash* dos dados que serão eliminados. Esse *hash* é imprescindível para o processo de auditoria posterior à eliminação dos dados.

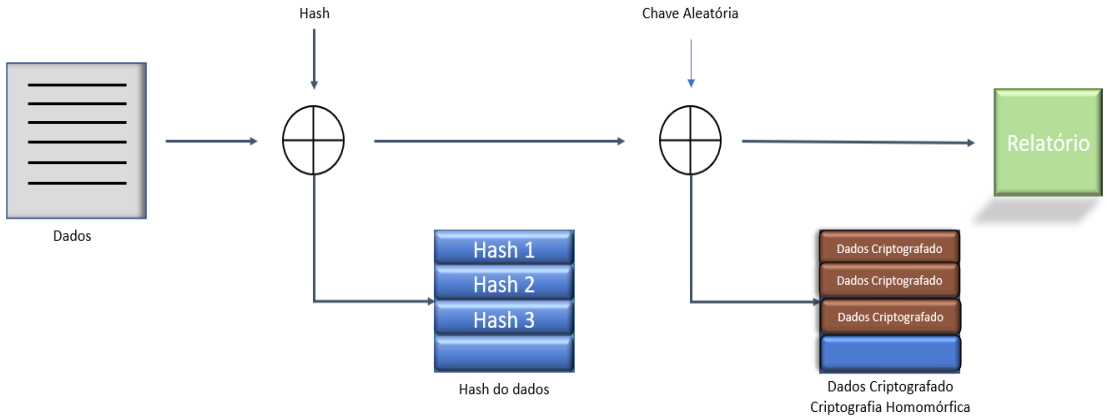

Figura 8 - Eliminação de dados por criptografia homomórfica
Fonte: Elaborado pelo Autor.

Na etapa seguinte, os dados serão cifrados pelo processo de criptografia homomórfica com uma chave gerada aleatoriamente. É importante ressaltar que a chave gerada não será armazenada, portanto, não será possível reverter o processo e obter os dados originais novamente.

O procedimento de criptografia homomórfica também será indispensável para a auditoria subsequente, pois serão utilizadas as propriedades desse método para evidenciar o processo da eliminação dos dados.

Em seguida será gerado um relatório com informações relacionadas à execução do processo desde o requerimento até a efetiva eliminação dos dados. Este relatório será a base para realização de auditoria e fiscalização na organização, acerca dos dados submetidos ao processo de eliminação definitiva.

9.2.1 Criptografia de dados para eliminação

A eliminação de dados por meio da criptografia, remove as chaves utilizadas para encriptação, tornando os dados inacessíveis. Esse método quando aplicado corretamente, garante extrema eficácia na eliminação dos dados.

Os pesquisadores Boneh e Lipton[229] foram uns dos primeiros a propor o uso de criptografia para resolver o problema de eliminação segura de dados: uma solução de descarte de dados que funciona criptografando todos os dados antes de salvá-los para o disco e, posteriormente, tornando os dados inacessíveis por meio do descarte da chave de encriptação.

Esta abordagem é especialmente desejável quando as duplicações de dados são feitas em locais distribuídos, tornando impossível eliminar todas as cópias. O uso de criptografia essencialmente muda o problema de exclusão de uma grande quantidade de dados para a exclusão de uma chave[230].

Quando a criptografia é usada para resolver o problema de eliminação de dados, o gerenciamento das chaves torna-se extremamente importante. Existem várias abordagens propostas na literatura para o gerenciamento das chaves criptográficas. Entretanto ainda é preciso definir o processo de exclusão das chaves utilizadas[231]. Para resolver essa questão, propõe-se a utilização de uma chave gerada aleatoriamente. Deste modo, ao criptografar os dados, automaticamente esses dados se tornam instantaneamente inacessíveis.

Portanto, é imprescindível que a chave utilizada para criptografar os dados seja aleatória, e em hipótese alguma seja armazenada no dispositivo. Deve-se utilizar uma das opções de geração de sequências aleatórias para obter uma chave segura.

Além disso, a criptografia utilizada no processo deverá possuir propriedades homomórficas. Deste modo, os dados criptografados poderão ser objeto de pesquisa para certificação e geração dos relatórios de auditoria.

---

[229] D. Boneh, R. Lipton, “A Revocable Backup System,” Proceedings 6th USENIX Security Conference, pp. 91-96, 1996.

[230] Idem.

[231] F. Hao, D. Clarke, and A. F. Zorzo, “Deleting secret data with publicverifiability, ”IEEE Transactions on Dependable and Secure Computing,vol. 13, no. 6, pp. 617–629, 2015.

9.2.2 Geração aleatória de chave

Na geração de chaves, a busca por sequências aleatórias é muito importante, principalmente no sentido de produção de chaves fortes. Para a geração das chaves aleatórias, recomenda-se a utilização de geradores de sequências aleatórias.

As sequências aleatórias são imprescindíveis no processo de criptografia. Sistemas criptográficos não devem ser totalmente secretos: possuem alguns elementos de conhecimento público – algoritmos, alguns elementos conhecidos apenas pelos integrantes do sistema – chaves, e, em alguns casos, devem possuir alguns elementos secretos e imprevisíveis[232] - números aleatórios.

A título de exemplo de evento aleatório, pode-se imaginar o seguinte cenário: duas pessoas fazem uma disputa de cara e coroa, com uma moeda não viciada, ou seja, a probabilidades de a moeda cair com cada um dos lados para cima é exatamente 50%. Neste cenário, ao escolher um dos lados, o jogador tem 50% de chance de ser o vencedor. Também é aleatório o lançamento de um dado não viciado, pois cada uma das faces tem a probabilidade de 16,66% de ser o número sorteado.

Além dessas fontes aleatórias, há outras fontes que podem prover sequências verdadeiramente aleatórias, por exemplo, o ruído do diodo, o material radioativo e, ainda, o dispositivo T12RNG. O ruído de um diodo pode servir para produzir sequências de números aleatórios. Isso acontece por causa de defeitos inerentes nos materiais semicondutores do diodo e seu comportamento na região limite de polarização. Os materiais radioativos também podem funcionar como um gerador de números aleatórios, uma vez que a radioatividade decai aleatoriamente. Um contador Geiger, conectado a um computador, pode ser usado para transformar estes dados em informação útil e administrável em criptografia. O T12RNG é um dispositivo capaz de gerar números aleatórios por meio de um hardware, que pode ser conectado a uma

[232] A. C. d. Araújo Neto, “Um algoritmo de criptografia de chave pública semanticamente seguro baseado em curvas elípticas,” 2006.

porta serial de computador. A origem da sua aleatoriedade na geração é devido ao ruído branco, propriedade dos semicondutores instalados no circuito[233].

No entanto, é muito difícil produzir algo perfeitamente aleatório em um computador, pois são máquinas determinísticas, ou seja, com um conjunto de dados de entrada, executam ações que geram sempre a mesma saída.

Assim sendo, para obter alguma aleatoriedade, deve-se misturar o resultado de elemento do mundo físico[234], como por exemplo a data e hora do sistema, o padrão de digitação do usuário, ou ainda dados como temperaturas, velocidade de rotação dos discos do HD, latência de tráfego na rede e movimentação do mouse pelo usuário, ou qualquer outro evento que aconteça no sistema frequentemente, mas de forma aleatória são fontes que, combinadas, podem gerar valores satisfatoriamente aleatórios.

A maioria dos geradores de sequência aleatória existentes na atualidade não são verdadeiramente aleatórios, mas para o propósito de auxílio aos processos criptográficos com a finalidade de apagamento dos dados, esses geradores são suficientes, são mais baratos de serem implementados e facilitam o gerenciamento das chaves geradas. O gerador de sequência de números pseudoaleatórios deve ter as seguintes características: assemelhar-se a uma sequência aleatória, ou seja, esta sequência deve passar em todos os testes estatísticos de aleatoriedade possíveis.

Uma sequência pseudoaleatória pode repetir. Essa repetição é definida como o período da sequência, mas estes períodos podem ser enormes e, por esse motivo, uma sequência pseudoaleatória é viável para se utilizar nos processos de criptografia visto que é mais fácil de implementar e possui um período muito grande, aumentando, assim, a facilidade de obtenção e de gerenciamento de chaves.

Existem vários métodos geradores de números pseudoaleatórios e a técnica mais comumente aplicada para a geração desses números faz uso de uma relação

---

[233] E. C. V. Borges Júnior, “Introdução a sistemas criptográficos e o uso de geradores de sequências de números aleatórios e pseudoaleatóorios,”2014.

[234] B. Schneier, “Protocol building blocks, Applied Cryptography, Second Edition: Protocols, Algorithms, and Source Code in C”, 2015.

recursiva na qual, o próximo número na sequência é uma função do último número gerado, isto é:

$$x_n = f(x_{n-1}, x_{n-2}, x_{n-3}, x_{n-4}, x_{n-5}, \ldots)$$

Para ser criptograficamente segura, uma sequência pseudoaleatória gerada deve possuir duas características: não pode ser reproduzida fielmente e ser imprevisível. A primeira característica determina que se um gerador for iniciado com a mesma entrada já utilizada, a saída deve ser uma sequência aleatória completamente independente da primeira. A segunda característica é a garantia de que seja computacionalmente impraticável prever o próximo bit aleatório da sequência, conhecendo todo o algoritmo, o hardware gerador ou toda a sequência anterior de bits.

Algoritmos de geração pseudoaleatórios servem para criar mecanismos auxiliares para a implementação de geradores de sequências de números aleatórios que podem ser utilizados na geração das chaves aleatórias para aplicação na eliminação e dados.

Existem vários métodos computacionais para geração de números pseudoaleatórios. Todos ficam aquém do objetivo da verdadeira aleatoriedade, embora possam cumprir, com sucesso variável, alguns dos testes estatísticos de aleatoriedade destinados a medir o quão imprevisível são seus resultados.

As primeiras tentativas de desenvolver um método computacional de geração de números aleatórios foram os métodos congruente linear[235], registrador de deslocamento de feedback linear[236] e o gerador geffe[237].

Atualmente, vários geradores de números pseudoaleatórios utilizados são modificações do método linear congruente. Este Método se baseia na seguinte relação: $x_{n+1} = (a.x_n + b)(mod\ m), n \geq 0$. O valor x_0 inicial é chamado semente, *a* é o multiplicador, *b* o incremento, *m* é a quantidade de número diferentes que se deseja

[235] B. Schneier, “Protocol building blocks”, Applied Cryptography, Second Edition: Protocols, Algorithms, and Source Code in C, 2015.
[236] E. C. V. Borges Júnior, “Introdução a sistemas criptográficos e o uso de geradores de sequências de números aleatórios e pseudoaleatóorios”,2014.
[237] Philip R. Geffe. How to protect data with ciphers that are really hard to break. Electronics, pages 99–101, 1973.

gerar, *n* é o índice dos números. A escolha adequada desses valores é fundamental para a determinação do período da sequência de saída.

Um registrador de deslocamento de *feedback* linear (*linear feedback shift register* - LSFR) é um registrador de deslocamento que assume uma função linear de um estado anterior como uma entrada para a geração de uma determinada saída. Mais comumente, essa função é um OU exclusivo (XOR). A geração de uma sequência aleatória é obtida da seguinte forma: a cada intervalo de tempo definido no gerador, um bit é requisitado e o bloco restante é deslocado para direita e um novo bit é inserido à esquerda. Este bit inserido é resultado de uma determinada operação realizada com os demais bits correntes do registrador.

O gerador de sequência aleatória Geffe usa uma combinação de três LFSRs de forma não linear, pois o LFSR não deve ser utilizado isoladamente, visto que, por ser linear, apresenta algumas fraquezas que podem ser exploradas. Por esse motivo o gerador Geffe utiliza-se do LFSR combinado com alguma função não linear, o que resulta em bons geradores pseudoaleatórios. Um gerador Geffe consiste em três LFSRs: dois LFSRs como multiplexadores de entrada e um LFSR como controlador.

No entanto, atualmente, existem geradores de números pseudoaleatórios cuidadosamente projetados e criptograficamente seguros, com recursos especiais projetados especificamente para uso em criptografia.

10 - CONCLUSÃO

Essa dissertação propôs a realização de uma análise descomplicada da origem, fundamentos e relacionamentos legais, bem como, investigar e apresentar de forma clara e objetiva, a aplicação de métodos e técnicas de anonimização, pseudonimização e eliminação de dados alinhados à Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018). Assim sendo, passamos a expor os principais conhecimentos obtidos pela investigação e análise.

A RGPD europeia serviu de arcabouço para os legisladores brasileiros conceberem a LGPD, que possui largo espectro de diálogo com leis que também garantem direitos individuais.

A inovação está no Decreto 10.474[238] de 26/08/2020, na forma como seu Anexo I, Capitulo I, Artigo 2º, estabeleceu o tratamento diferenciado dos dados pessoais de idosos, nos termos da Lei 10.471/2003[239], o Estatuto do Idoso.

O direito à privacidade e à autodeterminação são contemplados pela LGPD, na medida em que, diante do progresso tecnológico, garante a proteção dos dados pessoais do titular, para que ele consinta e tenha controle pleno sobre a coleta, a exatidão, a finalidade, sua forma de uso e a eliminação de seus dados.

Embora a Responsabilidade Civil e as Sansões Administrativas estejam claramente estabelecida na LGPD, a criminalização de más condutas envolvendo dados pessoais não foi contemplada. No entanto verificamos que, desde 2019, o Parlamento Brasileiro, tem em mãos o Anteprojeto de lei de Proteção de Dados para segurança pública e investigação criminal, que poderá corrigir tal deficiência.

Tal como a RGPD europeia, a LGPD estabelece que, para desidentificação dos dados, aplica-se a anonimização, em seguida restringe o uso da pseudoniminização e também trata da eliminação de dados. Todavia, em todas as situações foram subjetivas

---

[238] https://www.in.gov.br/web/dou/-/decreto-n-10.474-de-26-de-agosto-de-2020-274389226, acessado em 10/02/2021

[239] http://www.planalto.gov.br/ccivil_03/leis/2003/l10.741.htm, acessado em 10/02/2021.

em relação ao resultado esperado e absolutamente omissas no que tange à execução do processo.

Diante disso, cabe exclusivamente ao controlador decidir quais técnicas e métodos utilizar para atender a LGPD e, consequentemente, satisfazer os anseios do órgão de fiscalização, pois a ele caberá a interpretação e aplicação da lei.

Em relação à pseudonimização a LGPD, seu uso é limitado ao tratamento de dados sensíveis e sob o aspecto de privacidade e proteção, sendo indispensável sua conciliação com técnicas de anonimização, pois este tratamento utilizado isoladamente é alvo fácil para um ataque de reidentificação.

A análise pormenorizada das diferentes técnicas e métodos anonimização seus usos e combinações, elencados neste trabalho, constatou que, se aplicadas corretamente para o fim que se pretende, ou seja, a análise e o alinhamento da técnica, ao cenário e o objetivo desejado, trata-se de uma solução segura para proteção e privacidade de dados pessoais, pois garante a impossibilidade de identificar o titular dos dados, satisfazendo o objetivo da LGPD.

O processo de eliminação de dados pessoais por meio da utilização de geradores de chaves aleatórias e técnicas de criptografia com propriedades homomórficas, apresentou vantagens no que se refere à segurança ponta a ponta do processo, que são as bases para realização dos procedimentos de auditoria e fiscalização e fundamentalmente atender ao estabelecido na LGPD.

Após o processamento do hash do arquivo, a geração da chave aleatória é o que garante a irreversibilidade do processo de eliminação dos dados, que atende o estabelecido no Art. 5º, XIV da LGPD[240].

A geração de evidências da eliminação de dados consiste na possibilidade de sua certificação analítica e sintética, que pode ser obtida através da análise de

[240] Art. 5º Para os fins desta Lei, considera-se:
XIV - eliminação: exclusão de dado ou de conjunto de dados armazenados em banco de dados, independentemente do procedimento empregado;

metadados e da execução de algoritmos de busca no arquivo criptografado, atendendo ao disposto no Art.19 da LGPD[241].

Temos que a certificação do processo de eliminação segura do dado pode ser realizada mediante:

- ✓ a conferência do hash do arquivo processado;
- ✓ a aplicação de testes de aleatoriedade da chave criptográfica ;
- ✓ a certificação da lisura de todo o processamento criptografia do arquivo, através da preservação e perícia técnica;
- ✓ a aplicação de algoritmo de busca de registro no arquivo criptografado.

Este trabalho apresenta noções gerais precisas do que é exigido de um controlador de dados honesto para atender fielmente os comandos legais da LGPD, pelo menos intuitivamente quando espelhado nos entendimentos legais já pacificados na RGPD europeia.

Ainda oferece definições tecnicamente precisas para a exclusão de dados que representam possibilidades de interpretações do que a lei poderia razoavelmente esperar, e alternativas para o que versões futuras da lei poderiam explicitamente exigir.

Desta forma, é possível afirmar com razoável segurança que um controlador de dados está em conformidade com a LGPD, se ele atende as recomendações propostas

[241] Art. 19. A confirmação de existência ou o acesso a dados pessoais serão providenciados, mediante requisição do titular:
I - em formato simplificado, imediatamente; ou
II - por meio de declaração clara e completa, que indique a origem dos dados, a inexistência de registro, os critérios utilizados e a finalidade do tratamento, observados os segredos comercial e industrial, fornecida no prazo de até 15 (quinze) dias, contado da data do requerimento do titular.
§ 1º Os dados pessoais serão armazenados em formato que favoreça o exercício do direito de acesso.
§ 2º As informações e os dados poderão ser fornecidos, a critério do titular:
I - por meio eletrônico, seguro e idôneo para esse fim; ou
II - sob forma impressa.
§ 3º Quando o tratamento tiver origem no consentimento do titular ou em contrato, o titular poderá solicitar cópia eletrônica integral de seus dados pessoais, observados os segredos comercial e industrial, nos termos de regulamentação da autoridade nacional, em formato que permita a sua utilização subsequente, inclusive em outras operações de tratamento.
§ 4º A autoridade nacional poderá dispor de forma diferenciada acerca dos prazos previstos nos incisos I e II do caput deste artigo para os setores específicos.

neste trabalho, em especial quanto anonimização, pseudonimização e eliminação segura de dados.

Por fim, consideramos que a presente dissertação apresenta um claro e objetivo conjunto de noções e orientações, legais e tecnológicas sobre os princípios que norteiam os processos de anonimização, pseudonimização e eliminação de dados alinhados à Lei Geral de Proteção de Dados Brasileira (Lei n.º 13.709/2018).

11 - BIBLIOGRAFIA

_____, “Opinion 05/2014 on Anonymisation Techniques”. Article 29 **Data Protection Working Party** (European Commission), em 10 de Abril de 2014. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em: 10 ago 2020.

A. C. d. Araújo Neto, **Um algoritmo de criptografia de chave pública semanticamente seguro baseado em curvas elípticas.** Porto Alegre: PPGC da UFRGS, 2006.

ADLEMAN, R.; ADLEMAN, L.; DERTOUZOS, M. **On Data Banks and Privacy Homomorphisms**. In: Foundations of Secure Computation. [S.l.]: New York: Academic Press, 1978. p. 169–180.

AGGARWAL, C. C**. On k-anonymity and the curse of dimensionality**. In: Proceedings of the 31st international conference on Very large databases. [S.l.]: VLDB Endowment, 2005. p. 901–909.

ALMISHARI Mishari [et al.] - **Stylometric linkability of Tweets**. In: Proceedings of 13th Workshop on Privacy in the Electronic Society, 2016.

B. Ferreira, J. Rodrigues, J. Leitão, and H. **Domingos. Privacy-preserving content-based image retrieval in the cloud**. CoRR, abs/1411.4862, 2014.

B. Schneier, **Protocol building blocks Applied Cryptography**, Second Edition: Protocols, Algorithms, and Source Code in C, 2015.

BACEN, 2018, Resolução n. 4.658, de 26 de abr. de 2018. **Resolução que Dispõe sobre a política de segurança cibernética** e sobre os requisitos para a contratação de serviços de processamento e armazenamento de dados e de computação em nuvem a serem observados pelas instituições financeiras e demais instituições autorizadas a funcionar pelo Banco Central do Brasil. Disponível em: https://www.bcb.gov.br/pre/normativos/busca/downloadNormativo.asp?arquivo=/Lists/Normativos/Attachments/50581/Res_4658_v1_O.pdf; acesso em: 10 set 2020.

BASSO, T.; MATSUNAGA, R.; MORAES, R.; ANTUNES, N. **Challenges on anonymity, privacy, and big data**. In: IEEE. Dependable Computing (LADC), 2016 Seventh Latin-American Symposium on. [S.l.], 2016. p. 164–171.

BAUMAN, Zygmunt. **Vigilância líquida:** diálogos com David Lyon. Trad. Carlos Alberto Medeiros. Rio de Janeiro: Zahar, 2013.

BELLOVIN, Steven M. and Dutta, Preetam K. and Reitinger, Nathan, **Privacy and Synthetic Datasets** (August 20, 2018). Stanford Technology Law Review, Forthcoming. Available at SSRN: https://ssrn.com/abstract=3255766, 2018.

BERNSTEIN, Daniel J. **Introduction to post-quantum cryptography.** In Post-Quantum Cryptography. 2009.

BRASIL, 2014, Lei n. 12.965, de 23 de abr. de 2014. **Lei Brasileira que estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil**. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 10 set 2020.

BRASIL, **Lei n. 13.709, de 14 de ago. de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 10 set. 2020.

BRASIL, **Senado Federal**. Disponível em: <https://www12.senado.leg.br/assessoria-deimprensa/notas/nota-de-esclarecimento-vigencia-da-lgpd>. Acesso em 7 set. 2020.

BRASIL. **Código de Defesa do Consumidor, Lei 8.078 de 1990**. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm>. Acesso em: 10 jun. 2020.

BRASIL. **Constituição Federal de 1988.** Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 7 set. 2020.

BRASIL. **Direitos Autorais, Lei 9.610 de 1998.** Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9610.htm>. Acesso em: 10 de jun. 2020.

BRASIL. **Direitos das Organizações Criminosas, Lei 12.850 de 2013**. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/lei/l12850.htm>. Acesso em: 11 de jun. 2020.

BRASIL. **Direitos de Acesso à Informação, Lei 12.527 de 2011**. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em: 11 de jun. 2020.

BRASIL. **Direitos dos Softwares, Lei 9.609 de 1998.** Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9609.htm>. Acesso em: 10 de jun. 2020.

BRASIL. **Lei da Propriedade Industrial, Lei 9.727 de 1996**. Disponível em: <http://www.planalto.gov.br/ccivil_03/Leis/L9279.htm>. Acesso em: 13 de jun. 2020.

BRASIL. **Lei Geral de Proteção de Dados, Lei 13.709 de 2018**. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 11 de jun. 2020.

BRASIL. **Marco Civil da Internet, Lei 12.965 de 2014.** Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm>. Acesso em: 11 de jun. 2020.

BRASIL, 2018, Lei n. 13.709, de 14 de ago. de 2018. **Lei Geral de Proteção de Dados Pessoais** (LGPD). Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 10 set 2020.

BRASIL. Senado Federal. **Lei de Acesso à Informação no Brasil:** O que você precisa saber. Disponível em: <https://www12.senado.leg.br/transparencia/arquivos/sobre/cartilha-lai/>. Acesso em: 10 jun. 2020.

BRITO, Felipe Timbó. **Uma abordagem distribuída para preservação de privacidade na publicação de dados de trajetória**. 2016.

BURGER, John [et al.] - **Discriminating gender on Twitter**. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP '11), 2011.

CANCELIER, Mikhail Vieira de Lorenzi. **O Direito à Privacidade hoje: perspectiva histórica e o cenário brasileiro.** Disponível em: <http://www.scielo.br/pdf/seq/n76/2177-7055-seq-7600213.pdf>. Acesso em: 20 jul. 2020.

CANOTILHO, J. J. Gomes et al. (Coord.). **Comentários à Constituição do Brasil.** São Paulo: Saraiva, 2013.

CAO, J.; KARRAS, P. **Publishing microdata with a robust privacy guarantee**. Proc. VLDB Endow., v. 5, n. 11, p. 1388–1399, 2012.

CAO, N, C. Wang, M. Li, K. Ren and W. Lou. **Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data**, in IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 1, 2014.

CASTRO, Catarina Sarmento e. **Direito da Informática, Privacidade e Dados Pessoais**. Coimbra: Almedina, 2005.

COELHO, A. C. B. A. **Lei Geral de Proteção de Dados Pessoais Brasileira como meio de efetivação dos direitos da personalidade.** João Pessoa: [s.n.], 2019.

CORREIOS, B. **Estrutura CEP. [S.l.: s.n.],** 2018. https://www.correios.com.br/enviar-e-receber/ferramentas/cep/estrutura-do-cep. Acesso em: 13 ago. 2020.

COSTA, G. M. N. P. **Notificações de violações de dados: a mudança de paradigma com o regulamento geral de proteção de dados,** Master's thesis, Instituto Superior Técnico, 2018.

D. Boneh, R. Lipton, **A Revocable Backup System**, Proceedings 6th USENIX Security Conference, pp. 91-96, 1996.

D. J. Crandall, L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, and J. Kleinberg, **Inferring social ties from geographic coincidences**, Proceedings of the National Academy of Sciences, vol. 107, no. 52, pp.22 436–22 441, 2010.

DANIEL J. Bernstein. **Introduction to post-quantum cryptography**. In Post-Quantum Cryptography. 2009.

DE CAPITANI DI VIMERCATI, S. et al. Data privacy: **Definitions and techniques. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems**, v. 20, n. 06, p. 793-817, 2012. Disponível em: http://spdp.di.unimi.it/papers/ijufks2012.pdf. Acesso em: 17 ago. 2020.

DE MONTJOYE [et al.] -Unique in the Crowd: **The privacy bounds of human mobility**. In. Sci Rep, 2016. Acesso em jun. 2020.

DEITRICK, William [et al.] - **Gender Identification on Twitter Using the Modified Balanced** Winnow, In: Communications and Network, 2012. Acesso em jul.2020.

DONEDA, Danilo. **Considerações iniciais sobre os bancos de dados informatizados e o direito à privacidade.** Disponível em: <http://www.edtig.ipbeja.pt/Consideracoes.pdf>. Acesso em: 22 jul. 2020.

DONEDA, **Danilo. Da privacidade à proteção de dados pessoais.** Rio de Janeiro: Renovar, 2006.

E. C. V. Borges Júnior, **Introdução a sistemas criptográficos e o uso de geradores de sequências de números aleatórios e pseudoaleatóorios**, Brasília: Repositório UNB, 2014.

E.S.B.N.ONE-TIME, **Proposta de sistema eficiente e seguro deencriptação sequencial baseado no one-time pad**, Ph.D. dissertation, Instituto Militar de Engenharia, 2009.

ELMONGUI, Hicham G.; MORSY, Hader; MANSOUR, Riham - **Inference models for Twitter user's home location prediction**. IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA), 2015.

ENISA. **Pseudonymisation techniques and best practices em novembro de 2019**. Disponível em: https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices/at_download/fullReport. Acesso em: 10 ago. 2020.

EU (2014). **Opinion 05/2014 on Anonymisation Techniques**. Article 29 Data Protection Working Party (European Commission), em 10 de Abril de 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files /2014/wp216_en.pdf. Acesso em: 10 ago. 2020.

EU (2016). **Regulamento (UE) 2016/679**. Jornal Oficial da União Europeia. Legislação L119. 59 ano. 4 de maio. [Em linha]. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=OJ:L:2016:119:FULL&from=pt>. Acesso em: 10 set. 2020.

EU (2016). **Regulamento Geral sobre a Proteção de Dados** (RGPD). Regulamento (UE) 2016/679 do Parlamento Europeu e do conselho de 27 de abril de 2016, que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). Disponível em: https://www.privacy-regulation.eu/pt/4.htm. Acesso em: 19 ago. 2020.

EUA,2018, **California Consumer Privacy Act** (CCPA) Fact Sheet, State of California - Department of Justice - Office of the Attorney General. Disponível em: https://www.oag.ca.gov/system/files/attachments/press_releases/CCPA%20Fact%20Sheet%20%2800000002%29.pdf. Acesso em: 10 set 2020.

FABER, S. Jarecki, H. Krawczyk, Q. Nguyen, M. Rosu, and M. Steiner. **Rich queries on encrypted data:** Beyond exact matches. In Computer Security - ESORICS 2015 - 20th European Symposium on Research in Computer Security, 2015.

FERREIRA, B., J. Rodrigues, J. Leitão, and H. Domingos. **Privacy-preserving content-based image retrieval in the cloud**. CoRR, abs/1411.4862, 2014.

F. Hao, D. Clarke, and A. F. Zorzo, **Deleting secret data with publicverifiability**, IEEE Transactions on Dependable and Secure Computing,vol. 13, no. 6, pp. 617–629, 2015.

FUNG, B. C., Wang, K., Fu, A. W.-C., and Yu, P. S. **Introduction to Privacy Preserving Data Publishing: Concepts and Techniques**. Chapman & Hall/CRC, 1st edition. ISBN 978-1-4200-9148-9, 2010.

GAMBS, Sébastien; KILLIJIAN, Marc-Olivier; NÚÑEZ P. C., Miguel - De-anonymization attack on geolocated data. In: **Journal of Computer and System Science**. Elsevier, 2014.

GENTRY, C. **Fully homomorphic encryption using ideal lattices**. In: Proceedings of the 41st annual ACM Symposium on Theory of computing, pp. 169–178, 2009.

GERSTING, Judith L. **Fundamentos matemáticos para a ciência da computação**. LTC, 2001.

GLOBAL INTERNET LIBERTY CAMPAIGN. PRIVACY AND HUMAN RIGHTS: **AnInternational Survey of Privacy Laws and Practice.** Disponível em: <http://gilc.org/privacy/survey/intro.html>. Acesso em: 20 jul. 2020.

GRECO FILHO, Rogério. **Interceptação telefônica:** considerações sobre a Lei n.º 9.296, de 24 de julho de 1996. 2ª ed. São Paulo: Saraiva, 2009.

GYMREK M [et al.] - **Identifying personal genomes by surname inference**. In: Science,Vol. 339, Issue 6117, pp.321-324 DOI: 10.1126/science.1229566, 2013.

HAO, F.; Clarke, D.; Zorzo, A. F. **Deleting Secret Data with Public Verifiability**. IEEE Transactions on Dependable and Secure Computing, vol. 13 (6), 2016, pp. 617–29, 2016.

HOMER, Nils [et al.] - **Resolving individuals contributing trace amounts of DNA to highly complex mixtures using highdensity SNP genotyping microarrays**. In: PLoS Genet. 2008.

ISO - **International Organization for Standardization**, Disponível em: https://www .iso.org/home.html. Acesso em jul. 2020.

JOEL, R.; Srdjan, C., Basin, D. Data Node Encrypted File System: **Efficient Secure Deletion for Flash Memory**. In: Proceedings of the 21st USENIX Conference on Security Symposium, USENIX Association, 2012, pp. 17–17.

JUNIOR, Eliseu Castelo Branco, **Uma Estrategia Para Assegurar a Confidencialidade De Dados Armazenados Em Nuvem**. Disponível em: < http://repositorio.ufc.br/bitstream/riufc/23917/1/2017_tese_ecastelobrancojunior.pdf >. Acesso em: 05 set. 2020.

JÚNIOR, Marcos Ehrhardt. **Privacidade e proteção de dados pessoais durante a pandemia da COVID-19.** Disponível em: https://direitocivilbrasileiro.jusbrasil.com.br/artigos/824478175/privacidade-e-protecao-de-dados-pessoais-durante-a-pandemia-da-covid-19 . Acesso em: 5 mai. 2020.

LACOMBE, Francisco José Masset et al. **Administração – princípios e tendências.** São Paulo: Saraiva, 2003.

Lee, J.; Heo, J.; Cho, Y.; Hong, J.; Shin, S. **Secure deletion for NAND flash file system**. In Proceedings of ACM Symposium on Applied Computing, 2008, pp. 1710-1714.

LEMOS, R. G. L. **Marco Civil da Internet.** São Paulo: Altas, 2014.

LEMOS, Ronaldo. **O marco civil como símbolo do desejo por inovação no Brasil**. In: LEITE, George Salomão; LEMOS, Ronaldo (coord.). Marco Civil da Internet. São Paulo: Atlas, 2014.

Li, N., Li, T., and Venkatasubramanian, S. (2007). **t-closeness: Privacy beyond k-anonymity and l-diversity**. In 23th ICDE International Conference on Data Engineering (ICDE), pages 106–115.

LUZ, Pedro Henrique Machado da; LOUREIRO, Maria Fernanda Battaglin. **Privacidade e proteção de dados pessoais: os novos desafios na sociedade em rede**. Disponível em: <http://www.fumec.br/revistas/meritum/article/view/5811>. Acesso em: 8 jul. 2020.

MACAU, 2019, Governo da Região Administrativa Especial de Macau, Gabinete para a Protecção de Dados Pessoais, **Guia para Técnicas Básicas de Anonimização de Dados**.

MACHADO, Ulysses. **Direitos ao esquecimento, à privacidade e à informação:** como eles se relacionam? Disponível em: <serpro.gov.br/menu/noticias/noticias2020/direitoesquecimento-privacidade-lgpd>. Acesso em: 7 jul. 2020.

MACHANAVAJJHALA, A., Kifer, D., Gehrke, J., and Venkitasubramaniam, M. (2007). **l-diversity: Privacy beyond k-anonymity**. In ACM Transactions on Knowledge Discovery from Data (TKDD), 2007.

MARKOV, Ilia; BAPTISTA, Jorge; PICHARDO-LAGUNAS, Obdulia - **Authorship Attribution in Portuguese Using Character N-grams**. In: Acta Polytechnica Hungarica, 2017.

MOHAMMED, Noman & Fung, Benjamin & Hung, Patrick & Lee, Cheuk-kwong. (2009). **Anonymizing healthcare data: A case study on the blood transfusion service**. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 1285-1294. 10.1145/1557019.1557157, 2009.

MOLINARO, Carlos Alberto; SARLET, Ingo Wolfgang. **Breves notas acerca das relações entre a sociedade em rede, a internet e o assim chamado estado de vigilância**. In. LEITE, G. L., 2017.

N. Cao, C. Wang, M. Li, K. Ren and W. Lou, **Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data**, in IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 1, pp. 222-233, Jan. 2014, doi: 10.1109/TPDS.2013.45.

NARAYANAN, Arvind [et al.] - **On the feasibility of internet-scale author identification**. In: Proceedings of 33rd IEEE Symposium on Security and Privacy, 2012.

NARAYANAN, Arvind; SHMATIKOV, Vitaly - **Robust De-anonymization of Large Datasets**. In: SP '08 Proceedings of the 2008 IEEE Symposium on Security and Privacy, 2008.

NERGIZ, M. E., Atzori, M., and Clifton, C. (2007). **Hiding the presence of individuals from shared databases**. In Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data, SIGMOD ’07, pages 665–676, New York, NY, USA. ACM, 2007.

NERGIZ, M. E.; GÖK, M. Z. **Hybrid k-Anonymity. Computers & Security**, v. 44, p. 51-63, 2014.

OPINION 05/2014 on Anonymisation Techniques”. **Article 29 Data Protection Working Party (European Commission)**, em 10 de Abril de 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf. Acesso em: 10 ago. 2020.

OTTO Seppälä, Lauri Malmi & Ari Korhonen (2006) **Observations on student misconceptions—A case study of the Build – Heap Algorithm**, Computer Science Education, 16:3, 241-255, DOI: 10.1080/08993400600913523.

P. D. P. C. **Singapore, Guide To Basic Data Anonymization Techniques**, 2018. Disponível em: https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Other-Guides/ Guide-to-Anonymisation_v1-(250118).pdf. Acesso em: 10 ago. 2020.

P. Golle, **Revisiting the uniqueness of simple demographics in the us population**, in Proceedings of the 5th ACM Workshop on Privacy in Electronic Society, 2006, pp. 77–80.

PCI, **Padrão de Segurança de Dados**, Disponível em: https://pt.pcisecuritystandards .org/_onelink_/pcisecurity/en2pt/minisite/en/docs/PCI_DSS_v3.pdf; acesso em: 10 set 2020.

PACETE, Gustavo Luiz. **Multa aplicada ao Google é emblemática para a GDPR.** Disponível em: <https://www.meioemensagem.com.br/home/midia/2019/01/22/multa-aplicada-ao-google-e-divisor-de-aguas-para-a-gdpr.html>. Acesso em: 10 set. 2020.

PETERSON. Z.N.J.; Burns. R.; Herring. J.; Stubblefield. A.; Rubin. A.D**. Secure Deletion for a Versioning File System**. In: Proceedings of the 4th Conference on USENIX Conference on File and Storage Technologies (FAST), 2005, vol. 4, pp. 143-154.

PHILIP R. Geffe. **How to protect data with ciphers that are really hard to break**. Electronics, pages 99–101, 1973.

PHILLIP J. Chase. 1970. **Algorithm 383: permutations of a set with repetitions [G6]. Commun**. ACM 13, 6 (June 1970), 368–369. DOI:https://doi.org/10.1145/362384.362503.

QUEIROZ, M. J.; LINO, N. C.; MOTTA, G. **Uma ontologia de domínio para preservação de privacidade em dados publicados pelo governo brasileiro**. XII Simpósio Brasileiro de Sistemas de Informação. Florianópolis, SC, Brasil, 2016.

R L Rivest, L Adleman, and M L Dertouzos. **On data banks and privacy homomorphisms**. Foundations of Secure Computation. Academic Press, 1978.

RIVEST, R. L.; SHAMIR, A.; ADLEMAN, L. **A method for obtaining digital signatures and public-key cryptosystems**. CACM, l21, p. 120–126, 1978.

ROBERT Sedgewick. 1977. **Permutation Generation Methods**. ACM Comput. Surv. 9, 2 (June 1977), 137–164. DOI:https://doi.org/10.1145/356689.356692

S. Faber, S. Jarecki, H. Krawczyk, Q. Nguyen, M. Rosu, and M. Steiner. **Rich queries on encrypted data: Beyond exact matches**. In Computer Security - ESORICS 2015 - 20th European Symposium on Research in Computer Security, pages 123–145, 2015.

Sanjam Garg, Shafi Goldwasser, Prashant Nalini Vasudevan. **"Formalizando a Exclusão de Dados no Contexto do Direito de Ser Esquecido."** EUROCRYPT (2) 2020: 373-402.

SAMARATI, P. **Protecting respondents identities in microdata release**. Knowledge and Data Engineering, IEEE Transactions on, v. 13, n. 6, p. 1010–1027, 2001.

SCHREIBER, Anderson. **Manual de direito civil**: 3. ed. São Paulo: Saraiva Educação, 2020.

SEQUEIRA, C. S. A. “Identidade Digital – o **Espectro desde a Anonimização à Identificação**” Master’s thesis, Instituto Superior Técnico, 2019.

SILVA, Eugénio Alves da (2016). **Practical use of Partially Homomorphic Cryptography**. Monografia do curso de Mestrado Bolonha em Segurança de Informação e Direito no Ciberespaço, Universidade de Lisboa em parceria com o Instituto Superior Técnico e Escola Naval, Lisboa - PT, 2016, 67p.

SIMON, Pedro. **A impunidade veste colarinho branco. Brasília:** Senado Federal, 2010.

SOMBRA, Thiago Luis dos Santos. **Fundamentos da regulação da privacidade e proteção de dados pessoais:** pluralismo jurídico e transparência em perspectiva. São Paulo: Thomson Reuters Brasil, 2019.

STALLING, Willian, **Criptografia e Segurança de Redes: Princípios e Práticas** 4. Ed. Prentice Hall, Brasil, p. 17-36, 2007.

SUSAN, V.S.; Christopher, T. Anatomisation with slicing: **A new privacy preservation approach for multiple sensitive attributes**. Springer Plus 2016, 5, 964.

TAVARES, Letícia Antunes; ALVAREZ, Bruna Acosta. **Da proteção dos dados pessoais:** uma análise comparada dos modelos de regulação da Europa, dos Estados Unidos da América e do Brasil. Disponível em: <http://www.tjsp.jus.br/download/EPM/Publicacoes/ObrasJuridicas/ii%204.pdf?d=636680444556135606>. Acesso em: 7 set. 2020.

VIDOR, Daniel MARTINS, **LGPD:** origem e implicações. Disponível em:<http://mercurylbc.com/lgpd-origem-e-implicacoes/>. Acesso em: 8 set. 2020.

VIMERCATI, Sabrina de Capitani. FORESTI, Sara; LIVRAGA, Giovanni.; SAMARATI, Pierangela. **Data Privacy: Definicitions and Techiniques**. International Journal of Uncertainty, **Fuzziness and Knowledge-Based Systems** Vol. 20, No. 6 (2012) 793–817 World Scientific Publishing Company, 2012.

Xiao X, Tao Y (2006) **Anatomy: simple and effective privacy preservation**. In: Proceedings of international conference on very large data bases (VLDB), 1 September 2006.

YAN, Jinny; MATTHEWS, Suzanne - **Applying clustering algorithms to determine authorship of chinese twitter messages**, In: IEEE MIT Undergraduate Research Technology Conference (URTC), 2016.

ZHANG, X.; LIU, Q.; LIU, D.; XIE, W. **A survey on anonymity for privacy preserving data mining**. In: CRC PRESS. Information Science and Electronic Engineering: Proceedings of the 3rd International Conference of Electronic Engineering and Information Science (ICEEIS 2016), 4-5 January, 2016, Harbin, China. [S.l.], 2016. p. 343–346.