

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO**

[**MESTRADO EM HUMANIDADES DIGITAIS**](https://www.dcc.ufrrj.br/ppgihd/)

**PPGIHD – PROGRAMA DE PÓS-GRADUAÇÃO INTERDISCIPLINAR**

**A PRIVACIDADE NA *INTERNET* – UTILIZANDO O MÉTODO *LDA* PARA ANÁLISE DA LEI GERAL DE PROTEÇÃO DE DADOS E DAS POLÍTICAS DE PRIVACIDADE**

**ANDRÉ FERREIRA ZEFERINO**

**Rio de Janeiro**

**Maio – 2020**

**A PRIVACIDADE NA *INTERNET* – UTILIZANDO O MÉTODO *LDA* PARA ANÁLISE DA LEI GERAL DE PROTEÇÃO DE DADOS E DAS POLÍTICAS DE PRIVACIDADE**

André Ferreira Zeferino[[1]](#footnote-1)

Rodrigo de Souza Tavares²

Leandro Alvim³

**RESUMO**

O uso contemporâneo de concepções voltadas a conectividade 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 artigo 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, sendo assim, como isso vem 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. Desta forma, eles podem atuar aproveitando-se dos dados produzidos por cada usuário conectado.Diante desse fato, o artigo busca investigar se o conjunto de dados produzidos pelos usuários estão protegidos e utilizados de forma coerente, dentro da legislação vigente.

**Palavras-chave**: *Internet*, Privacidade, Dados, Informação.

***ABSTRACT***

*The contemporary use of conceptions focused on connectivity 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 article's main objective is to investigate the growth in the use of devices aimed at Internet use, the production of data that has grown massively with the ease and spread of its use, thus, how it has been 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. In this way, they can act taking advantage of the data produced by each connected user. Given this fact, the article seeks to investigate whether the set of data produced by users is protected and used in a coherent manner, within the current legislation.*

***Keywords****: Internet, Privacy, Data, Information.*

1. **INTRODUÇÃO**

A singularidade explícita e a inquirição dos seres humanos nos transferem a aprofundar em tudo que é relacionado às nossas origens e que igualmente nos direcionam e conduzem naquilo que devemos aguardar para as gerações que seguem. De fato, a nossa singularidade nos projeta a cada vez mais procurar o desenvolvimento dos métodos que nos guiam através da curiosidade, grau de complexidade e a inteligência.

A sociedade contemporânea alicerçada na influência tecnológica permitiu a todos nós, indivíduos, de gerações diversas, uma disparidade de ações com inúmeras possibilidades de formar, agir, pensar e produzir interação nos mais diversos movimentos com grupos sociais em que somos distribuídos diametralmente. O crescimento global que vivemos demonstra claramente que o desenvolvimento e a produção de novas informações, além do conhecimento, fazem parte do nosso cotidiano.

No decorrer dos séculos, nosso desenvolvimento cognitivo e intelectual nos aproximou das máquinas e dessa forma, produzimos, reinventamos e buscamos usar a nossa capacidade holística para direcionar esforços na produção de tecnologias nos mais variados campos de estudo, incluindo a tecnologia.

Os computadores, *tablets, smartphones* não são mais os únicos dispositivos que necessitam e utilizam as nossasinformações para traçar perfis ou características que definem o que somos, como agimos, com quem passamos parte do nosso dia, os jogos preferidos, o que gostamos de comer e até mesmo qual assunto tratamos em de forma privada.

A *Internet* trouxe praticidade e inovação para rotina dos usuários. Dessa forma, levanta-se a questão, se não existem mais limites e distâncias, podemos inferir que trocamos as ações praticadas em nossas vidas como barganha em troca de serviços que nos beneficiem, colocando em xeque a forma com que nossos dados são utilizados contra nós ou até mesmo a nosso favor.

A conexão com a rede mundial de computadores vai possibilitar o controle dessas ferramentas viabilizando que sejam utilizados e manipulados pelos seus provedores de serviços.

Estas novas competências de artefatos antes considerados limitados a exposição que a tecnologia mostrava em sua linha do tempo, produziram uma enorme quantidade de oportunidades nos mais diversos meios: acadêmico, industrial e pessoal. Entretanto, novos desafios foram impostos ao avanço da tecnologia, os dados produzidos por esses artefatos são utilizados para discriminar, gerar publicidade, acusar, mostrar sua localização etc. Os usuários são colocados em diferentes classes e o interesse em bens de mercado e serviços são utilizados de forma dessemelhante com base nessas categorias.

Dentre os aspectos demonstrados para a preservação da privacidade do usuário é necessário que as organizações obtenham o seu consentimento para poder manipular e utilizar seus dados, sendo garantida a revogação desta autorização.

* 1. **A LEI GERAL DE PROTEÇÃO DE DADOS**

A Lei Geral de Proteção da Dados (LGPD), [Lei N.º 13.709/2018](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm), 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.

A LGPD direciona à Autoridade Nacional de Proteção de Dados (ANPD) todo e qualquer tipo de incidentes de segurança que de alguma forma coloquem 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 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 da imagem indulgente dos usuários.

A política de privacidade é, de forma bem objetiva, uma declaração que descreve como as informações que foram coletadas do usuário são armazenadas, utilizadas ou distribuídas à terceiros. É importante salientar que a política de privacidade é exigida para todas as organizações que executam a coleta de dados pessoais do usuário.

De certo modo, a política de privacidade pode ser modificada pelos seus desenvolvedores de forma a adequar-se à legislação vigente no país, fazer atualizações e modificações a qualquer momento sem a obrigatoriedade de informar ao usuário. Podemos afirmar que todos os fatos descritos acima não são facilmente aplicados, os dados são produzidos a cada segundo e a evolução tecnológica acompanha todo o processo sem que haja qualquer disfunção de sua parte. A tecnologia auxilia o processo de armazenamento da informação com seus poderosos *Big Data* – que podem ser acessados de qualquer lugar do mundo.

De acordo com Noiret (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. (NOIRET, 2015, p. 28)

Diante de toda a realidade exposta sobre o aumento na produção e armazenamento do volume de dados, surge a necessidade da criação de uma legislação específica e que fosse equiparada aos padrões mundiais. Nesse sentido, é necessária uma análise criteriosa das políticas de privacidade disponibilizadas por seus desenvolvedores, para que saibamos se ela está de acordo com a legislação proposta em nosso país ou se existe a necessidade de adequação, dessa maneira, preservar o interesse do principal elemento desse conjunto, o usuário.

No decorrer do tempo, observamos que a produção de dados cresceu de forma notável, sendo assim, tornou-se um subproduto da informação, e seu valor para a humanidade é incalculável.

A fabricação de dados cresce duas vezes mais a cada dois anos, e o prognóstico é de que em 2020 sejam processados 350 *zettabytes* de dados, dessa forma, fica explícito que a exposição das informações dos usuários sem prévia autorização tornou-se uma incógnita.

As tecnologias emergentes propiciam essa produção e exposição, casas e carros conectados à *internet*, o uso de tecnologias incorporadas à indústria ou ao comércio, *smartcities* geram a otimização pela utilização de sensores que coletam dados, e os algoritmos processam os resultados para ajudar na tomada de decisões mais rápidas.

A maior parte dos usuários também desconsidera como é simples classificar os dados que antes eram tidos como anônimos. A coleta sem prévia autorização e sua exposição com diversas finalidades vai de encontro às legislações vigentes em todo o mundo, principalmente no eixo, Brasil, União Europeia e Estados Unidos.

Os governos e corporações coletam, armazenam e analisam a enorme quantidade de dados que divulgamos à medida que avançamos em nossas vidas digitalizadas. A inquietação com a proteção de dados também já se evidenciava no Marco Civil da Internet (Lei 12.965/2014), uma vez que é com a tecnologia da informação que a análise de dados pessoais adquire seu alcance atual.

A Lei Geral de Proteção de Dados (LGPD) estabelece como fundamentos o respeito à privacidade, àautodeterminação 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).

Dentre os fatos descritos, torna-se necessário uma ação de análise do *corpus* das Políticas de Privacidade, verificando através de uma metodologia sólida e com amplo espectro tecnológico, quais argumentos são mostrados através das palavras utilizadas pelos desenvolvedores dos artefatos e que tipo de mensagem deixaram explícita ou implícita nos *corpus* que fazem parte de sua estrutura textual.

1. **METODOLOGIA**

Neste artigo foi realizada uma revisão das ações ocasionadas pelo uso da *internet* a produção, o tratamento e a utilização dos dados coletados através das mais diversas plataformas, suas políticas de privacidade e das Leis que promovem a proteção e a privacidade dos dados do usuário. O mapeamento dos itens citados é de natureza intelectual, com direcionamento a fontes de pesquisas, análise quantitativa e qualitativa dos dados extraídos do *corpus* 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 a produção de algoritmos de processamento como a *NLP - Natural Language Processing* em *Python*. Dessa forma, foi possível investigar e tratar todo o *corpus* extraindo apenas o que era o objeto do experimento - sentenças, frases e palavras.

Com a utilização de plataformas como o *NLTK* – um pacote para o desenvolvimento de programas em *Python* com o objetivo de interpretar a linguagem humana, o *Gensim* – onde os modelos utilizam o *corpus* para buscar referências e assuntoscomuns 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: Alocação de *Dirichlet* Latente ([Distribuição Discreta Multivariada](https://pt.wikipedia.org/w/index.php?title=Distribui%C3%A7%C3%A3o_discreta_multivaridada&action=edit&redlink=1)), a busca por frequência de palavras, gráfico multidimensional, termos mais salientes e as similaridades entre termos que contribuíram para que fossem alcançados os resultados dessa pesquisa.

1. **OBJETIVOS**
   1. **OBJETIVO GERAL**

O objetivo dessa pesquisa é demonstrar que o uso constante de serviço da *internet*, 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 respeitar ou respeitam a legislação vigente no tocante a proteção e a privacidade dos dados do usuário. Dessa forma, foram coletadas 100 políticas de privacidade dos mais diversos segmentos, visando analisar e comparar sua base textual com a LGDP – Lei Geral de Proteção de Dados.

* 1. **OBJETIVOS ESPECÍFICOS**

1) Verificar a frequência de palavras dentro do *corpus* identificando os termos que têm relevância para a interpretação e comparação com a Lei Geral de Proteção de Dados;

2) Utilizar a Modelagem de Tópicos *(LDA)* para verificar e interpretar os resultados quando o algoritmo realiza a busca de 5 tópicos relevantes no *corpus* e qual é o seu comportamento quando alteramos o valor para 10 tópicos;

3) Interpretar através de palavras se, as Políticas de Privacidade estão próximas ou não do que é imposto pela Lei Geral de Proteção de Dados.

1. **ANALISANDO A LGPD – LEI GERAL DE PROTEÇÃO DE DADOS**

A Lei Geral de Proteção de Dados (LGPD) brasileira direciona seu eixo a privacidade e a segurança como pontos focais em sua narrativa. Aprovada em 14 de agosto de 2018, a Lei N.º 13.709 tenta direcionar e orientar 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 as informações de forma não-autorizada ou em circunstâncias consideradas imprevistas ou ilícitas, que ocasione a destruição, perda, modificação e exposição dos dados.

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”. (FILHO, 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 considerem 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.

* 1. **A FREQUÊNCIA DE PALAVRAS NO *CORPUS* DA LGPD**

O **Gráfico 1** e a **Tabela 1** representados abaixo, exibem as 10 palavras mais frequentes na LGPD – Lei Geral de Proteção de Dados:

|  |  |
| --- | --- |
| Tabela 1 - Frequência de Palavras na LGPD | |
| Palavras | **Frequência** |
| Dados | 250 |
| Pessoais | 152 |
| Tratamento | 143 |
| Titular | 77 |
| Nacional | 75 |
| Proteção | 62 |
| Autoridade | 54 |
| Controlador | 54 |
| Direitos | 35 |
| Consentimento | 35 |

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

Gráfico 1 – Frequência das palavras na LGPD

Tabela SEQ Gráfico\_-\_ \\* ARABIC 1 - Frequência de Palavras

Texto preto sobre fundo branco

Descrição gerada automaticamente

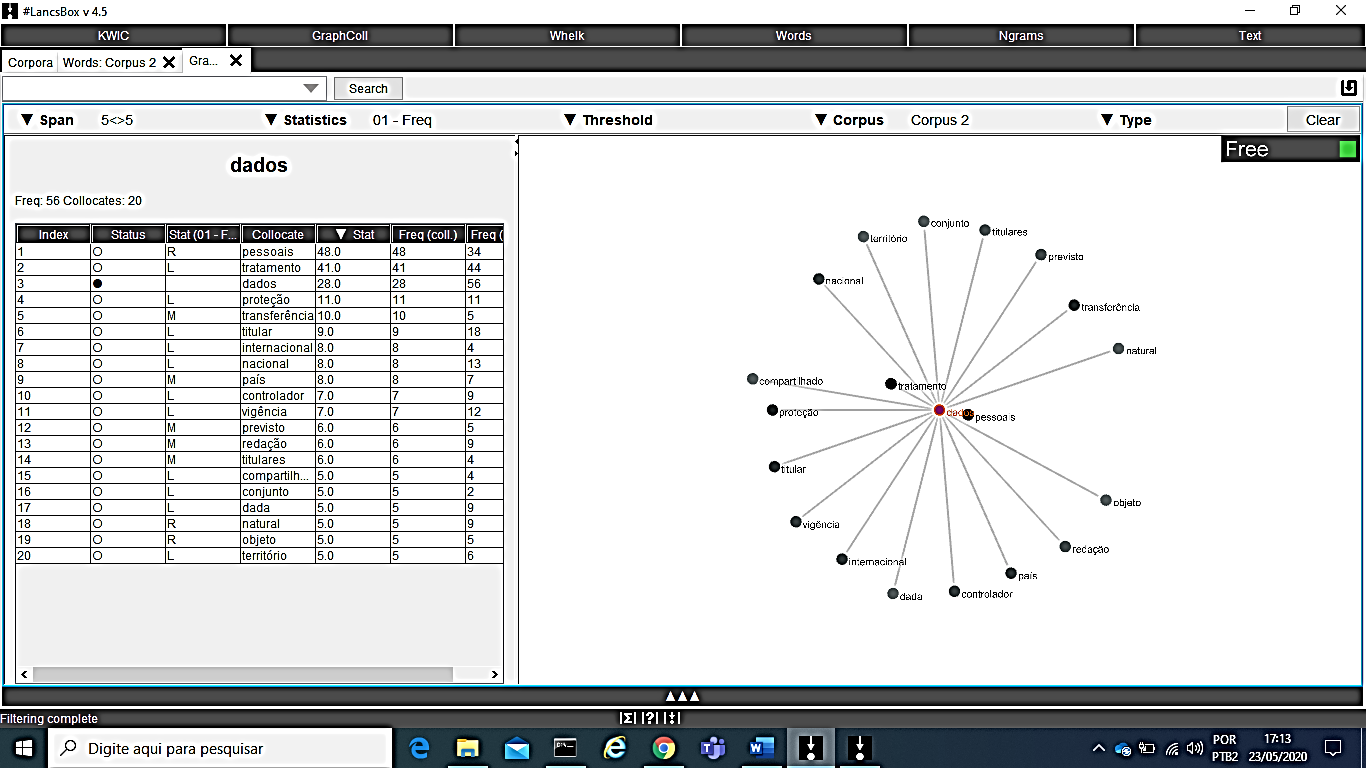
Gráfico SEQ Gráfico\_-\_ \\* ARABIC 2 - Frequência de Palavras

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

A LGPD – Lei Geral de Proteção de Dados, tem como premissa a proteção dos dados pessoais do usuário, dessa forma, observamos no **Gráfico 1,** a ocorrência massiva das palavras ‘*dados’*, ‘*pessoais’* e ‘*tratamento’* em relação às demais ocorrências.

O **Gráfico 2** exibido abaixo, mostra a proximidade dessas palavras na redação da Lei:

Gráfico 2 - Proximidade das palavras



Fonte: Resultado de processamento realizado no *software LancsBox.* <http://corpora.lancs.ac.uk/lancsbox/download.php>

O **Gráfico 2** só confirma a análise realizada anteriormente, e que produziu os resultados mostrados no **Gráfico 1**, dessa forma, as palavras com maior frequência na redação da LGPD, regularmente estão próximas umas às outras, dando maior credibilidade aos resultados.

* 1. **MODELAGEM DE TÓPICOS – *LATENT DIRICHLET ALLOCATION (LDA)***

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. Esse é um trabalho que robotiza a transcrição dos dados em um enorme conjunto de *corpus*, baseada em métodos estatísticos para avaliar e analisar os dados do documento retornando como saída os tópicos mais relevantes. Esse trabalho de modelagem de tópicos é caracterizado como não-supervisionado.

* + 1. **A MODELAGEM DE TÓPICOS NA LGPD**

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.

Cada um dos tópicos é caracterizado por uma distribuição de probabilidades sobre as palavras.

No processamento do algoritmo na redação da Lei Geral de Proteção de Dados, foi escolhido inicialmente uma matriz de 5 tópicos com 4 palavras, posteriormente, o mesmo processamento foi realizado com 10 tópicos com 4 palavras para que que fosse realizada uma análise mais aprofundada do *corpus*.

* + - 1. **A MODELAGEM DE TÓPICOS NA LGPD – 5 TÓPICOS**

Ao realizar o processamento do *corpus* na LGDP, o algoritmo, preparado para o retorno de 5 tópicos com 4 palavras, obteve o seguinte resultado na **Tabela 2**:

|  |  |  |
| --- | --- | --- |
| Tabela 2 – Resultado do LDA na LGPD | | |
| Tópico | **Palavra** | **Valor** |
| 0 | tratamento | 0.030 |
| dado | 0.029 |
| titular | 0.020 |
| pessoais | 0.020 |
| 1 | dado | 0.031 |
| nacional | 0.022 |
| pessoais | 0.018 |
| autoridade | 0.017 |
| 2 | dado | 0.048 |
| tratamento | 0.034 |
| pessoais | 0.024 |
| vigência | 0.012 |
| 3 | dado | 0.050 |
| pessoais | 0.034 |
| proteção | 0.021 |
| nacional | 0.013 |
| 4 | dado | 0.028 |
| pessoais | 0.018 |
| titular | 0.016 |
| controlador | 0.015 |

Tabela 2 – Resultado do LDA

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

No modelo *LDA*, os valores no início do resultado (0, 1, 2, 3 e 4) são a representação do quantitativo de tópicos escolhidos, ou seja, 5, 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*.

A **Tabela 3** exibida abaixo, mostra o resultado das análises realizada em cada um dos tópicos:

|  |  |  |
| --- | --- | --- |
| **Tabela 3 – Análise dos Tópicos na LGPD** | | |
| **Tópico** | **Palavras** | **Análise** |
| **0** | '0.030\*"tratamento" + 0.029\*"dado" + 0.020\*"titular" + 0.020\*"pessoais"' | 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.  **Maior Relevância:** Art. 1º e Art. 3º da Lei N.º 13.709/2018.  Pinheiro (2018, p. 49), 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”. Já no Art. 3º, Pinheiro (2018, p. 55), declara que “a LGPD objetiva delimitar a abrangência do tratamento de dados no que concerne à territorialidade, destacando que se considera a lei aplicável aos dados coletados dentro do território nacional”. (PINHEIRO, 2018, p. 49), |
| **1** | '0.031\*"dado" + 0.022\*"nacional" + 0.018\*"pessoais" + 0.017\*"autoridade" | Ao analisar esse tópico existe uma alusão à criação da ANPD – Autoridade Nacional de Proteção de Dados.  **Maior Relevância:** Art. 55-A da Lei N.º 13.709/2018.  Esse artigo foi VETADO.  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.” |
| **2** | '0.048\*"dado" + 0.034\*"tratamento" + 0.024\*"pessoais" + 0.012\*"vigência"' | Ao analisar esse tópico a referência diz respeito ao início da vigência da Lei. Inclui a Redação da Lei N.º 13.853/2019.  **Maior Relevância:** Parágrafo Único do Art. 1º da Lei N.º 13.709/2018.  Há uma referência à Lei N.º 13.853/2019 que altera a Lei nº 13.709, de 14 de agosto de 2018, para dispor sobre a proteção de dados pessoais e para criar a Autoridade Nacional de Proteção de Dados; e dá outras providências.  Dessa forma, o Art. 55-A foi restabelecido pela MP 869 de 2018 como nova redação. |
| **3** | '0.050\*"dado" + 0.034\*"pessoais" + 0.021\*"proteção" + 0.013\*"nacional"' | Ao analisar esse tópico existe o sentido da proteção dos dados pessoais do usuário e alusão a criação da ANPD e a amplitude da Lei que é em território nacional.  **Maior Relevância:** Art. 3º, Incisos I, II e III e Parágrafo 1º da Lei N.º 13.709/2018.  Pinheiro (2018, p. 55), declara que “a LGPD objetiva delimitar a abrangência do tratamento de dados no que concerne à territorialidade, destacando que se considera a lei aplicável aos dados coletados dentro do território nacional. A mesma lógica vale para os titulares dos dados coletados, assim como foi previsto pelo Marco Civil, em seu Art. N.º 11, Parágrafos 1º e 2º”. (PINHEIRO, 2018, p. 55), |
| **4** | '0.028\*"dado" + 0.018\*"pessoais" + 0.016\*"titular" + 0.015\*"controlador"' | Ao analisar esse tópico temos a criação dos responsáveis pelo tratamento dos dados dos usuários, o controlador.  **Maior Relevância:** Art.º 5, Inciso VI da Lei N.º 13.709/2018.  A Lei 13.709/2018 destaca em seu Art. 5º, Inciso VI, “O Controlador é pessoa natural ou jurídica, de direito público ou privado, a quem competem as decisões referentes ao tratamento de dados pessoais”. |

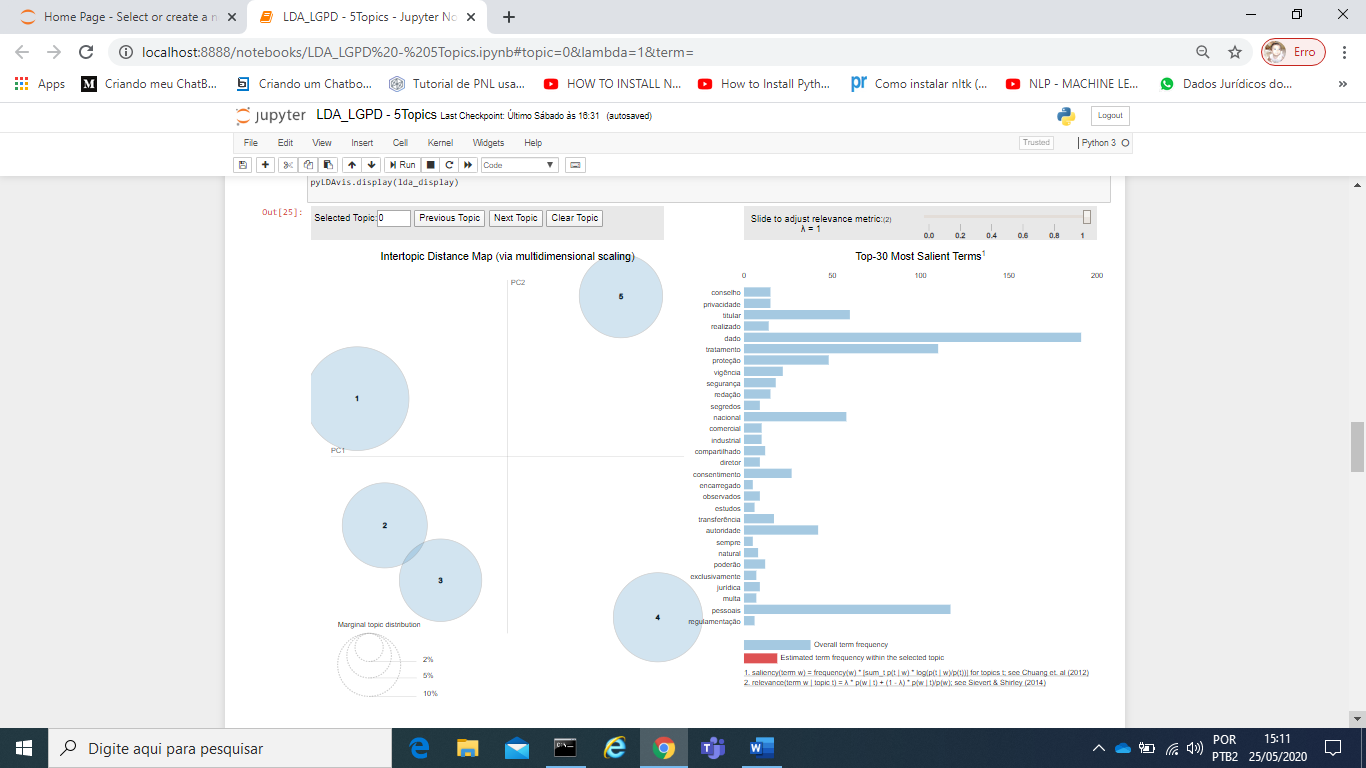
Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

Tabela 3 – Análise do LDA

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.

O **Gráfico 3**, multidimensional, exibe o grupo das 30 palavras mais salientes dentro da redação da LGPD, proporcionalidade e espaçamento entre os termos:

Gráfico 3 – Gráfico Multidimensional



Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

**Tópico 1 – 27,1% dos *tokens***

Este tópico 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. (MACIEL, 2019)

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

**Tópico 2 - 18,2% dos *tokens***

Este tópico 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.

De acordo com Pinheiro (2018, p. 61):

Apesar de alguns termos e expressões se diferenciarem, os papéis de cada sujeito ou processo são equivalentes no documento brasileiro quanto no europeu. É exemplo disso a questão do controlador/processador do GDPR, que na LGPD ganharam o nome de controlador/operador, cujas ações, funções e responsabilidades são equivalentes, só se modificando a nomenclatura adotada. (PINHEIRO, 2018)

**Tópico 3 - 17,1% dos *tokens***

O terceiro tópico traz palavras voltadas aos privilégios e os direitos que os usuários têm sobre seus dados. Esse tópico é diretamente ligado à LGPD. As palavras como *‘dado’* e *‘tratamento’* são o topo do ranking, desta forma, podemos inferir que o tópico fornece informações sólidas para o comparativo com à 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. 20):

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). (MACIEL, 2019)

**Tópico 4 - 20% dos *tokens***

O quarto tópico tem 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% do faturamento, cujo valo fica limitado a um total de R$ 50 milhões de reais – e diária, além da suspensão das atividades relativas ao banco de dados. (PINHEIRO, 2018)

**Tópico 5 - 17,9% dos *tokens***

O quinto tópico tem definições além das já expostas nos tópicos anteriores, existe referências as palavras *‘contratuais’* e *‘transferência’*, que fazem alusão ao consentimento realizado por cláusulas de contrato e o tráfego internacional de dados, caso seja necessário.

Segundo Vainzof (2018, p. 26):

[...]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. (VAINZOF, 2018)

* + - 1. **A MODELAGEM DE TÓPICOS NA LGPD - 10 TÓPICOS**

Ao realizar o processamento do corpus na LGDP, o algoritmo, preparado para o retorno de 10 tópicos com 4 palavras, obteve o seguinte resultado exibido na **Tabela 4**:

|  |  |  |
| --- | --- | --- |
| Tabela 4 – Resultado do LDA | | |
| Tópico | **Palavra** | **Valor** |
| 0 | dado | 0.026 |
| titular | 0.026 |
| pessoais | 0.021 |
| consentimento | 0.020 |
| 1 | dado | 0.053 |
| tratamento | 0.034 |
| pessoais | 0.026 |
| público | 0.022 |
| 2 | dado | 0.026 |
| titular | 0.026 |
| tratamento | 0.021 |
| pessoais | 0.013 |
| 3 | dado | 0.014 |
| disposto | 0.014 |
| poderá | 0.011 |
| nacional | 0.011 |
| 4 | transferência | 0.033 |
| internacional | 0.021 |
| contratuais | 0.012 |
| consentimento | 0.009 |
| 5 | saúde | 0.023 |
| dado | 0.022 |
| exclusivamente | 0.014 |
| pessoais | 0.013 |
| 6 | dado | 0.077 |
| pessoais | 0.054 |
| tratamento | 0.043 |
| nacional | 0.019 |
| 7 | tratamento | 0.015 |
| titular | 0.013 |
| autoridade | 0.013 |
| nacional | 0.012 |
| 8 | dado | 0.043 |
| proteção | 0.027 |
| pessoais | 0.026 |
| nacional | 0.024 |
| 9 | administrativo | 0.014 |
| processo | 0.012 |
| direitos | 0.012 |
| incisos | 0.010 |

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

Tabela 4 – Análise do LDA

A **Tabela 5**, exibida abaixo, mostra o resultado das análises realizada em cada um dos tópicos:

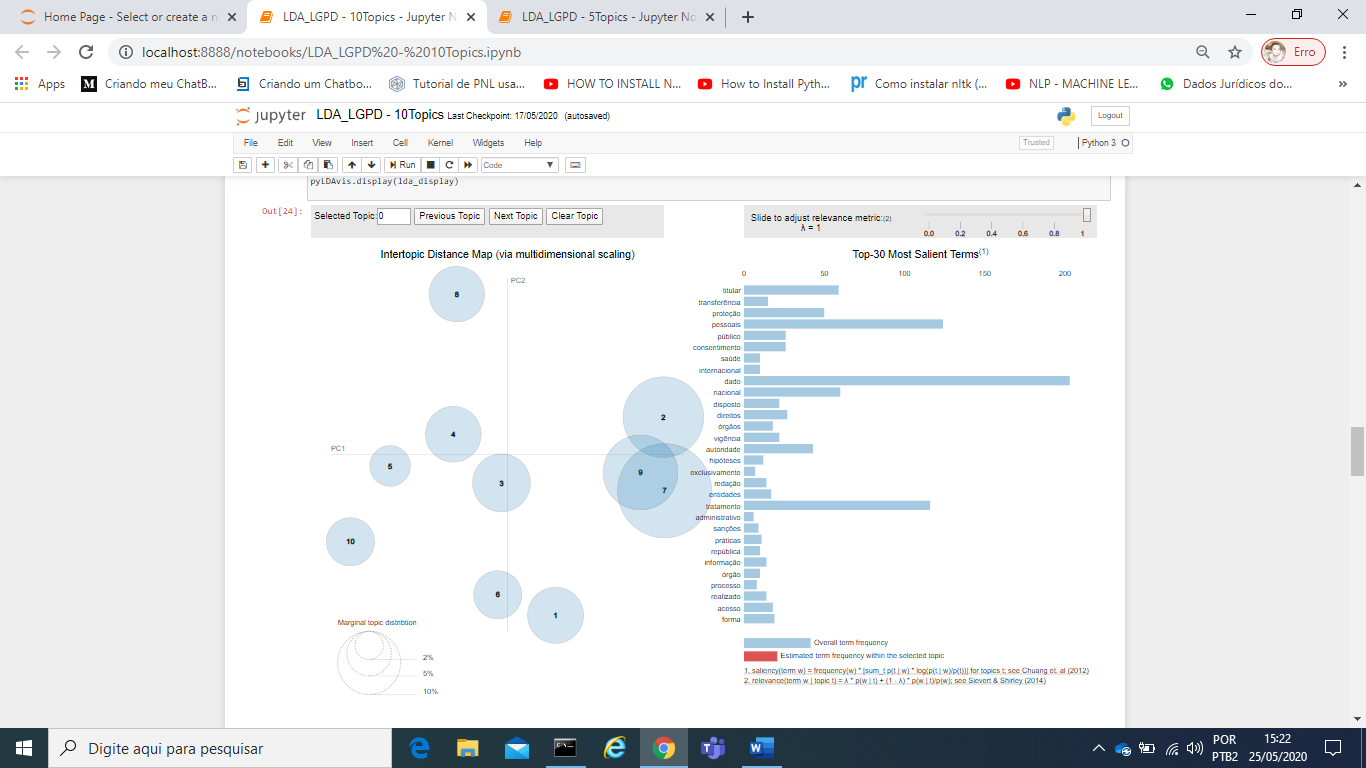
|  |  |  |
| --- | --- | --- |
| **Tabela 5 – Análise dos Tópicos** | | |
| **Tópico** | **Palavras** | **Análise** |
| **0** | '0.026\*"dado" + 0.026\*"titular" + 0.021\*"pessoais" + 0.020\*"consentimento"' | Ao analisar esse tópico podemos fazer referência a solicitação do consentimento do titular dos dados para qualquer tipo ou finalidade.  **Maior Relevância:** Art. 5º, Inciso XII da Lei N.º 13.709/2018 que em sua redação explana:  XII - consentimento: manifestação livre, informada e inequívoca pela qual o titular concorda com o tratamento de seus dados pessoais para uma finalidade determinada;  Cots (2018, p. 23) destaca que, “para que uma pessoa física ou jurídica trate dados será necessário que a mesma tenha uma base legal para tanto. A base mais comum será o consentimento, por ser a mais intuitiva em relação ao que já se praticava no Brasil”. (COTS, 2018) |
| **1** | '0.053\*"dado" + 0.034\*"tratamento" + 0.026\*"pessoais" + 0.022\*"público"' | Ao analisar esse tópico existe uma breve alusão as pessoas do direito público ou ao atendimento ao interesse público e ao tratamento dos seus dados.  **Maior Relevância:** Art. 3º, Art. 4º Parágrafos 1º e 2º da Lei N.º 13.709/2018.  Maciel (2019, p. 22) explana que, “Em relação ao tratamento de dados, a definição trazida pela LGPD tira do imaginário de muitos a premissa de que estariam fora do alcance da novel legislação, porquanto não utilizariam dados para performar seus negócios. Ledo engano, a lei é bem ampla e praticamente qualquer utilização de um dado pessoal, ainda que um acesso visual na máquina de outro estará sob o alcance da lei e deve ser devidamente levantado em um processo de conformidade, sem falar de que no ambiente entre empregado e empregador há tratamento de dados pessoais, não se eximindo o controlador, no caso o empregador, de seguir com as regras da lei”. (MACIEL, 2019) |
| **2** | '0.026\*"dado" + 0.026\*"titular" + 0.021\*"tratamento" + 0.013\*"pessoais"' | Ao analisar esse tópico podemos fazer referência em como a LGPD confirma o usuário como o titular dos dados pessoais, em território nacional, que venham a ser coletados e de que forma deve ser o tratamento dessas informações.  **Maior Relevância:** Art. 1º e Art. 3º da Lei N.º 13.709/2018.  Cots (2018, p. 23), explana sobre os direitos do titular, “Na esteira dos Fundamentos e Princípios, foram estabelecidos direitos dos titulares dos dados pessoais, entre os quais, destacamos: direito de ser informado sobre o tratamento, possibilidade de oposição ao mesmo, consulta e retificação de dados, etc”. (COTS, 2018) |
| **3** | '0.014\*"dado" + 0.014\*"disposto" + 0.011\*"poderá" + 0.011\*"nacional"' | Tópico sem relevância |
| **4** | '0.033\*"transferência" + 0.021\*"internacional" + 0.012\*"contratuais" + 0.009\*"consentimento"' | Ao analisar esse tópico temos a alusão do consentimento para a transferência de dados a terceiros ou de forma internacional.  **Maior Relevância:** Art. 7º, Inciso I, Art. 8º, Parágrafo I, Art. 13º, Parágrafo 2º e Art.º 33º da Lei N.º 13.709/2018.  Maciel (2019, p.55), “Um dos maiores desafios para a aplicação efetiva de uma legislação de proteção de dados pessoais é o controle sobre o fluxo internacional de dados. A liquidez dos dados permite que sejam transmitidos com uma velocidade enorme para outra jurisdição e o titular, na maioria das vezes, nem fica sabendo aonde suas informações foram parar. Da mesma forma, é impossível coibir o dinamismo da inovação que, muitas vezes utiliza-se de facilidades de *data centers* em outros países, beneficia-se de acordos comerciais com empresas estrangeiras, as quais muitas vezes são responsáveis por aportar capital de investimento, ou mesmo pela própria razão de que a *internet* não possui fronteiras delineadas, sendo que o dado de uma pessoa natural coletado no Brasil pode estar sendo obtido a partir de uma outra nação”. (MACIEL, 2019) |
| **5** | '0.023\*"saúde" + 0.022\*"dado" + 0.014\*"exclusivamente" + 0.013\*"pessoais"' | Ao analisar esse tópico temos o tratamento dos dados exclusivamente para a tutela de saúde.  **Maior Relevância:** Art. 11º, Alínea f.  Pinheiro (2018, p.70) diz que, “A importância do consentimento para a realização do tratamento de dados sensíveis é intrínseca à validade dessa ação, todavia há algumas situações em que tal consentimento pode ser relativizado (excetuado), como pontua o Art. 11º. Essas situações são relacionadas ao cumprimento de obrigações legais por parte do controlador, à garantia da segurança do titular, à prevenção à fraude, à execução de políticas públicas, à proteção da vida/incolumidade física, assim como à tutela da saúde”. (PINHEIRO, 2018) |
| **6** | '0.077\*"dado" + 0.054\*"pessoais" + 0.043\*"tratamento" + 0.019\*"nacional"' | Ao analisar esse tópico existe o sentido da proteção dos dados pessoais do usuário e uma breve alusão a criação da ANPD, além de qual a amplitude da Lei em território nacional.  **Maior Relevância:** Art. 3º, Incisos I, II e III e Parágrafo 1º da Lei N.º 13.709/2018.  Reani (2018) relata que, “Operações de tratamento de dados realizadas dentro do território brasileiro estão sujeitas a aplicação da LGDP. Além de operações de tratamento realizadas dentro do país, quando o tratamento tiver por objetivo a oferta ou fornecimento de bens ou serviços ou o tratamento de dados de indivíduos localizados no território brasileiro, a lei também pode se aplicar, ainda que a organização responsável por essa atividade esteja sediada ou localizada fora do país. Assim, o local onde os dados são tratados não é requisito único ou preponderante para aplicação da lei, sendo também importante identificar a localização do indivíduo cujos dados serão coletados”. (REANI, 2018) |
| **7** | '0.015\*"tratamento" + 0.013\*"titular" + 0.013\*"autoridade" + 0.012\*"nacional"' | Ao analisar esse tópico existe uma breve alusão a criação da ANPD – Autoridade Nacional de Proteção de Dados e o tratamento dos dados pela autarquia.  **Maior Relevância:** Art. 55-A da Lei N.º 13.709/2018.  Filho (2019, p. 39) destaca que, “A Lei ainda prevê a criação de um órgão responsável pela sua aplicação: a Autoridade Nacional de Proteção de Dados (ANPD), que terá a forma de uma autarquia especial vinculada ao Ministério da Justiça. (FILHO, 2018) |
| **8** | '0.043\*"dado" + 0.027\*"proteção" + 0.026\*"pessoais" + 0.024\*"nacional"' | Ao analisar esse tópico existe o sentido da proteção dos dados pessoais do usuário e uma breve alusão a criação da ANPD e a amplitude da Lei em território nacional.  **Maior Relevância:** Art. 3º, Incisos I, II e III e Parágrafo 1º da Lei N.º 13.709/2018.  Definição aproximada ao Tópico 6. |
| **9** | '0.014\*"administrativo" + 0.012\*"processo" + 0.012\*"direitos" + 0.010\*"incisos"' | Ao analisar esse tópico existe o sentido de como serão os procedimentos administrativos caso algum direito do titular dos dados seja violado. Faz alusão a criação da ANPD.  **Maior Relevância:** Art. 55J, Parágrafo 1º da Lei N.º 13.709/2018.  González (2019) afirma que “esse artigo cria a [Autoridade Nacional de Proteção de Dados (ANPD)](https://guialgpd.com.br/o-que-e-e-como-atua-a-autoridade-nacional-de-protecao-de-dados/), inicialmente vetada, mas depois retomada. O órgão da administração pública federal fará parte da Presidência da República e é de natureza transitória — ou seja, o Poder Executivo pode transformá-la em entidade da administração pública federal indireta, sendo então submetida a regime autárquico especial, mas mantendo o vínculo à Presidência”. GONZÁLEZ (2019) |

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

Tabela 5 – Análise dos tópicos

Ao expandir para 10 tópicos a modelagem trouxe maior granularidade aos tópicos, reforçando a criação da ANPD, questões de trâmite internacional de dados, tutela de saúde e processos administrativos em caso de violação dos direitos do titular. Ao analisar o **Gráfico 4**, multidimensional, temos a proporção dos grupos de palavras na LGPD e sua dimensão, a granularidade dos assuntos e espaçamento entre os termos:

Gráfico 4 – Gráfico Multidimensional



Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

**Tópico 1 – 7,9% dos *tokens***

Este tópico reflete o tratamento dos dados, menciona os mecanismos e hipóteses em que podem ser utilizados e indica a solicitação do consentimento para tal. A uma alusão a criação da ANPD e dos cargos que nela tem vital importância para manter a segurança das informações, além da criação do seu Conselho Diretor.

González (2019) destaca que:

O Conselho será responsável por contribuir na elaboração da Política Nacional de Proteção de Dados Pessoais e da Privacidade, publicar relatórios anuais avaliando a Política, sugerir ações a serem tomadas pela ANPD, estudar e realizar debates sobre proteção e privacidade de dados e fomentar informação e conhecimento sobre esses temas. (GONZÁLEZ, 2019)

São exibidos também, termos como *‘liberdades’* e *‘fundamentais’*, que flexionam ao Art. 1º da Lei Geral de Proteção de Dados.

Filho (2018, p. 14) evidência que:

O tratamento de dados pessoais necessário para atender ao interesse legítimo do controlador ou de terceiro é permitido pela LGPD, desde que tal tratamento não viole os direitos e liberdades fundamentais do titular dos dados e que medidas para garantir a transparência de tal tratamento sejam adotadas. (FILHO, 2018)

**Tópico 2 – 16,3% dos *tokens***

Este tópico destaca a criação da ANPD, seus cargos e direciona ao indicativo de procedimentos para o tráfego de dados do usuário. A palavra *‘finalidade’* que dá ao conjunto de *tokens* a referência ao consentimento e tratamento dos dados do usuário para determinado fim.

González (2019) salienta que:

A Lei lembra, mais uma vez, que o consentimento deve ser pedido para um determinado fim específico. É uma violação da LGPD pedir consentimento de forma não-específica e genérica; todo e qualquer dado pessoal solicitado e tratamento proposto deve ter uma finalidade clara. (GONZÁLEZ, 2019)

**Tópico 3 – 8,4% dos *tokens***

O terceiro tópico traz palavras voltadas e direcionadas novamente ao consentimento e a finalidade para que os dados sejam utilizados. Como escrito no Art. 9º Parágrafo 1º, as informações de consentimento não podem ser *‘inequívocas’*, um dos *tokens* exposto nesse tópico.

Filho (2018, p. 14) enfatiza que:

A LGPD estabelece que o consentimento é uma manifestação livre, informada e inequívoca que autoriza o tratamento de dados pessoais para uma finalidade determinada. Autorizações genéricas, isto é, autorizações que não têm como escopo uma finalidade específica, explícita e informada serão nulas. (FILHO, 2018)

**Tópico 4 – 7,8% dos *tokens***

O quarto tópico tem definições surpreendentes e traz uma referência a Lei N.º 12.527/2011, a Lei de Acesso à Informação, que em um pequeno trecho abrange as pessoas jurídicas de direito público.

Filho (2018, p. 10) ressalta que:

A LGPD regulamenta o tratamento de informações relacionadas a pessoas físicas apenas, de modo que não se aplica aos dados de pessoas falecidas e de pessoas jurídicas. Organizações do setor público e privado estão sujeitas à lei. Além disso, a LGPD regulamenta o tratamento de dados pessoais realizado por qualquer meio, dentro ou fora da internet, utilizando ou não meios digitais. (FILHO, 2018)

**Tópico 5 – 4,1% dos *tokens***

O quinto tópico tem definições além das já expostas nos tópicos acima, existem referências as palavras *‘contratuais’* e *‘transferência’*, que fazem alusão ao consentimento realizado por cláusulas de contrato e o tráfego internacional de dados, caso seja necessário. A palavra *‘anonimização’*, que é a utilização de meios para que seja protegida ou para que não seja associada a identificação do usuário. O tópico traz também *‘cooperação’* e *‘internacional’* que faz referência a trâmites jurídicos da relação Brasil e países que se preocupam com a proteção de dados do usuários, além das palavras *‘infrator’* e *‘sanções’* que traz a possibilidade de ampla defesa na forma gradativa, acumulada ou isolada, de acordo com o tipo de caso. Destaca-se muito pelo fato de direcionar as punições que podem ser aplicadas em caso de algum evento com os dados do usuário.

González (2019) salienta que:

Dados anonimizados não são considerados dados pessoais e, portanto, não são protegidos pelas diretrizes regulares da LGPD. Para que isso efetivamente aconteça, porém, a anonimização dos dados não pode, sob hipótese alguma, ser passível de reversão.

A lei determina que a reversão não pode ser possível “com esforços razoáveis”, o que deixa certa margem para interpretação. Mas, como explica o parágrafo 1º, a determinação de quais são os “esforços razoáveis” de um agente considera o tempo e o custo necessários para conseguir reverter a anonimização, assim como a tecnologia disponível. Ou seja, o total de esforço e a disponibilidade de ferramentas que possam reverter o processo. GONZÁLEZ (2019)

**Tópico 6 – 5,8% dos *tokens***

O sexto tópico destaca-se pelas palavras *‘saúde’, ‘tutela’* e *‘sanitária’* que fazem referência aos dados sensíveis do usuário e a tutela em caso de procedimentos realizados pelos profissionais da área da saúde ou de autoridade sanitária.

De acordo com Reuters (2019, p. 9):

O conceito de dado pessoal sensível é oferecido pela própria LGPD em seu art. 5º, II: “dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural”. (REUTERS, 2019)

**Tópico 7 – 22,2% dos *tokens***

O sétimo tópico novamente 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.

De acordo com Filho (2018, p. 20), cabe a organização que realiza o tratamento de dados:

Manter registro das operações de tratamento de dados pessoais que realize, podendo a autoridade nacional determinar que seja elaborado relatório de impacto à proteção de dados (pessoais ou sensíveis) referente às suas operações. (FILHO, 2018)

Existe a referência ao *‘controlador’*, cargo de grande importância na ANPD, pois compete a ele decidir sobre o tratamento nos dados dos usuários. Filho (2018, p. 8) enfatiza que o controlador é *“É a pessoa que tem competência para tomar decisões referentes ao tratamento de dados pessoais. Essa pessoa pode ser natural ou jurídica, de direito público ou privado”.*

**Tópico 8 – 7,7% dos *tokens***

O oitavo tópico reflete o tratamento dos dados, mencionando os mecanismos e hipóteses em que podem ser utilizados seus dados e indica a solicitação do *‘consentimento’* do usuário. Existe a referência a palavra *‘finalidade’*, que condiz também com a redação da Lei que expressa a *“realização do tratamento para propósitos legítimos, específicos, explícitos e informados ao titular”*. (Art. 6º, Inciso I da Lei N.º 13.709/2018).

**Tópico 9 – 14% dos *tokens***

O nono tópico traz *‘observância’* entre as 30 palavras consideradas mais relevantes, esse *token* demonstra a relação de responsabilização e prestação de contas para a indicação na proteção dos dados dos usuários. Existe também a palavra *‘pesquisa’*, indicador da permissão que órgãos de estudos tem em utilizar os dados desde que garantido o anonimato do usuário. Para González (2019) destaca que [...] *“órgãos de pesquisa também não precisam exigir consentimento, mas devem trabalhar com dados anonimizados sempre que possível — dessa forma, é possível ter acesso aos dados estatísticos sem que eles sejam conectados a um titular específico.”*

**Tópico 10 – 5,8% dos *tokens***

O tópico traz a palavra *‘último’*, que faz referência a Lei N.º 9.307/96 – Lei da Arbitragem, isso é um indicativo de que alguns conflitos podem ser resolvidos sem que haja a necessidade do ingresso de ações judiciais. A LGPD no seu Art. 7º, inciso VI, destaca que para o exercício regular de direitos em processo judicial, administrativo ou arbitral, esse último nos termos da [Lei nº 9.307, de 23 de setembro de 1996 (Lei de Arbitragem)](http://www.planalto.gov.br/ccivil_03/LEIS/L9307.htm).

* 1. **A FREQUÊNCIA DE PALAVRAS NO *CORPUS* EM POLÍTICAS DE PRIVACIDADE**

A frequência de palavras tem como base um *corpus* que é composto por 100 Políticas de Privacidade variadas, disponibilizadas em seus *sites* e *apps*, extraídas e colocadas em um único arquivo com extensão ***txt***. A análise foi realizada por um algoritmo desenvolvido em linguagem *Python*. No processamento foram utilizadas *stopwords*, ou seja, palavras sem relevância (artigos, preposições entre outros) que foram removidas do processamento para não afetar a frequência adquirida.

A **Tabela 6** e o **Gráfico 5** representados abaixo, exibem as 10 palavras mais frequentes encontradas no *corpus* das Políticas de Privacidade:

|  |  |
| --- | --- |
| Tabela 6 - Frequência de Palavras nas Políticas de Privacidade | |
| Palavras | **Frequência** |
| Informações | 12090 |
| Dados | 9792 |
| Serviços | 7343 |
| Pessoais | 5410 |
| Privacidade | 3904 |
| Usuário | 3670 |
| Conta | 2966 |
| Terceiros | 2748 |
| Política | 2694 |
| Cookies | 2246 |

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

Gráfico 5 – Frequência das palavras nas Políticas de Privacidade

Tabela SEQ Gráfico\_-\_ \\* ARABIC 1 - Frequência de Palavras

Tela de celular com texto preto sobre fundo branco

Descrição gerada automaticamente

Tabela 6 - Frequência de Palavras das Políticas de Privacidade

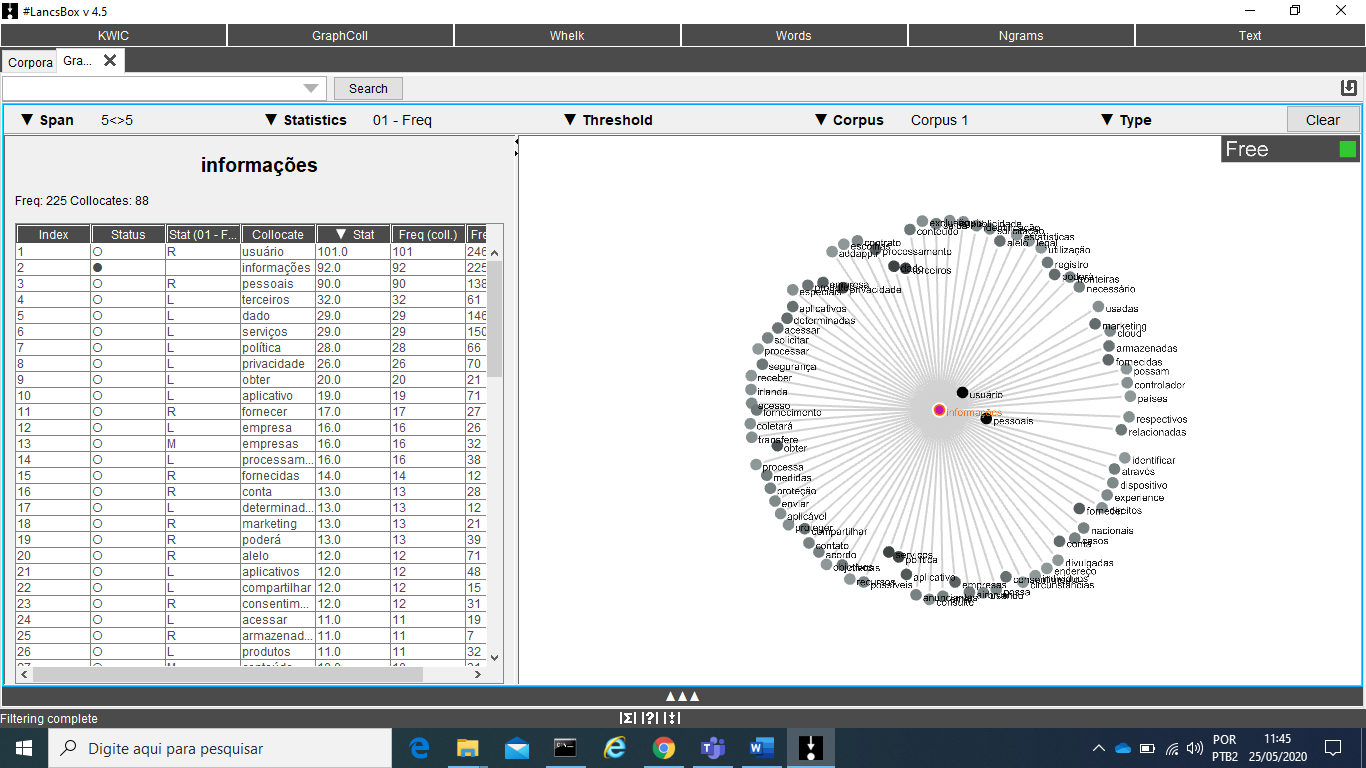
Gráfico 5 - Frequência de Palavras das Políticas de Privacidade

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

As palavras exibidas estão diretamente ligadas ao uso dos dados que os usuários disponibilizam nos *sites* e *apps* onde foram extraídas as Políticas de Privacidade e que podem ou não de forma implícita ser usados para outros fins. Nota-se a utilização de *cookies* que tem como função principal armazenar as ações dos usuários na *internet*. Destacam-se as palavras *‘informações’, ‘dados’* e *‘serviços’*, uma breve análise promove um direcionamento aos serviços ofertados pelos meios em que os dados são coletados e de que forma eles podem ser utilizados.

O **Gráfico 6** exibido abaixo, mostra a proximidade entre as palavras:

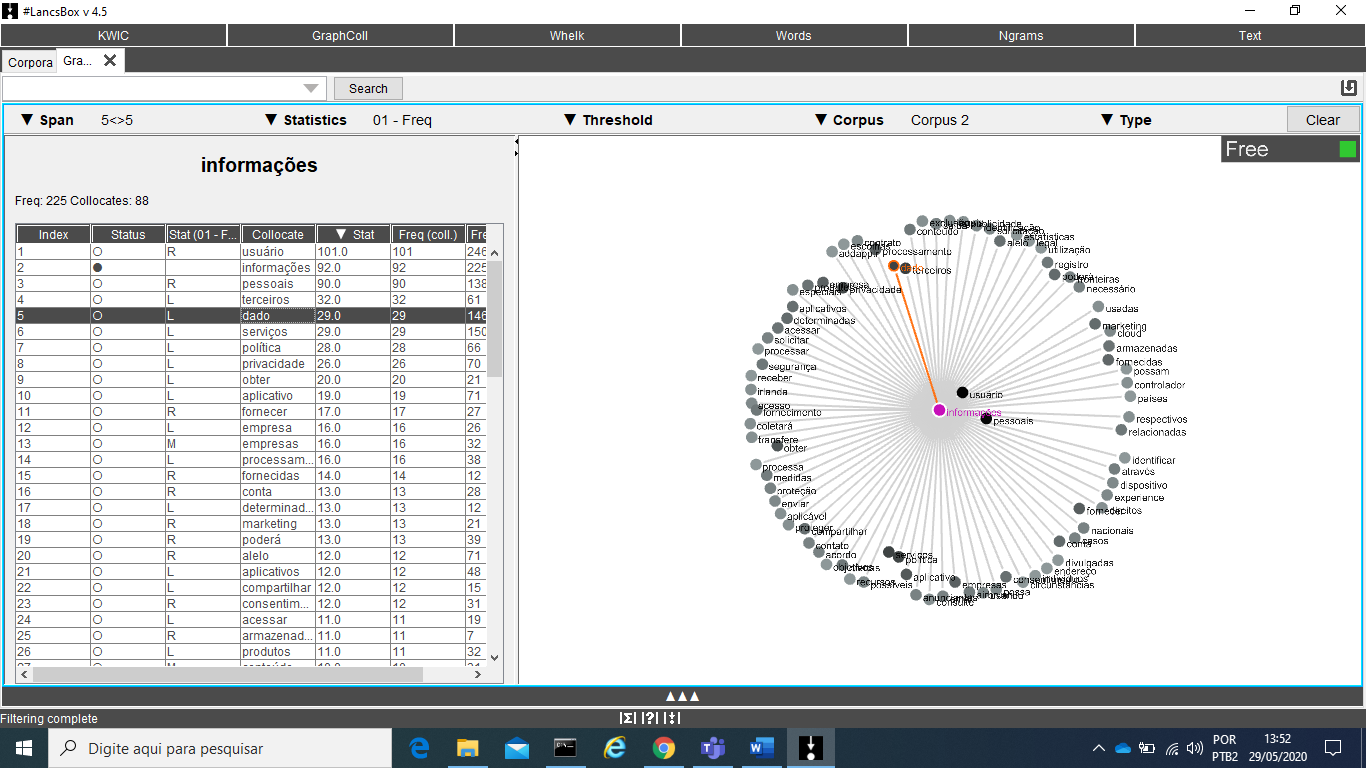
Gráfico 6 - Proximidade das palavras



Fonte: Resultado de processamento realizado no *software LancsBox.* <http://corpora.lancs.ac.uk/lancsbox/download.php>

Nota-se que a palavra *‘informações’* está mais próxima das palavras *‘usuários’* e *‘pessoais’*, isso demonstra que o *corpus* tem uma inclinação a coleta de informações pessoais do usuário. Ao comparar a proximidade da palavra *‘dados’*, palavra com maior frequência na LGPD e a palavra *‘informações’*, palavra com maior frequência no *corpus* das Políticas de Privacidade, observamos no **Gráfico 7** o distanciamento entre os termos:

Gráfico 7 - Proximidade das palavras *‘informações’* e *‘dados’*



Fonte: Resultado de processamento realizado no *software LancsBox.* <http://corpora.lancs.ac.uk/lancsbox/download.php>

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, os dados que expomos podem ser coletados sem o nosso conhecimento, normalmente, utilizados sem o nosso consentimento. Com base na coleta desses dados, pode-se inferir conclusões sobre como o usuário realiza suas atividades rotineiras.

A Revista Veja em matéria publicada em 2013 afirma:

A Agência Nacional de Segurança dos Estados Unidos e o FBI têm acesso direto aos servidores centrais de nove das maiores empresas de internet americanas. A inteligência americana pode consultar áudios, vídeos, fotografias, conteúdos de e-mails, arquivos transferidos e conexões dos usuários. O programa altamente secreto é chamado de Prisma e está em andamento desde 2007, informou o jornal *The Washington Post*. Ao longo desse período, nove companhias passaram a fazer parte da operação. A Microsoft foi a primeira a entrar, em novembro de 2007. O Yahoo passou a fazer parte em 2008. Google, Facebook e PalTalk, em 2009, YouTube em 2010, Skype e AOL, em 2011, e a última, a Apple, em outubro de 2012. (VEJA, 2013)

A maioria das informações que são divulgadas sobre vigilância governamental vem do escândalo *Edward Snowden*. Como contratado da *NSA* (Agência de Segurança Norte-Americana), *Snowden* divulgou milhares de documentos onde eram comprovadas as atividades da agência de vigilância.

No ano de 2012, o *New York Times* publicou uma reportagem sobre como as organizações analisam e tomam decisões estratégicas tendo em mãos dados de usuários para garantir e determinar formas de publicidade. *“O artigo revelou que a Target Corporation poderia determinar, a partir dos padrões de compra de uma mulher, que estava grávida e usaria essa informação para enviar anúncios e cupons para itens relacionados a bebês. A história incluía uma anedota sobre um homem de Minneapolis que reclamou com uma loja da Target que havia enviado cupons relacionados a bebês para sua filha adolescente, apenas para descobrir mais tarde que a Target estava certa.”*

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

* + 1. **A MODELAGEM DE TÓPICOS NAS POLÍTICAS DE PRIVACIDADE – 5 TÓPICOS**

Ao realizar o processamento do *corpus* das Políticas de Privacidade, o algoritmo, preparado para o retorno de 5 tópicos com 4 palavras, obteve o seguinte resultado exibido na **Tabela 7**:

|  |  |  |
| --- | --- | --- |
| Tabela 7 – Resultado do LDA | | |
| Tópico | **Palavra** | **Valor** |
| 0 | Serviços | 0.079 |
| informações | 0.026 |
| fornecer | 0.024 |
| produtos | 0.016 |
| 1 | conta | 0.044 |
| informações | 0.044 |
| usuário | 0.024 |
| conteúdo | 0.016 |
| 2 | dado | 0.103 |
| pessoais | 0.064 |
| informações | 0.042 |
| direitos | 0.018 |
| 3 | informações | 0.048 |
| privacidade | 0.046 |
| política | 0.037 |
| cookie | 0.030 |
| 4 | termos | 0.025 |
| usuário | 0.021 |
| segurança | 0.021 |
| utilização | 0.007 |

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

A **Tabela 8** exibida abaixo, retrata o conteúdo exibido nas Políticas de Privacidade após a análise de cada um dos *tokens* selecionados:

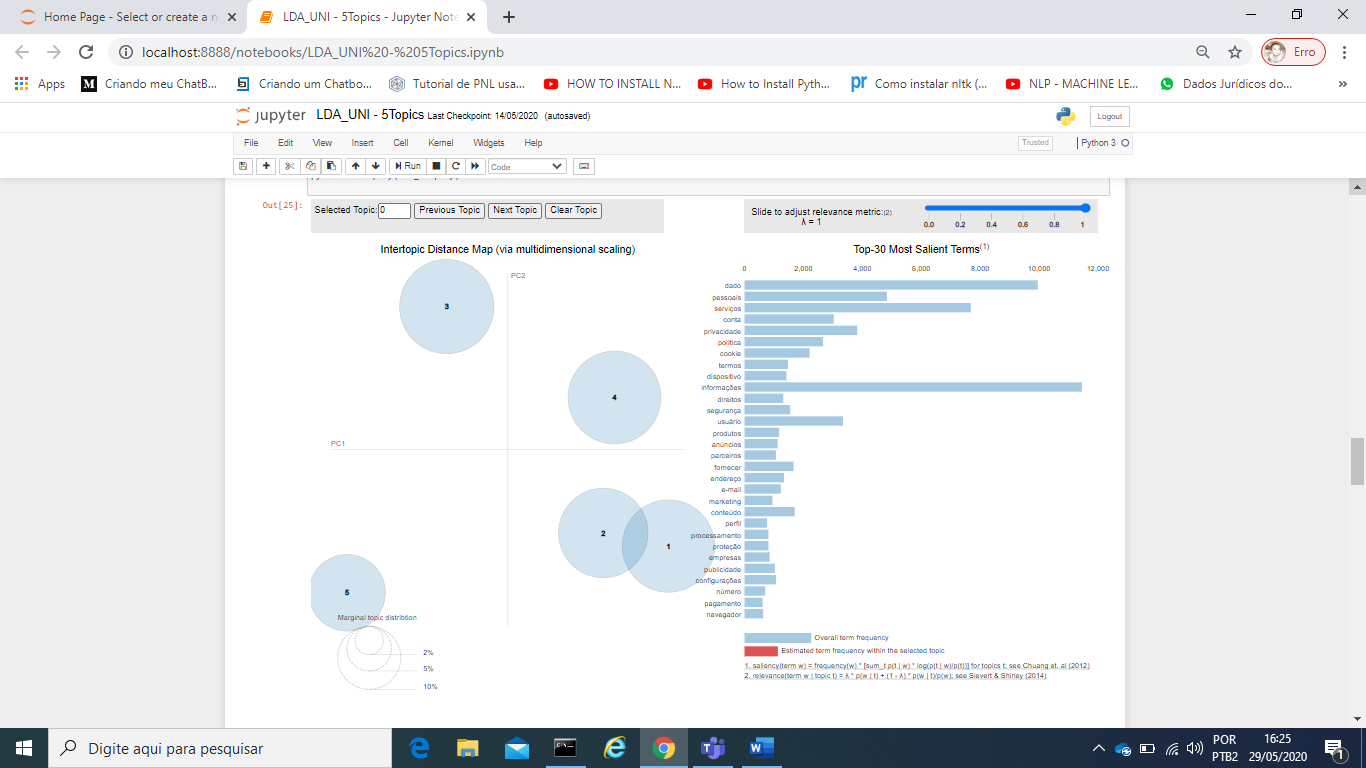
|  |  |  |
| --- | --- | --- |
| Tabela 8 – Análise das Políticas de Privacidade | | |
| Tópico | **Palavras** | **Análise** |
| 0 | '0.078\*"serviços" +  0.026\*"informações" +  0.018\*"fornecer" +  0.016\*"produtos"' | A exibição da palavra *‘serviços’* faz alusão ao acesso e compartilhamento de arquivos em *interfaces* de organizações terceirizadas, pode indicar também que para que os usuários tenham acesso a alguns recursos do *site* seja necessária a inclusão de dados pessoais através de formulários. Existe a utilização das palavras *‘serviços’* e *‘informações’*, ou seja, pode ser necessário que o usuário forneça algumas informações para que o *site, ‘forneça’* os *‘serviços’* solicitados. A palavra *‘produtos’* leva à utilização dos *‘serviços’* de suporte do *site*, que coleta os dados do usuário para futuramente exercer operações comerciais. |
| 1 | '0.044\*"conta" + 0.044\*"informações" + 0.024\*"usuário" + 0.016\*"conteúdo"' | A palavra *‘conta’* remete a confiabilidade das informações fornecidas pelo *‘usuário’*, onde, pode existir, a confirmação dos dados por telefone, *e-mail* ou outros meios. As *‘informações’* fornecidas pelo *‘usuário’* podem permitir que o acesso seja realizado em diversos dispositivos simultaneamente. Em alguns *sites* pode ser disponibilizado *‘conteúdo’* para parceiros ou terceirizados, assim como institutos acadêmicos e de pesquisa. |
| 2 | '0.103\*"dado" +  0.064\*"pessoais" +  0.042\*"informações" +  0.018\*"direitos"' | Ao analisar esse tópico destacamos as palavras *‘dados’*, *‘pessoais’* e *‘direitos’*. O tópico utiliza palavras voltadas para a LGPD. O tópico tem grande relevância para o estudo por apontar palavras muito utilizadas na redação da Lei. Apesar da proximidade com a legislação, as políticas encontram formas de impedir que o usuário transfira qualquer tipo de informação sem que haja o consentimento do *site*, além disso, existem políticas que de forma explícita informam ao usuário que ele não tem nenhum *‘direito’* sobre os conteúdos que possam ser oferecidos e contratados. |
| 3 | '0.048\*"informações" +  0.046\*"privacidade" +  0.037\*"política" +  0.030\*"cookie"' | Ao analisar o tópico surge a palavra *‘privacidade’* que é um dos elementos desse estudo. O contexto e utilização da palavra remete ao título ou a necessidade que os *sites* têm de informar que possuem uma Política de *‘Privacidade’*, salvo em algumas situações em que a redação informa que existe a preocupação com os dados nos termos de uso do *site.* 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. Além do uso de identificadores anônimos. |
| 4 | '0.025\*"termos" +  0.021\*"usuário" +  0.021\*"segurança" +  0.007\*"utilização"' | Esse tópico chama a atenção por exibir palavras como *‘termos’*, *‘segurança’* e *‘usuário’*. O indicativo de que existe uma preocupação com a segurança dos dados do usuário ao fazer seu cadastro. A palavra *‘utilização’* traz uma breve alusão a aceitação do usuário na utilização de *‘cookies’* em seu computador, além da forma como o usuário interage em redes sociais uns com os outros. A análise dentro do *corpus* demonstra que a palavra *‘utilização’* está ligada a publicidade com anunciantes, uma prática comum dentro dos termos de uso e políticas de privacidade. |

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

Em uma síntese geral, observamos que o algoritmo retornou itens considerados com grande relevância no *corpus* que reúne as Políticas de Privacidade, mas, não existe grande abrangência ou relevância no grupo de tópicos que faça uma redação mediana ao que sugere a Lei Geral de Proteção de Dados.

O **Gráfico 8**, multidimensional, exibe o grupo das 30 palavras mais salientes dentro do *corpus*, proporcionalidade e espaçamento entre os termos:

Gráfico 8 – Gráfico Multidimensional



Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

**Tópico 1 – 21,4% dos *tokens***

O tópico 1, baseado no gráfico multidimensional, traz os *tokens* voltados ao compartilhamento dos dados do usuário com outras empresas e parceiros terceirizados. É encontrado também palavras como *‘marketing’* e *‘anúncios’* que são direcionados para à atenção dos usuários com o objetivo de direcionamento em outros *sites,* podendo ofertar novos produtos ou serviços.

Segundo González (2019):

Utilizar os dados pessoais coletados para fins de e-mail marketing, por exemplo, entra dentro do legítimo interesse do controlador (conforme especificado no item I). Para essa e outras atividades de legítimo interesse, é importante que somente os dados estritamente necessários sejam utilizados. Assim, preserva-se tanto o legítimo interesse do controlador quanto a integridade dos dados do titular. (GONZÁLEZ, 2019)

**Tópico 2 – 20,1% dos *tokens***

Ao analisar esse tópico podemos inferir que ele faz a oferta dos seus serviços ao usuário que deve realizar seu cadastro, inserindo suas informações para ter acesso ao conteúdo do *site* ou *app*.

De acordo com Reuters (2019, p.3):

Nos últimos anos, o uso indevido de cadastros financeiros de consumidores provocou um número enorme de reclamações. Segundo levantamento feito pelo Instituto Brasileiro de Defesa do Consumidor (IDEC), queixas envolvendo problemas com transparência e uso inadequado de dados pessoais cresceram 1.134% entre 2015 e 2017. A principal reclamação, com 63% dos apontamentos, é referente à publicação, consulta ou coleta de dados pessoais sem autorização do consumidor. (REUTERS, 2019)

**Tópico 3 – 22,3% dos *tokens***

O terceiro tópico faz referência ao compartilhamento e a privacidade dos dados do usuário, pois ele tem direitos sobre seus dados, e para que seja repassado a terceiros, o usuário, deve dar consentimento para tal ação. O tópico é bem relacionado a LGPD.

Filho (2018, p. 2) relata que:

A LGPD impõe uma profunda transformação no sistema de proteção de dados brasileiro, em boa medida alinhada com a regulação europeia de proteção de dados (GDPR). É uma lei que estabelece regras detalhadas para a coleta, uso, tratamento e armazenamento de dados pessoais e afetará todos os setores da economia, inclusive as relações entre clientes e fornecedores de produtos e serviços, empregado e empregador, relações comerciais transnacionais e nacionais, além de outras relações nas quais dados pessoais sejam coletados, tanto no ambiente digital quanto fora dele. (FILHO, 2018)

**Tópico 4 – 21,6% dos *tokens***

O quarto tópico revela a utilização de *cookies*, indicando um possível rastreamento do que o usuário faz na *internet*. Essa palavra pode demonstrar que o usuário está dando a permissão ao portal, *site* ou *app*, para verificações de suas ações e muitas vezes sem o conhecimento daquilo que está aceitando. Ressalta a importância da Leitura da Política de Privacidade.

**Tópico 5 – 14,7% dos *tokens***

O quinto tópico indica ações voltadas a violação dos termos propostos pela plataforma, a responsabilidade legal, pode-se inferir que esse tópico direciona seus *tokens* ao respeito a redação da LGPD.

Na visão de González (2019):

[...] os termos de uso aos quais estamos acostumados, em que o uso dos dados pessoais aparece em meio a tantas outras cláusulas que quase nenhum usuário lê, não serão mais o bastante. É preciso investir em uma linguagem clara e acessível, garantindo que o titular saiba exatamente qual será o tratamento feito com seus dados pessoais. (GONZÁLEZ, 2019)

* + 1. **A MODELAGEM DE TÓPICOS NAS POLÍTICAS DE PRIVACIDADE – 10 TÓPICOS**

Ao realizar o processamento do *corpus* das Políticas de Privacidade, o algoritmo, preparado para o retorno de 10 tópicos com 4 palavras, obteve o seguinte resultado exibido na **Tabela 9**:

|  |  |  |
| --- | --- | --- |
| Tabela 9 – Resultado do LDA | | |
| Tópico | **Palavra** | **Valor** |
| 0 | dado | 0.183 |
| pessoais | 0.096 |
| direitos | 0.025 |
| proteção | 0.020 |
| 1 | serviços | 0.063 |
| conteúdo | 0.052 |
| produtos | 0.035 |
| melhorar | 0.022 |
| 2 | segurança | 0.052 |
| acesso | 0.018 |
| proteger | 0.018 |
| informações | 0.014 |
| 3 | informações | 0.266 |
| coletamos | 0.054 |
| pessoais | 0.045 |
| usamos | 0.023 |
| 4 | pagamento | 0.027 |
| dado | 0.021 |
| conta | 0.019 |
| usuário | 0.019 |
| 5 | cookie | 0.054 |
| dispositivo | 0.035 |
| dado | 0.025 |
| configurações | 0.020 |
| 6 | conta | 0.045 |
| informações | 0.027 |
| usuário | 0.027 |
| contato | 0.026 |
| 7 | serviços | 0.109 |
| terceiros | 0.051 |
| informações | 0.028 |
| dado | 0.026 |
| 8 | privacidade | 0.140 |
| política | 0.097 |
| termos | 0.036 |
| dado | 0.026 |
| 9 | serviços | 0.029 |
| marketing | 0.015 |
| personalizar | 0.014 |
| experiência | 0.012 |

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

A **Tabela 10** exibida abaixo, mostra o resultado da análise em cada um dos tópicos:

Tabela 9 – Resultado do LDA

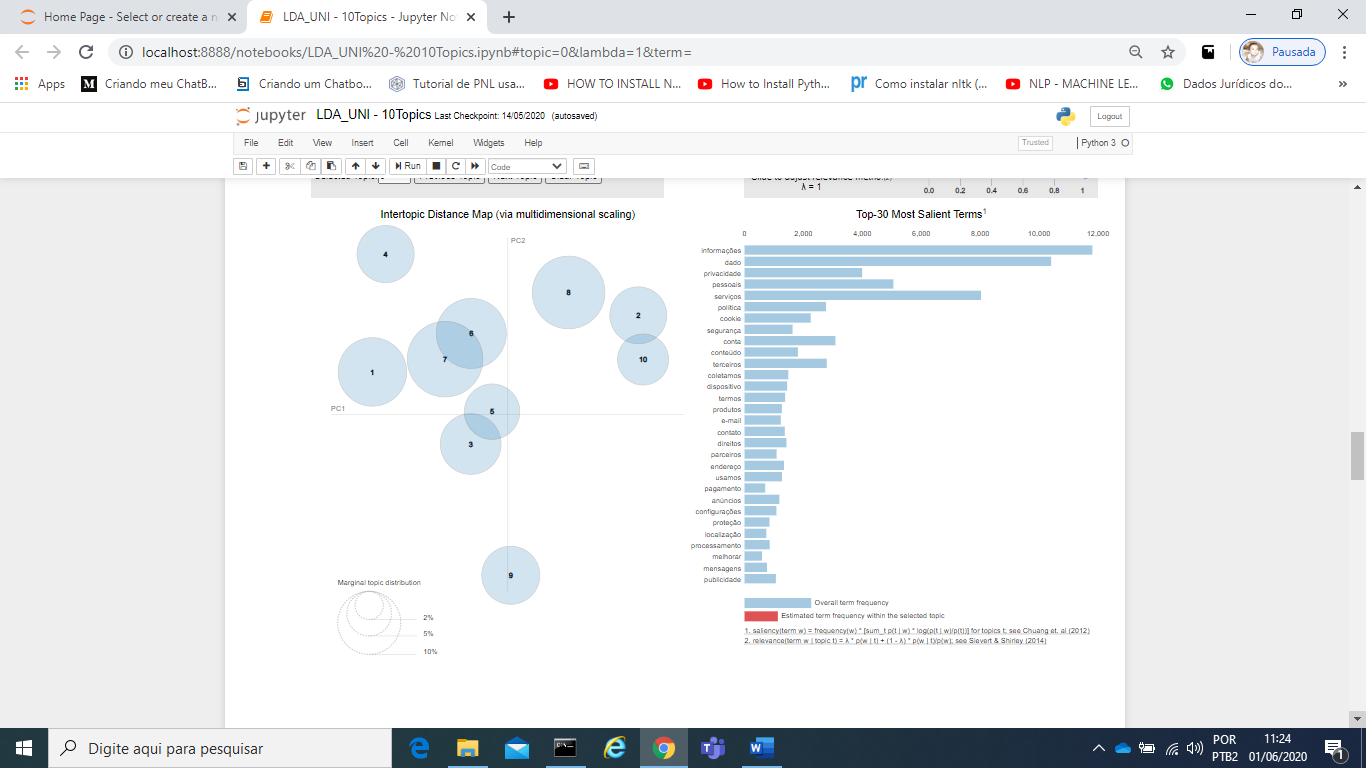
|  |  |  |
| --- | --- | --- |
| Tabela 10 – Análise dos Tópicos | | |
| Tópico | **Palavras** | **Análise** |
| 0 | '0.183\*"dado" +  0.096\*"pessoais" +  0.025\*"direitos" +  0.020\*"proteção"' | Ao analisar esse tópico destacamos as palavras *‘dados’*, *‘pessoais’* e *‘direitos’*. O tópico utiliza palavras voltadas para a LGPD. O tópico tem grande relevância para o estudo por apontar palavras muito utilizadas na redação da Lei. Apesar da proximidade com a legislação, as políticas encontram formas de impedir que o usuário transfira qualquer tipo de informação sem que haja o consentimento do *site*, além disso, existem políticas que de forma explícita informam ao usuário que ele não tem nenhum *‘direito’* sobre os conteúdos que possam ser oferecidos e contratados. A palavra *‘proteção’* geralmente faz alusão a Política de Privacidade do *site*, informando aos usuários as regras para a utilização e proteção dos dados pessoais. |
| 1 | '0.063\*"serviços" +  0.052\*"conteúdo" +  0.035\*"produtos" +  0.022\*"melhorar"' | A palavra *‘serviços’* demonstra na leitura das Políticas de Privacidade que o usuário pode ter as informações compartilhadas com terceiros, inclusive, outros usuários e empresas parceiras.  Dessa forma, Souza (2019, p. 23) esclarece que, *“A adequação de uma empresa aos critérios da LGPD é um processo complexo que envolve, principalmente, os ambientes de tecnologia, jurídico e governança. Seja um negócio ligado à produção de bens ou prestação de serviços, todos precisarão olhar para a segurança da informação e também revisar seus instrumentos contratuais observando cláusulas de tratamento de dados com clientes diretos, parceiros e colaboradores”.* (SOUZA, 2019)  A palavra *‘melhorar’* tem um vínculo com a palavra *‘cookie’*, já exibida em outro tópico. Dessa forma, as Políticas de Privacidade que o uso de *cookies* tem a finalidade de melhorar o acesso ao *‘conteúdo’*, a usabilidade e a experiência na utilização do *site*. |
| 2 | '0.052\*"segurança" +  0.018\*"acesso" +  0.018\*"proteger" +  0.014\*"informações"' | 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 *‘acesso’* no contexto das Políticas de Privacidade exibe uma forte pressão dos *sites* em coletar de forma automática dados como geolocalização, *ip* e registros de uso da plataforma.  De acordo com Filho (2018, p.17), *“A qualquer momento, o titular dos dados pessoais tem o direito de obter confirmação da existência de tratamento e acesso aos seus dados pessoais”.* (FILHO, 2018) |
| 3 | '0.266\*"informações" + 0.054\*"coletamos" + 0.045\*"pessoais" + 0.023\*"usamos"' | Esse tópico projeta ao usuário a coleta e utilização das informações do usuário. Não há indícios de solicitação de consentimento por parte do usuário.  Segundo Souza (2019, p. 5), sobre a LGPD é relatado que *“Com ela, empresas de todos os portes, públicas ou privadas, que coletam dados pessoais (sejam de clientes, fornecedores ou funcionários) devem estar de acordo com a legislação e o órgão regulatório será a ANPD (Autoridade Nacional de Proteção de Dados)”.* (SOUZA, 2019) |
| 4 | '0.027\*"pagamento" +  0.021\*"dado" +  0.019\*"conta" +  0.019\*"usuário"' | O tópico direciona para o *‘pagamento’* de taxas para a criação de uma conta para o usuário ou a utilização de serviços disponibilizados pelo *site*. |
| 5 | '0.054\*"cookie" + 0.035\*"dispositivo" + 0.025\*"dado" + 0.020\*"configurações"' | A análise desse tópico está voltada a utilização de *cookies*, que remete a verificar qual o comportamento do usuário na utilização do navegador. A palavra ‘configurações’ segue a linha de que o usuário pode fazer modificações em sua conta na hora que desejar, isso inclui a exibição ou não de anúncios. |
| 6 | '0.045\*"conta" + 0.027\*"informações" + 0.027\*"usuário" + 0.026\*"contato"' | O tópico remete a utilização das informações do usuário para algum tipo de contato, pode-se avaliar como *marketing*.  De acordo com Souza (2019, p. 7), *“A privacidade dos dados pessoais diz respeito à proteção das pessoas frente ao uso que é feito dessas informações, desde campanhas de marketing dirigidas, passando por autorização/negação de financiamento bancário ou ainda aceite (ou não) da participação em um plano de saúde”.* (SOUZA, 2019) |
| 7 | '0.109\*"serviços" + 0.051\*"terceiros" + 0.028\*"informações" + 0.026\*"dado"' | 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.  De acordo com González (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 punitivas”.* (GONZÁLEZ, 2019), |
| 8 | '0.140\*"privacidade" + 0.097\*"política" + 0.036\*"termos" + 0.017\*"alterações"' | O tópico direciona para que sejam realizadas alterações na política de privacidade sem que haja, possivelmente, a comunicação ao usuário.  Segundo Filho (2018, p. 21), *“Nas hipóteses em que o consentimento for exigido, o controlador deverá informar o titular caso haja alguma alteração na finalidade para a coleta de dados. Nesse momento, o titular poderá optar por renovar o consentimento ou revogá-lo.* (Filho, 2018) |
| 9 | '0.029\*"serviços" + 0.015\*"marketing" + 0.014\*"personalizar" + 0.012\*"experiência"' | O tópico releva a oferta de serviços para melhorar a experiência do usuário no ambiente. Uma estratégia de *marketing*.  De acordo com Prado (2019, p. 1), *“As atividades de marketing (online ou off-line) serão muito impactadas pela nova regulamentação de proteção de dados. O desafio será conciliar as tecnologias cada vez mais disruptivas com os princípios e bases legais da nova Lei, o que, com algumas cautelas, é plenamente viável. Na Europa, a maior multa da GDPR aplicada até agora foi de 50 milhões de euros por alegada falta de clareza sobre uso de dados para publicidade dirigida*”*.* (PRADO, 2019) |

Tabela 10 – Resultado do LDA

Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

O **Gráfico 10**, multidimensional, exibe o grupo das 30 palavras mais salientes dentro do *corpus*, proporcionalidade e espaçamento entre os termos:

Gráfico 10 – Gráfico Multidimensional



Fonte: Resultado de Algoritmo produzido pelo Autor em Análise *Python*

**Tópico 1 – 11,8% dos *tokens***

O tópico 1, baseado no gráfico multidimensional, traz os *tokens* voltados para a coleta e o consentimento do usuário com a finalidade da sua utilização para fins diversos. A palavra *‘proteção’* e *‘direitos’*, reforçam essa premissa. Esse tópico condiz plenamente com a redação da LGPD. Sobre a LGPD, Reuters (2019, p.7) esclarece que:

No Brasil, foram oito anos de discussão até se chegar à aprovação da Lei Geral de Proteção de Dados (LGPD), inspirada na GDPR, que foi sancionada em 14 de agosto de 2018. Aprovada em regime de urgência, por conta do mesmo incidente ocorrido com os dados de usuários do Facebook, a lei brasileira entra em vigor em 2020. A LGPD, assim como a GDPR, visa garantir maior controle dos cidadãos sobre suas informações pessoais. (REUTERS, 2019)

**Tópico 2 – 8,1% dos *tokens***

O tópico 2, baseado no gráfico multidimensional, traz os *tokens* voltados ao compartilhamento dos dados do usuário com outras empresas e parceiros terceirizados. São encontradas as palavras como *‘mídia’*, *‘serviços’* e *‘anúncios’* que são direcionadas para atrair à atenção dos usuários com o objetivo de fazer o direcionamento para *sites,* podendo ofertar novos produtos ou serviços. As empresas utilizam tecnologias de programação diversificadas com o objetivo de autenticar usuários, guardar preferências e configurações, verificar qual conteúdo visitado pelo usuário e verificar a força de campanhas de *marketing*.

**Tópico 3 – 9,3% dos *tokens***

A análise desse tópico demonstra que existe uma preocupação em se manter a segurança, evitar fraude, a violação e manter a proteção dos dados do usuário. O tópico tem muita relevância se comparado a redação da LGPD.

De acordo com González (2019), o Art. 11º da LGPD expõem que:

Em se tratando de dados sensíveis, um diferencial nas exceções é que, além de especificar a obediência às normas, a LGPD ainda destaca a prevenção à fraude e à segurança do titular. A cláusula é voltada especificamente para os cadastros digitais, ou seja, para o processo de *onboarding* digital. (GONZÁLEZ, 2019)

**Tópico 4 – 8,2% dos *tokens***

O tópico é um leva seu direcionamento para a utilização dos dados do usuário e a comercialização de produtos e serviços de parceiros.

Segundo Reuters (2019, p. 11), *“Cumprimento dos subcontratantes: a LGPD estende-se também aos subcontratantes de uma empresa, como fornecedores e parceiros de tecnologia. Eles também ficam sujeitos às obrigações e podem realizar pagamentos de indenização, por exemplo”.* (REUTERS, 2019)

**Tópico 5 – 7,7% dos *tokens***

O tópico traz palavras que direcionam ao entendimento para os serviços e produtos de uma empresa que podem ser prestados a um usuário. Direciona também para a faixa etária e as formas de pagamento pelos seus serviços.

De acordo como Reuters (2019, p. 8):

Para serviços da sociedade da informação, os pais ou responsáveis devem consentir na coleta e tratamento de dados de pessoas com até 16 anos de idade. Essa é uma inovação fundamental, pois a atividade de tratamento de dados no universo digital é caracterizada por ser ainda abstrata, pouco transparente e com alto grau de complexidade. Assim, a capacidade de observação e entendimento por crianças e adolescentes desse processo é extremamente limitada, sendo necessária uma mediação externa para equilíbrio desta relação comercial. (REUTERS, 2019)

**Tópico 6 – 12,3% dos *tokens***

A análise desse tópico está voltada 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. As empresas podem expor em suas políticas que podem ter colocado *cookies* de publicidade no seu computador, inferir que tem essa necessidade para que seus *sites* funcionem e ainda informar que coletam seus dados de uso através desse mecanismo.

**Tópico 7 – 14,3% dos *tokens***

O tópico direciona para o atendimento ao usuário para resolução de problemas, existe a palavra *‘excluir’*, que faz referência a um dos elementos da LGPD que é o direito a exclusão de seus dados em qualquer momento.

Nas palavras de González (2019) o Art. 6º promove:

Assim como a maior parte das leis, a LGPD prevê a boa-fé daqueles atingidos por ela. Isso é fundamental porque, quando falamos de certas regras da Lei — como a possibilidade de o titular solicitar a exclusão de seus dados ou um relatório completo de tratamentos —, nem sempre será possível fornecer provas absolutamente incontestáveis de que a Lei foi obedecida. (GONZÁLEZ, 2019)

**Tópico 8 – 13,3% dos *tokens***

O tópico 8 traz os *tokens* voltados ao compartilhamento dos dados do usuário com outras empresas e parceiros terceirizados. É encontrado também as palavras como *‘terceiros*, *‘serviços’* e *‘empresas’* que são direcionados para à atenção dos usuários com o objetivo de levá-los para outros *sites,* podendo ofertar produtos ou serviços.

Segundo a afirmação de Reuters (2019, p. 13):

Se houver mudança de finalidade ou repasse de dados a terceiros, um novo consentimento deverá ser solicitado. O usuário poderá, sempre que desejar, revogar a sua autorização, assim como pedir acesso, exclusão, portabilidade, complementação ou correção dos dados. (REUTERS, 2019)

**Tópico 9 – 8,5% dos *tokens***

O tópico direciona para que sejam 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 da plataforma. Em grande parte da Políticas observa-se a expressão: ‘*Esses termos podem ser modificados/alterados a qualquer momento, sem prévia comunicação’*.

**Tópico 10 – 6,5% dos *tokens***

O tópico releva a oferta de serviços para melhorar a experiência do usuário no ambiente. Uma estratégia de *marketing*. É exibida a palavra *‘europeia’*, que faz referência a *GDPR* - *General* *Database Practice Research*, ou seja, a legislação que direciona a proteção dos dados na União Europeia.

Segundo Reuters (2019, p. 5):

O GDPR é um projeto para proteção de dados e identidade dos cidadãos da União Europeia que começou a ser idealizado em 2012 e foi aprovado em 2016. Embora a região já tivesse leis relacionadas à privacidade, elas datavam de 1995 e, mesmo com atualizações, não correspondiam ao cenário tecnológico atual. (REUTERS, 2019)

1. **RESULTADOS E DISCUSSÕES**

Ao realizar o processamento da LGPD e do *corpus* das Políticas de Privacidade, observamos que, no que tange à similaridade entre termos existe um real disparate entre o que é solicitado pela legislação e o que é redigido pelas políticas de privacidade.

As Políticas de Privacidade têm uma inclinação importante no que diz respeito ao *marketing* e não existe de forma clara e objetiva, na maioria dos textos, um apontamento que direcione para a utilização dos dados dos usuários para fins diversos.

No comparativo relacionado ao Gráfico de Frequências de Palavras é observado um percentual de correspondência de 20%, ou seja, duas das dez palavras mais frequentes ocorrem nos dois gráficos. Esse número já cria uma distância entre a redação da LGPD e as Políticas de Privacidade.

No que diz respeito as análises com 5 tópicos as únicas palavras que são exibidas nos dois resultados são *‘dados’* e *‘pessoais’*, corroborando com a falta de alinhamento entre as Políticas de Privacidade e a LGPD. Quando a análise passa para uma amplitude maior, ou seja, dez tópicos, apenas as palavras *‘dados’*, *‘pessoais’* e *‘proteção’* surgem como palavras recorrentes, uma variação ainda muito baixa se comparado ao que é imposto pela redação da Lei e as Políticas de Privacidade.

Diante dos fatos apresentados e das análises realizadas em ambos os textos, podemos inferir que as Políticas de Privacidade estão em desacordo parcial com a redação da Lei Geral de Proteção de Dados, o que remete a uma preocupação – devido ao crescimento e exposição dos dados – em como serão tratados os dados dos usuários no Brasil.

1. **CONSIDERAÇÕES FINAIS**

O uso de dispositivos que atualmente dispõe da premissa direcionada da produção massiva de dados, e nos dias de hoje um hodierno modelo para os padrões da *Internet*, sendo lapidada em um grupo direcionado de novas e distintivas tecnologias em constante evolução. O acelerado desenvolvimento de autômatos conectados entre si, criam as mais diversas sensações na sociedade contemporânea. A privacidade e segurança dos dados do usuário ganham destaque e de fato, os mesmos, devem prestar à atenção devida em como suas informações são exploradas.

O iminente desafio em proteger os dados de usuários, vai de encontro com o revés econômico nesse momento de renovação tecnológica. É determinante que a legislação possa criar os 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.

1. **REFERÊNCIAS BIBLIOGRÁFICAS**

BIRD, Steven, Edward Loper e Ewan Klein (2009), Natural Language Processing with Python . O'Reilly Media Inc.

BRASIL. [Lei Nº. 13.709, de 14 de agosto de 2018.](http://legislacao.planalto.gov.br/legisla/legislacao.nsf/Viw_Identificacao/lei%2013.709-2018?OpenDocument) Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm>. Acesso em: 10 mar. 2020

COTS, Márcio. Impactos da Lei Geral de Proteção de Dados. Revista Conceito Jurídico. ZK Editora. 2018

FILHO, Demócrito Reinaldo. Lei de proteção de dados pessoais aproxima o Brasil dos países civilizados. Disponível em: <http://www.gandramartins.adv.br/project/ives-gandra/public/uploads/2019/02/21/22f3605085417p.pdf>. Acesso em: 08 jan. 2020

FILHO, Mattos. Guia para a Lei geral de proteção de dados. Disponível em: <https://www.legiscompliance.com.br/images/pdf/cartilha_lgpd_mattosfilho.pdf>. Acesso em: 11 jun. 2020

GONZÁLEZ, Mariana (2019). LGPD Comentada. Disponível em: <https://guialgpd.com.br/lgpd-comentada/>. Acesso em: 09 jun. 2020

GROWTH ENABLER (2017). Market pulse report, Internet of Things (IoT). Disponível em: https://growthenabler.com/reports/IOT.html. Acesso em: 06 jan. 2020.

MACIEL, Rafael Fernandes. Manual Prático sobre a Lei Geral de Proteção de Dados Pessoais (Lei nº 13.709/18). RM Digital Education. 1ª Edição. Goiânia – GO. 2019. ISBN: 9781093409420

NOIRET, Serge. “História Pública Digital”. Liinc em Revista. ​11, n. 1 (2015)

SCHNEIER, Bruce. Data and Goliath – The Hidden Battles to Collect Your Data and Control Your World;

HARTZOG, Woodrow. Privacy’s Blueprint: The Battle to Control the Design of New Technologies.

PINHEIRO, Patricia Peck. Proteção Dados Pessoais: Comentários à Lei N.º 13.709/2018 (LGPD)/Patricia Peck Pinheiro – São Paulo: Saraiva Educação, 2018.

PRADO, Luis Fernando. LGPD no Marketing. Disponível em: <https://www.daniel-ip.com/wp-content/uploads/2019/07/Infogr%C3%A1fico-LGPD-no-Marketing-002.pdf>. Acesso em: 15 jun. 2020

#### REANI, Valéria. Impactos da Lei Geral de Proteção de Dados para os negócios e as pessoas. Disponível em: <https://www.conjur.com.br/2018-out-25/valeria-reani-impactos-lei-protecao-dados-negocios>. Acesso em: 09 jun. 2020

REUTERS, Thomson. WHITEPAPER. Lei Geral de Proteção de Dados: impactos e mudanças no uso e na coleta de dados pessoais. Disponível em: <https://www.thomsonreuters.com.br/content/dam/openweb/documents/pdf/Brazil/white-paper/thomson-reuters-legal-whitepaper-lei-geral-de-protecao-de-dados.pdf>. Acesso em: 12 jun. 2020

SOUZA, Diego. Vexia. Adequação a LGPD impulsiona procura por prestação de serviços para adequação a LEI Nº 13.709/18. Disponível em: <https://rdstation-static.s3.amazonaws.com/cms%2Ffiles%2F28555%2F1581450471Ebook_LGPD_-_O_que__e_como_adequar_sua_empresa.pdf>. Acesso em: 15 jun. 2020

REVISTA VEJA, online. EUA: governo vasculha dados de fontes como Google e Facebook. Disponível em: <https://veja.abril.com.br/mundo/eua-governo-vasculha-dados-de-fontes-como-google-e-facebook/>. Acesso em: 06 jan. 2020

VAINZOF, Rony. Finalmente a Lei Geral de Proteção de Dados (LGPD) – Resumo dos pontos relevantes. Revista Conceito Jurídico. ZK Editora. 2018

1. Mestrando em Humanidades Digitais pela UFRRJ

   ²

   ³ [↑](#footnote-ref-1)