

A JUDICIALIZAÇÃO DA PANDEMIA DE COVID-19 NO BRASIL: ESTUDO DE CASO SOBRE A APLICAÇÃO DO MÉTODO DE MODELAGEM DE TÓPICOS PARA AGRUPAMENTO DE DOCUMENTOS JURÍDICOS¹

THE JUDICIALIZATION OF THE COVID-19 PANDEMIC IN BRAZIL: A CASE STUDY ON THE APPLICATION OF TOPIC MODELING FOR GROUPING LEGAL DOCUMENTS

Rodrigo de Souza Tavares²
Letícia Lobato Anicet Lisboa³
Renan Procópio Duarte⁴
Bruno Oliveira Valverde⁵

Resumo: O presente artigo pretende realizar uma investigação exploratória e empírica a respeito da judicialização da epidemia de Covid-19 no Brasil. Para realizar este objetivo foi criado um *corpus* de documentos jurídicos sobre o tema, que posteriormente passou pelo escrutínio de ferramentas de processamento de linguagem natural e linguística computacional. Ao fim da análise, encontrou-se grande predominância das discussões relativas à gestão da população dos sistemas prisionais e socioeducativos, sobretudo em função da grande repercussão alcançada pela Recomendação nº 62/2020 do CNJ. O estudo aponta para futuras linhas de investigação sobre o tema dos impactos da Covid-19 no sistema judiciário brasileiro e sinaliza o potencial existente na utilização da abordagem computacional para análise de grandes quantidades de dados jurídicos.

Palavras-chave: Judicialização; Covid-19; Processamento de Linguagem Natural; Linguística de Corpus; Modelagem de Tópicos.

Abstract: This article intends to carry out an exploratory and empirical investigation regarding the judicialization of the Covid-19 epidemic in Brazil. To achieve this objective, a corpus of legal documents on the subject was created, which subsequently went through the scrutiny of natural language processing and computational linguistic tools. At the end of the analysis, there was a great predominance of discussions regarding the management of the population of the prison and socio-educational systems, mainly due to the great repercussion achieved by Recommendation No. 62/2020 of the CNJ. The study points to future lines of investigation on the subject of the impacts of Covid-19 on the Brazilian judicial system and signals the potential that exists in using the computational approach to analyze large amounts of legal data.

Keywords: Judicialization; Covid-19; Natural Language Processing; Corpus Linguistics; Topic Modeling.

¹ Artigo submetido em 21/08/2020 e aprovado para publicação em 21/09/2020.

² Professor do Departamento de Ciências Jurídicas e do Programa de Pós-Graduação Interdisciplinar em Humanidades Digitais da UFRRJ. Doutor pela PUC-Rio. <https://orcid.org/0000-0002-7854-461X>.

³ Professora do Departamento de Ciências Jurídicas do Instituto Multidisciplinar da UFRRJ. Advogada. Doutora pela UERJ. <https://orcid.org/0000-0002-4627-0423/>.

⁴ Mestrando do Programa de Pós-Graduação Interdisciplinar em Humanidades Digitais da UFRRJ. Graduado em Ciência da Computação pela UFRRJ. <https://orcid.org/0000-0001-5537-4470>.

⁵ Mestrando do Programa de Pós-Graduação Interdisciplinar em Humanidades Digitais da UFRRJ. Graduado em Direito pela UNESA. <https://orcid.org/0000-0001-8737-2293>.

Introdução

No momento em que este artigo é escrito, mais de seiscentas mil pessoas em todo mundo morreram vítimas da COVID-19. No Brasil, o número de mortos passa de noventa mil e continua crescendo, de forma acelerada⁶. Medidas desencontradas nas políticas públicas, sobretudo nas áreas da saúde e economia, bem como uma gestão caótica da crise sanitária, nos colocaram no segundo lugar, atrás apenas dos E.U.A., no ranking dos países com maior número de vítimas da pandemia da COVID-19.

Essa tragédia projeta efeitos profundos em todas as direções, como ondas de choque que seguem ao impacto de um asteroide na terra. O sistema capitalista mundial vive a maior crise econômica do século. As fronteiras mundiais foram fechadas e a circulação de bens e pessoas limita-se ao mínimo necessário. Experiências de controle e monitoramento dos corpos e dados dos indivíduos prosperam em países com distinto grau de apreço à democracia. No Brasil, a pandemia age como catalizador de uma grave crise das instituições democráticas.

Contudo, ainda que todos sintamos os tremores dessa colisão descomunal, a distribuição dos ônus deste processo parece seguir a lógica desigual que perpassa nossas estruturas sociais. A tragédia pode ter origem natural, com o aparecimento de uma nova cepa do vírus sars-cov-2, mas a distribuição das vulnerabilidades frente à pandemia tem esteio em iniquidades sociais, políticas e econômicas enraizadas.

As instituições jurídicas não estão imunes nesta conjuntura, pelo contrário, são engrenagens ativas nos mecanismos de gestão e distribuição dos riscos decorrentes da epidemia de coronavírus. Desde a declaração pela Organização Mundial da Saúde da pandemia de COVID-19, diversos atos administrativos, leis e decisões judiciais buscam coordenar a vida em meio às circunstâncias extraordinárias.

Para o pesquisador do direito, é um desafio imenso tentar compreender, em tempo real, o funcionamento dessas dinâmicas sociais complexas. O volume das fontes e a velocidade dos acontecimentos atropelam a capacidade analítica e cognitiva daqueles que se dedicam a esta tarefa hercúlea. Não obstante essa dificuldade, estudos amplos de caráter qualitativo já foram realizados, apresentando recortes pertinentes sobre os desdobramentos jurídicos da pandemia nos contratos civis (ROSENVALD et. al, 2020), no direito

⁶ Painel da COVID-19 – OMS. Disponível em: <<https://covid19.who.int/>> Acesso em: 04/08/2020

administrativo (JUSTEN FILHO et. al., 2020) e nas relações de trabalho (POSSÍDIO e MARTINEZ, 2020).

Neste artigo pretendemos abordar o problema por outra perspectiva. Ao invés de analisar temas específicos, nosso objetivo é capturar numa grande angular o cenário da judicialização⁷ da pandemia no âmbito brasileiro. Propomos um afastamento metodológico para descrever, em grandes linhas, a maneira pela qual os conflitos atinentes aos impactos da COVID-19 foram transpostos para o Poder Judiciário.

Para cumprir esse objetivo, desenvolvemos uma aplicação em linguagem *Python* e usamos softwares de análise linguística para processar uma quantidade significativa de documentos jurídicos. Essa proposta de análise encontra-se na confluência de distintas tradições de pesquisa.

O uso de métodos quantitativos para realizar investigações em objetos normalmente afetos às ciências hermenêuticas, tais como textos literários, faz parte do que se convencionou chamar de linguística de corpus (SARDINHA, 2000). Essa leitura distanciada de textos visa normalmente encontrar padrões recorrentes em amostras extensas de fontes, ao invés de focar na tradicional leitura detida e na interpretação de documentos isolados (MORETTI, 2013). Após a formação de um *corpus* (grande conjunto de textos ou registros orais), alguns pesquisadores têm empregado ferramentas estatísticas, para obter conclusões a partir de parâmetros como frequência, distribuição e colocação de vocábulos.

Nos últimos anos, o campo da linguística de corpus incorporou recursos de inteligência artificial, notadamente da subárea do processamento de linguagem natural, para potencializar suas práticas. O emprego de métodos computacionais permite que quantidades massivas de documentos sejam analisados, em volume superior às capacidades humanas. Além disso, a detecção de padrões em grandes volumes de dados é tarefa na qual as máquinas apresentam performance excepcional. Obviamente, a forma pelas quais os computadores “leem” documentos difere essencialmente da leitura humana, oferecendo um modelo alternativo, que complementa a tarefa hermenêutica tradicional, ao invés de suplantá-la⁸.

⁷ Utilizamos o termo judicialização neste trabalho em sentido bastante amplo, equivalente a possibilidade de tradução de conflitos sociais e políticos em demandas jurídicas (BARROSO, 2018). Para uma visão crítica sobre o fenômeno da judicialização da esfera pública no Brasil ver: “A Judicialização da Política e das Relações Sociais no Brasil” (VIANNA, 1999). Para uma visão global ver: “The Judicialization of Politics. A World-Wide Phenomenon” (VALLINDER, 1994).

⁸ Para uma excelente introdução ao tema do processamento de linguagem natural aplicado ao campo jurídico e algumas reflexões filosóficas implicadas nesta abordagem, ver: “Ensinando um Robô a Julgar” (BOEING; ROSA, 2020).

Por fim, a digitalização de documentos jurídicos e sua ampla divulgação pela internet, quando associadas aos métodos da linguística de corpus, processamento de linguagem natural e mineração de textos, permite novas formas de pesquisa empírica no direito. Esse amálgama metodológico pode ser enquadrado dentro do campo da análise computacional do direito - *computational legal analysis* (LIVERMORE, ROCKMORE, 2019).

Em síntese, neste artigo iremos apresentar um estudo descritivo e quantitativo, realizado a partir da compilação e análise da ementa de cerca de dez mil decisões judiciais, abordando o problema da COVID-19. Esses dados, oriundos de diversos tribunais brasileiros, foram coletados e examinados através de dispositivos automatizados de sumarização e análise de documentos. A seguir, detalhamos a metodologia utilizada e os resultados obtidos.

1. Metodologia

O método utilizado no presente trabalho envolveu análise de dados com base em decisões judiciais. Após a obtenção dos resultados, foi necessária análise de legislação e doutrina relacionada àqueles produtos. Dessa forma, o presente capítulo versa sobre os processos utilizados na pesquisa, uma vez que foram fundamentais para a compreensão dos dados obtidos.

1.1. Criação do Corpus

Conforme mencionamos, um *corpus* pode ser definido como uma coleção de textos expressos em certa linguagem. Um corpus deve representar uma amostra da linguagem que se pretende analisar. Normalmente, ele é construído em algum suporte digital, para possibilitar sua manipulação através de mecanismos computacionais (ALUISIO; BARCELOS, 2006).

Em nosso trabalho reunimos um *corpus* com ementas de decisões judiciais a respeito do tema da pandemia da Covid-19. Para estruturá-lo criamos uma rotina em linguagem Python, que realiza uma busca pelo termo Covid-19 no site *Jusbrasil*⁹ e captura os resultados obtidos. Este site apresenta entre suas funcionalidades um *webcrawler* que realiza pesquisa por jurisprudência, de forma automatizada e simultânea, em muitos sites de tribunais

⁹ <https://www.jusbrasil.com.br/>

brasileiros¹⁰. Os resultados foram coletados e guardados num arquivo de texto para posterior análise. Ao total, foram reunidos dez mil, cento e trinta e cinco (10.135) registros, através da busca de jurisprudência pelo termo Covid, cada um equivalente a uma ementa de decisão judicial.

1.2. Pré-processamento do texto

Após a confecção do *corpus*, realizou-se o pré-processamento do texto. Esta etapa é muito importante nos trabalhos de mineração de textos que consiste em procedimentos técnicos de tratamento e limpeza dos dados, necessários para transformar o texto original num conjunto de caracteres apropriados ao processamento de linguagem natural por máquinas.

O primeiro procedimento realizado foi a tokenização do texto, ou seja, a transformação das palavras e sentenças do documento em unidades compreensíveis para as ferramentas de processamento de linguagem natural que serão aplicadas a seguir. Para entender esta tarefa, basta imaginar que computadores enxergam documentos de texto apenas como sequências de caracteres, separadas por espaços. Portanto, é preciso ensinar de alguma forma à máquina onde começam e terminam palavras e sentenças. Acidentes como junção de palavras e pontuação, hifenização e uso de caracteres especiais são questões facilmente identificáveis por humanos durante a leitura, mas representam um desafio para as máquinas que realizam processamento da linguagem natural.

Ademais, outra etapa importante consiste na remoção das chamadas *stopwords*. Estas são palavras muito frequentes, tais como preposições e artigos, que normalmente não agregam informações relevantes para a construção de modelos analíticos. Para facilitar o processamento e visualização dos vocábulos relevantes, as *stopwords* são usualmente retiradas do documento antes da criação do modelo. A ferramenta que empregamos para a tarefa foi a biblioteca NLTK, da linguagem *Python*, que dispõem de pacote específico de *stopwords* para a língua portuguesa. Todavia, após uma análise prévia dos resultados,

¹⁰ A ferramenta do site de busca do site Jusbrasil faz remissão aos seguintes Tribunais e órgãos públicos: STFJ, STJ, TSE, TST, STM CNJ, TCU, TRF-1, TRF-2, TRF-3, TRF-4, TRF-5, TRE-AC, TRE-AL, TRE-AM, TRE-AP, TRE-BA, TRE-CE, TRE-DF, TRE-ES, TRE-GO, TRE-MA, TRE-MG, TRE-MS, TRE-MT, TRE-PA, TRE-PB, TRE-PE, TRE-PI, TRE-PR, TRE-RJ, TRE-RN, TRE-RO, TRE-RR, TRE-RS, TRE-SC, TRE-SE, TRE-SP, TRE-TO, TRT-1, TRT-2, TRT-3, TRT-4, TRT-5, TRT-6, TRT-7, TRT-8, TRT-9, TRT-10, TRT-11, TRT-12, TRT-13, TRT-14, TRT-15, TRT-16, TRT-17, TRT-18, TRT-19, TRT-20, TRT-21, TRT-22, TRT-23, TRT-24, TJ-AC, TJ-AL, TJ-AM, TJ-AP, TJ-BA, TJ-CE, TJ-DF, TJ-ES, TJ-GO, TJ-MA, TJ-MG, TJ-MS, TJ-MT, TJ-PA, TJ-PB, TJ-PE, TJ-PI, TJ-PR, TJ-RJ, TJ-RN, TJ-RO, TJ-RR, TJ-RS, TJ-SC, TJ-SE, TJ-SP, TJ-TO, TCE-BA, TCE-MG, TCE-MS, TCE-PE, TCE-PR, TJM-MG, TJM-SP.

percebeu-se a necessidade de incorporar ao conjunto das *stopwords* pertencentes à biblioteca NLTK uma série de palavras próprias da linguagem jurídica que, embora muito frequentes no corpus, não adicionam qualquer potencial explicativo para análise do modelo construído.

1.3. Modelagem dos tópicos

Ultrapassada a fase de tratamento e limpeza dos dados, realizamos o processamento destes. Para tanto, empregamos um modelo estatístico visando agrupar os registros do corpus em um número pré-definido de tópicos, gerados de acordo com a frequência e distribuição das palavras - técnica conhecida como modelagem de tópicos.

Um tópico consiste num padrão de palavras que costumam ocorrer conjuntamente, ao longo dos vários documentos que compõem o corpus. Podemos ilustrar seu funcionamento a partir de uma metáfora. Imagine um leitor munido de diferentes canetas marca texto coloridas. Em cada documento analisado ele marca palavras-chave de diferentes assuntos, usando diferentes cores. Ao final da análise de vários documentos, cada cor de marca texto irá agrupar palavras-chave frequentemente repetidas nos vários documentos, em torno dos mesmos assuntos, esses conjuntos podem ser chamados de tópicos (BLEI,2012).

O modelo empregado neste trabalho realiza operação similar, de forma automatizada. A modelagem de tópicos com o uso de Alocação Latente de Dirichlet — LDA, enquadra-se entre os modelos não-supervisionados de aprendizado de máquina (*machine learning*). Em outras palavras, fornecemos uma entrada com os documentos (ementas de casos relativos à Covid-19) e informamos o número de tópicos nos quais deverão ser agrupados os documentos. O algoritmo não tem nenhum critério semântico para separar os assuntos. Ele trata cada documento como uma distribuição probabilística de palavras, em comparação à distribuição encontrada no corpus. Como retorno obtemos as palavras prevalentes em cada tópico, um gráfico representando o volume de documentos por tópico e a similaridade ou distância entre tópicos.

Atualmente é grande o interesse pela modelagem de tópicos como forma de descobrir a organização de assuntos em grandes coleções de texto, tais como coleções de periódicos médicos, interações em redes sociais, discursos políticos, entre outros temas (JELODAR; et. al., 2019). A abordagem parece pertinente para tratar do problema proposto, qual seja, investigar quais os temas decorrentes da judicialização da pandemia da Covid-19 no Brasil, a partir da análise de um grande número de ementas de decisões judiciais.

1.4. Análise linguística e avaliação qualitativa

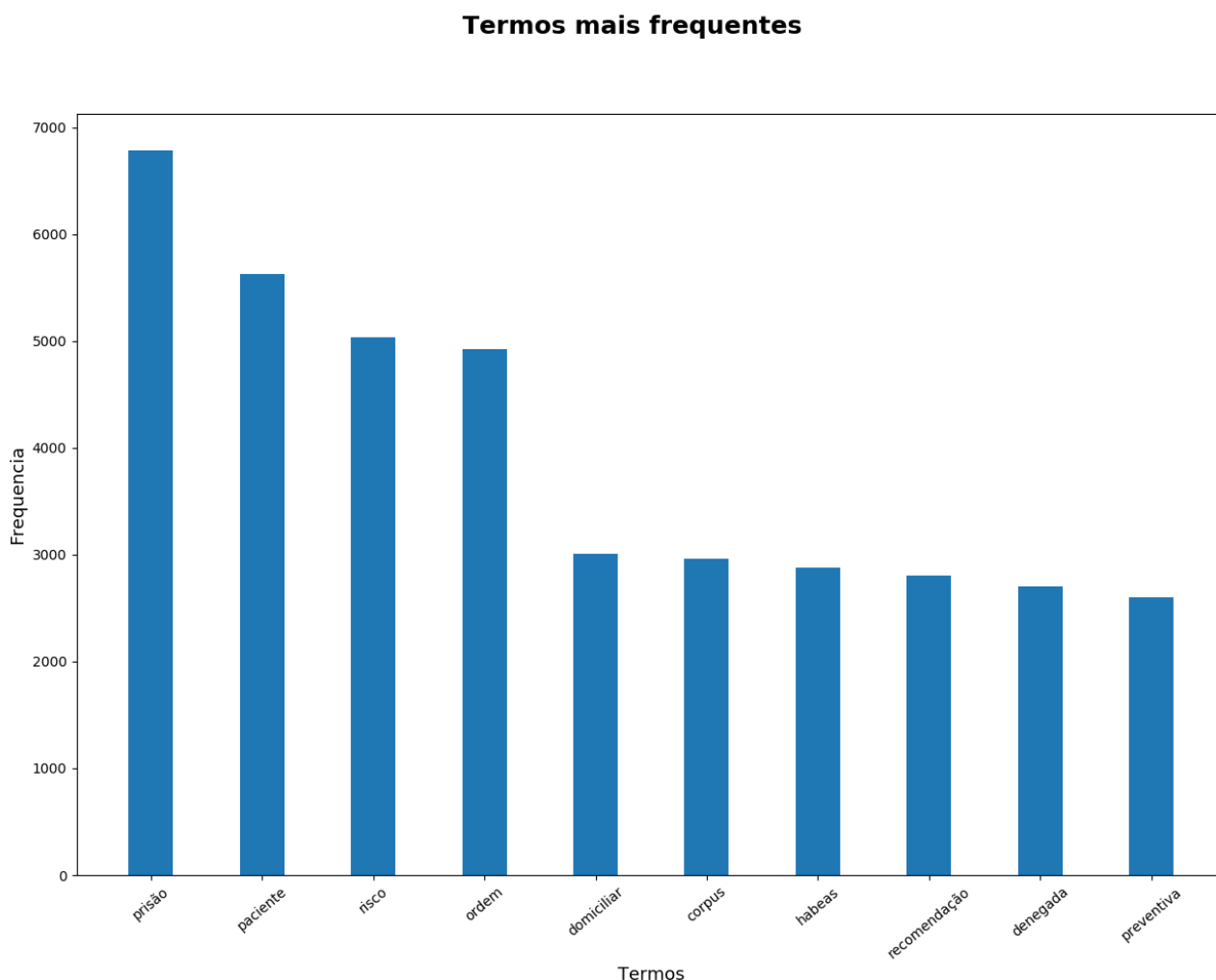
Por ser um método não-supervisionado de aprendizado de máquina, os resultados obtidos pela modelagem de tópicos dependem em grande medida da interpretação por especialistas com conhecimentos sobre o domínio ao qual se refere o corpus analisado. A partir de uma saída com as palavras-chave mais frequentes em cada tópico, cabe ao intérprete inferir a melhor explicação para aglutinar aquele conjunto de palavras, em suma, qual o tema subjacente ao tópico representado pelo modelo automatizado.

Para facilitar a interpretação dos tópicos, empregamos o seguinte procedimento. Com o auxílio da ferramenta de pesquisa linguística Lancsbox realizamos buscas para verificar as colocações das palavras-chave de cada tópico. O software permite que sejam visualizadas quais são as combinações preferenciais de termos escolhidos, no âmbito do corpus. Dessa forma, pudemos investigar os termos que circundam as palavras-chave e com isso obter uma melhor compreensão do contexto de cada tópico, conforme será apresentado a seguir.

2. Resultados

O primeiro resultado obtido após o processamento dos dados foi uma listagem com os dez termos mais frequentes na íntegra do corpus, já excluídas as palavras encontradas entre as chamadas *stopwords*. A lista pode ser conferida no gráfico abaixo:

Figura 1- Termos mais frequentes



Em seguida, apresentamos gráficos onde os documentos são separados em seis (6) e oito (8) tópicos. Os parâmetros com o número de tópicos são pré-estabelecidos no script *Python* elaborado para processamento dos dados. De forma geral, um número menor de tópicos gera resultados mais coerentes, porém, menos específicos. Na literatura sobre o tema, indicam-se duas vias para aferir a coerência dos tópicos: podem ser adotadas métricas estatísticas ou podemos submeter os tópicos gerados à avaliação por seres-humanos. Intérpretes com familiaridade na temática do Corpus podem estabelecer conexões de sentido e utilidade para a saída gerada pelo modelo estatístico (BOYD-GRABER et al., 2015). Optamos neste trabalho por trabalhar com os dois sistemas. Realizamos uma análise estatística para obter o número ideal de tópicos, visando coerência, a partir do corpus elaborado. O resultado de sucessivas iterações demonstrou a maior média de coerência para a estrutura de seis (6) tópicos e o maior valor absoluto para uma única iteração na estrutura de

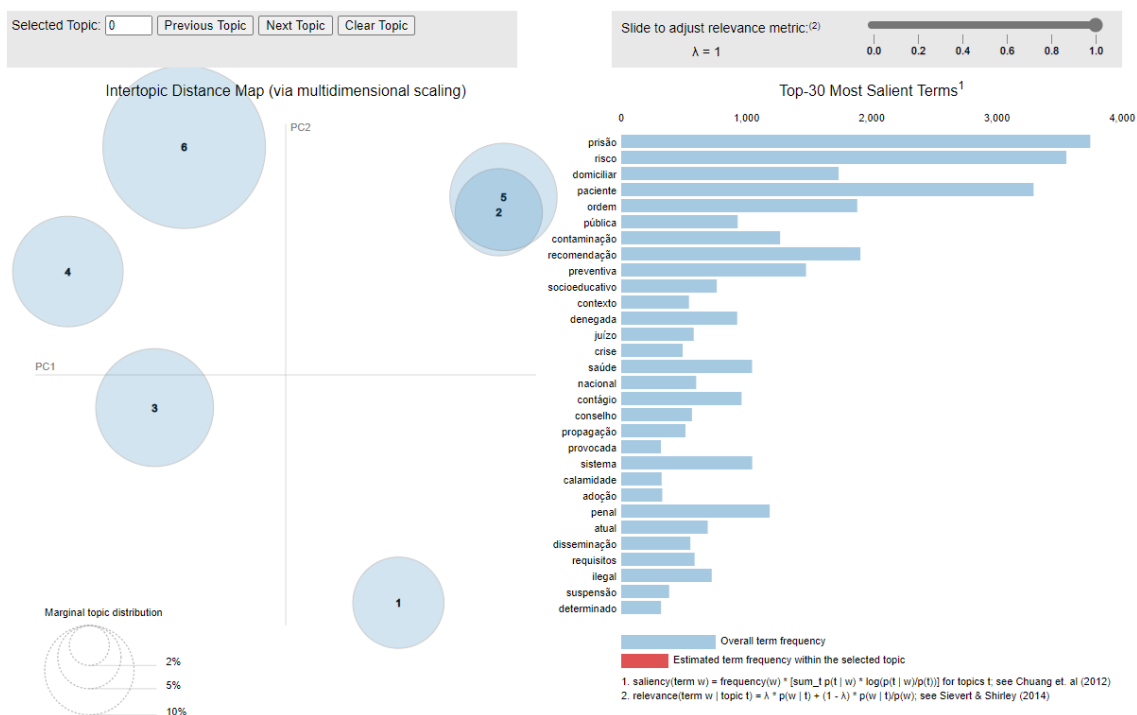
oito (8) tópicos. Posteriormente, realizamos análise manual do corpus, empregando software de análise linguística, para interpretar o escopo de cada tópico. Segue abaixo tabela com os resultados dos testes de coerência no intervalo de três (3) a dez (10) tópicos.

Figura 2 – Resultados de testes de coerência de três (3) a dez (10) tópicos

| Número de tópicos | 1ª Execução | 2ª Execução | 3ª Execução | 4ª Execução | 5ª Execução | 6ª Execução | 7ª Execução | 8ª Execução | 9ª Execução | 10ª Execução | Média |
|-------------------|-------------|-------------|------------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|--------------|--------|
| 3 | 0.62 | 0.51 | 0.48 | 0.52 | 0.56 | 0.55 | 0.58 | 0.52 | 0.56 | 0.53 | 0.5439 |
| 4 | 0.54 | 0.58 | 0.56 | 0.60 | 0.58 | 0.54 | 0.56 | 0.56 | 0.55 | 0.57 | 0.5627 |
| 5 | 0.59 | 0.53 | 0.55 | 0.47 | 0.53 | 0.56 | 0.53 | 0.56 | 0.60 | 0.51 | 0.5426 |
| 6 | 0.59 | 0.51 | 0.57 | 0.59 | 0.57 | 0.56 | 0.56 | 0.59 | 0.56 | 0.58 | 0.5677 |
| 7 | 0.57 | 0.55 | 0.54 | 0.57 | 0.57 | 0.53 | 0.54 | 0.57 | 0.52 | 0.55 | 0.5510 |
| 8 | 0.55 | 0.53 | 0.52 | 0.53 | 0.62 | 0.53 | 0.56 | 0.57 | 0.54 | 0.56 | 0.5495 |
| 9 | 0.53 | 0.49 | 0.59 | 0.53 | 0.53 | 0.54 | 0.52 | 0.56 | 0.52 | 0.49 | 0.5301 |
| 10 | 0.52 | 0.51 | 0.55 | 0.51 | 0.51 | 0.51 | 0.50 | 0.52 | 0.57 | 0.52 | 0.5231 |
| | | | | | | | | | | | |
| | | | 8 Tópicos, melhor valor individual | | | | | | | | |
| | | | 6 Tópicos, melhor valor em media | | | | | | | | |

Os gráficos gerados pelo pacote *Gensim*, através da utilização do algoritmo de LDA, apresentam um número de tópicos selecionados. A representação espacial com o tamanho e distância entre tópicos, bem como as trinta (30) palavras que mais se destacaram em cada tópico. Infelizmente, a ilustração colacionada adiante neste artigo não permite oferecer ao leitor a experiência interativa proporcionada pelo modelo, isto pois, a visualização gerada em HTML permite navegar entre cada tópico e verificar as palavras mais relevantes.

Por fim, apresenta-se uma lista com as quatro (4) palavras-chave mais salientes em cada um dos tópicos, conforme figura abaixo:

Figura 3 - Modelo com seis (6) tópicos e quatro (4) palavras-chave

(0, '0.025*"pública" + 0.023*"saúde" + 0.018*"propagação" + 0.017*"crise"')

(1, '0.018*"atual" + 0.013*"liminar" + 0.011*"sistema" + 0.010*"julgado"')

(2, '0.035*"risco" + 0.024*"paciente" + 0.018*"recomendação" + 0.018*"contaminação"')

(3, '0.054*"risco" + 0.052*"prisão" + 0.041*"domiciliar" + 0.040*"paciente"')

(4, '0.017*"juízo" + 0.013*"suspensão" + 0.010*"domiciliar" + 0.009*"prisao"')

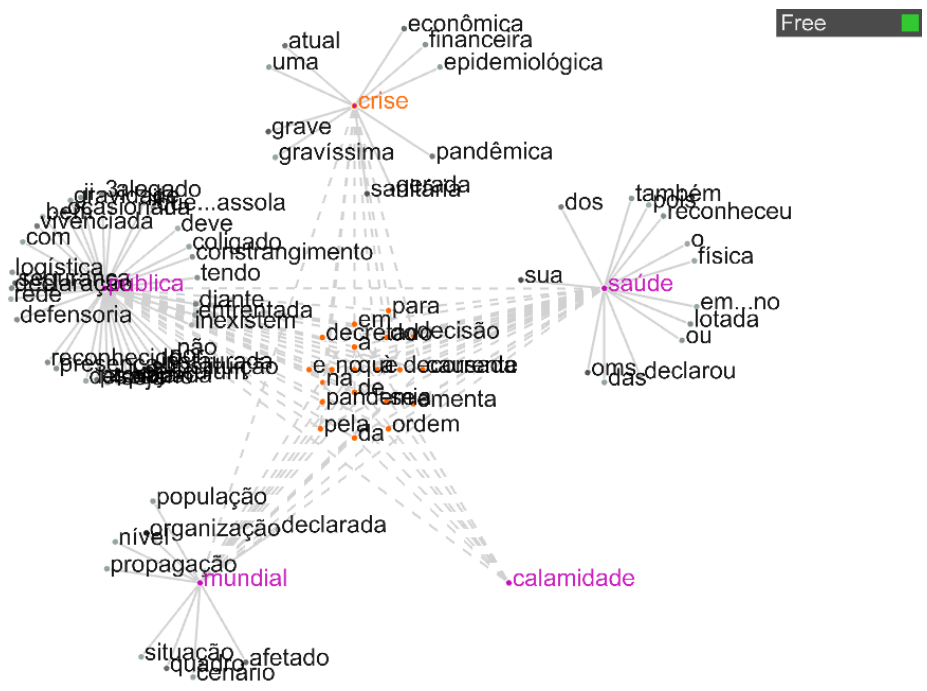
(5, '0.035*"prisão" + 0.029*"ordem" + 0.020*"paciente" + 0.019*"preventiva"')

3. Análise qualitativa dos resultados

Conforme verificou-se anteriormente, a pesquisa de dados do presente trabalho utilizou a ferramenta Lancsbox. Esse programa de computador realiza a análise linguística de dados e foi desenvolvida por pesquisadores da Lancaster University. O método executado consistiu no envio do “corpus”, àquela ferramenta.

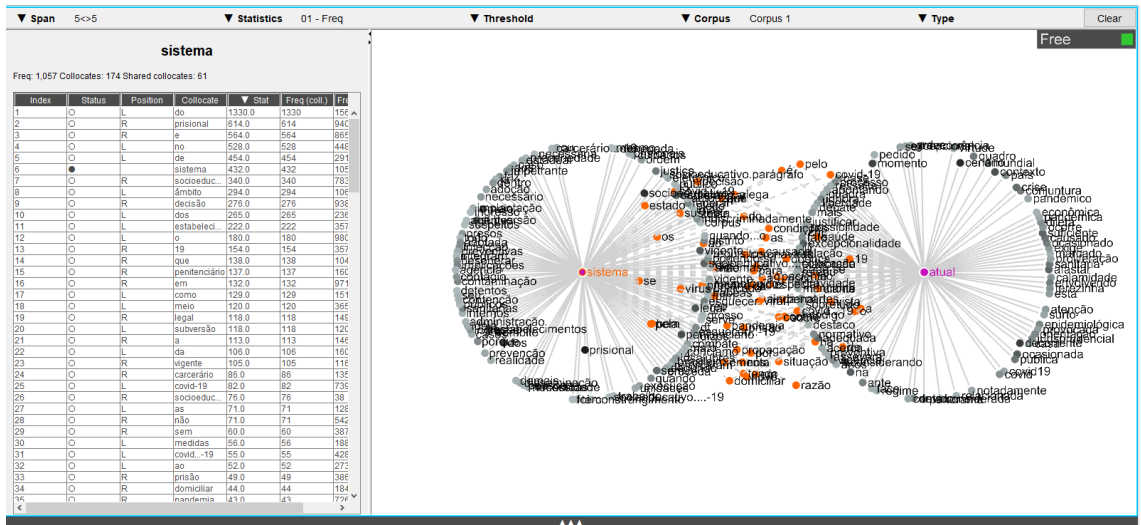
Após a obtenção dos resultados, foi realizada uma pesquisa minuciosa sobre as estruturas de tópicos geradas pelo sistema, mediante o exame da colocação das palavras-chave. Com auxílio do Lancsbox, foram verificadas as redes de conexões frequentes entre as palavras-chave no âmbito do corpus. Assim, foi possível a interpretar o conteúdo das decisões pela análise das proposições que conotam tais tópicos. Por exemplo, a palavra “prisão” era uma palavra-chave em vários tópicos e se verificou que estava comumente

Figura 5 – Conexões linguísticas entre “crise”, “propagação”, “pública e “saúde”



No segundo tópico (2) as palavras-chave são: “atual”, “liminar”, “sistema” e “julgado”. Pela análise do gráfico de colocações, o tópico parece estar referido à concessão de liminares em *habeas corpus*, solicitando relaxamento de prisão, em virtude dos riscos de contaminação inerentes às atuais condições do sistema carcerário brasileiro.

Figura 6- Conexões linguísticas entre “liminar”, “atual”, “sistema” e “julgado”



As expressões que cercam o termo “sistema” estão frequentemente relacionadas ao cumprimento de penas e medidas socioeducativas. Estas irão retornar em outros tópicos, portanto deixaremos para esmiuçá-las adiante.

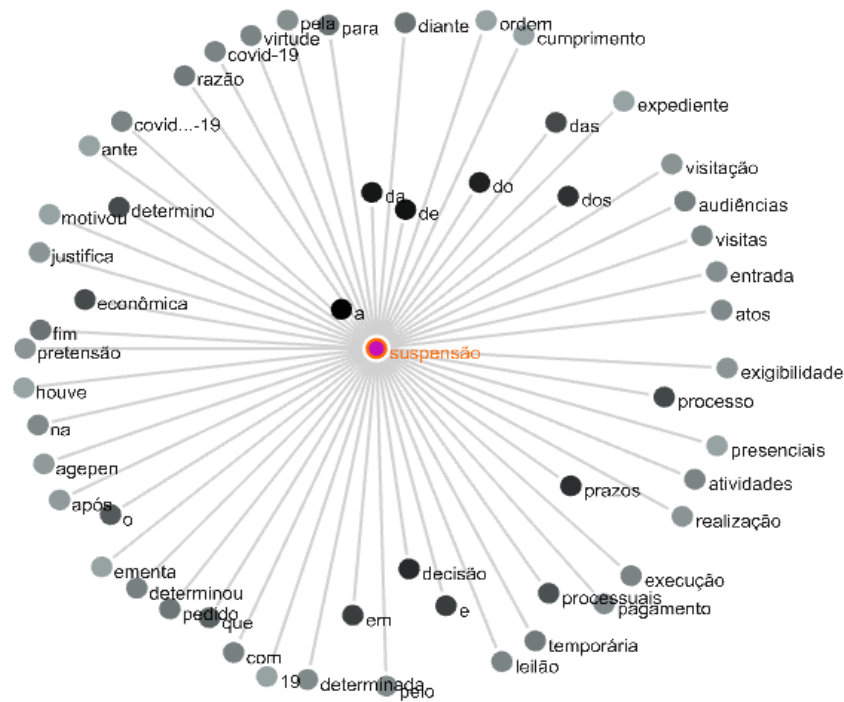
No terceiro tópico (3) as seguintes palavras-chave: “risco”, “paciente”, “recomendação” e “contaminação”. Com isso, percebe-se que se referem ao risco de contaminação aos presos e decisões judiciais que retratam o pedido de *habeas corpus* a pacientes integrantes de grupo risco. Este tópico apresenta a presença da Recomendação 62/2020 do Conselho Nacional de Justiça¹¹ na judicialização da pandemia de Covid-19. A referida recomendação estabelece no art. 1º a “adoção de medidas preventivas à propagação da infecção pelo novo coronavírus – Covid-19 no âmbito dos estabelecimentos do sistema prisional e do sistema socioeducativo”. Como podemos verificar em quase todos os tópicos, o tema dos reflexos judiciais oriundos das medidas propostas pelo CNJ, que visam evitar a contaminação pelo novo coronavírus no âmbito do sistema prisional e socioeducativo, é reiteradamente abordado, em diferentes perspectivas.

O quarto tópico (4) ressalta os termos: “risco”, “prisão”, “domiciliar” e “paciente”. Deste modo, interpretamos que o tópico diz respeito aos pedidos de concessão de prisão domiciliar, por meio de *habeas corpus*, em face do risco de contágio e propagação da Covid-19. Hipótese mencionada principalmente no art. 3º, da Recomendação nº 62/2020 do CNJ. Ademais, através do uso da ferramenta Lancsbox notou-se que a palavra “prisão”, além de fortemente associada ao termo “domiciliar”, encontra-se próxima de termos como “civil” e “tráfico”, dentre outros. A frequência dessas palavras parece indicar também a relevância de discussões envolvendo apenas pela prática do crime de tráfico de entorpecente, delito que responde por mais de 20% das incidências penais na população carcerária brasileira¹² e da prisão civil por dívida alimentícia, outra hipótese mencionada explicitamente no art. 6º da já mencionada Recomendação nº 62/2020 do CNJ.

¹¹ BRASIL. Conselho Nacional de Justiça. Recomendação nº 62, de 17 de março de 2020. Disponível em <https://www.cnj.jus.br/wp-content/uploads/2020/03/62-Recomenda%C3%A7%C3%A3o.pdf>. Acesso em 08 de junho de 2020.

¹² Informações do Departamento Penitenciário Nacional - DPEN. Disponível em: <https://app.powerbi.com/view?r=eyJrIjoiYWY5NjFmZjctOTJmNi00MmY3LThtMTetNWYwOTlmODFjYWQ5IiwidCI6ImViMDkwNDIwLTQ0NGMtNDNmNy05MmYyLTRiOGRhNmJmZThlMSJ9> Acesso em: 20 jul. 2020.

Figura 8 – Conexões da palavra “suspensão”



O sexto tópico (6) analisado apresentava as expressões: “prisão”, “ordem”, “paciente” e “preventiva”. Seguidamente, pesquisou-se no programa Lancsbox. Entre a palavra preventiva e prisão foram apresentados os seguintes tópicos correspondentes: paciente com 178 associações; e custódia com 94 associações. Dessa maneira, essa estrutura de tópicos representa as decisões que tratam dos pedidos de conversão de prisão preventiva em domiciliar e *habeas corpus*, nos termos do art. 5º, IV, da Recomendação nº 62/2020, do CNJ. O referido dispositivo recomendou aos magistrados com competência sobre execuções penais, a colocação em prisão domiciliar de pessoa presa com diagnóstico suspeito ou confirmado de Covid-19 na ausência de espaço de isolamento adequado, *in verbis*:

Art. 5º Recomendar aos magistrados com competência sobre a execução penal que, com vistas à redução dos riscos epidemiológicos e em observância ao contexto local de disseminação do vírus, considerem as seguintes medidas:

IV – colocação em prisão domiciliar de pessoa presa com diagnóstico suspeito ou confirmado de Covid-19, mediante relatório da equipe de saúde, na ausência de espaço de isolamento adequado no estabelecimento penal;

4. Breves considerações sobre o perfil institucional do CNJ e sua atuação na pandemia de Covid-19

Diante dessa análise, evidente o caráter ubíquo da Recomendação nº 62/2020 do CNJ no processo de judicialização da pandemia de Covid-19. Com base na modelagem de tópicos gerada pelo algoritmo, percebe-se a reiteração da temática relativa à gestão da população carcerária em meio à pandemia. Esta evidência demonstra a importância que a instituição alcançou desde sua criação pela Emenda Constitucional nº 45/2003.

Estudos já apontam que a amplitude da atuação e o desenho da institucional do CNJ estão sujeitos a dinâmicas internas de funcionamento, atreladas aos perfis de cada presidência e às divisões corporativas do órgão (FRAGALE, 2013). Por outro lado, a significativa referência à Recomendação nº 62/2020 demonstra também que fatores exógenos, tais como as práticas de aceitação e costumes constitucionais das autoridades do sistema jurídico brasileiro, serão fundamentais para consolidar o perfil do CNJ.

Para compreender esse processo, podemos nos valer do marco conceitual desenvolvido por H.L.A Hart, no “Conceito de Direito” (2005). Para o autor, os sistemas jurídicos contemporâneos são formados pela conjugação de regras primárias, cujo escopo principal é regular condutas, e regras secundárias, que dizem respeito à aplicação, alteração e reconhecimento destas regras primárias. As regras secundárias dependem de práticas sociais convergentes das autoridades públicas num determinado ordenamento jurídico. As regras secundárias adquirem corpo na medida em que são aceitas como pautas de ação por parlamentares, agentes do executivo e magistrados, dentre outras autoridades, que se sentem obrigados a segui-las. No nível do debate constitucional, sobretudo, encontramos uma pleora desta espécie de regras (ADLER; HIMMA, 2009).

A partir do arcabouço hartiano podemos ver como a judicialização da pandemia, fundada sobremaneira na Recomendação nº 62/2020 do CNJ, é um passo importante no reconhecimento e configuração dessa instituição dentro sistema jurídico brasileiro. Se entendemos a construção da arquitetura constitucional como uma empreitada orgânica, dependente de fatos sociais, concernentes na aceitação de regras secundárias por autoridades, entendemos que discussões sobre o alcance da atuação do CNJ ou a extensão de seus poderes não estão escritas em pedra, e que sua configuração é muito mais maleável e dependente da solução de controvérsias no âmbito dos tribunais.

Um estudo qualitativo pormenorizado das decisões aqui analisadas se faz necessário para determinar os contornos do perfil institucional em elaboração, podendo ter ele caráter restritivo ou expansivo das competências do CNJ. Contudo, vale ressaltar que o instrumento normativo da recomendação não é dotado de cogência e imperatividade, embora também vise dirigir comportamentos, assim como outras formas de prescrições normativas (HERDY, 2020). Dessa forma, apenas a reiterada menção a essa forma de *soft law* no âmbito das decisões parece indicar a adoção de posturas ampliativas em relação às competências do referido órgão e sua importância como fonte produtora de razões institucionais relevantes no âmbito da tomada de decisão jurídica (SCHECAIRA; STRUCHINER, 2016). Considerando a inconsistência das políticas públicas para enfrentamento da pandemia de Covid em nosso país, essa função diretiva se fazia necessária. Cabe verificar o exercício dessa atividade normativa em outros contextos. Novamente, pesquisas adicionais são necessárias para corroborar tal conjectura.

Conclusão

O presente estudo pretendeu obter um retrato instantâneo do processo de judicialização da epidemia de Covid-19 no Brasil. Para realizar a tarefa, escolhemos ferramentas de processamento de linguagem natural e análise linguística que permitiram a análise de uma quantidade de dados que seria impossível de ser realizada por métodos tradicionais de pesquisa jurídica, ao menos com os limites de tempo e recursos disponíveis para nossa equipe.

Essa escolha metodológica trouxe uma implicação epistêmica interessante. Abordamos o trabalho sem qualquer expectativa prévia em relação às fontes analisadas. Ainda que o ideal de objetividade oriente a empresa científica, sabemos que certo grau de intromissão do sujeito de pesquisa é inevitável na produção do conhecimento, ainda mais na seara das ciências sociais aplicadas, como o Direito. Todavia, o uso do algoritmo de LDA para elaboração da modelagem de tópicos a partir de documentos jurídicos subverte, em parte, a elaboração tradicional do conhecimento jurídico. Os resultados encontrados apontaram uma direção não antecipada, qual seja, a prevalência da gestão do sistema penitenciário e socioeducativo como foco principal de judicialização das demandas envolvendo a pandemia de Covid-19.

A ampla prevalência de tópicos vinculados à propagação da epidemia no sistema penitenciário e socioeducativo foi uma surpresa, porém, em retrospectiva, vemos que se tratava de uma tragédia anunciada. O sistema carcerário brasileiro é conhecido por suas deficiências, com destaque para a insalubridade e superlotação das celas, fatores que auxiliam na proliferação de epidemias e ao contágio de doenças. Enfermidades infectocontagiosas, tais como a tuberculose, a sarna, o HIV e a sífilis, são mais prevalentes nas unidades prisionais e correccionais do que no restante da população país (MINAYO; RIBEIRO, 2016).

Com base nos dados do último Infopen (Levantamento Nacional de Informações Penitenciárias), o Brasil possui, hoje, cerca de 726 mil pessoas em privação de liberdade, mas as vagas disponíveis somam apenas 436 mil. Do total da população carcerária, cerca de 250 mil têm algum tipo de doença. O Brasil ocupa o terceiro lugar no ranking mundial de países que mais prendem no mundo, ficando atrás apenas dos EUA¹³ e da China.

A violação contínua de direitos fundamentais da população carcerária brasileira já foi reconhecida inclusive na jurisdição constitucional brasileira, no âmbito da Arguição de Descumprimento de Preceito Fundamental nº 347, em tramitação no STF. Diante deste quadro, fácil imaginar a catástrofe iminente quando tratamos da situação de encarceramento no momento da pandemia de Covid-19.

Diante dos achados desta pesquisa, acreditamos que outras possibilidades de indagação se descortinam. A título de exemplo, quais as implicações dos fatos apresentados no contexto da proibição vedação constitucional de aplicação de penas cruéis e da garantia de respeito à integridade física e moral dos sujeitos privados de liberdade? Como analisar a judicialização da política carcerária, tendo em vista a extensão da Recomendação do 62 do CNJ, órgão de controle do judiciário, e a tibieza das medidas criadas por órgãos do poder executivo, tal o Ministério da Justiça, para enfrentamento do problema? Como tem sido o processo decisório, no que tange à apreciação individual dos casos envolvendo os pedidos de liberdade, com esteio na Recomendação 62/2020¹⁴? Abordamos em apertada síntese algumas implicações dos dados no que tange à consolidação do perfil institucional do CNJ, mas estas e outras questões merecem análise apartada e detida noutros trabalhos.

¹³ Em relação aos E.U.A. fontes demonstram que este país vive dificuldades similares em relação à gestão da política carcerária no âmbito da Covid-19. Diversas medidas foram tomadas para conter a propagação da doença no sistema penal e socioeducativo daquele país, com maior ou menor grau de eficácia. Para maiores detalhes conferir “Law in Time of Covid” (PISTOR, 2020).

¹⁴ Sobre a questão já se levantou interessante ponto em relação à seletividade na análise de provas e a falta de empatia de algumas das decisões que enfrentam a aplicação da Recomendação 62/20 do CNJ. VER (MATIDA, 2020). Outra análise interessante foi empreendida por uma equipe de pesquisadores da FGV-Rio (HARTMANN, et. al, 2020) empregando técnicas quantitativas e amostrais para compreensão das tendências decisórias no que tange aos Habeas Corpus impetradas perante o STF e STJ, no período da pandemia.

Em síntese, *habeas corpus* tendo por objeto a liberdade de personagens ilustres, como Fabrício Queiroz e Geddel Lima, apenas jogam luz sobre a ponta de um iceberg, cujos contornos se escondem sob águas profundas. A crise de Covid-19 aliada às condições odiosas de encarceramento no Brasil criam condições trágicas que merecem reflexão detida em diversos aspectos.

Esperamos que o presente trabalho, de caráter exploratório, contribua para o entendimento dos reflexos jurídicos da grave crise sanitária que vivemos. Ademais, esperamos ter demonstrado a validade e potência da utilização de mecanismos da ciência de dados na compreensão do Direito.

Referências

ALUÍSIO, Sandra Maria; DE BARCELLOS ALMEIDA, Gladis Maria. O que é e como se constrói um corpus? Lições aprendidas na compilação de vários corpora para pesquisa linguística. **Calidoscópico**, v. 4, n. 3, p. 156-178, 2006.

BARROSO, Luís Roberto. **A Judicialização da Vida e o Papel do Supremo Tribunal Federal**. Belo Horizonte: Forum, 2018.

BLEI, David M. Probabilistic topic models. **Communications of the ACM**, v. 55, n. 4, p. 77-84, 2012.44.

BOEING, Daniel H. Arruda; ROSA, Alexandre Morais da. **Ensinando um robô a julgar: pragmática, discricionariedade e vieses no uso de aprendizado de máquina no judiciário**. Florianópolis: EMais, 2020.

BOYD-GRABER, Jordan; MIMNO, David; NEWMAN, David. Care and feeding of topic models: Problems, diagnostics, and improvements. **Handbook of mixed membership models and their applications**, v. 225-255, 2014.

FILHO, Marçal Justen et al. **COVID-19 e o Direito brasileiro**. Curitiba: Justen, Pereira, Oliveira & Talamini, 2020.

FRAGALE FILHO, Roberto. Conselho Nacional de Justiça: desenho institucional, construção de agenda e processo decisório. *Dados*, Rio de Janeiro, v. 56, n. 4, p. 975-1007, Dec. 2013. Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0011-52582013000400008&lng=en&nrm=iso>. access on 01 Sept. 2020. <https://doi.org/10.1590/S0011-52582013000400008>.

HARTMANN, Ivar A.; MAIA, Natália; ABBAS DA SILVA, Lorena; MARPIN, Ábia; ALMEIDA, Guilherme. **Como STF e STJ decidem Habeas Corpus durante a Pandemia do COVID-19?** Uma Análise Censitária e Amostral (How Do the Brazilian Supreme Court and Superior Court of Justice Decide Habeas Corpus Writs during the COVID-19 Pandemic?)

A Census and Sample Based Analysis) Disponível em: <https://ssrn.com/abstract=3659624> Acesso em: 25 ago. 2020.

HERBERT, H. L.A. **O conceito de direito**. Lisboa: Fundação Calouste Gulbenkian, 2005.

HERDY, Rachel. A recomendação de uma instituição normativa serve para quê? **Consultor Jurídico**, 2020. Disponível em: https://www.conjur.com.br/2020-jul-31/limite-penal-recomendacao-instituicao-normativa-serve#_ftn5. Acesso em: 17 set. 2020.

JELODAR, Hamed et al. Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. **Multimedia Tools and Applications**, v. 78, n. 11, p. 15169-15211, 2019.

LIVERMORE, Michael A.; ROCKMORE, Daniel Nahum (Ed.). **Law as Data: Computation, Text, & the Future of Legal Analysis**. Santa Fé Institute Press, 2019.

MATIDA, Janaina. É preciso se dar fim à seletividade probatória. **Consultor Jurídico**, 2020. Disponível em: <https://www.conjur.com.br/2020-jul-17/limite-penal-preciso-dar-fim-seletividade-probatoria>. Acesso em: 17 jul. 2020.

MINAYO, Maria Cecília de Souza; RIBEIRO, Adalgisa Peixoto. Condições de saúde dos presos do estado do Rio de Janeiro, Brasil. **Ciênc. saúde coletiva**, Rio de Janeiro, v. 21, n. 7, p. 2031-2040, July 2016. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232016000702031&lng=en&nrm=iso. Acesso em 30 Jun. 2020. <http://dx.doi.org/10.1590/1413-81232015217.08552016>.

PISTOR, Katharina. **Law in the Time of COVID-19**. Columbia Law School, 2020.

POSSÍDIO, Cyntia; MARTINEZ, Luciano. **O trabalho nos tempos do Coronavírus**. São Paulo: Saraiva Educação SA, 2020.

ROSENVALD, Nelson et al. **Coronavirus e responsabilidade civil: impactos contratuais e extracontratuais**. Editora Foco, 2020.

SARDINHA, Tony Berber. Lingüística de corpus: histórico e problemática. **Delta: documentação de estudos em lingüística teórica e aplicada**, v. 16, n. 2, p. 323-367, 2000.

SHECAIRA, Fábio; STRUCHINER, Noel. **Teoria da argumentação jurídica**. Rio de Janeiro: PUC-Rio; Contraponto, 2016,

VALLINDER, Torbjörn. The Judicialization of Politics. A World-Wide Phenomenon: Introduction. **International Political Science Review/Revue internationale de science politique**, p. 91-99, 1994.

VIANNA, Luiz Werneck. **A judicialização da política e das relações sociais no Brasil**. Rio de Janeiro: Editora Revan, 1999.