

CORRELAÇÕES ENTRE OS DADOS LIGADOS E A LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS: CONSIDERAÇÕES SOBRE CENÁRIOS DE UTILIZAÇÃO

CORRELATIONS BETWEEN LINKED DATA AND THE GENERAL LAW FOR PERSONAL DATA PROTECTION: CONSIDERATIONS ON USAGE SCENARIOS

Herbert Alcântara Ferreira^a
Moisés Lima Dutra^b
Rodrigo de Sales^c

RESUMO

Objetivo: Este trabalho tem como objetivo geral analisar as correlações entre a Lei Geral de Proteção de Dados Pessoais (LGPD) e cenários de utilização de dados ligados.

Metodologia: Trata-se de um estudo exploratório, em que se utiliza de referências bibliográficas, documentais, além de normas jurídicas. Em um primeiro momento, fez-se uma revisão de literatura composta de três subseções: as duas primeiras analisaram o surgimento, o funcionamento e exemplos de uso dos dados ligados, enquanto a última subseção pesquisou o surgimento e a estrutura da LGPD. Posteriormente, foram analisadas as principais regras da referida lei aplicáveis a alguns cenários de uso de dados ligados, sob a perspectiva das fases do Ciclo de Vida de Dados: coleta, armazenamento, recuperação e descarte. **Resultados:** Percebeu-se que a complexidade desses sistemas é um desafio tecnológico diante da necessidade de se respeitar a lei e conservar a privacidade dos indivíduos. Nesse sentido, o conhecimento dessa nova norma e sua adaptação às organizações também são dilemas a serem superados. **Conclusões:** Logo, entendeu-se que a aplicabilidade da LGPD a cenários de dados ligados é variada, exigindo um trabalho de diversos atores profissionais e científicos para que empresas e instituições em geral possam se adequar às suas regras.

Descritores: Dados ligados. Direito à privacidade. Lei Geral de Proteção de Dados

^a Doutor em Ciência da Informação pela Universidade Federal de Santa Catarina (UFSC). Docente da Universidade Estadual de Montes Claros (Unimontes), Montes Claros, Brasil. E-mail: herbertalc@yahoo.com.br

^b Doutor em Ciência da Computação pela Universidade de Lyon 1. Docente do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina (UFSC), Florianópolis, Brasil. E-mail: moises.dutra@ufsc.br

^c Doutor em Ciência da Informação pela Universidade Estadual Paulista (UNESP). Docente do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina (UFSC), Florianópolis, Brasil. E-mail: rodrigo.sales@ufsc.br

Pessoais. Informação e Tecnologia.

1 INTRODUÇÃO

Nas mais repercutidas definições relativas à Ciência da Informação (CI), o imperativo tecnológico que viabiliza os fluxos informacionais está presente, bem como a interferência que tal imperativo pode causar à sociedade. Borko (1968), por exemplo, afirmou que a CI se tratava de uma disciplina que investigava as propriedades e o comportamento da informação, assim como as forças que governavam seu fluxo e os meios de processá-la para otimizar sua acessibilidade e uso. Wersig e Neveling (1975) chamaram a atenção para o fato de que a CI ganhou importância histórica na medida em que centrou seus esforços para os problemas informacionais que modificavam completamente a relevância da informação para a sociedade como um todo. Para Saracevic (1996), a CI sempre esteve inexoravelmente ligada à tecnologia da informação. Para o autor, o imperativo tecnológico impunha a transformação da sociedade moderna em sociedade da informação, ou sociedade pós-industrial. Nesse sentido, tratava-se (e ainda trata-se) de uma disciplina participante ativa e deliberada na evolução da sociedade da informação. O autor complementa afirmando que “a CI teve e tem um importante papel a desempenhar por sua forte dimensão social e humana, que ultrapassa a tecnologia” (SARACEVIC, 1996, p. 42).

Tocante à chamada sociedade da informação, Tarapanoff *et al.* (2000) afirmam que se trata de uma “mudança paradigmática da sociedade, que inicia um novo ciclo produtivo, centrado na informação e no conhecimento e que tem a informação como um bem econômico” (p. 93). Já Castells (1999) prefere entender que as tecnologias que constituem esta nova estrutura social possuem a própria informação como matéria-prima, e por esta razão adotou o termo sociedade informacional, em vez de sociedade da informação. Para o autor, o impacto das tecnologias de informação nas diversas atividades humanas e sociais faz da informação algo maior que um mero bem de consumo.

Nesse contexto de sociedade informacional, as questões atinentes à lida com os dados, ou com a comunicação e interligação de dados, ganham certo

protagonismo nos estudos de Ciência da Informação vinculados às perspectivas tecnológicas.

Com o desenvolvimento acelerado das tecnologias digitais, como a *Internet*, as infraestruturas de armazenamento de dados e os dispositivos digitais, o uso da informação tem se tornado cada vez mais recorrente em diversos setores da sociedade e da economia, o que implica na constante troca de informações entre máquinas, indivíduos e instituições. Nesse contexto, surgem as práticas de *Linked Data* (dados ligados), um paradigma de interconexão de dados, cuja inspiração originária deu-se no âmbito da proposta da *Web Semântica*. Atualmente, diversos sistemas de informação e fenômenos digitais utilizam *Linked Data* para facilitar o acesso a diferentes tipos de informações e na construção de conhecimento através de inferências de máquinas. *Linked Open Data* (dados abertos ligados), *Big Data* (ambiente de conjuntos de dados complexos e de difícil processamento pelas ferramentas tradicionais) e *Internet of Things* (*Internet* das Coisas) são exemplos atuais de fenômenos digitais potencialmente beneficiários da utilização de dados ligados.

Os dados ligados foram criados com a intenção de serem integráveis entre si, para que possam construir significado (por isso, diz-se que são estruturados semanticamente). Essa formação de sentido foi um avanço nas atividades de manuseio de dados, pois permite que os computadores façam inferências diante daqueles dados existentes. Para alcançar esse objetivo no ambiente da *Web*, utilizam-se modelos e linguagens computacionais específicos, tais como RDF e SPARQL. A saber, RDF é uma “arquitetura genérica de metadados que permite descrever recursos no contexto da *Web* através de metadados, visando o processamento por máquina” (LIMA; CARVALHO, 2005, p. 25). Já o SPARQL é uma linguagem de consulta que permite a recuperação de informações presentes em grafos RDF (VIDOTTI *et al.*, 2019). Devido a essa arquitetura computacional, sistemas de *Linked Data* têm sido usados por empresas, entidades públicas e instituições em geral que trabalham com semântica de dados em sistemas computacionais.

Dentro da proposta de *Linked Data*, também é possível trabalhar com dados abertos ligados (*Linked Open Data*), que possuem licença aberta. Ou seja,

aqueles dados que qualquer pessoa pode acessar e manusear. Um exemplo de seu uso é a possibilidade de um internauta encontrar diversas informações correlatas sobre um termo procurado em um *website* de pesquisa, o que é possível por causa da interconexão de dados provenientes de várias fontes da *Internet*. Outro exemplo são os grafos de conhecimento, utilizados por portais de conteúdo como a *Wikipédia* ou por empresas como a *Google*.

Contudo, há diversos problemas envolvendo os ambientes digitais, dos quais se destacam os riscos à violação de dados pessoais. Nesse sentido, eventuais ataques à integridade dos dados atingem “a garantia dos direitos da personalidade dos indivíduos, principalmente no que se reporta ao direito à privacidade” (DIAS; VIEIRA, 2013, p. 181). Em análise mais profunda, observam-se questões ainda mais específicas em relação à proteção de informações pessoais: a reutilização indiscriminada de algum dado coletado; o tratamento integrado de dados em sistemas que utilizam *Linked Data* sem que seus agentes sejam devidamente identificados (para o caso de responsabilização); bem como a utilização de dados compartilhados para fins obscuros ou ilícitos (RODRÍGUEZ-DONCEL *et al.*, 2016).

O compartilhamento indevido de dados de usuários do Facebook com a empresa Cambridge Analytica é um exemplo notório de ofensa a informações pessoais (CONFESSORE, 2018). Diante dos casos reais e, cada vez mais recorrentes, de violação de informações pessoais, vários países têm buscado regular e discutir a proteção de dados. Neste contexto, surge, no Brasil, a Lei Geral de Proteção de Dados Pessoais ou LGPD (BRASIL, 2018), sob forte inspiração das normas europeias sobre o tema. A lei brasileira, que entrou em vigor em 2020, passa a regulamentar a matéria, discriminando fundamentos, princípios, direitos e deveres, com a finalidade de proteger a privacidade e a liberdade dos titulares de dados.

O objetivo geral deste trabalho é analisar as correlações entre a LGPD e as aplicações de dados ligados. Neste estudo, são considerados os diversos cenários de utilização da *Linked Data*. Como apoio à consecução desse objetivo, recorre-se ao Ciclo de Vida de Dados (ou CVD), abordagem analítica que considera as atividades de tratamento de dados de acordo com as progressivas

etapas de sua “vida”, desde a coleta do dado, passando pelo seu armazenamento e recuperação, até a sua eventual exclusão (SANT’ANA, 2016). Entende-se que o CVD, por ser uma metodologia de abordagem aplicável a quaisquer dados, também pode ser adotada para análise de aplicação da LGPD através das etapas de tratamento de dados ligados.

Em relação aos objetivos específicos, objetiva-se: a) investigar a origem, o conceito, os usos e os exemplos de *Linked Data*; b) associar as principais regras da LGPD às fases do Ciclo de Vida de Dados; c) avaliar a possibilidade de aplicação da LGPD a cenários de *Linked Data*.

Trata-se de uma pesquisa exploratória que, para a consecução de seus objetivos, utilizou a abordagem bibliográfica e documental, recorrendo-se a artigos, teses, livros e *websites* para compreender, tanto sobre dados ligados, quanto sobre a LGPD e o CVD. Ademais, recorreu-se a algumas normas, que serviram de análise para o escopo jurídico abordado neste trabalho, em especial, à própria LGPD (Lei nº 13.709/2018). Diante da revisão de literatura, pautada em duas fases – *Linked Data* (origem, conceito, usos, exemplos e sua análise sob a perspectiva das fases no Ciclo de Vida dos Dados) e a LGPD (motivações, inspirações, surgimento e principais apontamentos) – este trabalho procedeu com a avaliação geral das principais hipóteses de aplicação da norma em sistemas que utilizam dados ligados.

Na investigação das principais regras previstas pela LGPD que possam ser aplicadas a cenários de dados ligados, optou-se por analisá-las sob a ótica do CVD, que compreende as fases de coleta, armazenamento, recuperação e descarte. É importante destacar que nem todos os cenários possíveis de utilização de dados ligados foram analisados por diversas razões.

Primeiramente, muitas questões previstas na lei necessitam de melhor regulamentação por parte da Autoridade Nacional de Proteção de Dados Pessoais (ANPD), entidade pública responsável pela aplicação da LGPD no Brasil. Dentre essas questões, cita-se: confronto de normas de proteção de dados entre diferentes países (possível de ocorrer entre datasets e sistemas internacionais), diversas regras específicas à Administração Pública, medidas de segurança computacional em aplicações de dados ligados e boas práticas de

governança informacional. Em relação a essas questões, ainda não se pode ter uma noção clara de como cenários de utilização de *Linked Data* podem se adequar a elas.

Ademais, a própria extensão do presente trabalho, bem como a amplitude de aplicação da lei (que ainda não é amplamente elucidada e/ou regulamentada pela entidade competente para tanto, como já dito), impedem a análise de todo o universo de aplicações da LGPD à *Linked Data*.

Assim, para viabilização deste estudo, optou-se por apoiar-se em um escopo jurídico e um conjunto de aplicações tecnológicas suficientes para analisar as repercussões da aludida lei a esses fenômenos.

2 LINKED DATA

Os dados ligados são uma especialização da *Web Semântica* aplicada a cenários mais específicos e, conseqüentemente, mais factíveis de se implementar e se gerenciar. Desde seu início, a *Web Semântica*, possui por objetivo “facilitar a interpretação e integração de dados na *Web*” (Cunha; Souza; Lóscio, 2011, p. 80). O *World Wide Web Consortium (W3C)*, organização internacional sem fins lucrativos que desenvolve, mantém e regula os padrões de dados na *Web*, disponibilizou diversas ferramentas para a *Web Semântica*, tais como RDF para a modelagem de dados e SPARQL para a sua recuperação (VIDOTTI *et al.*, 2019). A ideia por trás da *Web Semântica* é de que computadores possam desenvolver capacidade de extrair inferências a partir da leitura ambiental de informações, a fim de que as máquinas cooperem mais com os seus usuários (BERNERS-LEE; HENDLER; LASSILA, 2002). Berners-Lee (2006) apresenta os princípios que caracterizam os dados ligados. Apenas quando os quatro princípios, em conjunto, são observados é que podemos dizer que se está diante de um cenário de dados ligados, a saber:

- I. Utilizar URIs (Uniform Resource Identifiers) para nomear as coisas (os recursos).
- II. Utilizar o protocolo HTTP para que as pessoas possam acessar esses recursos apontados pelas URIs.
- III. Fornecer informações úteis através de padrões, como RDF e

SPARQL, quando alguém acessar um recurso via URI.

IV. Incluir *links* para outras URIs, de maneira que mais coisas se conectem e mais informações possam ser descobertas.

Essa interconexão de dados pode existir em ambientes digitais fechados (acesso restrito) ou abertos (acesso livre). Ela, na verdade, é uma estrutura de *links* referentes a dados existentes em diversas fontes, ou *datasets*. Os *datasets* são conjuntos lógicos de dados, modelos que indicam como os dados devem ser armazenados e consumidos. Nas aplicações de dados ligados, esses conjuntos são armazenados em repositórios físicos chamados *triplestores* (repositórios de triplas). Os *links* criados entre dados ligados (na verdade, triplas RDF) podem tanto se referir à *triplestore* local, quanto a outras remotas, acessíveis via *Internet*. Por meio desta estrutura de *links*, é possível se identificar a relação semântica entre os conceitos extraídos desses dados. Esse intercâmbio de fontes garante a interoperabilidade entre os *datasets*. Ou seja, constrói-se um espaço de compartilhamento constante de informações. O uso de vocabulários também é fundamental nessa estrutura, pois permite que se trabalhe com conceitos, relações e atributos pré-definidos referentes a cada recurso (NHACUONGUE; ROSZA; DUTRA, 2018). Um recurso é uma representação digital em forma de URI de algo concreto ou abstrato do mundo não-digital (documento, objeto, local, pessoa, organização, conceito, ideia, etc.).

Linked Open Data (LOD) representa um cenário de *Linked Data* em que todos os *datasets* possuem licenças abertas. Sua ideia final é possibilitar a criação de uma rede de informações online e acessível a qualquer usuário da *Web*. *LOD* trabalha com *links* que conectam recursos da *Web*, tais como artigos científicos localizados dentro de um mesmo repositório de trabalhos acadêmicos ou em repositórios distintos. Como se tratam de dados ligados, esses recursos também estão interconectados e têm a construção de significado como uma de suas finalidades (SANTARÉM SEGUNDO, 2015). Nessa iniciativa, o acesso a uma *triplestore* permite acessar outras *triplestores* relacionadas por meio dos *links* existentes entre os recursos. Em geral, essas ligações podem ser de relacionamento (apresentam informações relacionadas a um mesmo recurso, por exemplo, informações relacionadas sobre uma mesma pessoa); ligações de

identidade (indicam informações distintas originadas de várias fontes sobre um mesmo recurso, por exemplo, um *dataset* fornece informações geográficas sobre uma instituição e outro *dataset* fornece informações administrativas); ou ligações de vocabulário (apontam os mesmos conceitos em vocabulários distintos) (HEATH; BIZER, 2011; NHACUONGUE; ROSZA; DUTRA, 2018).

Com o passar do tempo, o fenômeno dos dados ligados se intercalou com outras tecnologias. Na atual geração digital, as aplicações de *Big Data* – grandes volumes de dados variados e em fluxo veloz – vêm disponibilizando cada vez mais funcionalidades aos seus usuários, visto que o seu universo de dados auxilia em diversas atividades contemporâneas. Além disso, permite que profissionais da tecnologia da informação manuseiem dados não-estruturados a uma grande velocidade. Por fim, destaca-se a riqueza informacional disposta por eles, cuja utilidade se aplica, por exemplo, a grandes bancos de dados nos quais a interconexão de informações seja necessária.

Também é notável o alcance da *Internet* das Coisas ou *Internet of Things* (*IoT*), que vem ganhando progressivo espaço no cotidiano das pessoas nos últimos anos. Ela é uma expressão que se refere a tecnologias nas quais dispositivos eletrônicos se conectam entre si, como forma de integrar a experiência digital do usuário. Por consequência, acabam por utilizar as informações cadastradas sobre os indivíduos, além de chegar a inferências a partir das experiências do usuário com dispositivos integrados entre si. Ou seja, são potenciais produtores e utilizadores de dados ligados, com o objetivo de aumentar sua curva de aprendizado e desenvolver melhor suas atividades, utilizando-se de armazenamento tanto local quanto em nuvem. Fazem parte desse contexto os dispositivos eletrônicos de computação pervasiva ou vestível: casas “inteligentes”, roupas que “aprendem” sobre os hábitos de seus usuários, relógios que compartilham informações com celulares e computadores, além de outras possibilidades (MAGRANI, 2018), todos geradores de dados com informação semântica associada.

Além disso, há sistemas que coletam, tratam e compartilham dados pessoais, tanto em empresas (a exemplo da prática de *e-commerce*, que salva informações de seus consumidores em seus *websites* por meio de *cookies* de

navegador), quanto em órgãos públicos (a exemplo do Portal Gov.br, do governo brasileiro, que integra informações pessoais de diversos órgãos federais). Sem falar nas *triplestores* do tipo *LOD*, que armazenam *datasets* para finalidades diversas, especialmente científicas e acadêmicas (ROZSA; DUTRA; NHACUONGUE, 2017).

3 O CICLO DE VIDA DOS DADOS

O comportamento e o tratamento de dados ligados podem ser analisados por meio da metodologia do Ciclo de Vida de Dados (CVD), como se adota neste estudo, abordagem da “vida” de um dado, compreendendo etapas progressivas (SANT’ANA, 2016). Essa abordagem foi escolhida por também se aplicar aos dados ligados, que são coletados, armazenados e disponibilizados em *triplestores* e outras aplicações similares.

A primeira fase é a coleta, na qual o dado é integrado ao *dataset* da instituição ou da pessoa que irá utilizá-lo. Em seguida, se prosseguirá com a fase de armazenamento desse dado para sua posterior recuperação (por onde se alcançará a finalidade de uso dessa informação). Por fim, haverá o descarte do dado - ainda que, nem sempre, a exclusão do dado ocorra, conforme preconiza a LGPD.

4 LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS

O fenômeno dos dados ligados, seja em ambientes abertos ou fechados, compreende a interoperabilidade de *datasets*. Essa interconexão de dados também possui implicação na coleta, no armazenamento e na utilização de informações pessoais presentes nesses sistemas. Quando se trata de dados abertos ligados, também se deve considerar o acesso compartilhado entre diversos indivíduos ou instituições, o que agrava ainda mais a exposição de informações sensíveis. Essas tecnologias podem servir de instrumento para que uma informação íntima sobre alguém seja vazada em um portal de dados abertos, por exemplo. Também é possível que empresas parceiras ou de um mesmo grupo empresarial compartilhem entre si os dados de seus respectivos

clientes, sem a devida autorização dos consumidores. Essas situações possíveis para as tecnologias de dados ligados violam a privacidade do indivíduo e o seu direito de ter controle sobre seus dados. Fato é que o uso de dados ligados já é uma realidade na prestação de serviços públicos e nas atividades empresariais. Entretanto, a sua utilização não pode atingir negativamente a vida do indivíduo, de modo que sua privacidade e autonomia sejam violadas.

Vários países, nos últimos anos, têm elaborado leis que visam garantir o direito à privacidade e à liberdade, adaptadas ao cenário mundial atual de produção e consumo de dados. Até 2018, mais de cem países já possuíam alguma legislação sobre dados pessoais em vigor, como Chile, Argentina, Estados Unidos, Japão e China (CONSUMERS INTERNATIONAL, 2018). Assim, o Brasil optou por também desenvolver sua legislação específica, a exemplo de outros países. No início da década de 2010, já havia algumas leis que citavam a proteção de dados pessoais, por exemplo: o Código de Defesa do Consumidor (Lei nº 8.078, de 1990), a Lei de Acesso à Informação (Lei nº 12.527, de 2011) e o Marco Civil da *Internet* (Lei nº 12.965, de 2014). Ainda assim, essas normas não foram suficientes para resolver as diversas problemáticas relativas ao tema. Um projeto específico sobre o assunto foi apresentado em 2012, sob a identificação de Projeto de Lei (PL) nº 4.060/2012, que, após anos de discussões, gerou a Lei nº 13.709/2018, conhecida como Lei Geral de Proteção de Dados Pessoais (LGPD). Verifica-se que a maior inspiração para a LGPD foi o Regulamento Geral de Proteção de Dados (*GDPR*, na sigla original) (UNIÃO EUROPEIA, 2016) aprovado em 2016, mas que só entrou em vigor em 2018 (IRAMINA, 2020). É curioso observar que muitos elementos da GDPR são adotados pela LGPD, por exemplo: a definição de dados pessoais, os princípios acerca da temática, previsão de sanções e de autoridades nacionais de proteção de dados (IRAMINA, 2020).

4.1 PRINCIPAIS ASPECTOS DA LGPD

A LGPD é uma lei federal que tem como finalidade definir regras gerais, direitos e deveres visando a proteção de dados pessoais relativos a pessoas naturais (pessoas físicas). Dentro da noção de dados pessoais, pode-se citar:

nome, Cadastro de Pessoa Física, número de identidade, endereço, contato, dados de usuário em aparelho digital, informações de saúde, dentre múltiplas possibilidades. De acordo com o artigo 1º da lei, ela regula o tratamento de dados pessoais (de pessoa natural) realizado tanto por outras pessoas naturais, quanto por pessoas jurídicas de direito privado (empresas e instituições privadas) e de direito público (órgãos e entidades públicas). É uma norma de interesse nacional e, por isso, deve ser obedecida em todos os entes da Federação (União, Estados, Distrito Federal e Municípios).

O artigo 2º define fundamentos sobre a disciplina de proteção de dados pessoais. Dentre esses princípios, pode-se destacar alguns: respeito à privacidade (inciso I); autodeterminação informativa, que é a soberania do cidadão sobre seus próprios dados (inciso II); intimidade, honra e imagem como direitos invioláveis (inciso III); direitos humanos (inciso VII). A lei também se destaca pelo seu didatismo, pois explica conceitos importantes para a sua própria aplicação. O artigo 5º, por exemplo, define dezenove termos. Dentre esses, é importante conhecer os seguintes termos, abordados no decorrer deste estudo:

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; [...]

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

X - tratamento: toda operação realizada com dados pessoais [...] (BRASIL, 2018, artigo 5º).

No artigo 6º, são definidos alguns princípios que devem orientar as atividades de tratamento de dados pessoais. Além do princípio da boa-fé, lista-se, em síntese:

Quadro 1 – Fundamentos da LGPD

Fundamento	Definição
Finalidade (inciso I)	Definir fins para o uso de dados
Adequação (inciso II)	Tratar os dados apenas para os fins informados pelo titular
Necessidade (inciso III)	Tratamento de dados apenas para o que é minimamente necessário
Livre acesso (inciso IV)	Direito do titular de consultar de forma facilitada, integral e gratuita os seus dados pessoais
Qualidade dos dados (inciso V)	As informações coletadas devem estar claras e atualizadas para a consecução dos fins a que se destinam
Transparência (inciso VI)	As informações estejam compreensíveis ao titular, quando ele quiser acessá-las
Segurança (inciso VII)	Adoção de medidas de segurança informacional
Prevenção (inciso VIII)	Adoção de medidas para prevenir qualquer ofensa à violação dos dados
Não-discriminação (inciso IX)	Proibição do manuseio de dados para fins discriminatórios ilícitos ou abusivos
Responsabilização e prestação de contas (inciso X)	O agente de tratamento de dados deve deixar claro que adota medidas de preservação da integridade dos dados pessoais do titular

Fonte: Lei Geral de Proteção de Dados Pessoais (BRASIL, 2018)

Em relação à aplicação territorial da lei, o artigo 3º (incisos I ao III) aponta que ela será aplicada quando cumprir algumas dessas situações, alternativamente: a) o tratamento dos dados tenha sido feito em território brasileiro; b) a finalidade do tratamento dos dados seja ofertar/fornecer bens ou serviços no território nacional; c) o titular dos dados esteja em território nacional; d) os dados tenham sido coletados no Brasil. Observa-se que a coleta de dados pessoais em território nacional ocorre quando seu titular encontra-se no Brasil

no momento da coleta. Depreende-se que a LGPD também pode ser aplicada no estrangeiro, caso cumpra algum desses requisitos. No Direito, esse fenômeno é conhecido como extraterritorialidade. Por exemplo, caso seu tratamento se enquadre em alguns dos requisitos do referido artigo da lei, “[...] o dado pessoal tratado por uma empresa de serviço de *cloud computing* que armazene o dado fora do país terá que cumprir as exigências da LGPD” (PINHEIRO, 2018, não paginado). O mesmo valeria para qualquer outro sistema de dados ligados. Por vezes, a questão da extraterritorialidade exige que o agente de tratamento conheça as normas de proteção de dados pessoais de todos os países onde ele operar. Isso exigirá um apoio jurídico para que ele entenda as complexas contradições que podem existir entre normas de diferentes países.

Por outro lado, o artigo 4º declara a não-aplicação da LGPD, caso se enquadre em algum desses requisitos: a) tratamento realizado por pessoa natural para fins exclusivamente particulares e não econômicos; b) para fins jornalísticos e artísticos; c) para fins de segurança pública; d) quando o uso compartilhado de dados não se dê no território nacional ou não esteja ligado a agentes de tratamento brasileiros (incisos I ao IV).

Também é essencial a figura da Autoridade Nacional de Proteção de Dados Pessoais (ANPD), regulada no artigo 55 da lei e que foi inicialmente vetada quando da sua sanção (Valente, 2018) e só voltou a integrar o ordenamento jurídico a partir da Lei nº 13.853/2019, que alterou a LGPD e definiu sua entrada em vigor. Atualmente, a ANPD é definida como uma entidade autônoma (goza de autonomia) e tem como funções precípuas: zelar pela aplicação da LGPD, elaborar orientações complementares à lei (no limite de seus poderes jurídicos), fiscalizar e aplicar sanções.

5 LGPD, CICLO DE VIDA DOS DADOS E CENÁRIOS DE DADOS LIGADOS

A expressão “uso compartilhado de dados” pode ser entendida como um gênero de utilização de dados em fluxo, do qual os dados ligados são um subtipo. Tal conceito indica ações em que dados pessoais são comunicados, difundidos, transferidos, interconectados ou submetidos a situações afins. (artigo 5º, inciso XVI). Portanto, a análise das situações de aplicação da LGPD a dados ligados é

efetuada neste trabalho de acordo com essa expressão. Ainda, importante notar que as normas gerais da lei também produzem efeitos nesse cenário tecnológico específico, semelhante ao que ocorre no contexto das normas europeias (RODRÍGUEZ-DONCEL *et al.*, 2016).

Para análise de como a LGPD se aplica, especialmente a cenários de dados ligados, recorre-se à abordagem do Ciclo de Vida de Dados (CVD), anteriormente apresentada, considerando suas quatro fases: coleta, armazenamento, recuperação e descarte. Como essa perspectiva analítica se aplica a dados em geral, logicamente também pode ser utilizada para abordar o comportamento e o tratamento de dados em ambientes de *Linked Data*.

5.1 FASE DA COLETA DE DADOS

Na primeira fase do CVD, a coleta é o momento em que o agente de tratamento “toma posse” de dados do titular. No sistema da LGPD, a coleta é geralmente realizada mediante consentimento da pessoa (artigo 7º, inciso I), que deve ser uma “manifestação livre, informada e inequívoca” (artigo 5º, XII), sendo “fornecido por escrito ou por outro meio que demonstre a manifestação de vontade do titular” (artigo 8º, caput). Nesse momento, é de responsabilidade do agente esclarecer à pessoa acerca dos fins a que se destina o tratamento, limitando-se esse aos objetivos propostos, conforme anuência do titular. Ainda, o titular goza da prerrogativa de exigir revogação do consentimento, “mediante manifestação expressa do titular, por procedimento gratuito e facilitado”, e requerer a eliminação dos tratamentos “realizados sob amparo do consentimento anteriormente manifestado” (artigo 8º, §5º).

Contudo, o tratamento de dados pessoais também pode ocorrer em outras hipóteses (justificativas para o tratamento) em que não se exige o consentimento formal do titular, de acordo com o artigo 7º:

- a. Quando necessário para que o controlador cumpra alguma obrigação legal (inciso II).
- b. Para execução de políticas públicas por parte da Administração Pública (inciso III).
- c. Na realização de estudos por órgão de pesquisa, utilizando sempre

- dados anonimizados, sempre que possível (inciso IV).
- d. Quando for necessário para cumprir contrato ou atos preliminares a ele, desde que os dados coletados sejam de pessoa que é parte no contrato e que haja pedido por parte dela para a coleta (inciso V).
 - e. Para o exercício regular (legítimo) de direito em processo judicial, administrativo ou em arbitragem (inciso VI).
 - f. Quando necessário para a proteção da vida e da integridade física do titular dos dados ou de terceiros (inciso VII) ou para a tutela da saúde por profissional competente, a exemplo da coleta de informações de um paciente em um hospital (inciso VIII).
 - g. Para atender interesses legítimos do controlador ou de terceiros, desde que não ofenda os direitos e as liberdades fundamentais do titular dos dados (inciso IX). Por exemplo, retenção de informações de clientes para ações de *marketing*.
 - h. Nas hipóteses legais de proteção de crédito, a exemplo do cadastro positivo (inciso X).

Em relação aos dados que remetem a crianças e adolescentes, não é permitido o compartilhamento de informações de menores com terceiros sem o devido consentimento de seus responsáveis, exceto para contatar os responsáveis pelo menor, a fim de coletar a sua autorização (artigo 14 e parágrafos seguintes). Sobre os dados pessoais sensíveis (que serão conceituados mais à frente), vale observar que o artigo 11 permite o seu uso sem necessidade de consentimento do titular, em hipóteses semelhantes às do artigo 7º. Um exemplo dessa dispensa é a hipótese de coleta para “tratamento compartilhado de dados necessários à execução, pela administração pública, de políticas públicas previstas em leis ou regulamentos” (art. 11, inciso II, item ‘b’).

5.2 FASES DO ARMAZENAMENTO E RECUPERAÇÃO DE DADOS

O tratamento dos dados (do qual o armazenamento é uma de suas atividades) remete ao princípio da finalidade, referente à “realização do tratamento para propósitos legítimos, específicos, explícitos e informados ao

titular, sem possibilidade de tratamento posterior de forma incompatível com essas finalidades” (artigo 6º, inciso I). Nessa lógica, a pessoa que cede suas informações tem o direito de saber se os seus dados serão compartilhados via *Linked Data*, em aplicações de *Big Data* ou em qualquer outro meio em que dados se interconectem. Por sua vez, os “dados pessoais sensíveis” recebem proteção especial da lei. Esses são definidos como:

[...] 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 (BRASIL, 2018, artigo 5º, inciso II).

A LGPD permite que a ANPD crie regulamento que proíba o compartilhamento de dados sensíveis para fins econômicos (artigo 11, §3º). A pretensão do texto mostra a inviabilidade do compartilhamento de dados sensíveis entre *triplestores* de organizações distintas, considerando as intenções da lei em proteger a privacidade dos envolvidos nessas relações.

Além disso, caso haja efetivo compartilhamento de dados, o titular tem direito de saber como essa atividade se realizou e qual a sua finalidade, de maneira “clara, adequada e ostensiva” (artigo 9º, caput, e inciso V). Na mesma lógica, a lei afirma, em outro momento, o direito do titular em requerer informações sobre com quem os seus dados são compartilhados (artigo 18, caput e inciso VII). A informação sobre o tratamento de dados inclui o conhecimento sobre atividades de dados ligados (como numa empresa que disponibiliza dados de seus consumidores numa rede compartilhada). Em cenários de *Linked Data*, a observância dessa regra é imprescindível, ante a capacidade estrutural dessa tecnologia de permitir o acesso de dados por terceiros. No uso compartilhado de dados, o responsável precisa informar aos demais operadores sobre qualquer anonimização, correção, eliminação ou bloqueio, para que todos os envolvidos no tratamento possam proceder igualmente (artigo 18, §6º).

5.3 FASE DO DESCARTE DE DADOS

O encerramento do tratamento pode ocorrer nas seguintes situações

(artigo 15): i) verificação da consecução dos fins destinados ao tratamento ou da desnecessidade de uso; ii) término do período estabelecido para tratamento; iii) por exigência do titular; iv) por exigência da ANPD, na forma da lei. Quando qualquer dessas hipóteses ocorrerem, os agentes de tratamento deverão eliminar os dados, da maneira mais segura possível e independentemente do procedimento utilizado para tanto (artigo 5º, inciso XIV). “O descarte pode e deve ocorrer, já que é um direito”, nominado de direito ao apagamento ou direito ao esquecimento (PINHEIRO, 2018, não paginado). Porém, os dados podem ser conservados nas seguintes hipóteses:

I - cumprimento de obrigação legal ou regulatória pelo controlador;

II - estudo por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais;

III - transferência a terceiro, desde que respeitados os requisitos de tratamento de dados dispostos nesta Lei; ou

IV - uso exclusivo do controlador, vedado seu acesso por terceiro, e desde que anonimizados os dados. (Brasil, 2018, artigo 16) (grifo nosso)

5.4 LGPD EM CENÁRIOS DE UTILIZAÇÃO DE DADOS LIGADOS

Expostas as principais regras da lei aplicáveis à Linked Data, à luz do CVD, apontam-se algumas exigências que são ainda mais convenientes para a proteção de informações pessoais em fenômenos específicos de dados ligados. Em razão da dimensão de aplicação da Linked Data, que vem se ampliando constantemente graças às latentes inovações neste campo da tecnologia, preferiu-se por adotar uma lista não-exaustiva de cenários de dados ligados, agrupados por categoria, considerando os principais fenômenos estudados pela literatura especializada.

Quadro 2 - Cenários de dados ligados agrupados por categoria

Categoria	Cenário <i>Linked Data</i>
Conteúdo proprietário	Criação de triplas/ <i>datasets</i> RDF com vocabulário ou ontologias proprietários criados no Brasil Criação de triplas/ <i>datasets</i> RDF com vocabulário ou ontologias proprietários criados em qualquer lugar Criação de triplas/ <i>datasets</i> RDF que modelam dados e/ou contextos proprietários
Conteúdo de terceiros	Criação de triplas/ <i>datasets</i> RDF com vocabulário ou ontologias de terceiros (brasileiros ou estrangeiros) criados no Brasil Criação de triplas/ <i>datasets</i> RDF com vocabulário ou ontologias de terceiros (brasileiros ou estrangeiros) criados em qualquer lugar Criação de triplas/ <i>datasets</i> RDF que modelam dados e/ou contextos de terceiros (brasileiros ou estrangeiros)
Armazenamento	<i>Triplestores</i> armazenadas em servidores brasileiros <i>Triplestores</i> armazenadas em servidores estrangeiros <i>Triplestores</i> armazenadas parcialmente ou totalmente em servidores estrangeiros
Acesso	<i>Triplestores</i> armazenadas no Brasil e acessadas por <i>queries</i> (buscas) brasileiras <i>Triplestores</i> armazenadas no Brasil e acessadas por <i>queries</i> (buscas) internacionais <i>Triplestores</i> armazenadas fora do Brasil e acessadas por <i>queries</i> (buscas) brasileiras
Dados abertos	Relação entre a licença <i>Creative Commons</i> (e outras do tipo) com a LGPD Princípio da livre distribuição dos dados abertos por terceiros (<i>Open Data</i>)

Fonte: Elaborado pelos autores

Diante das categorias de cenários de dados ligados, indicados pelo Quadro 2, pode-se analisar como ocorre a aplicação da LGPD a cada uma delas.

Primeiramente, em relação ao conteúdo proprietário, a primeira repercussão da LGPD diz respeito à necessidade de que *datasets* e triplas RDF proprietários que agreguem dados pessoais se submetam à referida norma, quando se enquadrarem em alguma das hipóteses do artigo 3º e desde que não incorram em qualquer das situações de dispensa de aplicação da lei, previstas no artigo 4º. Nesse campo de aplicação, também estão compreendidos vocabulários e ontologias que trabalham com dados pessoais, desde que se encontrem de acordo com os critérios de incidência da lei, já mencionados. Destaca-se que, se dados e contextos proprietários não utilizam informações pessoais ou são utilizados por pessoa física para fins meramente particulares e não-econômicos, não se aplica a LGPD (artigo 4º, inciso I).

Ainda sobre contextos proprietários, o controlador pode criar *datasets* ou triplas RDF com dados pessoais sem o consentimento de seus respectivos titulares, quando justificado por legítimo interesse, sendo, por outro lado, vedado o compartilhamento desses dados com outras instituições (artigo 10). Nesse cenário, o legítimo interesse pode ser alegado, por exemplo, no aproveitamento de dados de clientes para ações de *marketing* em uma rede de lojas.

Diante do teor do artigo 13 da lei, órgãos de pesquisa podem criar *datasets* com informações pessoais para realização de seus estudos, mas, via de regra, não podem compartilhar esses dados com terceiros. Caso o compartilhamento dos *datasets* realmente se faça necessário, recomenda-se que os dados sejam anonimizados ou pseudoanonimizados.

À guisa de conclusão da análise da categoria conteúdo proprietário, vê-se que, considerando os artigos 23 a 25 da LGPD, órgãos públicos podem criar *datasets* e outros sistemas de *Linked Data* para execução de políticas públicas em interoperabilidade com outras entidades estatais.

Em relação ao conteúdo de terceiros, ao criar *datasets* ou triplas RDF com vocabulários e ontologias não-proprietários, deve-se verificar a existência de informações pessoais coletadas ou tratadas de acordo com alguma das hipóteses do artigo 3º. Caso haja, aplica-se a LGPD. Porém, a lei não será aplicada aos dados adstritos às possibilidades de dispensa do artigo 4º. Por exemplo, aplica-se a aludida norma a vocabulário cujos termos são nomes de

consumidores de uma empresa que fez a coleta desses dados no Brasil, mesmo que a tripla que o utiliza fosse criada por instituição estrangeira.

Ainda sobre conteúdo de terceiros, é possível a transferência (compartilhamento) internacional de dados pessoais por tecnologias *Linked Data*, desde que os países envolvidos possuam níveis satisfatórios de proteção de dados, de acordo com critérios estabelecidos pela ANPD (artigo 34).

Deve-se observar que, ainda que o compartilhamento de dados pessoais com terceiros ocorra dentro de um mesmo país ou com estrangeiros, sempre haverá a responsabilização (obrigação de indenizar) de controladores e operadores que descumprirem as normas de segurança estabelecidas pela ANPD (artigo 47). Tanto aqueles que criaram os sistemas de *Linked Data* quanto os que deles se aproveitaram podem ser responsabilizados.

Quanto ao cenário de armazenamento de dados ligados, por meio de *triplestores*, também são observadas algumas especificidades. *Triplestores* (brasileiras ou estrangeiras) que armazenarem dados pessoais coletados ou tratados de acordo com os artigos 3º e 4º devem obedecer a LGPD. Se a *triplestore* estiver armazenada parcialmente ou totalmente em servidores estrangeiros e contiver informações coletadas ou tratadas conforme os artigos 3º e 4º, são aplicadas as regras da LGPD e dos outros países envolvidos. O contrário implica na mesma lógica.

Previsão legal com importante repercussão nas *triplestores* refere-se à dispensa de aplicação das regras da LGPD quando dados armazenados estiverem totalmente anonimizados, isto é, registrados de tal forma que não seja mais possível conhecer a pessoa a quem a informação se refere (artigo 12).

Em relação ao acesso de dados ligados, observa-se que, assim como em outros contextos de *Linked Data*, aplica-se a LGPD às *triplestores* (estrangeiras ou brasileiras) acessíveis por *queries* (nacionais ou internacionais) que trabalharem com dados pessoais, satisfizerem os critérios do artigo 3º e não incorrerem nas dispensas do artigo 4º. Fontes de *queries* (consultas) precisam manter seus sistemas de acordo com a LGPD, se contiverem informações pessoais e estiverem de acordo com os artigos 3º e 4º. Devem ser coletadas e tratadas de acordo com alguma das hipóteses dos artigos 7º (se dado pessoal

comum) ou 11 (se dado pessoal sensível).

Faz-se a ressalva de que dados manifestamente (notoriamente) públicos podem ser usados como fontes de consulta sem necessidade de consentimento do seu titular, desde que sua utilização no sistema respeite as demais regras previstas na lei (artigo 7º, §4º).

Por fim, diante da categoria dados abertos, vê-se que licenças abertas (como *Creative Commons*) precisam respeitar a LGPD, se os conteúdos compartilhados contiverem dados pessoais coletados ou tratados em alguma das circunstâncias do artigo 3º. Quando se trabalha com dados pessoais, os princípios da livre distribuição de dados abertos devem ser pautados pelos limites jurídicos da LGPD, ante a existência de algum dos requisitos do artigo 3º e da inexistência de qualquer das hipóteses de dispensa do artigo 4º.

Via de regra, dados pessoais podem estar em sistema de LDO, desde que seja permitido pelo titular dessas informações por meio de seu consentimento. Porém, informações manifestamente públicas não precisam de permissão do titular para serem compartilhadas, desde que respeitadas as demais regras previstas na lei (artigo 7º, §4º). A exceção que se faz a essa dispensa é em relação a dados pessoais sensíveis, que, em razão da vulnerabilidade a discriminações injustas e abusivas, não podem ser compartilhadas em contextos de *Linked Open Data*.

Em todos esses cenários de dados ligados, a aplicação territorial pode ser um entrave para que os tratamentos de dados se adequem não apenas à LGPD, mas a legislações estrangeiras. Postos os requisitos para aplicação da LGPD, previstos no artigo 3º da lei (veja subseção 3.1), pode haver convergências e divergências entre legislações de diversos países, além de que há outros Estados com lacunas jurídicas quanto a essa matéria (por ainda não terem legislações próprias). Existe, ainda, a dificuldade técnica de levantar quais são os países envolvidos em tratamento de dados pessoais em cenários *Linked Data*, pois nem sempre é possível fazê-lo. Somente a análise de cada caso concreto pode dizer quais legislações devem ser observadas.

6 CONCLUSÕES E PERSPECTIVAS FUTURAS

Os dados ligados surgiram há 15 anos como proposta inovadora para se atribuir semântica aos dados disponíveis na *Web*, dentro do contexto da *Web Semântica*. Considerados inicialmente como uma simplificação dos modelos RDF-RDFS/XML, baseados em ontologias, os dados ligados focaram numa representação mais enxuta por meio de triplas (recurso-propriedade-recurso ou recurso-propriedade-valor), de maneira a facilitar a construção de um modelo de dados que fosse um elo entre o que é, ao mesmo tempo, computacionalmente processável e humanamente compreensível. Apesar desse intuito inicial, que se mostrou bem-sucedido ao aproximar o “mundo do humanamente compreensível” do “mundo do computacionalmente processável”, novas abordagens para a produção e consumo de dados ligados foram sendo propostas ao longo dos anos. Entre os sistemas que mais se beneficiaram da sua utilização estão os que trabalham com a interoperabilidade de dados e com a representação do conhecimento. Por meio de buscas SPARQL, por exemplo, é possível se construir motores de inferência bastante acurados para aplicação em *triplestores*. Mais recentemente, têm sido propostos uma série de projetos que visam o uso combinado de dados ligados com ferramentas e técnicas de processamento de linguagem natural. *Triplestores*, ambientes *Big Data*, *Internet of Things* e aplicações diversas da *Linked Open Data* (como a licença aberta *Creative Commons*) também são outros exemplos de fenômenos que se fundamentam e/ou aproveitam da tecnologia de dados ligados.

Todavia, é importante se ressaltar que o insumo básico de todos esses cenários continua sendo o mesmo, os dados *Web*. Para cada recurso ou propriedade de qualquer tripla RDF criada, há uma URI correspondente. Ou seja, dados ligados significam dados disponíveis na *Web*. É neste momento, portanto, que a temática *Linked Data* se encontra com a Lei Geral de Proteção de Dados Pessoais (LGPD).

Apoiando-se na abordagem analítica do Ciclo de Vida dos Dados, este trabalho propôs-se a analisar correlações entre a LGPD e cenários de utilização de dados ligados. Especificamente, buscou-se analisar a aplicação da aludida

lei a fenômenos de *Linked Data* diversos, tais como: i) conteúdos produzidos (proprietários ou de terceiros); ii) armazenamento das *triplestores*; iii) acesso às triplas e aos resultados de *queries* SPARQL; e iv) relação entre licenças abertas comuns na *Web* e a LGPD. Além destas, há ainda questões derivadas relativas a privacidade de dados, direitos autorais, consumo de dados, informações sigilosas, dados sensíveis, conteúdo modelado pelos *datasets*, uso compartilhado de dados, entre outros. A LGPD também procura regulamentar o uso de dados que gerem informações de diversas naturezas (financeiras, públicas, privadas, de escopos organizacionais, etc.). No entanto, é no que tange aos dados pessoais que a lei procura resguardar mais fortemente a privacidade e liberdade individual na criação e utilização de dados, interesses de cunho íntimo e proteção de informações pessoais, especialmente nos artigos que tratam de condutas ofensivas a estes critérios. Para tais situações, a lei prevê diversos níveis de sanções, indo de advertências e orientações para quem ainda busca se adequar à norma até a aplicação de multas em casos mais graves.

A LGPD é também bastante específica com relação a empresas estrangeiras que atuam no território brasileiro. A lei prevê que essas empresas sejam notificadas para se adequarem à legislação brasileira. Entretanto, fica claro que há uma série de circunstâncias que demandam um olhar mais cuidadoso neste cenário, como por exemplo, todas as possíveis situações de produção, consumo e intercâmbio de dados entre matrizes localizadas no exterior e filiais em território nacional. Há ainda a questão do armazenamento “físico” dos dados, ou seja, onde se localizam geograficamente os chamados *data centers*. E como a LGPD deve ser interpretada com relação a eles?

Devido à extensão do texto da LGPD (65 artigos no total), somada à grande quantidade de possíveis cenários *Linked Data* (que continua em constante expansão), optou-se por relacionar os principais fenômenos de dados ligados às regras jurídicas que já denotam alguma clareza à sua aplicação. Assim, questões legais que carecem de melhor regulamentação pela ANPD (especialmente diante da realidade dos dados ligados), como conflito de normas de diferentes países, regras mais profundas sobre a proteção de dados pessoais em entidades públicas, ações de segurança computacional, dentre outras, não

foram abordadas no presente trabalho.

Finalmente, para além da questão legal, há ainda os desafios tecnológicos existentes com relação a produção e consumo de dados na *Web*. A grande pergunta que precisará ser respondida num futuro próximo é: como garantir a aplicação da lei se esta poderá ser burlada por estratégias tecnológicas? Acredita-se que trabalhos futuros devem investir na análise dessas questões.

REFERÊNCIAS

- BERNERS-LEE, T. Linked Data. Design Issues, 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 10 maio 2021.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web: a new form of Web content that is meaningful to computers. **Scientific American Special Online Issue**, [S. l.], p. 24-30, Apr. 2002. Disponível em: http://csis.pace.edu/~marchese/CS835/Lec9/112_SemWeb.pdf. Acesso em: 10 maio 2021.
- BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked Data – The Story so Far. **International Journal on Semantic Web and Information System**, [S. l.], Jul. 2009. Disponível em: <http://eprints.epwp.eprints-hosting.org/id/eprint/92/>. Acesso em: 10 maio 2021.
- BORKO, H. Information science: what is it? **American Documentation**, v. 19, n. 1, p. 3-5, 1968.
- BRASIL. **Lei nº 8.078, de 11 de setembro de 1990**. [Código de Defesa do Consumidor (CDC)]. Dispõe sobre a proteção do consumidor e dá outras providências. Brasília: Presidência da República. 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 10 mar. 2021.
- BRASIL. Câmara dos Deputados. **Projeto de Lei nº 4.060/2012, de 13 de junho de 2021**: ficha de tramitação. Dispõe sobre o tratamento de dados pessoais, e dá outras providências. Brasília: Câmara dos Deputados, 2012. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=548066>. Acesso em: 10 maio 2021.
- BRASIL. Câmara dos Deputados. **Projeto de Lei nº 5.276, de 15 de maio de 2016**: ficha de tramitação. Dispõe sobre o tratamento de dados pessoais para a garantia do livre desenvolvimento da personalidade e da dignidade da pessoa natural. Brasília: Câmara dos Deputados, 2016. Disponível em:

<https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2084378>. Acesso em: 10 maio 2021.

BRASIL. **Governo Federal**. Portal Gov. [20--]. Disponível em:
<https://acesso.gov.br>. Acesso em: 21 mar. 2021.

BRASIL. **Lei nº 12.527/2011 [Lei de Acesso à Informação]**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília: Presidência da República, 2011. Disponível em:
http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm.
Acesso em: 14 abr. 2021.

BRASIL. **Lei nº 12.965, de 23 de abril de 2014 [Marco Civil da Internet]**. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Brasília: Presidência da República. 2014. Disponível em:
http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm.
Acesso em: 10 mar. 2021.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018 [Lei Geral de Proteção de Dados Pessoais (LGPD)]**. Dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado. Brasília: Presidência da República. 2018. Disponível em:
http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm.
Acesso em: 14 abr. 2021.

CASTELLS, M. A sociedade em rede. 6. ed. São Paulo: Editora Paz e Terra, 1999.

CONFESSORE, N. Cambridge Analytica and Facebook: The Scandal and the Fallout so far. **The New York Times**, 4 abr. 2018. Disponível em:
<https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html>. Acesso em: 10 mar. 2021.

CONSUMERS INTERNATIONAL. **The state of data protection rules around the world**: a briefing for consumer organisations. [S. l.]: Consumers International, 2018. Disponível em:
<https://www.consumersinternational.org/media/155133/gdpr-briefing.pdf>.
Acesso em: 10 mar. 2021.

CUNHA, D. R. B.; SOUZA, D.; LÓSCIO, B. F. Linked Data: da web de documentos para a web de dados. In: SANTANA, A. M. *et al.* (org.). **Livro texto dos minicursos ERCEMAPI**. Teresina: 2011. p. 79-99. Disponível em:
https://www.cin.ufpe.br/~daise/arquivos/publications/2011/2011_Cap%204%20Linked%20Data%20-%20MC2.pdf. Acesso em: 10 maio 2021.

DIAS, G. A.; VIEIRA, A. A. N. Big Data: questões éticas e legais emergentes. **Ciência da Informação**, Brasília, v. 42, n. 2, p. 174-184, 2013. Disponível em: <http://revista.ibict.br/ciinf/article/view/1380>. Acesso em: 10 maio 2021.

IRAMINA, A. RGPD V. LGPD: Adoção estratégica da abordagem responsiva na elaboração da Lei Geral de Proteção de Dados do Brasil e do Regulamento Geral de Proteção de Dados da União Europeia. **Revista de Direito, Estado e Telecomunicações**, Brasília, v. 12, n. 2, p. 91-117, out. 2020. Disponível em: <https://core.ac.uk/reader/337598120>. Acesso em: 10 maio 2021.

LEME, R. S.; BLANK, M. Lei Geral de Proteção de Dados e segurança da informação na área da saúde. **Cadernos Ibero-Americanos de Direito Sanitário**, Brasília, v. 9, n. 3, p. 210-224, jul./set. 2020. Disponível em: <https://www.cadernos.prodisa.fiocruz.br/index.php/cadernos/article/view/690>. Acesso em: 10 maio 2021.

LIMA, J. C.; CARVALHO, C. L. **Resource Description Framework (RDF)**. Goiânia: Universidade Federal de Goiás, 2005.

MAGRANI, E. J. G. **A Internet das Coisas: Privacidade e Ética na Era da Hiperconectividade**. 2018. 333 f. Tese (Doutorado em Direito) – Departamento de Direito, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2018. Disponível em: <https://www.maxwell.vrac.puc-rio.br/colecao.php?strSecao=resultado&nrSeq=36942@1>. Acesso em: 10 maio 2021.

NHACUONGUE, J. A.; ROZSA, V.; DUTRA, M. L. Linked Data e Ciência da Informação: diretrizes para a publicação de datasets institucionais abertos. **Biblios**, [S. l.], n. 73, p. 20-34, Oct./Dic. 2018. Disponível em: http://www.scielo.org/pe/scielo.php?script=sci_arttext&pid=S1562-47302018000400002. Acesso em: 02 dez. 2021.

PINHEIRO, P. P. **Proteção de dados pessoais: comentários à Lei n. 13.709/2018 (LGPD)**. São Paulo: Saraiva, 2018. *E-book*. Não paginado.

RODRÍGUEZ-DONCEL, V.; SANTOS, C.; CASANOVAS, P.; GÓMEZ-PÉREZ, A. Legal Aspectsof Linked Data. The European Framework. **Computer Law & Security Review**, [S. l.], v. 32, n. 6, p. 799-813, Dec. 2016. Disponível em: https://www.researchgate.net/publication/306022669_Legal_aspects_of_linked_data_-_The_European_framework. Acesso em: 10 maio 2021.

SANT'ANA, R. C. G. Ciclo de vida de dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, Londrina, v. 21, n. 2, p. 116-142, 2016. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/27940>. Acesso em: 21 mar. 2021.

SANTAREM SEGUNDO, J. E. Web semântica, dados ligados e dados abertos: uma visão dos desafios do Brasil frente às iniciativas internacionais.

Tendências da Pesquisa Brasileira em Ciência da Informação, [S. l.], v. 8, n. 2, p. 219-239, jul./dez. 2015. Disponível em:
<https://brapci.inf.br/index.php/res/download/119690>. Acesso em: 22 mar. 2021.

SARACEVIC, T. Ciência da informação: origem, evolução e relações.
Perspectivas em Ciência da Informação, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

TARAPANOFF, K.; ARAÚJO JÚNIOR, R. H.; CORNIER, P. M. J. A sociedade da informação e inteligência em unidades de informação. **Ciência da Informação**, Brasília, v. 29, n. 3, p. 93, set./dez. 2000. Disponível em:
<https://www.scielo.br/pdf/ci/v29n3/a09v29n3.pdf>. Acesso em: 12 abr. 2021.

UNIÃO EUROPEIA. **Regulamento 2016/679, de 4 de maio de 2016** [General Data Protection Regulation (GDPR)]. Bruxelas (Bélgica): Parlamento Europeu, 2016. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504&qid=1532348683434>. Acesso em: 10 maio 2021.

VALENTE, J. Temer sanciona lei de proteção de dados mas veta órgão regulador. **Agência Brasil**, 14 ago. 2018. Disponível em:
<https://agenciabrasil.ebc.com.br/geral/noticia/2018-08/temer-sanciona-lei-de-protecao-de-dados-mas-veta-orgao-regulador>. Acesso em: 10 mar. 2021.

VIDOTTI, S. A. B. G.; CONEGLIAN, C. S.; ROA-MARTINEZ, S. M.; VECHIATO, F. L.; SANTARÉM SEGUNDO, J. E. Web, Web Semântica e Web Pragmática: um posicionamento da Arquitetura da Informação. **Informação & Sociedade: Estudos**, João Pessoa, v. 29, n.1, p. 195-214, jan./mar. 2019. Disponível em: <https://brapci.inf.br/index.php/res/download/112083>. Acesso em: 10 maio 2021.

W3C. **W3C**. Disponível em: <https://www.w3.org/>. Acesso em: 10 maio 2021.

WERSIG, G.; NEVELING, U. The phenomena of interest to information science. **Information Scientist**, [S. l.], v. 9, p. 127-140, 1975.

CORRELATIONS BETWEEN LINKED DATA AND THE GENERAL LAW FOR PERSONAL DATA PROTECTION: CONSIDERATIONS ON USE SCENARIOS

ABSTRACT

Objective: The general objective of this work is to analyze the correlations between the General Law for Personal Data Protection (LGPD) and the scenarios for using linked data. **Methodology:** This is an exploratory study, using bibliographic and documentary references, as well as legal standards. Initially, a literature review was carried out consisting of three subsections: the first two analyzed the emergence, functioning and examples of use of linked data, while the last subsection researched the emergence and

structure of the LGPD. Subsequently, the main rules of the aforementioned law were proven, provisions on some limited data use scenarios, from the perspective of the Data Life Cycle phases: collection, storage, recovery and disposal. **Results:** It was noticed that the complexity of these systems is a technological challenge given the need to respect the law and preserve the privacy of individuals. In this sense, knowledge of this new standard and its adaptation to organizations are also dilemmas to be overcome. **Conclusions:** Therefore, it was understood that the applicability of the LGPD to linked data scenarios is varied, requiring the work of several professional and scientific actors so that companies and institutions in general can adapt to its rules.

Descriptors: Linked data. Right to privacy. General Law for Personal Data Protection. Information and Technology.

CORRELACIONES ENTRE DATOS ENLAZADOS Y LA LEY GENERAL DE PROTECCIÓN DE DATOS PERSONALES: CONSIDERACIONES SOBRE ESCENARIOS DE USO

RESUMEN

Objetivo: El objetivo general de este trabajo es analizar las correlaciones entre la Ley General de Protección de Datos Personales (LGPD) y los escenarios de uso de datos vinculados. **Metodología:** Se trata de un estudio exploratorio, utilizando referencias bibliográficas y documentales, así como normas legales. Inicialmente se realizó una revisión de la literatura compuesta por tres subapartados: los dos primeros analizaron el surgimiento, funcionamiento y ejemplos de uso de datos vinculados, mientras que el último subapartado investigó el surgimiento y estructura de la LGPD. Posteriormente, se dispusieron las principales normas de la citada ley, disposiciones sobre algunos escenarios limitados de uso de datos, desde la perspectiva de las fases del Ciclo de Vida de los Datos: recolección, almacenamiento, recuperación y disposición. **Resultados:** Se percibió que la complejidad de estos sistemas es un desafío tecnológico dada la necesidad de respetar la ley y preservar la privacidad de las personas. En este sentido, el conocimiento de este nuevo estándar y su adaptación a las organizaciones también son dilemas a superar. **Conclusiones:** Por lo tanto, se entendió que la aplicabilidad de la LGPD a escenarios de datos vinculados es variada, requiriendo el trabajo de varios actores profesionales y científicos para que las empresas e instituciones en general puedan adaptarse a sus normas.

Descriptores: Datos vinculados. Derecho a la privacidad. Ley General de Protección de Datos Personales. Información y Tecnología.

Recebido em: 17.02.2022

Aceito em: 17.01.2024