

UFRRJ
INSTITUTO MULTIDISCIPLINAR
CURSO DE PÓS-GRADUAÇÃO EM HUMANIDADES
DIGITAIS

DISSERTAÇÃO

A Privacidade está *on*?

Uma avaliação nos textos de Políticas de Privacidade e Termos de Uso de artefatos *IoT* em relação à Lei Geral de Proteção de Dados

ANDRÉ FERREIRA ZEFERINO

2021



INSTITUTO MULTIDISCIPLINAR CURSO DE PÓS-GRADUAÇÃO EM HUMANIDADES DIGITAIS

A Privacidade está *on*?

Uma avaliação nos textos de Políticas de Privacidade e Termos de Uso de artefatos *IoT* em relação à Lei Geral de Proteção de Dados

ANDRÉ FERREIRA ZEFERINO

Sob a Orientação do Professor
Rodrigo de Souza Tavares, Dr.

e Coorientação do Professor
Marcelo Panaro de Moraes Zamith, Dr.

Dissertação submetida como requisito parcial para obtenção do grau de **Mestre em Humanidades Digitais** no Programa de Pós-graduação Interdisciplinar em Humanidades Digitais, Área de Concentração em Análise Qualitativa e Quantitativa de Dinâmicas Sociais.

Nova Iguaçu, Rio de Janeiro

Outubro de 2021

Universidade Federal Rural do Rio de Janeiro
Biblioteca Central / Seção de Processamento Técnico

Ficha catalográfica elaborada
com os dados fornecidos pelo(a) autor(a)

F41p

Ferreira Zeferino, André, 1979-

A Privacidade está on? Uma avaliação nos textos de Políticas de Privacidade e Termos de Uso de artefatos IoT em relação à Lei Geral de Proteção de Dados / André Ferreira Zeferino. - Rio de Janeiro, 2021.
63 f.

Orientador: Rodrigo de Souza Tavares.

Coorientador: Marcelo Panaro de Moraes Zamith.
Dissertação(Mestrado). -- Universidade Federal Rural do Rio de Janeiro, PPGIHD - Programa de Pós Graduação Interdisciplinar, 2021.

1. Dados. 2. LGPD. 3. Privacidade. 4. Proteção. I. de Souza Tavares, Rodrigo, 1978-, orient. II. Panaro de Moraes Zamith, Marcelo, 1976-, coorient. III Universidade Federal Rural do Rio de Janeiro. PPGIHD - Programa de Pós-Graduação Interdisciplinar. IV. Título.



**INSTITUTO MULTIDISCIPLINAR
PROGRAMA DE PÓS-GRADUAÇÃO
INTERDISCIPLINAR EM HUMANIDADES DIGITAIS**



ANDRÉ FERREIRA ZEFERINO

Dissertação/Tese submetida como requisito parcial para obtenção do grau de **Mestre em Ciências**, no Curso de Mestrado em Humanidades Digitais, área de Concentração Análise Qualitativa e Quantitativa de Dinâmicas Sociais.

DISSERTAÇÃO APROVADA EM 13/10/2021

RODRIGO DE SOUZA TAVARES. Dr., UFRRJ (Orientador)

LEANDRO GUIMARAES MARQUES ALVIM. Dr., UFRRJ

MARCELO PANARO DE MORAES ZAMITH. Dr., UFRRJ

CAITLIN SAMPAIO MULLHOLAND. Dra., PUC - RJ

ATA DE DEFESA DE TESE N° 59/2023 - PPGIHD (11.39.00.16)

(N° do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 17/03/2023 10:09)
LEANDRO GUIMARAES MARQUES ALVIM
PROFESSOR DO MAGISTERIO SUPERIOR
DeptCC/IM (12.28.01.00.00.83)
Matricula: ###008#2

(Assinado digitalmente em 20/03/2023 09:00)
MARCELO PANARO DE MORAES ZAMITH
PROFESSOR DO MAGISTERIO SUPERIOR
DeptCC/IM (12.28.01.00.00.83)
Matricula: ###810#1

(Assinado digitalmente em 22/03/2023 13:53)
RODRIGO DE SOUZA TAVARES
PROFESSOR DO MAGISTERIO SUPERIOR
DeptCJ/IM (12.28.01.00.00.85)
Matricula: ###922#0

(Assinado digitalmente em 22/03/2023 11:31)
CAITLIN SAMPAIO MULHOLLAND
ASSINANTE EXTERNO
CPF: ###.###.587-##

(Assinado digitalmente em 17/03/2023 08:10)
ANDRÉ FERREIRA ZEFERINO
DISCENTE
Matricula: 2019#####3



Emitido em 12/05/2023

TERMO Nº 529/2023 - PPGIHD (11.39.00.16)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 12/05/2023 18:11)

RICARDO CORDEIRO CORREA
COORDENADOR CURSOS GRADUACAO - TITULAR
PPGIHD (11.39.00.16)
Matrícula: ###07#9

Visualize o documento original em <https://sipac.ufrrj.br/documentos/> informando seu número: **529**, ano: **2023**, tipo: **TERMO**, data de emissão: **12/05/2023** e o código de verificação: **d9cc0fab4**

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. *This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.*

Agradeço à Deus e aos Orixás que permitiram que esta dissertação fosse desenvolvida.

A minha esposa Ingrid Duarte que não permitiu que meu sonho fosse deixado de lado.

Te Amo!

As minhas filhas Luísa e Sofia, que essa entrega sirva de exemplo para mostrar que quando você quer algo, você pode conquistar.

Aos orientadores Rodrigo Tavares e Marcelo Zamith que foram fundamentais, de maneiras distintas, e nos momentos oportunos para que este trabalho fosse desenvolvido de maneira ímpar. Agradeço pela paciência com minha ansiedade e pelo aprendizado.

A memória de meu Pai Fábio Zeferino que infelizmente não poderá compartilhar este momento comigo.

A minha Mãe Cinira Zeferino que nunca compreendeu o motivo de tanta dedicação, mas que sempre me apoiou.

Ao Ilè Asè Ògùnjà, meu refúgio, minha fortaleza.

A minha família espiritual e carnal, Julio Cesar, José Antônio, Lena Duarte, Silvia Dias, Bruna Duarte (minha irmã) , Diego Duarte, Sandra, Josinaldo e Gustavo.

Aos amigos Beatriz, Laura, Marcos, Manuela e Flávio que em suas demonstrações de carinho me ajudaram a chegar neste momento.

Aos meus irmãos Patrícia e Roberto Zeferino.

Aos professores Ricardo e Leandro, que talvez não tenham percebido o seu grau de importância, nesta dissertação.

RESUMO

O uso contemporâneo de artefatos voltados à conectividade disposta nos mais diversos dispositivos é uma ramificação designada como *Internet*. Assim, inspirada na união desses artefatos do nosso habitual senso de conectividade, a rede virtual se torna indispensável ao nosso cotidiano. Dessa forma, este experimento tem como principal objetivo investigar o crescimento no uso de dispositivos voltados ao uso *Internet*, a produção de dados que cresceu massivamente com a facilidade e a propagação do seu uso, afetando a privacidade de cada usuário. A *Internet* se destaca como um mecanismo capaz de produzir, catalogar, direcionar e gerenciar as tomadas de ações estratégicas em qualquer ramo da sociedade, mas, todos os benefícios expostos produzem questionamentos de como nossa privacidade pode ser monitorada pelas organizações e governos, entre outros. Consoante a isso, os dados produzidos e que são coletados de cada usuário conectado levantam alguns questionamentos: existem mecanismos para exibir os textos dos Termos de Uso e as Políticas de Privacidade para o usuário de modo que haja uma fácil leitura e interpretação, os dados coletados estão devidamente protegidos e são utilizados de forma coerente, com o consentimento do usuário e dentro da legislação vigente no Brasil, sobretudo à LGPD. A NLP – em paralelo com outras técnicas – foi utilizada com o propósito de produzir às respostas necessárias para os questionamentos realizados juntamente com uma análise qualitativa dos resultados que foram obtidos.

Palavras-chave: Dados. LGPD. Privacidade. Proteção.

ABSTRACT

The contemporary use of artifacts aimed at connectivity arranged in the most diverse devices is a branch designated as the Internet. Thus, inspired by the union of these artifacts of our usual sense of connectivity, the virtual network becomes indispensable to our daily lives. Thus, this experiment has as its main objective to investigate the growth in the use of devices aimed at Internet use, the production of data that grew massively with the ease and spread of its use, affecting the privacy of each user. The Internet stands out as a mechanism capable of producing, cataloging, directing and managing the taking of strategic actions in any branch of society, but all the benefits exposed produce questions about how our privacy can be monitored by organizations and governments, among others. Accordingly, the data produced and collected from each connected user raises some questions: there are mechanisms to display the text of the Terms of Use and the Privacy Policies for the user so that the collected data can be easily read and interpreted they are duly protected and used consistently, with the user's consent and within the current legislation in Brazil, especially the LGPD. NLP - in parallel with other techniques - was used with the purpose of producing the necessary answers to the questions made together with a qualitative analysis of the results that were obtained.

Keywords: Data. LGPD. Privacy. Protection.

LISTA DE FIGURAS

Figura 1.1 – Site Google, adaptado.....	15
Figura 1.2 – Equipamento utilizado para acesso à <i>Internet</i> , adaptado	17
Figura 1.3 – Crescimento da <i>IoT</i> , adaptado.....	19
Figura 1.4 – <i>IoT</i>	20
Figura 1.5 – Características de soluções baseadas em Internet das Coisas, adaptado.....	21
Figura 1.6 – Os 3 V's do <i>Big Data</i>	24
Figura 1.7 – Relacionamento entre <i>IoT</i> e <i>Big Data</i>	25
Figura 2.1 – Linha do tempo até a LGPD	27
Figura 3.1 – Frequência das Palavras na LGPD	40
Figura 3.2 – Proximidade das palavras ‘dados’ e ‘pessoais’	41
Figura 3.3 – Frequência das Palavras nas Políticas de Privacidade	42
Figura 3.4 – Frequência das Palavras nos Termos de Uso	43
Figura 3.5 - Gráfico Multidimensional da LGPD.....	50
Figura 3.6 - Gráfico Multidimensional das Políticas de Privacidade	51
Figura 3.7 - Gráfico Multidimensional dos Termos de Uso.....	52
Figura 3.8 – Mapa de Calor: Políticas de Privacidade X LGPD	54
Figura 3.9 – Mapa de Calor: Termos de Uso X LGPD	55

LISTA DE TABELAS

Tabela 1.1: Uso da <i>Internet</i> pela população mundial	16
Tabela 1.2: Uso de <i>Internet</i> em Países da América do Sul	16
Tabela 3.1: LGPD – Resultados do Tópico 0.....	44
Tabela 3.2: LGPD – Resultados do Tópico 1	44
Tabela 3.3: LGPD – Resultados do Tópico 2.....	45
Tabela 3.4: Políticas de Privacidade – Resultados do Tópico 0.....	45
Tabela 3.5: Políticas de Privacidade – Resultados do Tópico 1	46
Tabela 3.6: Políticas de Privacidade – Resultados do Tópico 2.....	46
Tabela 3.7: Termos de Uso – Resultados do Tópico 0.....	47
Tabela 3.8: Termos de Uso – Resultados do Tópico 1	47
Tabela 3.9: Termos de Uso – Resultados do Tópico 2.....	48
Tabela 3.10: Políticas de Privacidade e Capítulos da LGPD	53
Tabela 3.11: Termos de Uso e Capítulos da LGPD	53

LISTA DE ABREVIATURAS E SIGLAS

ANPD	Autoridade Nacional de Proteção de Dados
ARPA	Advanced Research Projects Agency
CPF	Cadastro de Pessoa Física
DPO	Data Protection Officer
GDPR	General Data Protection Regulation
IBGE	Instituto Brasileiro de Geografia e Estatística
IP	Internet Protocol
IoB	Internet of Behaviors
IoT	Internet of Things
LDA	Latent Dirichlet Allocation
LGPD	Lei Geral de Proteção de Dados
NLP	Natural Language Processing
NSA	National Security Agency
PLC	Projeto de Lei da Câmara
RG	Registro Geral
SERPRO	Serviço Federal de Processamento de Dados
SRI	Stanford Research Institute
TJUE	Tribunal de Justiça da União Europeia

SUMÁRIO

INTRODUÇÃO	12
CAPÍTULO 1	14
1 DA CONECTIVIDADE À HIPERCONECTIVIDADE	14
1.1 CONECTIVIDADE: O ADVENTO DA <i>INTERNET</i>	14
1.2 A <i>INTERNET</i> DAS COISAS E A HIPERCONECTIVIDADE.....	18
1.3 BIG DATA.....	23
CAPÍTULO 2.....	27
2 RECURSOS PARA A PROTEÇÃO DE DADOS PESSOAIS: A LGPD E A PRIVACIDADE	27
2.1 LGPD – LEI GERAL DE PROTEÇÃO DE DADOS	27
2.2 O DIREITO CHAMADO DE PRIVACIDADE.....	32
2.2.1 <i>Aviso de Privacidade</i>	36
2.2.2 <i>Políticas de Privacidade</i>	36
2.2.3 <i>Termos de Uso</i>	37
CAPÍTULO 3.....	39
3 AVALIAÇÃO EMPÍRICA DA PROTEÇÃO DE DADOS CONFERIDA POR TERMOS DE USO E POLÍTICAS DE PRIVACIDADE	39
3.1 A FREQUÊNCIA DE PALAVRAS NO CORPUS DO EXPERIMENTO.....	39
3.2 AS INFERÊNCIAS SOBRE PRIVACIDADE NA MODELAGEM DE TÓPICOS	43
3.2.1 <i>Representação Gráfica da Modelagem de Tópicos</i>	48
3.3 SIMPLIFICANDO A VISUALIZAÇÃO: MAPA DE CALOR.....	52
CONCLUSÃO.....	55
REFERÊNCIAS	57

INTRODUÇÃO

A hiperconectividade é uma realidade em nossa sociedade, não damos nenhum passo sem que haja qualquer tipo de monitoramento, seja por nossos familiares, governos ou organizações. Tudo isso ocorre para que se faça entender como agimos a cada minuto. Na era dos veículos conectados, consoles que criam perfis e avatares, propagandas direcionadas, cidades e casas inteligentes surge o questionamento, onde está a nossa privacidade?

Nesta dissertação vamos abordar os artefatos *IoT – Internet of Things* ou simplesmente *Internet* das Coisas que nos cercam diariamente, provocando uma explosão de dados trafegando e expondo com quem estamos, onde estamos ou o que fazemos enquanto estamos conectados.

De acordo com Danilo Doneda ¹(2018, p.11):

A internet das coisas, à parte maiores rigores semânticos, é um termo que acaba evocando o aumento da comunicação entre máquinas pela internet (M2M, ou *machine-to-machine*, que recentemente ultrapassou em volume a comunicação interpessoal pela internet), o desenvolvimento de diversos utensílios (desde os prosaicos exemplos das geladeiras ou torradeiras ligadas à internet), além de microdispositivos, como sensores que, dispostos das mais diversas maneiras para captar dados.

Diante deste conceito e do surgimento do novo marco regulatório chamado de Lei Geral de Proteção de Dados, foram criados mecanismos para, em tese, preservar a integridade e a privacidade de cada pessoa que possua artefatos que estejam conectados à *Internet*. Este estudo está dividido de forma conceitual, prática e analítica, são três Capítulos com a seguinte estrutura:

No Capítulo I relata o início da conectividade entre as pessoas até a hiperconectividade, uma proposta histórica, que visa relacionar o crescimento do uso de artefatos conectados à *Internet* com a exposição dos dados produzidos pelos usuários. O Capítulo II faz referência à LGPD – Lei Geral de Proteção de Dados e os conceitos embarcados nos artefatos *IoT*, como Políticas de Privacidade, Avisos de Privacidade e Termos de Uso, suas definições e como estes instrumentos são expostos para a clareza do usuários na exposição dos seus dados. Nesse ponto, se faz necessários compreender os termos técnicos utilizados também pela lei e como é a sua aplicação prática. O Capítulo III vem munido de estratégias de comparação entre a LGPD com as Políticas de Privacidade, Avisos de Privacidade e Termos

¹ DONEDA, Danilo. Magrani, Eduardo. A internet das coisas / Eduardo Magrani. — Rio de Janeiro : FGV Editora, 2018. 192 p

de Uso. Os algoritmos criados através da linguagem *Python* e outros *softwares* de apoio ajudaram no desenvolvimento prático em encontrar similaridades entre a redação dos textos dos artefatos com a legislação vigente.

A conclusão responde ao título desta dissertação e apresenta os números e métricas que ajudam a compreender como estão sendo utilizados os dados que produzimos diariamente através dos artefatos *IoT – Internet of Things*. Será que nossa privacidade está sob nosso controle?

CAPÍTULO 1

1 DA CONECTIVIDADE À HIPERCONECTIVIDADE

1.1 Conectividade: O advento da *Internet*

A troca de informações tornou-se uma tarefa tão habitual que não podemos mais encarar o mundo sem que haja a comodidade na transmissão de todo e qualquer tipo de dado. A praticidade, a resolução de divergências, a diversão, entre outros, realizam-se pela transmissão de dados digitais. Essa é uma realidade que avança rapidamente.

No decorrer dos séculos, o nosso desenvolvimento intelectual nos permitiu a produção de novos artefatos e dispositivos. Dessa forma, estruturamos, reinventamos e buscamos usar toda a nossa capacidade na elaboração de tecnologias que revolucionaram nossos meios de comunicação.

De acordo com Kurose e Ross (2013, p. 11), “*a Internet é uma tecnologia revolucionária e inovadora e podem constatar que ela está provocando uma profunda transformação em nosso mundo.*”

A cada dia que passa, ferramentas e aplicações são desenvolvidas para criar experiências únicas para o usuário. Dessa forma, é possível comparar a *Internet* da era contemporânea com a sua forma original e primitiva. A *Internet* proporciona atualmente aos usuários uma mescla muito maior de comodidades onde criam um ambiente propício à interação de forma virtual e um controle sobre o condicionamento prático de como nossas atividades serão direcionadas.

Nas décadas de 1962 até 1995, Castells (2001) descreve o pressuposto início das atividades de troca de dados entre redes de computadores. A ARPA - *Advanced Research Projects Agency* foi elaborada com o objetivo primário de realizar pesquisas com o aval do Departamento de Defesa dos Estados Unidos. Em pormenores, era uma estratégia de cunho militar para competir com a extinta União Soviética. Ao mesmo tempo, os Centros Universitários disponibilizavam em sua rede, os “nós” , com o objetivo de realizar testes e promover pesquisas científicas, destacando-se a Universidade da Califórnia em Santa Clara e

em Los Angeles e a Universidade de Utah com o projeto SRI (*Stanford Research Institute*). (CASTELLS, 2001)

Os projetos citados trazem apenas algumas das ideias que deram origem aos serviços de *Internet* que utilizamos diariamente, sendo essenciais para o funcionamento da sociedade. As diferenças mais expressivas estão em como as páginas *web* foram modificadas através de linguagem de programação, que foram aperfeiçoadas no decorrer das décadas, e o desenvolvimento de aplicações *mobile*, com o objetivo de simplificar a vida do usuário.

A Figura 1.1 mostra a evolução do site *Google*, a esquerda é exibida a página lançada em 1996 e a direita a página atual em 2021. As mudanças são evidentes, com relação ao esmero em melhorar a experiência de navegação e interação do usuário.

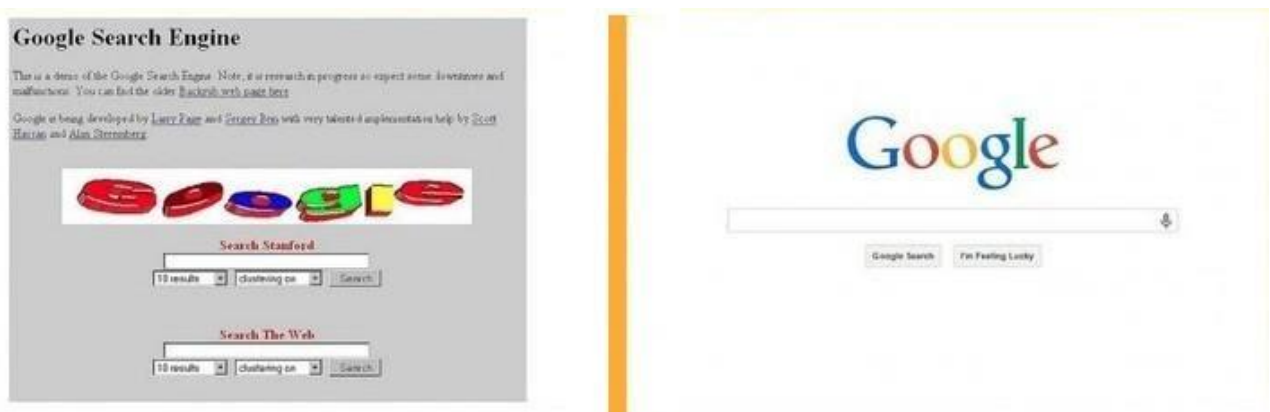


Figura 1.1 - *Site Google*, adaptado

Fonte: <https://noctulachannel.com/sites-famosos-historia/pagine-site-google-no-inicio/>

Desde a sua popularização, a *Internet* é responsável por diversas mudanças ocorridas em nossas vidas. As modificações abrangem campos que não são mais vistos sem a conectividade das redes, velocidade e responsividade promovidos por ela.

Segundo Anderson (2006, p. 8), “A única grande área em crescimento acelerado é a *Internet*, mas nesse caso trata-se de um oceano sem categoria própria, com milhões de destinos, cada um desafiando, à sua maneira, a lógica convencional da mídia e do marketing.”

A Tabela 1.1 exibe uma estimativa de uso de *Internet*² da população mundial no ano de 2020.

² Fonte: Internet World Stats, adaptado

TABELA 1.1: USO DA *INTERNET* PELA POPULAÇÃO MUNDIAL

Região	Pop. Total	% Pop. Mundial	Usuários, dez/2020
África	1,357,198,684	17.3	633,8856,924
Ásia	4,309,503,789	55.0	2,563,503,922
Europa	835,700,837	10.7	727,848,547
América Latina/Caribe	658,382,700	8.4	477,824,732
Oriente Médio	263,933,993	3.4	184,856,813
América do Norte	370,146,066	4.7	332,910,868
Oceania/Austrália	43,138,089	0.6	29,066,532

A Tabela 1.2 exibe os 14 países da América do Sul com seu quantitativo de usuários conectados à *Internet*. Em dezembro de 2017 o Brasil³ era o 3º do ranking atrás da Bolívia e da Argentina.

TABELA 1.2: USO DE *INTERNET* EM PAÍSES DA AMÉRICA DO SUL

Países	População	Uso de Internet, 31/12/2017	% População
Argentina	44,688,864	41,586,960	93.1
Bolívia	11,215,674	7,570,580	67.5
Brasil	210,867,954	149,057,635	70.7
Chile	18,197,209	14,108,392	77.5
Colômbia	49,464,683	31,275,567	63.2
Equador	16,863,425	13,476,687	79.9
Ilhas Malvinas	2,922	2,900	99.2
Guiana Francesa	289,763	120,000	41.4
Guiana	782,225	395,007	50.5
Paraguai	6,896,908	6,177,748	89.6
Peru	32,551,815	22,000,000	67.6
Suriname	568,301	340,000	59.8
Uruguai	3,469,551	3,059,727	88.2
Venezuela	32,381,221	17,178,743	53.1

Apesar de encontrarmos no Brasil um número expressivo de pessoas que não possuem renda para conseguir ter o acesso à *Internet*, o estudo demonstra que existe uma certa relevância no quantitativo de pessoas que conseguem ter o mínimo de acesso. As desigualdades sociais são um enorme empecilho no desenvolvimento tecnológico do país.

³ Fonte: Internet World Stats - América do Sul, adaptado

Segundo pesquisa realizada pelo Instituto Brasileiro de Geografia e Estatística – IBGE em 2019, o acesso à *Internet* no Brasil é realizado de forma majoritária por telefone móvel celular.



Figura 1.2 – Equipamento utilizado para acesso à *Internet*, adaptado
 Fonte: <https://educa.ibge.gov.br/jovens/materias-especiais/20787-uso-de-internet-televisao-e-celular-no-brasil.html#subtitulo-1>

Entre 2018 e 2019, no País, houve redução do uso do microcomputador e do *tablet* para acessar a *Internet* (3,0 e 1,4 p.p., respectivamente), ao passo que aumentou em 8,4 p.p. o uso da televisão. O uso do telefone móvel celular para acessar a *Internet* permaneceu estável no período. (IBGE⁴, 2019)

A sociedade contemporânea, alicerçada na influência tecnológica, permitiu a todos nós, indivíduos de gerações diversas, inúmeras possibilidades de agir, pensar e interagir com diversos grupos sociais. Paradoxalmente, de acordo com Gates (1995), o tempo de sociabilidade entre as pessoas será reduzido. (GATES, 1995)

⁴ Fonte: PNAD Contínua: https://biblioteca.ibge.gov.br/visualizacao/livros/liv101794_informativo.pdf, p. 6

Na visão de Geus e Melo (2015, p. 35) é necessário:

(...) refletir sobre a segurança dos dispositivos que estamos conectando à *Internet*. Tais dispositivos podem armazenar e transmitir informações altamente sensíveis sobre seus usuários, como hábitos de chegada e saída em casa, locais por onde a pessoa correu, passeou com sua criança ou animal de estimação e, mesmo hoje, geladeiras já avisam da falta de algum alimento, em breve também fazendo automaticamente o pedido de reposição ao e-supermercado.

O crescimento econômico e demográfico mundial demonstrou claramente nas últimas décadas que o desenvolvimento e a produção de novas informações, além do conhecimento, sofreram modificações. De forma detalhada, Prahalad (2009) acrescenta que o potencial participativo de usuários, além da facilidade de inclusão convergem como fatores que chamam a atenção de diversas organizações para aumentar de forma exponencial a sua visibilidade na rede. Os índices aumentam a cada ano, usuários e organizações trafegando seus dados de forma colaborativa e com interesse comercial. (PRAHALAD, 2009)

1.2 A *Internet* das Coisas e a Hiperconectividade

A evolução da *Internet* trouxe diversas oportunidades de modificações nas Redes de Computadores, além de equipamentos e do surgimento de novas técnicas para que os dados trafegassem mais rápido, o conceito das Redes de Computadores foi sendo alterado com o surgimento da *Internet* das Coisas (*Internet of Things - IoT*). De acordo com Tanenbaum (2002) as Redes de Computadores são um grupo de *workstations* que conseguem trocar informações através do serviço de compartilhamento. E segundo sua própria indicação, Ashton (2009) teria sido o primeiro na utilização do termo *IoT*, em 1999.

A *Internet* das Coisas, em linhas gerais, é o aumento significativo da conectividade de dispositivos à *Internet*, o que nos permite utilizar os artefatos do nosso dia a dia, aproveitando sua propensão computacional e de aproximação, quando conectados a uma rede *web*. Segundo a *Cisco Internet Business Solutions Group* (2011) a *IoT* é claramente o ponto em que coisas ou objetos são mais numerosos em relação a conectividade do que propriamente o quantitativo de usuários. Diante de utensílios que apresentam aos usuários funcionalidades simplificadas das tarefas rotineiras, a *IoT* trouxe uma atribuição primordial para a evolução da tecnologia. Estes objetos agora trazem em suas funções a possibilidade de enviar e receber dados por meio de sensores associados à amplitude do processo mais importante, a comunicação. (IBSG, 2011)

Segundo Almeida (2015, p. 8):

A informação estará disponível em qualquer lugar, a qualquer momento, em qualquer situação, de forma integrada às atividades do cotidiano, no seu relógio, pulseira, óculos, dentre outros *wearables*. Por outro lado, segurança e privacidade passam a ser as principais preocupações: tudo conectado, tudo acessível, tudo observado!

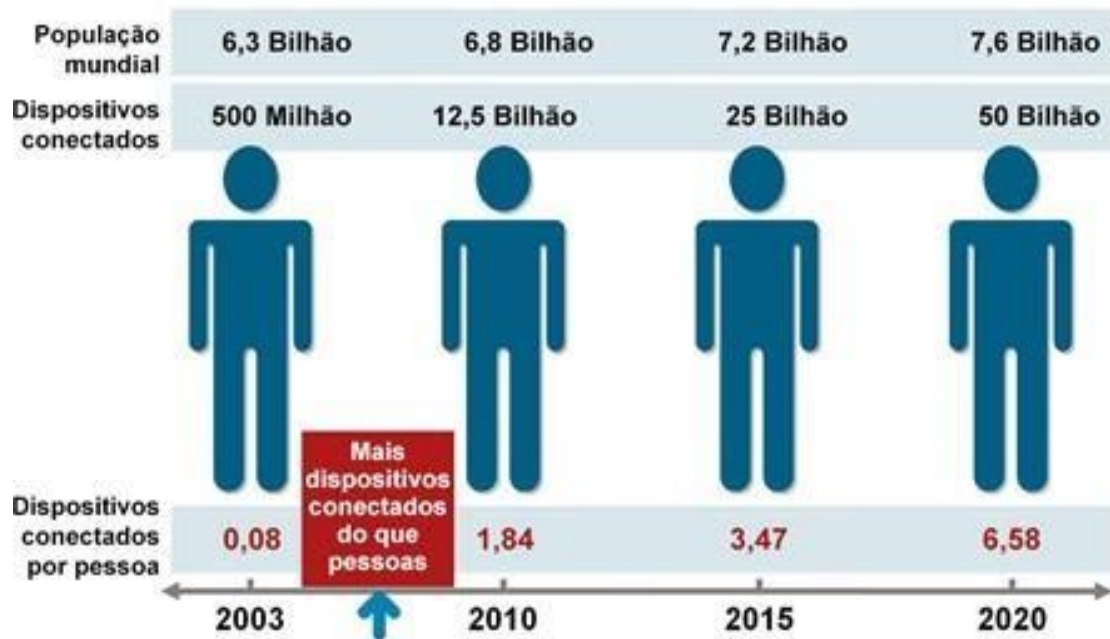


Figura 1.3 – Crescimento da *IoT*, adaptado
 Fonte: *Cisco Internet Business Solutions Group* (2011)

De fato, é importante esclarecer que o grande crescimento da *IoT* precisa ser observado e de maneira cautelosa. De acordo com a Figura 1.3, o número da população mundial cresceu, mas não acompanhou o ritmo dos dispositivos conectados. Pode-se inferir, observando a Figura 1.3, que a quantidade de dispositivos conectados em relação a população mundial é um pouco maior que 6 dispositivos por pessoa no ano de 2020. Enquanto a população cresceu em média 18% em 17 anos.

Ao analisar os dados acima podemos destacar que, com o surgimento da *IoT*, as empresas que trabalham diretamente com tecnologia criaram um nicho a ser explorado, mas podem ter esquecido do quesito segurança.

Diante dos fatos expostos, podemos afirmar que as tecnologias que foram desenvolvidas para ser rotuladas como *IoT*, podem expor seu melhor resultado desde que tomados os devidos cuidados. De acordo com Magrani (2018) a *Internet das Coisas* é uma das

tecnologias digitais emergentes que mais surpreendeu os usuários e ganhou um grande destaque entre as empresas privadas e governos de todo o mundo. (MAGRANI, 2018)

A Figura 1.2 mostra a relação entre o surgimento da *Internet* das Coisas com o passar dos anos, além do aumento de usuários, destacando o crescimento do número de dispositivos conectados.

Sobre *IoT*, em o artigo de 1999 para o Jornal de RFID, Ashton escreveu:

Se tivéssemos computadores que soubessem tudo sobre as coisas em geral - usando dados que coletassem sem a nossa ajuda - seríamos capazes de rastrear e contar tudo, e reduzir bastante o desperdício, a perda e os custos. Nós saberíamos quando é necessário substituir, reparar ou fazer um recall de um produto, e se estão novos ou ultrapassados. Precisamos capacitar os computadores com seus próprios meios de coletar informações, para que possam ver, ouvir e cheirar o mundo sozinhos, com toda a sua glória aleatória. O RFID e a tecnologia de sensores capacitam os computadores a observar, identificar e entender o mundo sem as limitações dos dados inseridos pelos humanos.

Em linhas gerais, o progresso emergente da tecnologia da *Internet* das Coisas trouxe um dinamismo para a sociedade e a economia em todo o mundo. Como citado anteriormente, diversas áreas foram e ainda serão impactadas (social, ambiental, educacional, agricultura, entre outros) pelas diversas aplicações do fenômeno *IoT*. As finalidades concretas nos dispositivos *IoT* que possam atender as demandas de organizações em todo o mundo podem trazer vantagens competitivas, além de alavancar empreendimentos de maneira congruente e ajudar a eliminar ou agilizar processos.



Figura 1.4 – *IoT*

Fonte:

https://cdn.shopify.com/s/files/1/1871/4015/articles/a_internet_das_coisas_580x.jpg?v=1494472163

Um dos fatores que vem alarmando a sociedade é a grande quantidade de dispositivos *IoT* conectados, coletando e trafegando informações pessoais. Com o aumento das novas tecnologias digitais foi criada uma sociedade pautada em uma cultura digital, onde a base de

conhecimento está voltada à hiperconectividade e ao conjunto de dados que produzimos a cada segundo.

Magrani (2018, p. 15) destaca que “*O atual cenário de hiperconectividade é, portanto, baseado na estreita relação entre seres humanos, objetos físicos, sensores, algoritmos, big data, inteligência artificial (computacional), cloud computing, entre outros elementos.*”

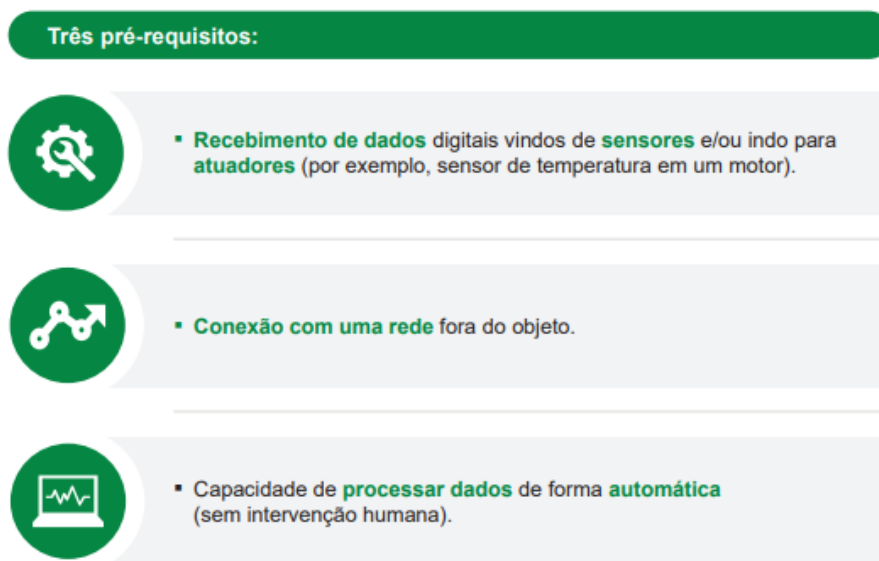


Figura 1.5 – Características de soluções baseadas em Internet das Coisas, adaptado
 Fonte: <https://www.bndes.gov.br/wps/wcm/connect/site/db27849e-dd37-4fbd-90466fda14b53ad0/produto-13-cartilha-das-cidades-publicada.pdf?MOD=AJPERES&CVID=m7tz8bf>

De acordo com Knopik (2018, p. 8):

Com mais pessoas vivendo nas cidades é certo que uma série de problemas surgirão, envolvendo questões ligadas à mobilidade urbana, segurança pública, saúde, sustentabilidade ambiental, infraestruturas de energia e de saneamento, aprofundamento da desigualdade social, entre outros. Como forma de enfrentamento a estes desafios os gestores públicos e empresas privadas tem analisado o emprego de políticas públicas ligadas à Tecnologia da Informação e Comunicação (TIC), trazendo à tona a ideia da *Smart City* ou simplesmente Cidade Inteligente.

Como já dito, a alta demanda na utilização de artefatos conectados à *Internet*, seja utilizado por pessoas ou mesmos distribuídos em espaços de uso coletivo, cria um impasse entre questões como, à coleta de dados, a supervisão sobre os locais de uso público, com quem estamos, que horas chegamos e saímos, qual a forma de pagamento, entre outros. De modo

geral, podemos afirmar que essas práticas são uma forma de controle de nossas ações e reduzem a perda da privacidade das pessoas.

Segundo Lacerda (2020, p. 63) *apud* Yan et al. (2008):

IoT deve ser uma plataforma de rede orientada às pessoas, na qual as comunidades criem valor a partir de múltiplas contribuições individuais, de modo que o ambiente futuro é aquele que vai depender não apenas de grandes centros de supercomputação e do governo, mas das práticas de várias matrizes de indivíduos em constante mudança, voltados à concepção de novos ambientes.

Dada importância do volume em que os dados são produzidos, coletados e armazenados nas ações que utilizam dispositivos *IoT*, devem ser levadas em consideração demandas relevantes e muitas vezes ignoradas por pessoas que usam artefatos conectados: a coleta que foi permitida pela pessoa sem violar o seu direito à privacidade e a segurança dessas informações contra vazamentos e invasões.

Cabe ressaltar que não são coletados dados em todos os artefatos *IoT*. Um caso bem interessante é o de aplicação de sensores que são usados para a contagem de pessoas trafegando em determinados ambientes, seu principal objetivo é de auxiliar no desenvolvimento de planos de mobilidade urbana. Mas é válido reforçar que essa aplicação é uma das infinitas possibilidades de aplicação nos artefatos. No caso de uso pessoal, podemos citar o exemplo dos *smartwatches* que coletam informações como localização, frequência cardíaca, quantidade de passos dados em uma simples caminhada, criando assim, um perfil de seu proprietário.

Em documento publicado pelo BNDES⁵ (2017, p. 12) objetiva-se, “*Acelerar a implantação da Internet das Coisas como instrumento de desenvolvimento sustentável da sociedade brasileira, capaz de aumentar a competitividade da economia, fortalecer as cadeias produtivas nacionais, e promover a melhoria da qualidade de vida*”. Sendo assim, podemos afirmar que o uso da *IoT*, pode trazer diversos benefícios desde que sejam regulados os processos de coleta de dados, entre outros.

⁵Fonte: <https://www.bndes.gov.br/wps/wcm/connect/site/269bc780-8cdb-4b9b-a297-53955103d4c5/relatorio-final-plano-de-acao-produto-8-alterado.pdf?MOD=AJPERES&CVID=m0jDUok&CVID=IXysvoX&CVID=IXysvoX&CVID=IXysvoX&CVID=IXysvoX&CVID=IXysvoX&CVID=IXysvoX&CVID=IXysvoX>

É de grande valia observar que os apontamentos voltados para a segurança de dados em dispositivos *IoT* vem ao encontro da necessidade de preservar a privacidade, um Direito Constitucional e que agora conta com proteção em legislação específica, a Lei Geral de Proteção de Dados.

1.3 Big Data

“O crescimento do ambiente online está gerando um aumento no fluxo de dados produzidos e o Big Data, como campo de estudo e comportamento, reúne esses dados e através do IoT estimula a criação de dispositivos cada vez mais complexos.”⁶

No decorrer do tempo, observamos que a produção de dados cresceu de forma notável, sendo assim, tornou-se um subproduto da informação, o seu valor para a humanidade é incalculável. A fabricação de dados cresce duas vezes mais a cada dois anos, e a estimativa é de que em 2020 tenham sido processados cerca de 350 *zettabytes* de dados, dessa forma, fica explícito que a exposição das informações dos usuários sem prévia autorização tornou-se uma incógnita.

A intensificação de uso de artefatos no universo *online*, fez com que as pessoas encontrassem um processo irreversível na produção massiva de dados. Esse movimento criou a necessidade de tecnologias com alto padrão de resposta e disponibilidade. O *Big Data* tornou-se um instrumental valioso para o desenvolvimento de novos nichos de estudo, seja de hábitos ou preferências ou apenas o armazenamento para consultas de grande quantidade de informações.

As tecnologias emergentes propiciam essa produção e exposição. Casas e carros conectados à *Internet*, o uso de tecnologias incorporadas à indústria ou ao comércio, cidades inteligentes geram a otimização pela utilização de sensores que coletam dados, e os algoritmos processam os resultados para ajudar na tomada de decisões mais rápidas. Diante da produção e coleta de dados processados massivamente surge a necessidade de uma forma de armazenamento que possa dar suporte para um grande volume de informações, a disponibilidade para que sejam armazenados dados de quaisquer tipos e a velocidade para que a resposta às solicitações seja rápida.

De acordo com Chen, Chiang e Storey (2012, p. 36) na década de 1950 já era postulada a grande quantidade de dados em áreas de conhecimento como a computação e a administração,

⁶<https://www.macnicadhw.com.br/noticias/2020-06-19-uniao-do-big-data-e-iot-geram-dispositivos-ainda-mais-inteligentes>

sendo assim o processo evolutivo para a grande capacidade de armazenamento era uma questão de tempo.

O surgimento de tecnologias que precisavam abarcar a produção de informações não ficou estagnado no mundo corporativo, o *Big Data*, segundo TOTVS (2021), “*é um termo que descreve o grande volume de dados que inunda uma empresa no dia a dia. Assim, o big data pode ser analisado em busca de insights que levam a melhores decisões e movimentos estratégicos de negócios.*” (TOTVS, 2021)

Porto, Ziviani e Ogasawara (2015, p. 26) indicam que:

Fica claro que estamos em face de mais uma grande (r)evolução na computação impulsionada pela *Internet*, e potencializada pela integração de dispositivos à rede. A autonomia na produção de dados retira a restrição da capacidade humana como agente da produção de dados e geração de conhecimento para tomada de decisão. Nesse novo contexto, novos componentes computacionais serão necessários, bem como revisitados serão aqueles hoje tidos como estado da arte. Parece que o termo *Big Data* ganha uma nova dimensão como uma nova Ciência de Dados.

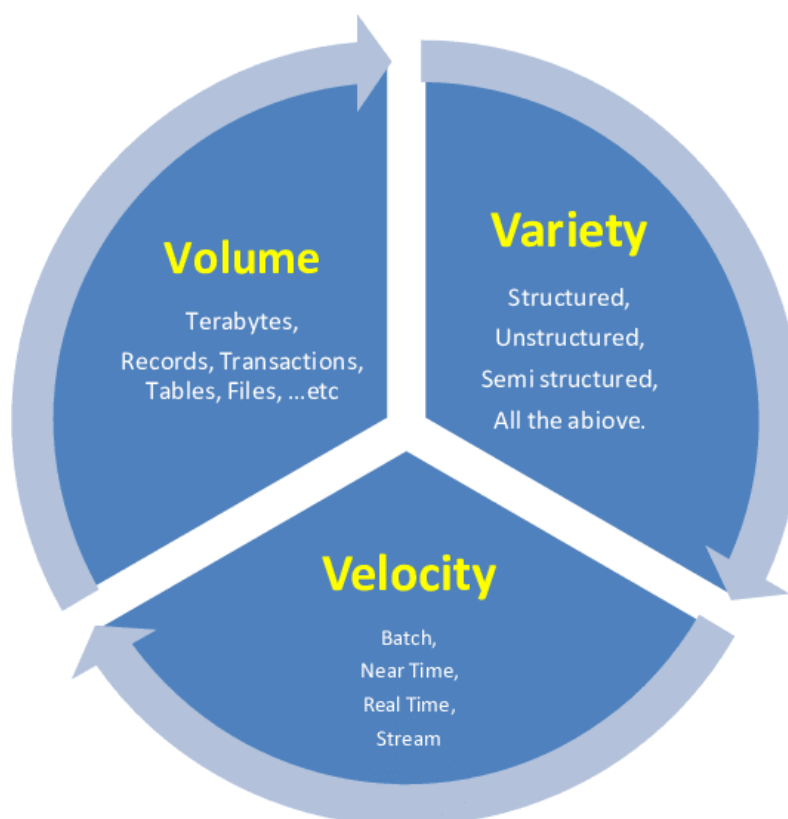


Figura 1.6 – Os 3 V's do *Big Data*

Fonte: https://www.researchgate.net/figure/BM-Big-Data-characteristics-3V-Adopted-from-Zikopoulos-and-Eaton-2011_fig1_258247680

De acordo com o Guia *Intel IT Center* (2014, p. 3):

Big Data refere-se a grandes conjuntos de dados que estão em ordens de magnitudes maiores (volume), mais diversificadas, incluindo dados estruturados, semiestruturados e não estruturados (variedade), e chegando mais rapidamente (velocidade) do que você ou sua organização já viram. Essa avalanche de dados é gerada por dispositivos conectados, de PCs e *smartphones* a sensores, como leitores de RFID e câmeras de trânsito. Além disso, é heterogênea e vem em muitos formatos, incluindo texto, documentos, imagens, vídeos, *weblogs*, transações e muito mais.

Para toda organização que deseja se manter em pleno funcionamento existe a necessidade de entender e aplicar ações voltadas à utilização de tecnologias emergentes. Podemos afirmar que a crescente produção de dados acompanha a constante evolução tecnológica ocorrida nos últimos anos.

De acordo com Reidy (2018), “*IoT e Big Data são combinados. A IoT conduz Big Data, fornecendo mais informações, de várias fontes diferentes em tempo real, e permite obter perspectivas completamente novas sobre o ambiente ao nosso redor.* REIDY (2018),”, tradução nossa.

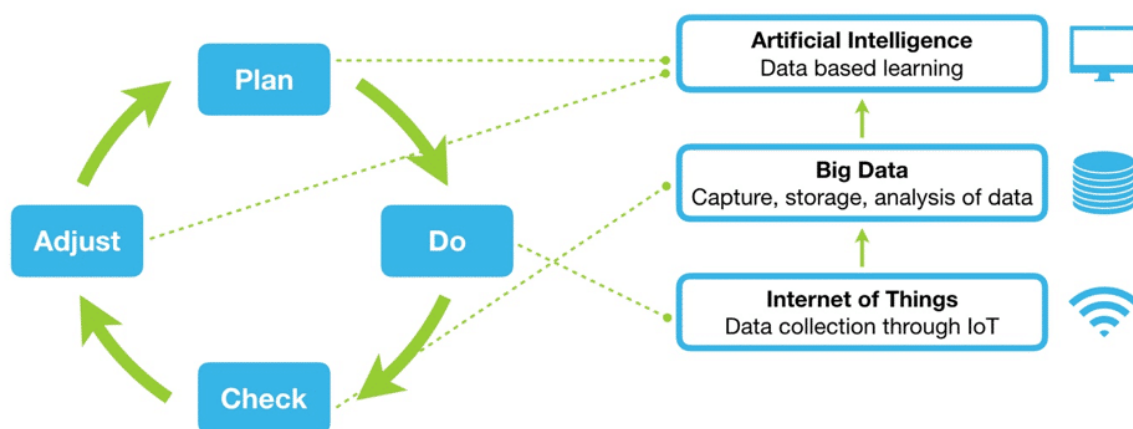


Figura 1.7 – Relacionamento entre *IoT* e *Big Data*, reproduzido
 Fonte: <https://arviem.com/iot-and-big-data-supply-chain-industry>

A tecnologia auxilia o processo de armazenamento e análise das informações a partir dos poderosos recursos disponibilizados pelo *Big Data* – que podem ser acessados de qualquer lugar do mundo, mas estes carecem de tratamento adequado para sua interpretação com valor. Em concordância com Kampakis (2020), o *Big Data* é uma coleção de dados geralmente

desorganizados, incompletos e que necessitam de *softwares* específicos para que seja realizado o seu tratamento. (KAMPAKIS, 2020)

Segundo a publicação de Serge Noiret na Liinc em Revista, (2015, p. 28):

A virada digital na história reformulou nossa documentação, transformou as ferramentas usadas para armazenar, tratar e acessar a informação, e, por vezes, adiantou novas questões epistemológicas juntamente com novas ferramentas criadas para responder por elas. Ainda assim, no momento, não há uma metodologia sistemática desenvolvida para abordar de forma crítica essas ferramentas digitais, analisar o deslocamento dos “*Big Data*” e compreender a nova capacidade pública para todos trabalharem com o passado.

Em colaboração, Burke (2021, p. 4) exhibe seu pensamento em relação ao comportamento das pessoas que corrobora com a produção massiva de dados:

Os dados podem vir de uma variedade de fontes, comerciais, dados do cliente para mídia social, para reconhecimento facial e como mais e mais dados se tornam disponíveis, o *IoB* - Internet de Comportamentos, irá capturar quantidade crescente de informações. Tradução nossa.

Segundo Chen (2018) essa nova forma de examinar o comportamento do usuário, propicia o desenvolvimento de estratégias de estudo nas áreas de conhecimento de ciências sociais e humanas. O objetivo é determinar novos pontos de vista da sociedade e desvendar novas formas de interpretar conceitos. (CHEN, 2018)

Conforme o Guia Brasileiro de Análise de Dados (2021), os dados não são adquiridos facilmente. Sendo assim, uma diversidade de tratamentos deve ser administrada para que seja extraída o melhor dessa análise. (SHIKIDA; MONASTERIO; NERY, 2021)

CAPÍTULO 2

2 RECURSOS PARA A PROTEÇÃO DE DADOS PESSOAIS: A LGPD E A PRIVACIDADE

2.1 LGPD – Lei Geral de Proteção De Dados

A Lei Geral de Proteção da Dados (LGPD), Lei N.º 13.709/2018, define diretrizes para a guarda dos dados dos usuários. Entre os mais relevantes aspectos, é considerado que as organizações tenham o total consentimento do usuário para utilizar seus dados que foram coletados, sendo observado o direito de revogação. Diante disso, é explícito o direito ao conhecimento do volume de dados armazenado pela organização, entidade ou governo, e se de alguma forma eles foram compartilhados com terceiros.

Filho (2018) relata que, “*Embora tardiamente o Brasil editou sua Lei geral de proteção de dados pessoais. Essa é uma Lei que traz ao mundo do direito a importância que os dados pessoais já possuem para a economia digital, onde são considerados “o novo petróleo”.*”

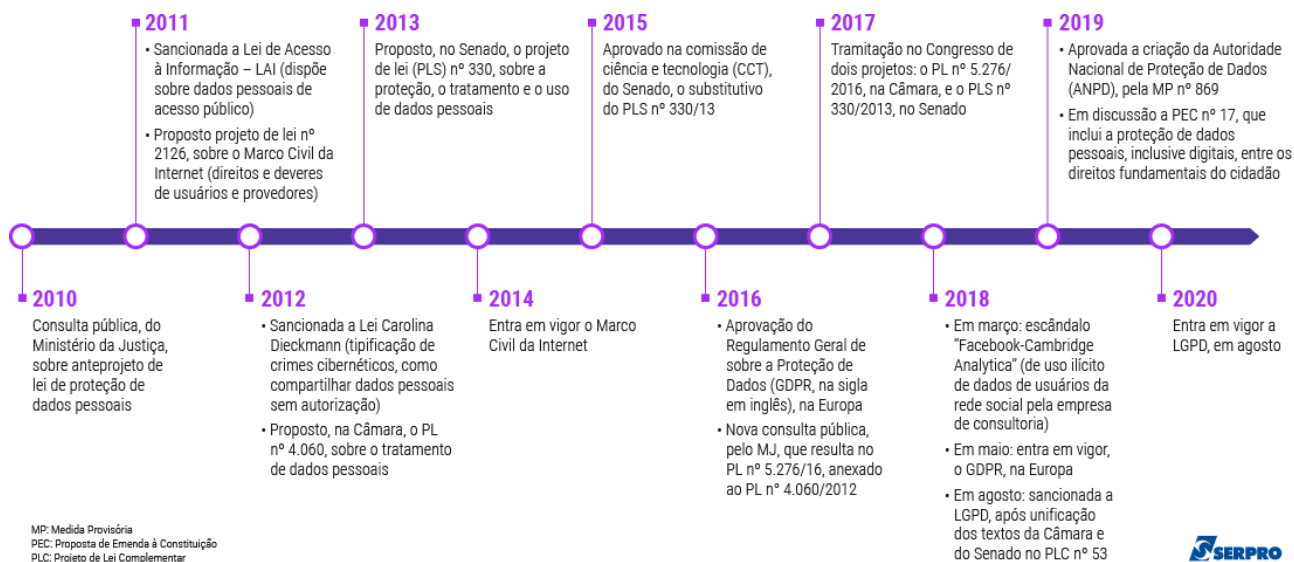


Figura 2.1 – Linha do tempo até a LGPD

Fonte: <https://www.serpro.gov.br/lgpd/menu/arquivos/linha-do-tempo/view>

A nova legislação chega em momento em que a inexistência de fronteiras permite que o usuário fique exposto a diversas ações irregulares e a exteriorização de sua privacidade. Sendo

assim, há obrigatoriedade de que cada *site*, *app*, aplicativo ou quaisquer que sejam os mecanismos utilizados para a coleta de dados dos usuários exiba sua Política de Privacidade e Termos de Uso, como uma forma de preservação dos direitos da personalidade dos usuários.

A LGPD direciona à Autoridade Nacional de Proteção de Dados (ANPD) todo e qualquer tipo de incidentes de segurança que de alguma forma coloque em risco os dados dos usuários, além de uma variedade de regulamentações que estejam em conformidades, dando ainda o aspecto mais rigoroso para as sanções administrativas. O fato é que muitas organizações, entidades e o próprio governo não estão preparados para a mudança cultural que a nova legislação exigirá. Atualmente os dados podem estar espalhados por vários sistemas, podem estar em poder de parceiros ou ainda sendo tratados para várias finalidades distintas dentro da mesma organização.

A Lei Geral de Proteção de Dados (LGPD) estabelece como fundamentos o respeito à privacidade, à autodeterminação informativa, à liberdade de expressão, de informação, de comunicação e de opinião, a inviolabilidade da intimidade, da honra e da imagem, ao direito de livre desenvolvimento da personalidade, ao desenvolvimento econômico e tecnológico, à livre iniciativa, à livre concorrência e a defesa do consumidor (Artigo 2º da Lei N.º 13.709/2018).

Um aspecto bem importante segundo a Lei, é ofertar grandes princípios que sejam incluídos no processo de desenvolvimento das tecnologias direcionadas ao usuário e que se proporcione artefatos mais seguros de fomento a construção da segurança e proteção dos seus dados. Definir um ponto de equidade sobre a idealização de tecnologias que sejam mais seguras, dentro de eventuais diretrizes que regem o seu desdobramento.

A definição de privacidade e segurança deve fazer parte do plano de desenvolvimento de algoritmos e *softwares* que prezam pelo potencial sentido de se preservar os dados de usuários em suas plataformas, consideradas digitais. Sendo assim, proporcionando a máxima segurança em informações que naquele momento estão sendo confiadas a terceiros.

A LGPD - Lei Geral de Proteção de Dados Pessoais é uma tentativa de desenvolver uma regulamentação que possa promover a preservação dos dados dos usuários e por consequência a sua privacidade. Promulgada em 14 de agosto de 2018 pelo então presidente da república Michel Temer, teve imensa repercussão para as instituições de caráter privado e

público. O esforço em produzir uma legislação de proteção de dados no Brasil surge mesmo antes do acesso de forma mais comum pela população brasileira à *Internet*.

A LGPD teve sua origem através do PLC - Projeto de Lei da Câmara Nº. 53/2018 por iniciativa do Deputado Federal Milton Monti. O projeto tinha a premissa de alterar a lei Nº 12.965, de 23 de abril de 2014, conhecida como o Marco Civil da *Internet*.

A nova lei possuía um prazo de dezoito meses para adequação às suas diretrizes nas organizações públicas ou privadas, não importando o segmento e o grau de relevância para a sociedade, entretanto, devido ao estado de pandemia iniciado em meados de 2020, a aplicação de suas penalidades fora postergada para agosto de 2021, o que permitiu um tempo maior para que as organizações pudessem adotar as medidas necessárias ao atendimento da norma.

A obrigação de existir uma lei bem alinhada em uma sociedade chamada de digital torna-se uma necessidade para a proteção dos dados do cidadão que nela existe. Além desse resgate, existe um paralelo com a implementação dos Direitos Fundamentais, visto no capítulo anterior, e a Declaração Universal dos Direitos Humanos que preza em seu Artº 12 que nenhuma pessoa será lesada em sua vida privada, entre outros direitos que protegem a dignidade humana.

A Lei Geral de Proteção de Dados (LGPD) brasileira direciona seu eixo à privacidade e à segurança como seus pontos focais, e orientando os princípios de uma sociedade cada vez mais definida como digital. Dessa forma, o tratamento das informações, realizadas pelos agentes de tratamento de dados, forçam os mesmos a definirem estratégias de cunho técnico e fundamentalmente administrativas, voltadas para a segurança dos dados de cada usuário. No que diz respeito ao acesso às informações de forma não-autorizada ou em circunstâncias consideradas imprevistas ou ilícitas, que ocasionem a destruição, perda, modificação e exposição dos dados.

(...)a LGPD é um manual para tratamento de dados pessoais, essa é a diferença. Quando observamos regulações setoriais, por exemplo, elas falam sobre o tratamento de dados pessoais para uma finalidade específica desse setor regulado. Já, com a chegada da LGPD, o cidadão passa a ter um poder maior nas mãos, pois uma das premissas da nova lei é que o cidadão é o titular dos dados, e essa titularidade atrai uma concepção de quase "propriedade". Ou seja, o titular voltou a ser o "dono" do dado. Essa titularidade, garantida pela LGPD, é incontestável, e um ganho para a sociedade. (MOTA, 2019)

Em seus dez Capítulos e sessenta e cinco Artigos, a lei descreve o entendimento sobre conceitos como Dados Pessoais, Dados Sensíveis, Tratamento e Consentimento. A Lei Geral de Proteção de Dados Pessoais é dispensada caso haja quaisquer relações que envolvam a segurança pública ou ações que envolvam a segurança nacional. O penúltimo artigo do Capítulo I, Art. 5º, traz a descrição dos termos utilizados para o melhor entendimento da LGPD, são eles:

- **Dados Pessoais:** é considerado algum elemento como nome e prenome, RG, CPF, entre outros que possa fazer o reconhecimento de uma pessoa. Pode-se incluir dados como localização, acadêmicos etc. Cabe ressaltar que está direcionado a pessoa viva;
- **Dado pessoal sensível:** é aquele que de alguma forma possa ser empregado para a discriminação de pessoa natural viva, seja por raça/etnia, religião, sexualidade, entre outros.
- **Dados anonimizados:** são aqueles pertencentes a uma pessoa que não consiga ser mencionado e necessite de estratégias técnicas para o seu tratamento.
- **Banco de dados, de acordo com o Portal GSTI (2021):**

“Um banco de dados é uma coleção organizada de dados (esquemas, tabelas, consultas, relatórios, exibições e outros objetos). Os dados são normalmente organizados para modelar aspectos da realidade de uma forma que suporta os processos de pedidos de informação, tais como modelar a disponibilidade de quartos em hotéis de uma forma que permita encontrar um hotel com vagas.” (GSTI, 2021)

- **Titular:** aquele que é proprietário dos dados que serão tratados.
- **Controlador:** é a pessoa natural ou jurídica, de natureza pública ou privada, que possui a incumbência pelas deliberações referentes ao tratamento dos dados de uma pessoa.
- **Operador:** é a pessoa que procede o tratamento dos dados em função das decisões tomadas pelo Controlador ou por determinação legal.

De acordo com Doneda (2019):

Sempre vai haver um controlador, que nada mais é do que a empresa que determina o consentimento da base legal. A lei tem essa diferenciação porque a responsabilidade do operador, nesse sentido, é menor. A diferença entre os dois é que se alguém toma as decisões sobre como os dados são tratados, ela é controladora. Se são apenas funções técnicas, ela é operadora.

- **Encarregado:** a principal função do encarregado é promover a comunicação entre as partes, titulares, controladores e a ANPD - Autoridade Nacional de Proteção de Dados.

Além disso, entre as suas atribuições estão a de receber queixas, conceder explicações e direcionar providências às demandas. O Encarregado na GDPR - *General Data Protection Regulation*, regulamento geral de proteção de dados da União Europeia, tem o nome de DPO - *Data Protection Officer*.

- Agentes de tratamento: é a forma utilizada para designar as tarefas e atribuições do Controlador e do Operador, já abordadas anteriormente.
- Tratamento: toda atividade executada com os dados de um Titular.
- Anonimização: com a essa aplicação os dados pessoais são totalmente desassociados dos seus titulares, esse método permite de forma absoluta que seja improvável o acesso ao titular.

Segundo matéria publicada por TOTVS⁷ (2020) :

A anonimização visa desatrelar os dados pessoais de seu titular, assegurando sua integridade e protegendo a identidade das pessoas, já que as empresas não podem expor essas informações ou usá-las de maneira indevida. Não é possível rastrear um dado anonimizado, o que infringiria as regras da Lei Geral de Proteção de Dados.

- Consentimento: é a autorização dada de maneira irrefutável pelo titular dos dados para que seja realizado o seu tratamento.

De acordo com Schrappe e Becker⁸ (2021) sobre consentimento:

(...)no artigo 5º da lei, que funciona como glossário ou termos de referência, consentimento, dentro do microsistema de proteção de dados e privacidade, como “manifestação livre, informada e inequívoca pela qual o titular concorda com o tratamento de seus dados pessoais para uma finalidade determinada.

- Bloqueio: é a interrupção do tratamento de dados de um titular que precisa ser replicada a todos os agentes de tratamento.
- Eliminação: entende-se como a exclusão dos dados do titular.
- Transferência Internacional de Dados: é o ato de deslocar os dados do titular, sob custódia e proteção da LGPD para outro país, assegurando que os dados terão a mesma proteção e efeitos da lei brasileira.

⁷ <https://www.totvs.com/blog/negocios/anonimizacao/>

⁸ <https://www.jota.info/opiniao-e-analise/artigos/consentimento-e-o-consentimento-na-lgpd-03012021>

- Uso compartilhado de dados: é a divisão dos dados do titular entre as entidades públicas ou privadas, respeitando as bases legais e o consentimento ofertado por ele.
- Relatório de impacto à proteção de dados pessoais: é tarefa obrigatória do controlador. Nele deve constar todas as ameaças que podem ocorrer no tratamento de dados e como podem ser desenvolvidas ações que possam de alguma forma reduzi-las.
- Órgão de pesquisa: são entidades que possuem no texto da lei formas diferenciais de tratar e solicitar o consentimento para tratamento de dados.
- Autoridade Nacional: órgão responsável pela regulação e cumprimento da lei.

De acordo com Doneda⁹ (2018):

É necessária para que diversas atividades da Autoridade que não são de cunho repressivo, como as de caráter educativo, de orientação, o estabelecimento de parâmetros e outras possam ser implementadas, concretizando a aplicação do princípio da prevenção, previsto pela LGPD.

2.2 O Direito chamado de Privacidade

“(...) é que as quatro paredes da propriedade privada de uma pessoa oferecem o único refúgio seguro contra o mundo público comum – não só contra tudo o que nele ocorre, mas também contra a sua própria publicidade, contra o fato de ser visto e ouvido. Uma existência vivida inteiramente em público, na presença de outros, torna-se, como se diz, superficial. Retém a sua visibilidade, mas perde a qualidade resultante de vir à luz a partir de um terreno mais sombrio, que deve permanecer oculto a fim de não perder sua profundidade em um sentido muito real, não subjetivo. O único modo eficaz de garantir a escuridão do que deve ser escondido da luz da publicidade é a propriedade privada, um lugar possuído privadamente para se esconder¹⁰.”

A privacidade tornou-se um grande foco de problemas para a nossa sociedade. Manter em segurança todas as nossas ações e informações sem que haja a interferência ou o desprazer de terceiros é uma tarefa complexa e que requer muita atenção. A definição para privacidade é pluralista, e de certo modo isso define a sua importância e seus valores. Conforme temos em nossas vidas a adição de novas tecnologias, cria-se uma diversidade de conceitos para a privacidade.

Segundo a definição de Solove (2008), a privacidade é o “*mais compreensivo dos direitos e o direito mais valioso para os homens civilizados*”. De acordo com o autor, é

⁹ <https://www.jota.info/opiniao-e-analise/artigos/o-que-esta-em-jogo-com-a-nova-autoridade-nacional-de-protecao-de-dados-13082018>

¹⁰ ARENDT, Hannah. A condição humana, p. 87

inestimável que haja uma visão geral sobre o termo, visto que, existem uma série de dificuldades em se encontrar uma definição adequada, principalmente ao mundo virtual. O autor descreve que as distinções históricas e culturais promoveram essa diversidade de definições sobre privacidade. (SOLOVE, 2008), tradução nossa.

A Constituição Federal Brasileira de 1988 traz em sua redação no seu Art. 5º, inciso X que “*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.*” Desse modo, podemos inferir que com o advento da comunicação, que se tornou praticamente instantânea, e dos recursos tecnológicos, o espírito protetor da dignidade humana presente no texto constitucional preocupou-se em explicitar a proteção da inviolabilidade à privacidade de cada cidadão.

Em sua obra *The Digital Person: Technology And Privacy In The Information Age*, Solove (2004) recomenda uma modificação nas leis para que se possa dar liberdade ao usuário no mundo digital, sem ferir o seu direito à privacidade. (SOLOVE, 2004)

Posner (2010) esclarece que o indivíduo tem o direito de se excluir da sociedade e manter assim sua privacidade resguardada, porém essa ação poderia ser onerosa e não ter bons olhos diante da sociedade. (POSNER, 2010)

Na visão de Mota (2019):

(...) a proteção de dados pessoais tem uma gênese na privacidade. A privacidade foi o primeiro elemento com o qual a humanidade se deparou como sendo um valor social que precisava de uma proteção jurídica e, assim, foi criado o direito à privacidade. Isso foi só o início. Temos, agora, uma construção baseada no desenvolvimento científico-tecnológico. Com esse desenvolvimento, os entes públicos e privados começaram a ter mais condições de processar dados. E aí surgiu a necessidade de criar obrigações relacionadas ao processamento dos dados. E essas obrigações foram traduzidas no que se conformou como sendo o direito à proteção de dados. A conclusão foi que não basta respeitar a privacidade, é preciso também observar a forma como os dados são processados. E isso para preservar tanto a privacidade como qualquer outra liberdade civil que o cidadão detenha. Ou seja, a necessidade de constitucionalizar a proteção de dados pessoais surge desta visão de que privacidade é um elemento, e a proteção de dados é outro, sendo uma evolução demandada pela própria humanidade. E no meio dessa evolução, que partiu dos aspectos tecnológico e econômico, também tinha que acontecer uma evolução jurídica.

Na era em que praticamente tudo o que fazemos está direcionado ao mundo digital, a privacidade do indivíduo precisa ser preservada em quaisquer que sejam as esferas. As compras que são realizadas de forma *on-line*, *apps* que são instalados em dispositivos *mobile* tem praticamente acesso a diversas informações de natureza bancária, de localização, informações

acadêmicas, e várias outras voltadas à intimidade do usuário. A pessoa deve ter a prerrogativa de que seus dados estejam inteiramente sob sua custódia e usá-los da forma que lhe for mais pertinente. (RODOTÁ, 2008)

Pode-se inferir que a privacidade é o tema de maior discussão nos últimos anos, parte desse contexto deve-se ao fato da Tecnologia da Informação ter ficado mais próxima das classes ditas menos favorecidas, gerando um fluxo de informações muito maior, que perpassa classes sociais.

Os governos e organizações diariamente coletam, armazenam e analisam uma enorme quantidade de informações que, muitas vezes, o usuário expõe de forma involuntária em suas redes sociais, entre outros meios. Na proporção que avançamos em nossas vidas digitalizadas, corremos o risco de que os dados que expomos sejam coletados sem o nosso conhecimento e utilizados sem o nosso consentimento. Com base na coleta desses dados, pode-se inferir conclusões sobre como o usuário realiza suas atividades rotineiras.

No evento de exposição de dados sigilosos da NSA realizado por *Edward Snowden* em meados de 2013, gigantes do mundo digital como *Google, Dell, Facebook e Microsoft*, além do próprio Governo americano, foram apontados como responsáveis pela espionagem de dados de usuários e de outros Governos. Como contratado da NSA (Agência de Segurança Norte-Americana), *Snowden* divulgou milhares de documentos onde eram comprovadas as atividades da agência de vigilância, além de trazer informações sobre o Programa *PRISM*¹¹ Programa de Vigilância dos Estados Unidos.

Segundo o artigo publicado por *Lily Hay Newman* no site *wired.com*:

Parte da série *WikiLeaks* de revelações de *hackers* da CIA detalham um método que a agência usa para localizar os computadores e as pessoas que os utilizam. A agência infecta dispositivos-alvo com *malware* que podem verificar quais redes *Wi-Fi* públicas um determinado computador pode se conectar em um determinado momento, bem como a intensidade do sinal dessas redes. A partir daí, o *malware* compara a lista de opções de *Wi-Fi* disponíveis com bancos de dados de redes públicas de *Wi-Fi* para descobrir aproximadamente onde o dispositivo está. (NEWMAN, 2017), tradução nossa.

De fato, é observado que os meios utilizados para determinar quais ações serão promovidas em prol de direcionar o usuário a padrões de compra, comportamento ou determinar

¹¹ O *PRISM* é concebido como uma das ferramentas mais utilizadas pelo governo americano com o objetivo de conseguir acesso a dados privados e públicos.

se ele pode praticar crimes como pedofilia ou invasão de privacidade, são comprovadamente praticados por governos e corporações de todo o mundo.

Segundo Cate (1997) os dados dos usuários são coletados, tratados, terçados e assim são criados modelos de nossos comportamentos. Essa exposição criou um mercado muito promissor para atrair usuários que são tratados como mercadorias valiosas. (CATE, 1997) Diante dos fatos apresentados foi criada a necessidade de garantir que o usuário tenha suas valiosas informações cuidadosamente tratadas e que tudo seja realizado de forma transparente.

O uso de dispositivos que atualmente dispõe da premissa direcionada da produção massiva de dado, é nos dias de hoje um padrão da *Internet*, sendo aplicados em novas tecnologias sujeitas a constante evolução. O acelerado desenvolvimento de autômatos conectados entre si, criam as mais diversas transformações na sociedade contemporânea. A privacidade e segurança dos dados do usuário ganham destaque e de fato, eles devem prestar a atenção devida em como suas informações serão exploradas.

O iminente desafio em proteger os dados de usuários, vai de encontro com os interesses econômicos preponderantes nesse momento de renovação tecnológica. É determinante que a legislação possa criar subsídios para a proteção das informações dos usuários e que governos e corporações estejam engajados em encontrar alternativas para que não haja exposição indevida de dados.

É inegável a importância que o Direito tem para a sociedade. Através dos seus poderes é possível manter a regulação de condutas sociais. No atual contexto de complexidade organizacional a modelagem de ações construída pela via do direito assume papel essencial na condução das transformações que a sociedade sofre devido à inovação tecnológica.

De acordo com Reuters (2019, p. 1):

O Direito sempre atuou em contextos de alta complexidade organizando as relações e os fluxos. Contudo, diante do volume massivo de dados, informações, da rápida transformação dos conhecimentos, e de transformações contínuas em mercados, essa tarefa tem exigido cada vez mais. O fato é que o futuro do setor jurídico sem tecnologia e conectividade integradas já não é uma opção, ficar para trás hoje é uma escolha.

Segundo Thompson (2012, p. 32) a *Internet* ainda possui diversos espaços sem lei e onde ainda impera um direito próprio e absoluto:

As configurações da *Internet*, como ensina a doutrina, são maleáveis. Não comportam, portanto, somente o grito libertário que não conhece limites. Não demandam a

neutralização de tudo que se ponha no caminho de usuários de liberdade infinita. A Internet será a imagem precisa das sociedades que queremos ser. Remove, sim, ditadores, e deve fazê-lo. Mas não pode, no caminho da democracia, extinguir-lhe a razão de ser – o igual valor, a dignidade de cada um dos integrantes do povo.

O Direito Digital de maneira simplificada é uma divisão entre todas as que já são encontradas no Direito. A demanda relacionada a sua criação vem ao encontro de uma sociedade que passou a produzir muitas informações, realizar negócios e desenvolver uma cultura digital. Acerca do exposto, Pinheiro(2009) demonstra preocupação em relação aos profissionais de Direito sobre como lidar com as demandas produzidas por essa nova cultura. (PINHEIRO, 2009)

2.2.1 *Aviso de Privacidade*

“De nada adianta um Aviso de Privacidade que não reflita de maneira fidedigna as operações de tratamento de dados aos titulares e torne-se obstáculo ao exercício de seus direitos¹².”

Os avisos de privacidade são manifestações que informam as pessoas sobre como serão coletados ou tratados seus dados enquanto utilizam artefatos ou visitam *sites* da *Internet*. De acordo com as boas práticas, estes avisos devem estar em evidência com vocabulário de fácil compreensão e ser apresentado antes de quaisquer tipo de coleta ou tratamento.

De acordo com Logarezzi (2016, p. 7):

Todos os órgãos e entidades públicas devem oferecer uma plataforma *online* para receber pedidos de informação, com protocolo. As páginas eletrônicas dos órgãos devem oferecer linguagem e ferramentas fáceis e claras para leigos, mas também oferecer facilidades para aqueles que manipulam dados de forma mais complexa.

Os avisos de privacidade não são objeto de estudo neste experimento.

2.2.2 *Políticas de Privacidade*

“É importante garantir que a política esteja facilmente disponível. Dessa forma, a organização demonstra profissionalmente seu compromisso com a transparência no tratamento dos dados pessoais. E o usuário deve demonstrar seu expresso consentimento e concordância com os termos da política antes do início desse tratamento¹³”

¹² <https://www.privacyacademy.com.br/politica-ou-aviso-de-privacidade.html>

¹³ <https://www.serpro.gov.br/lcpd/noticias/2019/elabora-politica-privacidade-aderente-lcpd-dados-pessoais>

É extremamente importante que no desenvolvimento das políticas de privacidade sejam compreendidos os conceitos pertinentes da Lei Geral de Proteção de Dados – LGPD, dessa forma será possível uma escrita que atenda todos os requisitos expressos na lei.

Nas palavras de Freitas (2019):

A política de privacidade é um dos instrumentos de implementação do *privacy by design*¹⁴ e faz parte da estrutura de documentos para a proteção de dados. A política objetiva dar visibilidade ao tratamento de dados pessoais em um determinado serviço, atendendo princípios da Lei Geral de Proteção de Dados Pessoais. (FREITAS, 2019)

De certo modo, a política de privacidade é “ignorada” pelas pessoas, muitas vezes é dado o consentimento para a coleta e o tratamento dos dados sem que haja qualquer preocupação na leitura do texto e as implicações podem ser irreversíveis caso surjam problemas referentes ao uso indevido de suas informações. Em alguns casos fica difícil para as pessoas conseguirem ler o documento, principalmente em artefatos *IoT*, visto que, você deve adquirir o produto e somente após a compra e ativação do artefato será permitido a leitura da sua política de privacidade.

Na visão de Schultz (2020):

Política de privacidade é um documento que informa ao usuário sobre os dados coletados de forma direta ou indireta durante o acesso a um site, aplicativo ou sistema. Esse documento também deve esclarecer o motivo pelo qual essas informações estão sendo guardadas. (SCHULTZ, 2020)

Uma das principais particularidades das políticas de privacidade faz relação com o processo de identificação do usuário, dessa forma, a LGPD tem mecanismos que auxiliam na prevenção e na proteção dos dados dos usuários.

2.2.3 Termos de Uso

De maneira geral os termos de uso têm a finalidade fornecer aos usuários o regramento para a utilização de *sites*, *apps* ou artefatos *IoT*. Podemos afirmar que é uma espécie de acordo entre quem fornece o serviço e quem faz uso dele. Sendo assim, é criada um anteparo jurídico para ambas as partes.

¹⁴ Nas palavras de Freitas o *privacy by design* ou privacidade desde a concepção, é uma abordagem ligada à Engenharia de Sistemas e que preza pela privacidade do usuário durante todo o processo de construção de uma solução.

Na visão de Bonani¹⁵ (2020), “(...) esse documento costumeiramente visto como um aborrecimento, do qual rapidamente se foge com um clique em “aceito”, é, na realidade, do interesse de quem quer que se coloque frente a quaisquer dos lados da tela”.

No Capítulo 3, será realizado a análise dos Termos de Uso e Políticas de Privacidade de dispositivos *IoT*, além da comparação com a LGPD.

¹⁵ Disponível em: <https://www.bonani.adv.br/termos-de-uso-o-que-sao-e-para-que-servem>, Acesso em 28 jul. 2021.

CAPÍTULO 3

3 AVALIAÇÃO EMPÍRICA DA PROTEÇÃO DE DADOS CONFERIDA POR TERMOS DE USO E POLÍTICAS DE PRIVACIDADE

Neste experimento foi realizado uma revisão das ações ocasionadas pelo uso da *Internet* à produção, o tratamento e a utilização dos dados coletados através das mais diversas plataformas *IoT*, seus termos de uso; políticas de privacidade e da LGPD que visa a proteção e a privacidade dos dados das pessoas. O mapeamento dos itens citados é de natureza intelectual, com direcionamento a fontes de pesquisas, análise quantitativa e qualitativa dos dados extraídos do *corpus*¹⁶ processado com a sua possível disparidade em relação à LGPD - Lei Geral de Proteção de Dados Pessoais.

A construção do projeto é baseada em pesquisa descritiva de cunho experimental com o uso de ferramentas direcionadas à produção de algoritmos de processamento como a *NLP* - *Natural Language Processing*¹⁷ em linguagem de programação *Python*. Dessa forma, foi possível investigar e tratar todo o *corpus* extraindo apenas o que é o objeto do experimento – sentenças, frases e palavras.

Com a utilização de plataformas como o *NLTK* - um pacote para o desenvolvimento de algoritmos em *Python* com o objetivo de interpretar a linguagem humana, o *Gensim* - onde os modelos utilizam o *corpus* para buscar referências e assuntos comuns e o *Spacy* - que possibilita disponibilizar termos comuns entre os documentos. Dessa forma, foi possível criar a *Topic Modeling Latent Dirichlet Allocation (LDA)* ou Modelagem de Tópicos: Distribuição Discreta Multivariada, a busca por frequência de palavras, gráfico multidimensional, termos mais salientes e as similaridades entre termos que contribuíram para que fossem alcançados os resultados desta pesquisa.

3.1 A frequência de palavras no corpus do experimento

¹⁶ Segundo o site Dicio.com, *Corpus* é “coletânea; reunião dos textos ou documentos sobre um assunto ou tema”

¹⁷ *Natural language processing (NLP) is a field that focuses on making natural human language usable by computer programs. NLTK, or Natural Language Toolkit, is a Python package that you can use for NLP.* Traduzido: O processamento de linguagem natural (PNL) é um campo que se concentra em tornar a linguagem humana natural utilizável por programas de computador. *NLTK*, ou *Natural Language Toolkit*, é um pacote *Python* que você pode usar para NLP. Fonte: <https://realpython.com/nltk-nlp-python/>. Acesso em: 29 jul. 2021.

O objetivo desta pesquisa é demonstrar que o uso constante de serviço da *Internet* das Coisas, proporciona a produção de dados de forma grandiosa, a observância geral pode ser extremamente abusiva de forma que é preciso investigar se as organizações e o governo estão preparados para a legislação vigente no tocante a proteção e a privacidade dos dados do usuário. Dessa forma, foram coletadas 9 Políticas de Privacidade e seus respectivos Termos de Uso dos artefatos *IoT*, visando analisar e comparar sua base textual com a LGPD - Lei Geral de Proteção de Dados.

A definição de privacidade e segurança deve fazer parte do plano de desenvolvimento de algoritmos e *softwares* que considerem pelo potencial sentido de se preservar os dados de usuários em artefatos *IoT*. Sendo assim, proporcionando a máxima segurança em informações que naquele momento estão sendo confiadas a terceiros. Em uma breve análise da frequência de palavras na Lei Geral de Proteção de Dados, podemos inferir que a redação é fortemente voltada à preservação da integridade das pessoas.

A Figura 3.1 representada abaixo, exibe as 10 palavras mais frequentes na LGPD - Lei Geral de Proteção de Dados:

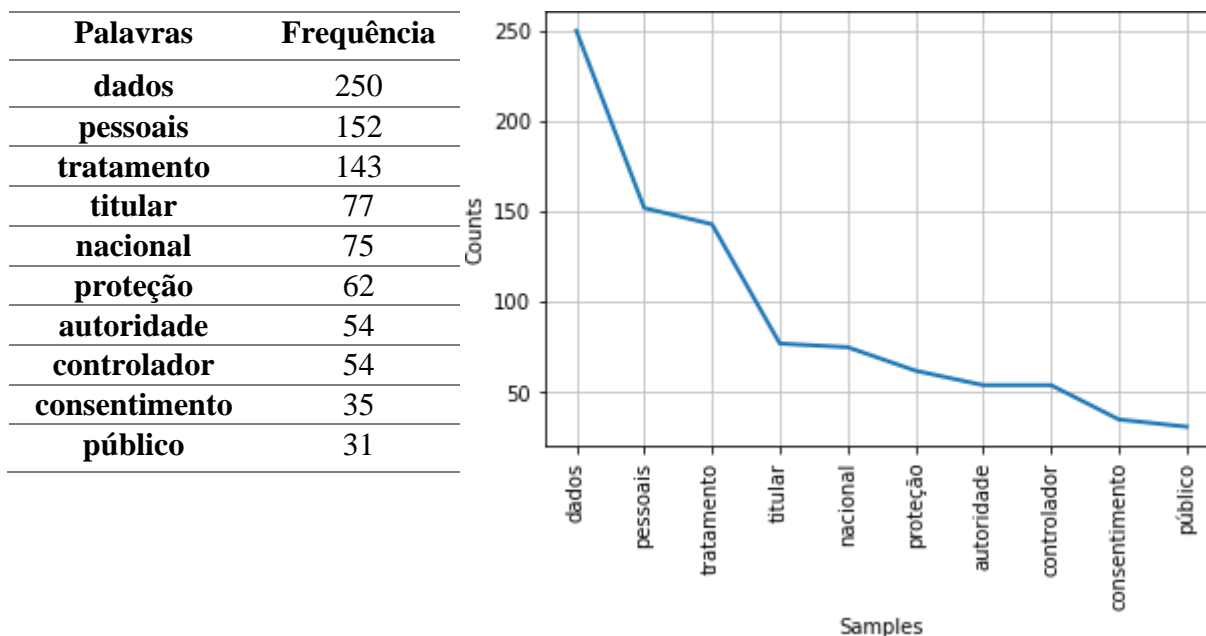


Figura 3.1 – Frequência das Palavras na LGPD
Fonte: Algoritmo desenvolvido pelo Autor

A LGPD - Lei Geral de Proteção de Dados, tem como premissa a proteção dos dados pessoais do usuário, dessa forma, observamos no Figura 3.1 a ocorrência consideravelmente superior das palavras “dados, pessoais e tratamento” em relação às demais palavras.

De acordo com Bortolossi¹⁸ (2019):

Com o advento da *internet*, o avanço da ciência e a interação entre povos de línguas diferentes, mais e mais dados são gerados em forma de texto. Extrair informações relevantes e identificar padrões nesta massa de dados constituem desafios tecnológicos atuais. Problemas desta natureza são estudados em campos de pesquisa como mineração de texto (*text mining*) e linguística computacional. Neste contexto, a estatística se estabelece como uma ferramenta fundamental e indispensável (outras ferramentas como banco de dados, teoria dos grafos, álgebra linear numérica e inteligência artificial também são usadas).

A Figura 3.2 só confirma a análise realizada anteriormente. Dessa forma, as palavras com maior frequência na redação da LGPD regularmente estão próximas.

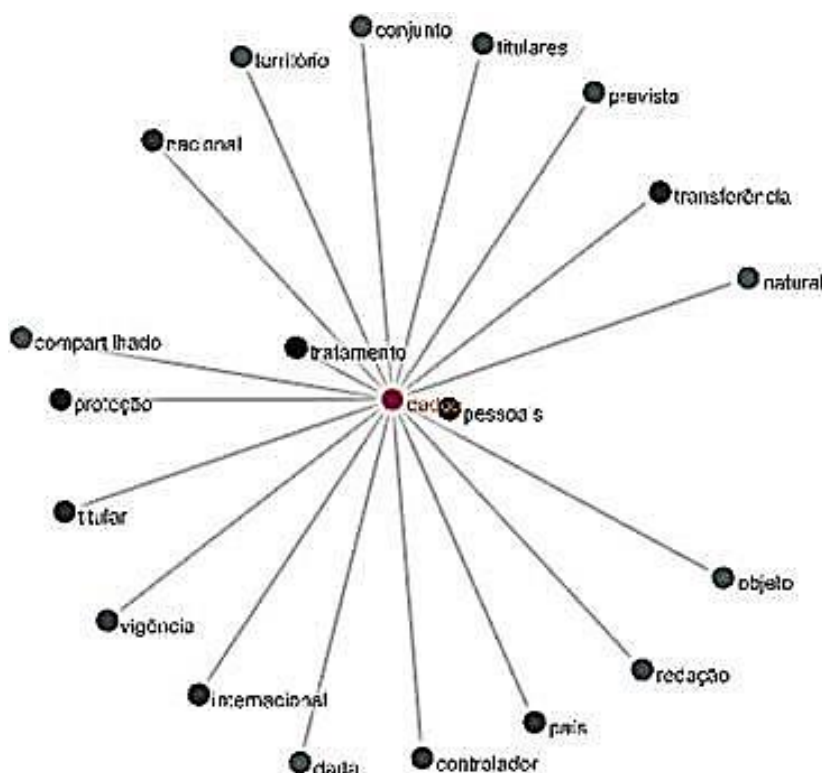


Figura 3.2 – Proximidade das palavras ‘dados’ e ‘pessoais’
 Fonte: Resultado obtido no *software Lancsbox*,
<http://corpora.lancs.ac.uk/lancsbox/download.php>

¹⁸ https://rdstation-static.s3.amazonaws.com/cms/%2Ffiles/%2F28555/%2F1581450471Ebook_LGPD_-_O_que_e_como_adequar_sua_empresa.pdf

A frequência e proximidade maior destas palavras era esperada dado o escopo da lei, que visa justamente garantir a proteção de dados pessoais.

As Políticas de Privacidade dos artefatos *IoT* demonstram em sua frequência de palavras uma grande variedade de vocábulos. Podemos observar um conjunto em que palavras como ‘serviços’ – indicando os serviços prestados, ‘voz’ – onde podemos afirmar que os artefatos podem fazer a captura das vozes das pessoas em sua volta para uma possível personalização de padrões, ‘terceiros’ – onde o contexto pode ser do compartilhamento, ou não, dos dados pessoais com outras organizações.

É de grande valia ressaltar que a frequência de palavras traz também os termos utilizados preferencialmente na LGPD, são eles: ‘dados’, ‘privacidade’ e ‘pessoais’. Isso é um indicativo de preocupação com as multas que podem ser aplicadas, caso sejam descumpridos os regramentos da lei.

Palavras	Frequência
informações	436
serviços	299
dados	294
privacidade	146
pessoais	140
grupo	116
usuário	106
software	101
voz	85
terceiros	82

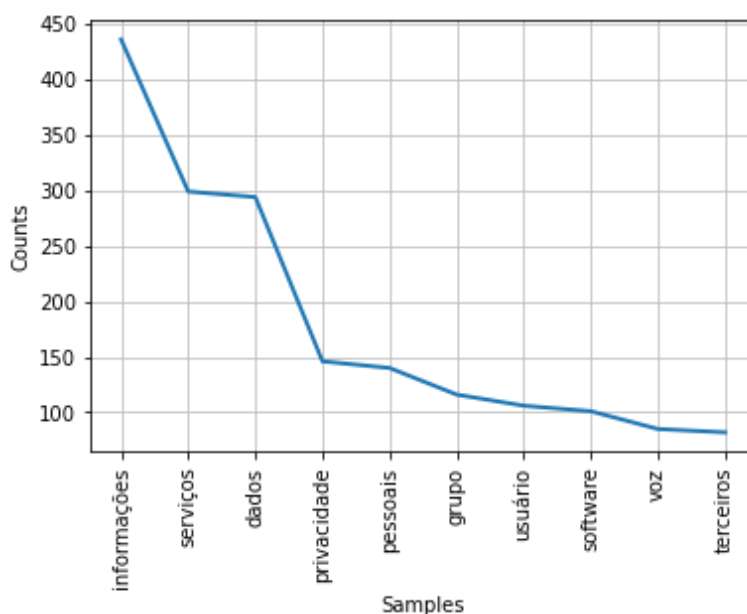


Figura 3.3 – Frequência das Palavras nas Políticas de Privacidade
Fonte: Algoritmo desenvolvido pelo Autor

Nos Termos de Uso dos artefatos encontramos termos voltados a questões contratuais ou fazendo com que a pessoa ceda seus dados em troca da utilização dos serviços que são ofertados pelo artefato. Como dito anteriormente, os Termos de Uso são uma espécie de contrato celebrado entre as partes com o objetivos que tragam vantagens para uma ou ambas as partes. Os textos possuem elementos chave que impedem que os serviços sejam acessados caso não seja aceito suas regras, ou seja, as pessoas ficam sujeitas as normas e regramentos

estipulados na redação dos termos. Diferente das Políticas de Privacidades, os termos considerados mais importantes como ‘dados’, ‘pessoais’ e ‘privacidade’ não estão entre os dez mais frequentes dentro do *corpus* dos Termos de Uso.

Palavras	Frequência
serviços	420
contratante	294
conta	229
usuário	229
conteúdo	216
informações	179
contrato	161
acesso	126
terceiros	121
pagamento	117

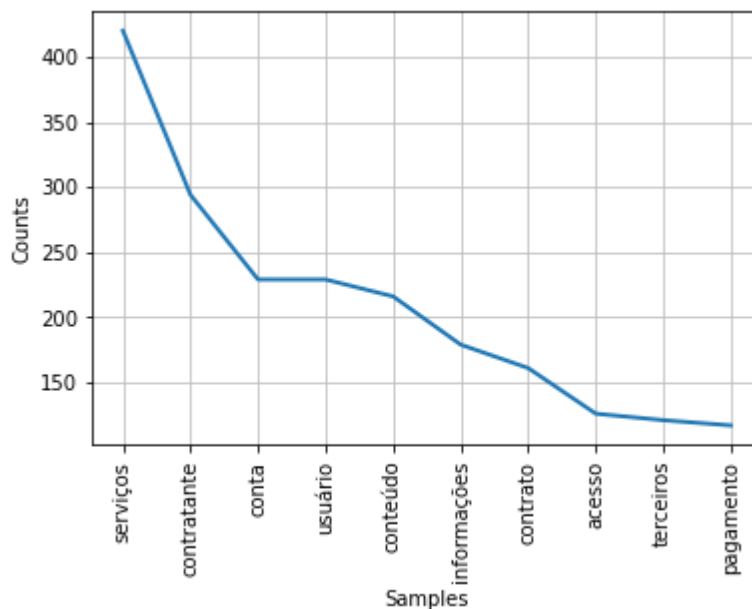


Figura 3.4 – Frequência das Palavras nos Termos de Uso
Fonte: Algoritmo desenvolvido pelo Autor

3.2 As inferências sobre Privacidade na Modelagem de Tópicos

A modelagem de tópicos tem o propósito de sintetizar, a partir de um conjunto de *corpus*, os tópicos mais relevantes e que retratam os temas mais inquiridos no grupo utilizado. Essa é uma tarefa que robotiza a transcrição dos dados em um *corpus*, baseada em métodos estatísticos para avaliar e analisar os dados do documento retornando como saída os tópicos mais relevantes. A modelagem de tópicos foi realizada em *LDA*, uma técnica de *NLP* sólida e que combina métodos probabilísticos. O *corpus* foi processado para o levantamento de ‘x’ tópicos com ‘y’ palavras (*tokens*) cada, dessa forma, o conceito é que ele seja representado por combinações diversificadas e latentes de palavras, ou seja, que não são conhecidas.

No processamento do algoritmo na redação da Lei Geral de Proteção de Dados, foi escolhido inicialmente uma matriz de 3 tópicos com 7 palavras, método que também foi aplicado no *corpus* das Políticas de Privacidade e os Termos de Uso. Esses valores foram utilizados por apresentarem o melhor conjunto de palavras por tópico, sem que houvesse a repetição dos *tokens* agrupados. O algoritmo foi processado diversas vezes até que fosse apresentados o

melhor conjunto e um conjunto de observações como: não verificar a classe gramatical da palavra, não utilizar somente o radical e sim a palavra completamente similar para através da métrica de frequências chegar as conclusões apresentadas.

No modelo *LDA*, os valores no início do resultado (0, 1 e 2) são a representação do quantitativo de tópicos escolhidos, ou seja, 3, já o valor representado por um decimal a esquerda da palavra é a sua relevância dentro do tópico gerado em relação às demais palavras. A partir do resultado gerado, pode-se inferir o que o conjunto de palavras tem a dizer e qual a sua relevância dentro do *corpus*. Na LGPD foram obtidas as seguintes análises:

TABELA 3.1: LGPD – RESULTADOS DO TÓPICO 0

<i>dado</i>	0.057
<i>pessoais</i>	0.035
<i>tratamento</i>	0.029
<i>titular</i>	0.019
<i>proteção</i>	0.018
<i>nacional</i>	0.016
<i>autoridade</i>	0.010

Ao analisar esse tópico existe a proposta de criação da Autoridade Nacional de Proteção de Dados, assim como, confirma que o usuário é o titular dos dados pessoais que venham a ser coletados em território nacional. Maior Relevância: Art. 55-A da Lei N.º 13.709/2018. Esse artigo inicialmente foi VETADO, logo sendo revogado e dessa forma criada à ANPD. Segundo a Mensagem N.º 451 de 14 de agosto de 2018, as razões do veto são: “Os dispositivos incorrem em inconstitucionalidade do processo legislativo, por afronta ao artigo 61, § 1º, II, ‘e’, cumulado com o artigo 37, XIX da Constituição.”

TABELA 3.2: LGPD – RESULTADOS DO TÓPICO 1

<i>dado</i>	0.036
<i>tratamento</i>	0.024
<i>pessoais</i>	0.024
<i>vigência</i>	0.010
<i>segurança</i>	0.008

<i>saúde</i>	0.007
<i>redação</i>	0.007

Ao analisar esse tópico podemos fazer referência a como a LGPD confirma que o usuário é o titular dos dados pessoais, em território nacional, que venham a ser coletados e de que forma deve ser o tratamento dessas informações. São exibidas as palavras ‘vigência, ‘saúde’ e ‘segurança, tratadas na LGPD em Artigos muito bem redigidos e que protegem o direito ao anonimato diante de certas circunstâncias. Maior Relevância: Art. 1º e Art. 3º da Lei N.º 13.709/2018. Pinheiro (2018), destaca que no Art. 1º, “a LGPD surge com o intuito de proteger os direitos fundamentais como privacidade, intimidade, honra e direito de imagem e dignidade”.

TABELA 3.3: LGPD – RESULTADOS DO TÓPICO 2

<i>dado</i>	0.017
<i>controlador</i>	0.015
<i>nacional</i>	0.015
<i>tratamento</i>	0.012
<i>titular</i>	0.012
<i>autoridade</i>	0.011
<i>pessoais</i>	0.009

Ao analisar esse tópico temos a criação dos responsáveis pelo tratamento dos dados dos usuários, o controlador. Existe o direcionamento para a ANPD e o reforço em que os dados pessoais são um bem que pertence ao titular dos dados. Na análise realizada nas Políticas de Privacidade em artefatos *IoT* os resultados foram:

TABELA 3.4: POLÍTICAS DE PRIVACIDADE – RESULTADOS DO TÓPICO 0

<i>informações</i>	0.036
<i>serviços</i>	0.034
<i>dado</i>	0.026
<i>cookie</i>	0.022
<i>dispositivo</i>	0.014
<i>publicidade</i>	0.010

<i>conteúdo</i>	0.010
-----------------	-------

Esse tópico direciona a utilização dos dados e informações do usuário para o compartilhamento com terceiros, não há como avaliar se o usuário permitiu esse trâmite. O tópico chama a atenção por exibir a palavra *cookie* – que são formas de verificar qual o comportamento do usuário no navegador. Este artifício promove uma falsa impressão ao usuário de que a utilização de cookies traz somente benefícios ao usuário.

De acordo com González¹⁹ (2019):

(...) isso também vale para o detalhamento quanto ao tratamento a ser feito, presente na solicitação do consentimento ao titular. Até que surjam evidências do contrário, o titular deve presumir que o controlador realmente está utilizando seus dados pessoais somente para os fins acordados. Caso apareçam evidências do contrário, aí sim, caberá à Autoridade Nacional de Proteção de Dados tomar as devidas providências punitiva.

TABELA 3.5: POLÍTICAS DE PRIVACIDADE – RESULTADOS DO TÓPICO 1

<i>informações</i>	0.034
<i>serviços</i>	0.027
<i>conta</i>	0.027
<i>usuário</i>	0.025
<i>privacidade</i>	0.016
<i>e-mail</i>	0.011
<i>terceiros</i>	0.010

A análise desse tópico demonstra que existe uma preocupação em se manter a segurança e proteção das informações de acesso do usuário. O tópico tem relevância se comparado a redação da LGPD. A palavra ‘conta’ no contexto das Políticas de Privacidade exibe uma forte pressão dos *sites* em coletar de forma automática dados como registros de uso do artefato.

TABELA 3.6: POLÍTICAS DE PRIVACIDADE – RESULTADOS DO TÓPICO 2

<i>dado</i>	0.068
<i>pessoais</i>	0.045
<i>informações</i>	0.038

¹⁹ Fonte: <https://guialgpd.com.br/lgpd-comentada/>, acesso em: 10 jul. 2021.

<i>privacidade</i>	0.019
<i>segurança</i>	0.013
<i>serviços</i>	0.012
<i>processamento</i>	0.008

A análise desse tópico demonstra que existe uma preocupação em se manter a segurança e proteção das informações de acesso do usuário. O tópico tem relevância se comparado a redação da LGPD. Na análise realizada nos Termos de Uso em artefatos *IoT* os resultados encontrados foram:

TABELA 3.7: TERMOS DE USO – RESULTADOS DO TÓPICO 0

<i>segurança</i>	0.018
<i>usuário</i>	0.017
<i>informações</i>	0.010
<i>acesso</i>	0.008
<i>pagamento</i>	0.008
<i>dado</i>	0.008
<i>legais</i>	0.008

Neste tópico observa-se o entendimento para os serviços e produtos encontrados no artefato e que podem ser prestados as pessoas. Direciona também as formas de pagamento pelos serviços adicionais que são prestados pelo artefato.

TABELA 3.8: TERMOS DE USO – RESULTADOS DO TÓPICO 1

<i>informações</i>	0.029
<i>conta</i>	0.024
<i>usuário</i>	0.013
<i>dispositivo</i>	0.013
<i>endereço</i>	0.012
<i>e-mail</i>	0.011
<i>contato</i>	0.010

A palavra ‘informações’ remete a confiabilidade das informações fornecidas pelo usuário, onde pode existir a confirmação dos dados por telefone, e-mail ou outros meios. As tecnologias fornecidas pelo artefato podem, em parte, permitir que o acesso seja realizado em diversos ‘dispositivos’ simultaneamente.

TABELA 3.9: TERMOS DE USO – RESULTADOS DO TÓPICO 2

<i>dado</i>	0.065
<i>informações</i>	0.057
<i>serviços</i>	0.045
<i>pessoais</i>	0.038
<i>privacidade</i>	0.018
<i>terceiros</i>	0.018
<i>cookie</i>	0.017

A utilização de *cookies*, que remete a verificar qual o comportamento do usuário na utilização do navegador, suas preferências e os *sites* que mais visita. Além disso, podem ser realizadas alterações na política de privacidade sem que haja, previamente uma comunicação ao usuário, isso pode causar transtornos e inconvenientes como exposição indevida da vida pessoal dos membros dos artefatos.

3.2.1 Representação Gráfica da Modelagem de Tópicos²⁰

²⁰ De acordo com a documentação oficial reproduzida nesta nota o LDAvis é uma visualização interativa baseada na web de tópicos estimados usando LDA. Ele fornece uma visão global dos tópicos (e como eles diferem uns dos outros), enquanto ao mesmo tempo permite uma inspeção profunda dos termos mais associados a cada tópico individual. O pacote extrai informações de um modelo de tópico LDA ajustado para informar uma visualização interativa baseada na web. A visualização possui duas peças básicas. O painel esquerdo visualiza os tópicos como círculos no plano bidimensional cujos centros são determinados calculando a divergência de Jensen – Shannon entre os tópicos e, em seguida, usando escala multidimensional para projetar as distâncias entre tópicos em duas dimensões. A prevalência geral de cada tópico é codificada usando as áreas dos círculos. O painel direito representa um gráfico de barras horizontais cujas barras representam os termos individuais mais úteis para interpretar o tópico selecionado atualmente à esquerda. Um par de barras sobrepostas representa a frequência de todo o corpus de um determinado termo, bem como a frequência de tópico específico do termo. O controle deslizante λ permite classificar os termos de acordo com a relevância do termo. Por padrão, os termos de um tópico são classificados em ordem decrescente de acordo com sua probabilidade específica do tópico ($\lambda = 1$). Mover o controle deslizante permite ajustar a classificação dos termos com base em muitos critérios discriminatórios (ou "relevantes") para o tópico específico. Disponível em: <http://bl.ocks.org/AlessandraSozzi/raw/ce1ace56e4aed6f2d614ae2243aab5a5/#topic=0&lambda=0.61&term=>>, Acesso em: 10 ago. 2021

Em uma síntese geral, observamos que o algoritmo retornou itens considerados com grande relevância na redação Lei, como a criação da ANPD, seus cargos, a abrangência em território nacional, incluindo a proteção dos dados dos usuários. A Figura 3.5 , multidimensional, exibe o grupo das 30 palavras mais salientes dentro da redação da LGPD, proporcionalidade e espaçamento entre os termos:

O Gráfico reflete o tratamento dos dados, menciona os mecanismos e hipóteses em que podem ser utilizados e indica a solicitação do consentimento para tal.

Segundo Maciel (2019, p. 29):

O consentimento sempre é visto, por muitos, como a panaceia para tratamento de dados pessoais. Com consentimento pode-se quase tudo. Bem verdade, o consentimento é a autorização expressa dada pelo titular ao controlador para que ele possa tratar os dados da forma desejada.

A uma referência a criação da ANPD e dos cargos que nela tem vital importância para manter a segurança das informações.

Destaca a criação da ANPD, seus cargos e direciona ao indicativo de procedimentos para o tráfego de dados do usuário. A palavra ‘impacto’ dá ao conjunto de tokens alguma referência ao vazamento das informações e como isso deve ser administrado. Existe a referência ao ‘controlador’, cargo de grande importância na ANPD, pois compete a ele decidir sobre o tratamento nos dados dos usuários.

São exibidas palavras voltadas aos privilégios e os direitos que os usuários têm sobre seus dados. As palavras como ‘dado’ e ‘tratamento’ são o topo do ranking, desta forma, podemos inferir que o gráfico fornece informações sólidas para o comparativo com a LGPD. São exibidas as palavras ‘anonimização’, ‘saúde’ e ‘encarregado’, tratadas na LGPD com Artigos muito bem redigidos e que protegem o direito ao anonimato diante de certas circunstâncias, sendo o ‘encarregado’ a pessoa responsável pela comunicação entre o controlador e o operador.

Segundo Maciel (2019, p. 22):

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. Encarregado (DPO): pessoa indicada pelo controlador para atuar como canal de comunicação entre o controlador, os titulares dos dados e a Autoridade Nacional de Proteção de Dados (MP 869).

As definições que direcionam aos princípios da LGPD e apontamento para possíveis multas em caso de não cumprimento da legislação pelas entidades. Novamente, faz-se à referência a criação da ANPD e a indicação do cargo de ‘diretor’, membro Conselho Diretor da ANPD.

De acordo com Pinheiro (2018, p. 109):

A imputação de sanções administrativas faz com que os entes responsáveis pelo tratamento de dados pessoais atentem-se à garantia da segurança das informações que estão utilizando. Dessa forma, observa-se que a LGPD busca estimular a aplicação de seus dispositivos em caráter preventivo. As sanções vão desde advertências até imputação de multa simples – que pode chegar a 2 por cento do faturamento, cujo valor fica limitado a um total de 50 milhões de reais, além da suspensão das atividades relativas ao banco de dados.

Além dos termos já expostos nos anteriormente, existem referências às palavras ‘contratuais’ e ‘transferência’, que fazem referência ao consentimento realizado por cláusulas de contrato e o tráfego internacional de dados, caso seja necessário.

Segundo Vainzof (2018, p. 21):

(...) há anos se discutia no Brasil um marco legal em proteção de dados pessoais, diante da sua relevância para o nosso país, principalmente para trazer maior segurança jurídica mediante a harmonização de conceitos, elevando a proteção aos direitos individuais das pessoas e ao fomento da economia digital, bem como, com um nível de legislação compatível com outros países, da facilitação ao fluxo de transferência internacional de dados.

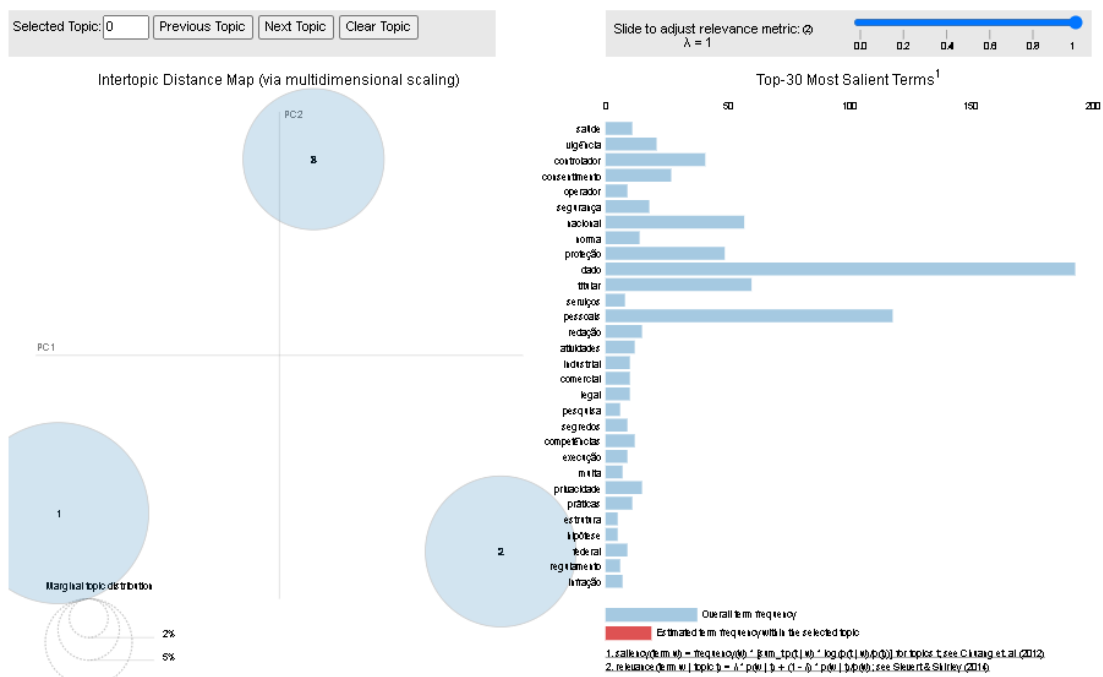


Figura 3.5 – Gráfico Multidimensional da LGPD
 Fonte: Algoritmo desenvolvido pelo Autor

Nas Políticas de Privacidade existe um grande direcionamento pelo uso da palavra ‘publicidade’, que leva seu direcionamento para a utilização dos dados do usuário e a comercialização de produtos e serviços de parceiros. Observa-se o entendimento para os serviços e produtos de uma empresa que podem ser prestados a um usuário. Direciona também a faixa etária e formas de pagamento pelos seus serviços.

Com base no gráfico multidimensional, traz os *tokens* voltados para a coleta de dados do usuário com a finalidade da sua utilização para fins diversos. As palavras usuário e acesso, reforçam essa premissa. O compartilhamento dos dados do usuário com outras empresas e parceiros terceirizados estão compreendidos quando da utilização das palavras terceiros, coletados e direitos.

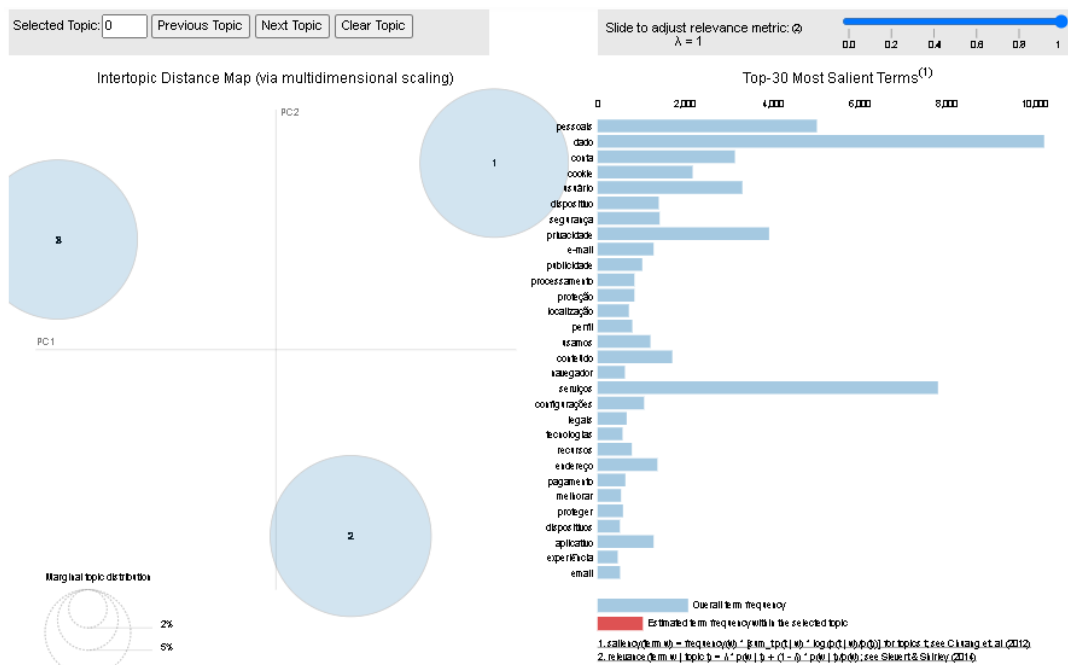


Figura 3.6 – Gráfico Multidimensional das Políticas de Privacidade
 Fonte: Algoritmo desenvolvido pelo Autor

Em relação aos Termos de Uso a análise nos traz a falta impressão de que no seu aceite as pessoas estão protegidas contra qualquer vazamento ou falha na segurança. O que pode ser dito é que esse ‘contrato’, forma pelo qual podemos definir os Termos de Uso, permitem as organizações que produzem os artefatos *IoT* que usem seus dados para criar estratégias de ‘publicidade’, verificar como é o comportamento diário dos portadores da tecnologia, conhecendo assim suas rotinas. Um fato interessante é que, assim como as Políticas de

Privacidade, os Termos de Uso, obrigatoriamente, devem ser aceitos pelos usuários antes do uso dos artefatos *IoT*.

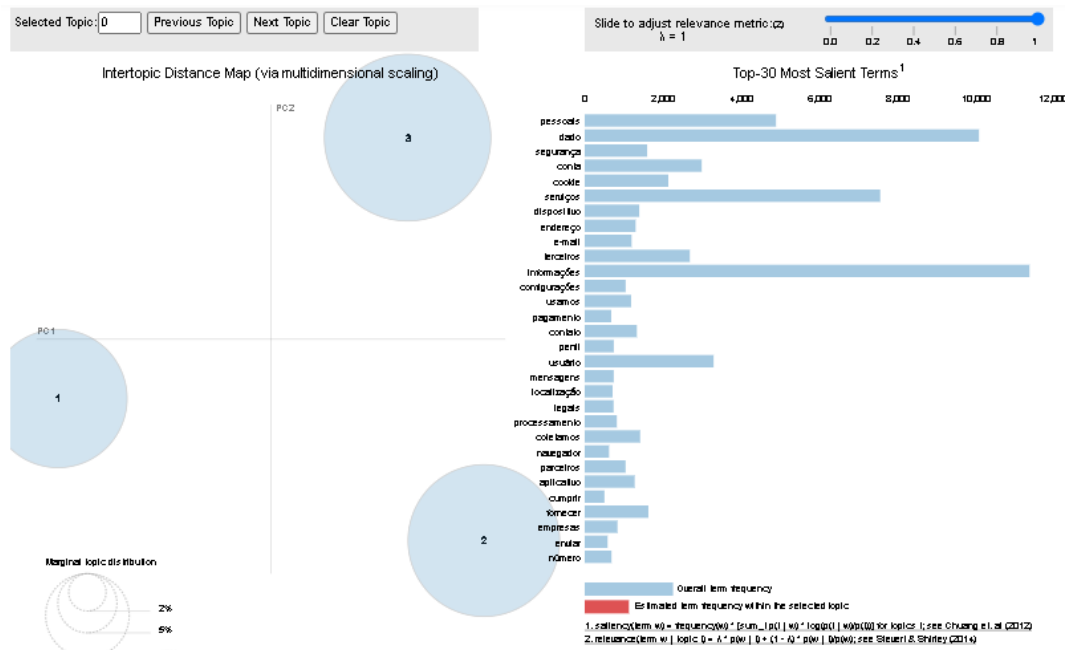


Figura 3.7 – Gráfico Multidimensional dos Termos de Uso
Fonte: Algoritmo desenvolvido pelo Autor

De acordo com Eduardo Magrani ²¹(2017), “*Os termos de serviços de plataformas digitais são documentos longos, repletos de termos técnicos e legais, às vezes vagos e pouco claros, e que poucas pessoas se dão ao trabalho de ler.*” (MAGRANI, 2017)

No Mapa de Calor exibido na próxima seção, podemos facilmente interpretar quais Políticas de Privacidade e Termos de Uso possuem maior similaridade com os Capítulos da LGPD - Lei Geral de Proteção de Dados. A legenda mostra que, quanto mais clara for a cor, menor é a similaridade do Capítulo comparado.

3.3 Simplificando a visualização: Mapa de Calor

²¹ MAGRANI, Eduardo. Termos de serviço indicam desrespeito a direitos humanos em plataformas *online*. Disponível em: <http://eduardomagrani.com/termos-de-servico-indicam-desrespeito-direitos-humanos-em-plataformas-online/>. Acesso em: 10 ago. 2021.

Especialmente ao processar os documentos para *NLP*, geralmente os vetores são sempre positivos e o resultado da aplicação de cossenos semelhantes está entre 0 e 1. De forma geral, o entendimento parte da premissa que existem dois vetores, vetor **a** e vetor **b**. O vetor **b** é projetado no vetor **a**. O tamanho da sombra em dependência ao vetor **a** indica a sua similaridade. Quando o valor da projeção é igual a 1, indica que o vetor **b** está exatamente no topo do vetor **a**, ou seja, os dois são paralelos e iguais.

TABELA 3.10: POLÍTICAS DE PRIVACIDADE E CAPÍTULOS DA LGPD

Capitulo	cap1	cap2	cap3	cap4	cap5	cap6	cap7	cap8	cap9	cap10
Política										
politicass/politica_privacidade_amazon_alexa.txt	0.047008	0.062535	0.067245	0.058654	0.033945	0.046164	0.064483	0.021610	0.057634	0.026050
politicass/politica_privacidade_apple_watch_watchos7.txt	0.131528	0.152913	0.120525	0.123763	0.089482	0.113369	0.122064	0.060800	0.084819	0.079809
politicass/politica_privacidade_chevrolet_onstar.txt	0.447613	0.558293	0.555610	0.452441	0.273921	0.469406	0.414552	0.180993	0.338221	0.267572
politicass/politica_privacidade_google_nest.txt	0.216883	0.227900	0.231045	0.220994	0.123140	0.215554	0.208267	0.082030	0.179609	0.126153
politicass/politica_privacidade_hyundai_blue_link.txt	0.653858	0.668910	0.615457	0.619867	0.342895	0.626593	0.570683	0.259755	0.500184	0.384716
politicass/politica_privacidade_lg_smart_tv.txt	0.124839	0.186187	0.156135	0.168443	0.102673	0.135087	0.180703	0.050974	0.146429	0.062732
politicass/politica_privacidade_nintendo_switch.txt	0.103028	0.166129	0.104818	0.152984	0.095600	0.099508	0.153843	0.039543	0.106446	0.051151
politicass/politica_privacidade_pagseguro_maquina_de_cartao.txt	0.346154	0.373619	0.352208	0.346209	0.189399	0.347664	0.319780	0.181627	0.269773	0.209724
politicass/politica_privacidade_sony_playstation.txt	0.098266	0.147274	0.121947	0.147889	0.087652	0.108248	0.153779	0.039801	0.095017	0.052843

TABELA 3.11: TERMOS DE USO E CAPÍTULOS DA LGPD

Capitulo	cap1	cap2	cap3	cap4	cap5	cap6	cap7	cap8	cap9	cap10
Termos										
termos\termo_de_uso_amazon_alexa.txt	0.067947	0.095927	0.083637	0.076232	0.055158	0.064648	0.086730	0.043915	0.055592	0.042554
termos\termo_de_uso_apple_watch_watchos7.txt	0.037603	0.051831	0.043567	0.042075	0.029164	0.036683	0.051319	0.031401	0.041402	0.026761
termos\termo_de_uso_chevrolet_onstar.txt	0.158221	0.188729	0.179813	0.178088	0.103150	0.156957	0.143915	0.092371	0.108360	0.132303
termos\termo_de_uso_google_nest.txt	0.123979	0.142218	0.119554	0.118935	0.067648	0.109260	0.118740	0.052549	0.085830	0.074892
termos\termo_de_uso_hyundai_blue_link.txt	0.057940	0.069211	0.062056	0.066138	0.043830	0.041525	0.057780	0.034737	0.040988	0.039832
termos\termo_de_uso_lg_smart_tv.txt	0.087748	0.124082	0.077454	0.094322	0.066685	0.073026	0.091525	0.054042	0.054837	0.045232
termos\termo_de_uso_nintendo_switch.txt	0.095012	0.124550	0.084690	0.102846	0.070766	0.076202	0.100559	0.065886	0.068344	0.052876
termos\termo_de_uso_pagseguro_maquina_de_cartao.txt	0.081302	0.097774	0.104085	0.090076	0.086795	0.074453	0.078556	0.053382	0.068458	0.056984
termos\termo_de_uso_sony_playstation.txt	0.079662	0.112870	0.086988	0.096415	0.075356	0.071736	0.091291	0.068312	0.058295	0.058978

As análises exibidas nas Tabela 3.2 e Tabela 3.2, resultado do processamento do algoritmo criado, indica a comparação realizada entre as nove Políticas de Privacidade e Termos de uso dos artefatos *IoT* em relação a cada um dos 10 Capítulos da Lei Geral de Proteção de Dados. Para facilitar a leitura foi explorada a utilização de um Mapa de Calor, dando ao resultado uma expressiva facilidade de interpretação.

O Mapa de Calor é um relatório gerado por ferramentas que usam a cor como referência para promover o entendimento das informações. O Mapa auxilia a compreender melhor o arranjo de dados com base nos pontos com os quais existe maior similaridade ou maior frequência. Desta forma, a compreensão é mais precisa e pode-se verificar quais partes de um texto merecem uma análise mais criteriosa, ou seja, para que trecho ou elementos ineficientes possam ser claramente identificados.

No Mapa de Calor exibido abaixo, podemos facilmente interpretar quais Políticas de Privacidade possuem maior similaridade com os Capítulos da LGPD - Lei Geral de Proteção de Dados. A legenda mostra que, quanto mais clara for a cor, menor é a similaridade do Capítulo comparado.

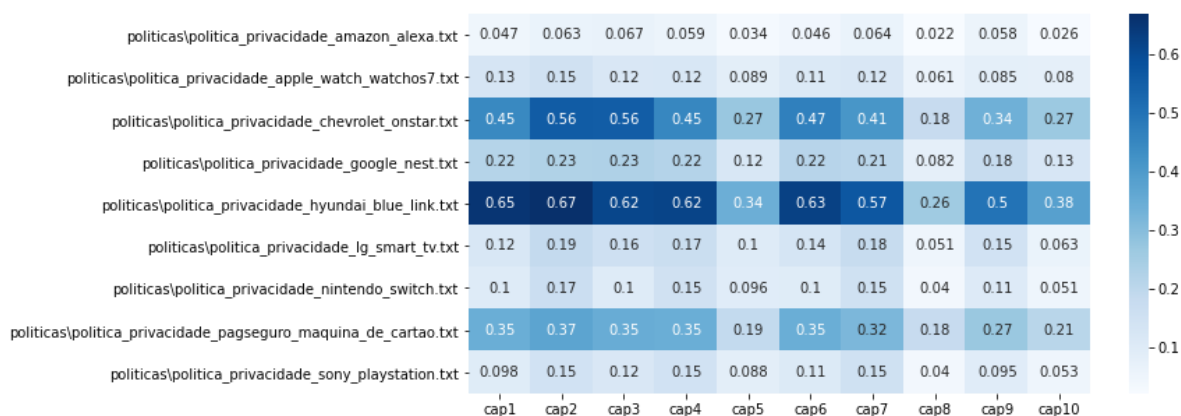


Figura 3.8 – Mapa de Calor: Políticas de Privacidade X LGPD

Fonte: Algoritmo desenvolvido pelo Autor

Dentre as Políticas de Privacidade que possuem maior similaridade com a LGPD, podemos destacar:

- Artefato *Hyunday Blue Link*, que possui valores de similaridades entre 0.26 e 0.67;
- Artefato *Chevrolet OnStar*, que possui valores de similaridades entre 0.18 e 0.56;
- Artefato *PagSeguro – Máquina de Cartão*, que possui valores de similaridades entre 0.18 e 0.37.

Em outros artefatos que não estão adequados à legislação brasileira estão:

- Artefato *Amazon Alexa*, que possui valores de similaridades entre 0.02 e 0.06;

- Artefato *Nintendo Switch*, que possui valores de similaridades entre 0.04 e 0.17;
- Artefato *Apple Watch*, que possui valores de similaridades entre 0.08 e 0.15.

Em relação ao Termos de Uso os resultados foram:

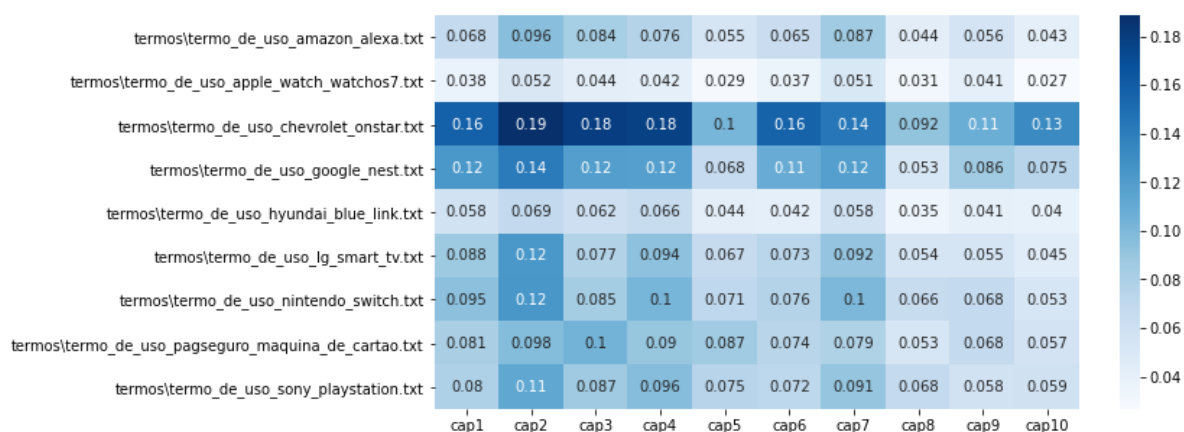


Figura 3.9 – Mapa de Calor: Termos de Uso X LGPD

Fonte: Algoritmo desenvolvido pelo Autor

O único resultado que apresenta maior similaridade com a redação da lei está *Chevrolet OnStar* e *Google Nest*, enquanto grande parte dos Termos de Uso estão completamente fora do que se espera para a utilização com privacidade dos artefatos.

CONCLUSÃO

Neste trabalho foram analisadas a Lei Geral de Proteção de Dados, as Políticas de Privacidade e Termos de Uso de artefatos *IoT*. Os resultados obtidos fazem parte de um conjunto multidisciplinar entre Direito e Tecnologia, onde através da investigação do *corpus* de documentos, foram encontradas a maior frequência de palavras, o grupo de palavras mais relevante entre os documentos – modelagem de tópicos – e a similaridade entre os Capítulos da LGPD e o conjunto de documentos criado. Quando analisada, a LGPD, criou uma matriz que serviu como base para inferir o percentual de palavras mais frequentes das listas nas Políticas de Privacidade e Termos de Uso, uma observação distante e não qualitativa dos textos utilizados no processamento.

Quando comparamos os resultados obtidos sobre a Lei Geral de Proteção de Dados e as Políticas de Privacidade, o percentual de similaridade das duas listas é de apenas 20%, ou

seja, não há similaridade relevante entre a redação da lei e o *corpus* de artefatos *IoT*. Ao submetermos o experimento entre a LGPD e os Termos de Uso o resultado é de 0% de similaridade, desta forma, podemos inferir que no primeiro experimento – frequência de palavras – ambos, Políticas de Privacidade e Termos de Uso de artefatos *IoT* não estão de acordo com a legislação brasileira. Para efeitos demonstrativo, ao comparar as Políticas de Privacidade com seus Termos de Uso o resultado de similaridade é de 40%, o que nos remete a textos que não possuem grande coerência entre si.

Na segunda fase do experimento utilizamos a Modelagem de Tópicos, desta forma criando 3 tópicos com 7 palavras para cada um. Na relação entre LGPD com Políticas de Privacidade a modelagem trouxe os seguintes resultados: Tópico 0, o índice de similaridade de 30,1%, no Tópico 1, a similaridade foi de 35,4% e o Tópico 2, com similaridade de 34,6% dos *tokens*. Sendo assim, podemos afirmar que os arranjos dos Tópicos não tornam nenhum deles um destaque para confirmar a similaridade entre os textos da LGPD e o *corpus* das Políticas de Privacidade, reafirmando que os textos não estão adequados à redação lei.

Dando continuidade a Modelagem de Tópicos, usando a relação LGPD com Termos de Uso, os resultados foram: Tópico 0, o índice de similaridade de 27,3%, no Tópico 1, a similaridade foi de 33% e o Tópico 2 com similaridade de 39,7% dos *tokens*. Com esses resultados podemos replicar a conclusão dada na relação LGPD e Políticas de Privacidade, ou seja, o *corpus* de ambos é inadequado ao que rege a legislação vigente.

Na terceira e última parte do experimento foi realizada a relação de cada um dos documentos de Política de Privacidade e Termos de Uso com os Capítulos da LGPD separadamente, formando um Mapa de Calor através do método de similaridades por cosseno, sendo assim, quanto mais próximo de 1, mais próximo da redação do Capítulo da LGPD esta aquele texto.

Os artefatos que mais demonstram semelhança entre os documentos de Política de Privacidade e Capítulos da LGPD foram: Artefato *Hyunday Blue Link*, que possui valores de similaridades entre 0.26 e 0.67; Artefato *Chevrolet OnStar*, que possui valores de similaridades entre 0.18 e 0.56; Artefato *PagSeguro – Máquina de Cartão*, que possui valores de similaridades entre 0.18 e 0.37.

Dentre os que possuem baixíssima similaridade temos: Artefato *Amazon Alexa*, que possui valores de similaridades entre 0.02 e 0.06; Artefato *Nintendo Switch*, que possui valores de similaridades entre 0.04 e 0.17; Artefato *Apple Watch*, que possui valores de similaridades entre 0.08 e 0.15.

Entre os Termos de Uso podemos afirmar que os que demonstram algum resultado expressivo são: *Chevrolet OnStar* e *Google Nest*.

Diante de todos os fatos apresentados neste experimento, conclui-se que a redação de Políticas de Privacidade e Termos de Uso de Dispositivos *IoT*, feita sem uma análise qualitativa, não possuem similaridade com a LGPD que começa a aplicar sanções no mês de agosto de 2021. É necessário que as empresas que fabricam os artefatos estejam atentas a essa demanda e façam as devidas correções o mais breve possível. Para trabalhos futuros sugiro a reavaliação do *corpus*, uma abordagem qualitativa dos textos, com a leitura integral dos documentos e verificar possíveis adequações à legislação vigente.

REFERÊNCIAS

- ALMEIDA, H. Internet das Coisas: Tudo conectado. 2015. Disponível em: <https://www.sbc.org.br/images/flippingbook/computacaobrasil/computa_29_pdf/comp_brasil_2015_4.pdf>. Acesso em: 13 fev. 2021.
- ALVES, B. H. B.; PINTO, G. R.; OLIVEIRA, R. S. de. A proteção de dados pessoais de crianças consumidoras de smart toys: análise da legislação brasileira (lei nº 13.709/2018) sobre a hello barbie doll. Santa Maria, RS: UFSM - Universidade Federal de Santa Maria, 2019.
- ANDERSON, C. A Cauda Longa. Rio de Janeiro, RJ: Editora Campus, 2006.
- ASHTON, K. Uma introdução à Internet da Coisas (IoT). 1999. Disponível em: <https://www.cisco.com/c/dam/global/pt_br/assets/brand/iot/iot/pdfs/lopez_research_an_introduction_to_iiot_102413_final_portuguese.pdf> Acesso em: 9 fev. 2021.
- ASHTON, K. That ‘internet of things’ thing. RFIJ Journal. [S.l.]: RFIJ Journal, 2009.
- BIONI, B.; FAVARO, I.; RIELLI, M. Por que a transferência internacional de dados tem que ser segura?2020. Disponível em: <<https://www.observatorioprivacidade.com.br/2020/10/05/por-que-a-transferencia-internacional-de-dados-tem-que-ser-segura/>>. Acesso em: 01 abr. 2021.
- BLUM, R. O. et al. Melhores práticas de Governança e Conformidade com a LGPD. São Paulo: Opice Blum, 2020.
- BOBBIO, N. A Era dos Direitos. Rio de Janeiro: Campus, 1992.
- BORTOLOSSI, H. J. Vexia. Adequação a LGPD impulsiona procura por prestação de serviços para adequação a LEI Nº 13.709/18. 2019. Disponível em: <<https://rdstation-static.s3.amazonaws.com/cms/%2Ffiles/%2F28555>>

%2F1581450471Ebook_LGPD_-_O_que_e_como_adequar_sua_empresa.pdf.> Acesso em: 21 abr. 2021.

BRASIL. Constituição Federal. Brasília, DF: Centro Gráfico, 1988. BRASIL. Lei de Acesso à Informação. Lei 12.527. 2011. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm>. Acesso em: 5 fev. 2021.

BRASIL. Lei de Crimes Informáticos. Lei 12.737. 2012. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/112737.htm>. Acesso em: 9 fev. 2021.

BRASIL. Marco Civil da Internet. Lei 12.965. 2014. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm>. Acesso em: 4 fev. 2021.

BRASIL. Lei Geral de Proteção de Dados Pessoais (LGPD). Lei 13.709. 2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 2 fev. 2021.

BURKE, B. Top Strategic Technology Trends for 2021. [S.l.]: Gartner, 2021.

CASSANTI, M. d. O. Crimes Virtuais, Vítimas Reais. Rio de Janeiro: Brasport, 2014. v. 1. ed.

CASTELLS, M. The Internet Galaxy: Reflections on Internet, Business and Society. Oxford University: Pearson Education do Brasil, 2001.

CATE, F. H. Privacy in the Information Age. Washington: [s.n.], 1997.

COELHO, I. M. B.; GONET, P. G. Curso de direito constitucional. São Paulo: Saraiva, 2008. v. 3 ed. rev. atual.

CHAVES, L. F. P. LGPD - Lei Geral De Proteção De Dados - Comentada - Coordenação Maldonado, Viviane Nóbrega e Blum, Renato Opice. São Paulo: Fiscosoft - Thomson Reuters, 2019. v. 2.

CHAVES, M. M. Direito Digital Aplicado 3.0. São Paulo: Editora Revista dos Tribunais, 2018.

CHEN, H.; CHIANG, R. H.; STOREY, V. C. Business Intelligence and Analytics: From Big Data to Big Impact. [S.l.]: MIS Quarterly, 2012.

CHEN, S.-H. Big Data in Computational Social Science and Humanities. [S.l.]: Springer, 2018.

CRISTÓVAM, J. S. d. S. Colisões entre princípios constitucionais. Curitiba: Juruá, 2006. v. 12.

DONEDA, D. JOTA: O que está em jogo com a nova Autoridade Nacional de Proteção de Dados. 2018. Disponível em: <<https://www.jota.info/opiniao-e-analise/artigos/o-que-esta-em-jogo-com-a-nova-autoridade-nacional-de-protecao-de-dados-13082018>>. Acesso em: 22 fev. 2021.

DONEDA, D. Danilo Doneda representará a Câmara dos Deputados no Conselho Nacional da ANPD. 2019. Disponível em: <<https://teletime.com.br/17/10/2019/danilo-doneda-representara-a-camara-dos-deputados-no-conselho-nacional-da-anpd/>>

>. Acesso em: 4 mar. 2021.

DONEDA, D. Por dentro da LGPD. 2019. Disponível em: <<https://www.consumidormoderno.com.br/2019/04/25/dentro-lgpd-coautor-lei-faz-projecoes-para-o-brasil/>>. Acesso em: 19 fev. 2021.

FACHINI, Tiago. Direitos e garantias fundamentais: conceito e características. 2021. Disponível em: https://www.projuris.com.br/o-que-sao-direitos-fundamentais#direitos_fundamentais_na_constituicao_federal. Acesso em 13 mai. 2021

FILHO, D. R. Lei de proteção de dados pessoais aproxima o Brasil dos países civilizados. 2019. Disponível em: <<http://www.gandramartins.adv.br/project/ives-gandra/public/uploads/2019/02/21/22f3605085417p.pdf>>. Acesso em: 9 fev. 2021.

FILHO, E. T. Marco Civil da Internet: uma lei sem conteúdo normativo. 2016. Disponível em: <<https://www.revistas.usp.br/eav/article/view/115093>>. Acesso em: 4 fev. 2021.

FREITAS, Carla. Como elaborar uma política de privacidade aderente à LGPD? Disponível em: <<https://www.serpro.gov.br/lgpd/noticias/2019/elabora-politica-privacidade-aderente-lgpd-dados-pessoais/#sdfootnote1sym>>. Acesso em: 27 jul. 2021.

GATES, B. A Estrada para o Futuro. [S.l.]: Companhia das Letras, 1995.

GEUS, P. L. d.; MELO, A. G. “Coisinhas” da Internet das Coisas. 2015. Disponível em: <https://www.sbc.org.br/images/flippingbook/computacaobrasil/computa_29_pdf/comp_brasil_2015_4.pdf>. Acesso em: 13 fev. 2021.

GOMES, M. Privacidade diferencial e anonimização. 2019. Disponível em: <<https://privacytech.com.br/artigos/privacidade-diferencial-e-anonimizacao,319897.jhtml>>. Acesso em: 31 mar. 2021.

GONZALEZ, M. LGPD Comentada. 2019. Disponível em: <<https://guialgpd.com.br/lgpd-comentada/>>. Acesso em: 02 abr. 2021.

GONZÁLEZ, M. LGPD Comentada. 2019. Disponível em: <<https://guialgpd.com.br/lgpd-comentada/>>. Acesso em: 10 abr. 2021.

GSTI. O que é Banco de Dados? 2021. Disponível em: <<https://www.portalgsti.com.br/banco-de-dados/sobre/>>. Acesso em: 19 fev. 2021.

IBSG, C. I. B. S. G. A Internet das Coisas. Como a próxima evolução da Internet está mudando tudo. [S.l.: s.n.], 2011.

INTEL. Introdução à Big Data Como avançar com uma implantação bem-sucedida. 2014. Disponível em: <<https://www.intel.com.br/content/dam/www/public/lar/br/pt/documents/articles/e7-big-data-planning-guide-webready-por.pdf>>. Acesso em: 12 fev. 2021.

JIMENE, C. d. V. LGPD - Lei Geral De Proteção De Dados - Comentada - Coordenação Maldonado, Viviane Nóbrega e Blum, Renato Opice. São Paulo: Fiscosoft - Thomson Reuters, 2019. v. 2.

KAMPAKIS, S. The Decision Maker's Handbook for Data Science: a Guide for Non Technical Executives, Managers and Founders. London: Apress, 2020.

KNOPIK, M. J. V. Políticas públicas e a Internet das Coisas: sugestões para a operacionalização jurídica das cidades hiper conectadas. Revista Jurídica da Escola Superior de Advocacia da OAB-PR, v. 3, n. 3, dez. 2018.

KUROSE, J. F.; ROSS, K. W. Redes de computadores e a Internet: uma abordagem top-down. Tradução. São Paulo: Pearson Education do Brasil, 2013. v. 6ª Edição.

LACERDA, Flávia. Análise ex ante do plano nacional de internet das coisas (*IOT*): Ambiente Cidades Inteligentes. Texto para discussão / Instituto Serzedello Corrêa. – Brasília: ISC/TCU, 2020-.

LOGAREZZI, L. Guia prático da lei de acesso à informação [livro eletrônico]. São Paulo: [s.n.], 2016.

MAGRANI, Eduardo. Entre dados e robôs : ética e privacidade na era da hiperconectividade / Eduardo Magrani. - Rio de Janeiro : Konrad Adenauer Stiftung, 2018.

MACIEL, R. F. Manual Prático sobre a Lei Geral de Proteção de Dados Pessoais. .Goiânia – GO: RM Digital Education, 2019. v. 1ª Edição.

MAGRANI, E. O modelo de consentimento falhou. 2017. Disponível em: <<http://eduardomagrani.com/o-modelo-de-consentimento-falhou/>>. Acesso em: 24 fev. 2021.

MAGRANI, E. A internet das coisas. Rio de Janeiro: FGV Editora, 2018. MENDES, G. F.;

MENDES, M. E. G.; VIEIRA, N. B. Os crimes cibernéticos no ordenamento jurídico brasileiro e a necessidade de legislação específica. 2012. Disponível em: <<http://www.gcpadvogados.com.br/artigos/os-crimes-ciberneticos-no-ordenamento-juridico-brasileiro-e-a-necessidade-de-legislacao-especifica-2>>. Acesso em: 9 fev. 2021.

MOTA, F. da. Proteção de dados pessoais é a evolução da privacidade. 2019. Disponível em: <<https://www.serpro.gov.br/lgpd/noticias/protecao-dados-evolucao-privacidade>>. Acesso em: 30 mar. 2021.

NEWMAN, L. H. Wikileaks dump reveals how the CIA can track your exact location. 2017. Disponível em: <<https://www.wired.co.uk/article/wikileaks-cia-tracking>>. Acesso em: 24 mar. 2021.

NOIRET, S. História Pública Digital. [S.l.]: Liinc em Revista, 2015. v. 11. PINHEIRO, P. P. Direito Digital. São Paulo: Saraiva, 2009. v. 3ª Ed.

PINHEIRO, P. P. Proteção Dados Pessoais: Comentários à Lei N.º13.709/2018. São Paulo: Saraiva Educação, 2018. POLETIZE. INCISO X – INTIMIDADE. 2019. Disponível em: <<https://www.politize.com.br/artigo-5/intimidade>>. Acesso em: 13 fev. 2021.

PORTO, F.; ZIVIANI, A.; OGASAWARA, E. O. Big data: um novo desafio à nossa porta. 2015. Disponível em:

<https://www.sbc.org.br/images/flippingbook/computacaobrasil/computa_29_pdf/comp_brasil_2015_4.pdf>. Acesso em: 13 fev. 2021. POSNER, R. A. A economia da justiça. São Paulo: WMF Martins Fontes, 2010.

PRAHALAD, C. K. The fortune at the bottom of the pyramid. Kindle: Companhia das Letras, 2009. Revised and updated 5th anniversary edition eradicating poverty through profits.

REIDY, Stefan. How IoT and Big data are accelerating the supply chain industry? Disponível em: <<https://arviem.com/iot-and-big-data-supply-chain-industry/>>. Acesso em: 26 jul 2021.

REUTERS, T. WHITEPAPER. Introdução ao Direito Digital: Tudo que o advogado precisa saber. 2019. Disponível em: <<https://www.thomsonreuters.com.br/content/dam/openweb/documents/pdf/Brazil/white-paper/pdf-introducao-ao-direito-digital.pdf>>. Acesso em: 05 fev. 2021.

RODOTÁ, S. A vida na sociedade de vigilância: a privacidade hoje. Rio de Janeiro: Renovar, 2008.

SANCHES, P.; LAMOSA, E. O Direito à Privacidade dos Dados na Seara do Direito das Famílias. 2020. Disponível em: <<https://ibdfam.org.br/artigos>>. Acesso em: 31 mar. 2021.

SCHRAPPE, C.; BECKER, D. JOTA: Consentimento e o consentimento na LGPD. 2021. Disponível em: <<https://www.jota.info/opiniao-e-analise/artigos/consentimento-e-o-consentimento-na-lgpd-03012021>>. Acesso em: 22 fev. 2021.

SERPRO. COMO CUMPRIR A LGPD? 2018. Disponível em: <<https://www.serpro.gov.br/lgpd/empresa/como-cumprir-a-lgpd>>. Acesso em: 01 abr. 2021.

SERPRO. Conheça os princípios e as bases legais que dão suporte à Lei Geral de Proteção de Dados Pessoais. 2018. Disponível em: <<https://www.serpro.gov.br/lgpd/menu/tratamento-dos-dados/principios-da-lgpd>>. Acesso em: 22 fev. 2021.

SERPRO. Checar os dados, de forma responsável, ajuda a proteger os cidadãos. 2019. Disponível em: <<https://www.serpro.gov.br/lgpd/noticias/che-car-dados-cidadaos-lgpd-lei-geral-de-protecao-de-dados-pessoais>>. Acesso em: 30 mar. 2021.

SHIKIDA, C. D.; MONASTERIO, L.; NERY, P. Guia brasileiro de análise de dados: armadilhas & soluções. Brasília: Enap, 2021.

SILVA, M. et al. Guia para o programa de conformidade LGPD. [S.l.]: Marcelo Silva / W2CONN, 2020. SOLOVE, D. J. The Digital Person: Technology And Privacy In The Information Age. [S.l.: s.n.], 2004.

SOLOVE, D. J. Understanding privacy. Cambridge, London: Harvard University Press, 2008. Kindle edition.

SOUZA, D. Estatísticas, Letras, Palavras e Períodos. 2009. Disponível em: <<http://www.cdme.im-uff.mat.br/lpp/lpp-html/lpp-br.html>>. Acesso em: 01 abr. 2021.

SCHULTZ, Felix. Política de Privacidade: o que é e como montar uma. Disponível em: <<https://milvus.com.br/politica-de-privacidade-o-que-e/>>. Acesso em: 27 jul. 2021.

TANENBAUM, A. Computer Networks. Prentice Hall Professional Technical Reference. [S.l.: s.n.], 2002. v. 4th edition.

TASSO, F. A. LGPD - Lei Geral De Proteção De Dados - Comentada – Coordenação Maldonado, Viviane Nóbrega e Blum, Renato Opice. São Paulo: Fiscosoft – Thomson Reuters, 2019. v. 2.

TEFFÉ, C. S. D.; VIOLA, M. Tratamento de dados pessoais na LGPD: estudo sobre as bases legais. [S.l.]: civilistica.com, 2020. v. 9.

THOMPSON, M. Marco civil ou demarcação de direitos? Democracia, razoabilidade e as fendas na internet do Brasil. Rio de Janeiro: Revista de Direito Administrativo, 2012.

TOTVS. Anonimização de dados e LGPD: Conheça melhor. 2020. Disponível em: <<https://www.totvs.com/blog/negocios/anonimizacao/>>. Acesso em: 22 fev. 2021.

VAINZOF, R. Finalmente a Lei Geral de Proteção de Dados (LGPD) - Resumo dos pontos relevantes. [S.l.]: Revista Conceito Jurídico. ZK Editora, 2018.

VAINZOF, R. LGPD - Lei Geral De Proteção De Dados - Comentada – Coordenação Maldonado, Viviane Nóbrega e Blum, Renato Opice. São Paulo: Fiscosoft – Thomson Reuters, 2019. v. 2.