

PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM
POLÍTICAS PÚBLICAS EM SAÚDE
ESCOLA FIOCRUZ DE GOVERNO
FUNDAÇÃO OSWALDO CRUZ

JULIANE APARECIDA ALVES

**Painel de Monitoramento e Análise de Informações da Produção
Legislativa em Saúde na Câmara dos Deputados: Uma Abordagem de
Sistematização de Dados Abertos**

BRASILIA
2022

JULIANE APARECIDA ALVES

**Painel de Monitoramento e Análise de Informações da Produção
Legislativa em Saúde na Câmara dos Deputados: Uma Abordagem de
Sistematização de Dados Abertos**

Trabalho de Dissertação apresentado à Escola
Fiocruz de Governo como requisito parcial para
obtenção do título de mestre em Políticas Públi-
cas em Saúde.

Orientadora: Prof. A Dra. Fernanda Maria Du-
arte Severo

BRASILIA
2022

A474p

Alves, Juliane Aparecida.

Painel de monitoramento e análise de informações da produção legislativa em saúde na Câmara dos Deputados: uma abordagem de sistematização de dados abertos / Juliane Aparecida Alves. -- 2022.

66 f. : il.color, fotos.

Orientadora: Fernanda Maria Duarte Severo.

Dissertação (Mestrado Profissional em Políticas Públicas em Saúde) - Fundação Oswaldo Cruz, Gerência Regional de Brasília, Escola de Governo Fiocruz Brasília, Brasília, DF, 2022.

Bibliografia: f. 56-62.

1. Dados Abertos. 2. Legislação em Saúde. 3. Câmara dos Deputados. 4. Mineração de Dados. I. Título.

CDD 614

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da Rede de Bibliotecas da Fiocruz com os dados fornecidos pelo(a) autor(a).

Bibliotecário responsável pela elaboração da ficha catalográfica: Livia Rodrigues Batista - CRB-1/3443
Biblioteca Fiocruz Brasília

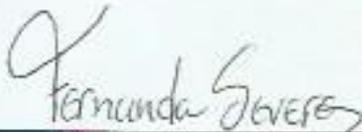
Juliane Aparecida Alves

Sistematização dos dados abertos da Câmara dos Deputados: Painel de monitoramento e análise de informações da produção de matérias legislativas de saúde.

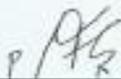
Dissertação apresentada à Escola de Governo Fiocruz como requisito parcial para obtenção do título de mestre em Políticas Públicas em Saúde, na linha de pesquisa Saúde e Justiça Social.

Aprovado em 16/12/2022.

BANCA EXAMINADORA



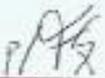
Dra. Fernanda Maria Duarte Severo- Orientadora - Fundação Oswaldo Cruz - Fiocruz Brasília



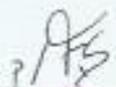
Dr. Armando Martinho Bardou Raggio – Membro Interno - Fundação Oswaldo Cruz - Fiocruz Brasília



Dra. Silvia Badim Marques– Membro Externo - Universidade de Brasília - UnB



Dra. Alethele Oliveira Santos – Suplente 1 - Conselho Nacional de Secretário de Secretários de Saúde



Me. Leonardo Vinícius Dias da Silva – Suplente 2 - Universidade Federal de Minas Gerais

AGRADECIMENTOS

Agradeço a todos os que fizeram parte desta jornada e tornaram possível a conclusão desta dissertação. A minha orientadora, Dra. Fernanda Severo, expressei minha gratidão pela paciência, sabedoria e incentivo constante durante todo o processo. Você foi uma luz nessa jornada, obrigada por sua dedicação e apoio incansável e por segurar minha mão até o final.

A equipe de trabalho em que faço parte, minha profunda gratidão, Marcus Vinícios, Núbia Rocha, Felipe Giuliane, Felipe Ferré, por sua colaboração e contribuição significativa neste projeto.

Expressei meus agradecimentos a banca examinadora, Dra. Silvia Badim, Dr Armando Raggio e M.Sc Leonardo Vinícios, por sua experiência e valiosos comentários que ajudaram a aprimorar ainda mais meu trabalho. De maneira singular agradeço a Dra. Alethele Oliveira Santos, que para além das suas preciosas contribuições, seu apoio e inspiração para desenvolver sobre essa temática.

Sou imensamente grata minha família, meus irmãos por seu amor e apoio incondicional. Vocês foram minha fonte de força durante todo o processo e estou agradecida por estarem ao meu lado. Em memória do meu pai, Gumercindo, e minha Avó Dimellia, quero agradecer por tudo o que fizeram por mim e por suas influências inestimáveis em minha vida.

Gostaria de agradecer especialmente à minha mãe, Maria Divina, por sua coragem e determinação em lutar por uma vida melhor para si e para todos os seus filhos. Sou a oitava filha desta mulher que não teve a oportunidade de estudar, mesmo assim, superou os limites impostos pela desigualdade social. Lutando incansavelmente para que todos os seus filhos pudessem percorrer um caminho diferente baseados na educação. Sou grata por ter uma mãe guerreira e corajosa.

Quero reconhecer e agradecer aos meus ancestrais, que enfrentaram desafios e adversidades impensáveis, tudo que inclui os mais de trezentos anos de escravidão que o povo negro enfrentou nesse país. Mesmo assim, conseguiram transmitir sua força, sabedoria e herança cultural para as gerações futuras. Sinto-me honrada por carregar essa essência e por ter a oportunidade de perpetuar sua memória.

Por fim, gostaria de agradecer à Fiocruz Brasília e ao Conselho Nacional de Secretários de Saúde por serem minhas instituições formadoras e por me darem a oportunidade de aprender e crescer profissionalmente. E também a todos os meus colegas de trabalho por seus ensinamentos e amizade.

Estou profundamente grato a todos vocês e espero que este trabalho possa contribuir para a melhoria da saúde pública no Brasil.

“Costuro o infinito sobre o peito.” Hilda Hilst

Resumo

Este estudo se concentra em tornar a produção legislativa em saúde na Câmara dos Deputados (CD) mais acessível e compreensível para os gestores estaduais do Sistema Único de Saúde (SUS). Abordando a questão da obtenção e análise de informações relevantes sobre questões legislativas de saúde na Câmara dos Deputados (CD) no Brasil. Com base na Lei de Acesso à Informação (LAI), que garante aos cidadãos o direito de solicitar e obter informações em posse da administração pública, este trabalho desenvolveu um mecanismo para extrair e modelar dados da CD. O resultado é um painel dinâmico de informações que pode ser utilizado por gestores para produzir relatórios valiosos e tomar decisões oportunas. A análise de dados gerada tem um papel crucial na formulação e monitoramento de políticas públicas eficazes na área da saúde. O trabalho contribui com uma abordagem eficiente para extrair e visualizar informações valiosas a partir de fontes governamentais abertas, contribuindo para a transparência e responsabilidade pública na área da saúde.

Palavras-chave: Dados abertos; Câmara dos Deputados; Visualização de informações; Matérias legislativas em saúde; Mineração de dados.

Abstract

This study focuses on making legislative production in health in the Chamber of Deputies (CD) more accessible and understandable for state managers of the Unified Health System (SUS). Addressing the issue of obtaining and analyzing relevant information on health legislative matters in the Chamber of Deputies (CD) in Brazil. Based on the Access to Information Law (LAI), which guarantees citizens the right to request and obtain information in the possession of the public administration, this work developed a mechanism to extract and model data from the CD. The result is a dynamic panel of information that can be used by managers to produce valuable reports and make timely decisions. The analysis of data generated plays a crucial role in formulating and monitoring effective public policies in the health sector. The work presents an efficient approach to extract and visualize valuable information from open government sources, contributing to public transparency and responsibility in the health area.

Keywords: Open data; Chamber of Deputies; Information visualization; Health legislative matters; Data mining.

Lista de ilustrações

Figura 1 – Fluxograma de tramitação.	24
Figura 2 – Portal de buscas por proposições	28
Figura 3 – Etapas do processo KDD	33
Figura 4 – Estrutura Básica OpenAPI	35
Figura 5 – Exemplo de Arquivo Json	36
Figura 6 – Grafo da Relação dos Endpoints	39
Figura 7 – Modelagem de Dados CD	46
Figura 8 – Seleção de Registros das Tabelas Presentes no Banco de Dados da Câmara. .	47
Figura 9 – Número de Registros das Tabelas Presentes no Banco de Dados da Câmara.	47
Figura 10 – Etapas de elaboração do protótipo	49
Figura 11 – Painel de Matérias legislativas em Saúde no Congresso Nacional.	50
Figura 12 – Pesquisa por tema	51
Figura 13 – Painel de Matérias legislativas em Saúde: Sessão Covid-19	52
Figura 14 – Painel de Matérias legislativas em Saúde: Proposições e Óbitos por COVID-19	53
Figura 15 – Painel de Matérias legislativas em Saúde: Sessão Covid-19: Busca por Ementas	54

Lista de tabelas

Tabela 1 – Relação no Endpoint	40
Tabela 2 – Tipos de Dado	42

Lista de abreviaturas e siglas

AI	Arquitetura da Informação
API	Application Programming Interface
BDTD	Biblioteca Digital de Teses e Dissertacoes
BI	Business Intelligence
CD	Câmara do Deputados
CF	Constituição Federal
CONASS	Conselho Nacional de Secretários de Saúde
CSV	Comma-separated values
ETL	Extract Transform Load
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
JSON	JavaScript Object Notation
KDD	Processo de Descoberta de Conhecimento
LAI	Lei de Acesso à Informação
MPOG	Ministério do Planejamento, Orçamento e Gestão
OAS	Organization of American States ("Organização dos Estados Americanos")
OGP	Open Government Partnership
ONU	Organização das Nações Unidas
SGBD	Sistema de Gerenciamento de Banco de Dados
SUS	Sistema Único de Saúde
URL	Uniform Resource Locator (Localizador Padrão de Recurso)
UX	User Experience

Sumário

1	INTRODUÇÃO	12
2	OBJETIVOS	15
2.1	OBJETIVOS ESPECÍFICOS	15
3	REVISÃO DA LITERATURA	16
4	DISCUSSÃO	18
4.1	Política de Dados Abertos	18
4.2	Arquitetura da Informação	21
4.3	Processo Legislativo Brasileiro	22
4.4	A Informação Legislativa	25
4.5	Matérias legislativas no Contexto do Sistema Único de Saúde Brasileiro	25
5	MÉTODO, TÉCNICA E PROJETO APLICADO	30
5.1	Identificação de Método e Técnicas	30
5.2	Integração dos Dados da Câmara dos DeputadosAPI (Application Program- ming Interface)	33
5.3	OpenAPI	34
5.4	Json	35
5.5	Processo de Desenvolvimento	36
5.5.1	Dados Disponíveis	36
5.5.2	Estrutura de disposição dos dados	36
5.5.3	Ferramenta de integração	40
5.6	Técnicas de Pré-Processamento e Modelagem Aplicadas ao Banco de dados da API da CD.	41
5.6.1	<i>Python</i>	41
5.6.1.1	Biblioteca Pandas	43
5.6.1.2	Preparação dos Dados	43
5.6.1.3	Limpeza de dados	43
5.6.1.4	Redução dos dados	44
5.6.2	Modelagem de banco de dados	45
5.6.3	Sistema de Banco de Dados	45
5.7	Desenvolvimento do modelo	45
5.7.1	Carregando o Dataset	46
6	PROTÓTIPO DO PAINEL DE MONITORAMENTO DE PROPOSIÇÕES LEGISLATIVAS	49

CONSIDERAÇÕES FINAIS	55
Referências	56
ANEXOS	66

1 INTRODUÇÃO

Desde o início do século XXI, vê-se um aumento no interesse de governos em tornar seus processos institucionais mais transparentes e acessíveis ao público. Essa tendência, que tem ganhado força ao nível global, levou à implementação de diversos projetos voltados para essa finalidade. Em 2011, a Câmara dos Deputados (CD), enquanto representante legítima do povo brasileiro, decidiu aderir a essa tendência ao implementar seu próprio projeto de dados abertos.¹ Ao fazê-lo, a CD deu um importante passo na direção de uma administração pública mais transparente e responsiva às necessidades da sociedade.¹ O repositório de bases do legislativo² em formato digital desburocratizado e acessível permite a reutilização dessas informações sem restrições por outros órgãos públicos, entidades da sociedade civil, pesquisadores e qualquer pessoa interessada nesse conteúdo, seguindo as premissas de dados abertos.^{2,3}

As bases de dados do legislativo, disponibilizadas conforme a Lei de Acesso à Informação (LAI) n.º 12.527/2011. A Lei garante ao cidadão o direito de ter acesso a informações públicas sobre os atos e decisões do poder público.⁴ Através do Portal de Transparência da Câmara dos Deputados, é possível encontrar uma ampla gama de informações legislativas, desde proposições e pautas de órgãos até dados de deputados e votações em plenário. Além disso, o Portal também disponibiliza informações sobre temas de grande importância para a sociedade, como saúde, meio ambiente e educação, entre outros.³

Para publicizar esses dados na internet, a CD utiliza alguns formatos, como o de uma *API*--Application Programming Interface (Interface de Programação de Aplicação).³ É possível encontrar no seu endereço da *web*, os tutoriais de utilização da API e outros projetos da comunidade voltados para a análise dos materiais disponibilizados.^{6,7,3}

Em que pese, a Câmara dos Deputados cumpra seu dever perante a sociedade, consolidando premissas democráticas, com a publicização do conteúdo legislativo. Foram encontradas algumas dificuldades, como a deste estudo contendo avaliação da qualidade dos 'datasets' da CD.⁸

Evidenciou que, em 88% dos dados, foram detectados alguns tipos de opacidades, lacunas e outras disfunções.⁸ Apontando para um panorama em que diferentes usuários podem ter algum embaraço, que impeça ou dificulte a sua obtenção.

Naturalmente, a apropriação de 'Dados Abertos' em todo seu potencial está no uso de *softwares* ou algoritmos para gerar conhecimento a partir do uso de informações estruturadas e isso tem relação direta com o tipo de usuário e que não se trata do cidadão comum.^{9,1} O usuário-padrão (tomado como base para esta análise) é aquele que tem habilidades limitadas para lidar com processamento de dados digitais, usando ferramentas de Ciência de Dados.³

¹ <https://dadosabertos.camara.leg.br/>

² <https://dadosabertos.camara.leg.br/swagger/api.html>

³ *Application Programming Interface* (Interface de Programação de Aplicação) é um conjunto de normas, que possibilita a comunicação entre plataformas através de uma série de padrões e protocolos, auxiliando na consolidação em repositórios de dados.⁵

Em termos práticos, são os cientistas de dados, jornalistas de dados, programadores e outros especialistas que lidam com dados digitais. Por meio, de ferramentas tecnológicas em locais com estrutura, capazes de abarcar-las como: os centros de pesquisas em universidades, núcleos tecnológicos governamentais e empresariais e demais instituições.^{3,10}

Nesse sentido, um conjunto de dados com padrões de baixa qualidade para reutilização, prejudica o usuário, principalmente que detém poucos recursos ou experiência para driblar tais problemas. Esse fator compromete a eficácia e eficiência do conhecimento que pode ser gerado, afetando a qualidade dos resultados e insumos para modelos futuros ou tomada de decisão organizacional.^{11,6,7,3}

Assim, investir na qualidade dos 'datasets' pensando no seu valor e precisão, abrange vários aspectos, como de singularidade; completude; validade e consistência. Explica-se que, elementos não padronizados aumentam o custo é reuso (porque demandam mais energia do usuário em uniformiza-los). Outras disfunções, como variáveis opacas com lacunas, precisam ser mensuradas podendo influenciar consideravelmente o valor, tempo e reposta do projeto.^{3,10}

A ausência de inteligibilidade de variáveis, a existência de células vazias, são pequenos exemplos, capazes de inutilizar parte dos dados. Impossibilitando cruzamentos suficientes para responder uma pergunta. Elementos aparentemente técnicos também têm impactos reais no processo de apropriação social. Por isso, é importante gerenciar a qualidade dos dados, buscando padrões de confiabilidade que possibilitem decisões mais rápidas baseada em evidência.^{8,11,12}

Como ocorrido na pandemia de COVID-19, (iniciada em fevereiro de 2020 no Brasil). Houve um aumento da velocidade de implementação de diversas tecnologias, ao modo que muitas dessas soluções inovadoras foram baseadas em consumo de banco de dados.¹³. Responsáveis por alimentar algoritmos aplicados em várias áreas, desde genômica até teleconsultas. Nesse período, na Câmara dos Deputados ocorreu uma massiva publicação de matérias legislativas, relativas à pandemia. Em sua maioria, projetos de leis e de emendas constitucionais.¹⁴

Esse movimento, evidenciou a necessidade de utilização de estratégias eficientes para armazenamento e modelagem dessa base de dados. Pois quando dispostos descentralizadamente, tornam difíceis e morosas as análises realizadas por seres humanos.^{8,2} Tanto na sumarização de informações, identificação, compreensão de sua estruturação, quanto na disponibilidade ou recuperação na fonte.^{14 8}

Diante deste cenário de emergência sanitária global, e com isso, o aumento substancial de volumetria dos dados do legislativo. O Conselho Nacional de Secretários de Saúde (CONASS)⁴, enquanto instituição reconhecida legalmente (art. 14-B da Lei 8.080/90) representa a gestão estadual do SUS, aumentou o investimento de recursos para o monitoramento das legislações transversais ao SUS.¹⁴

A produção legislativa em saúde é crucial para a tomada de decisões informadas pelos gestores estaduais do Sistema Único de Saúde (SUS). Visando apoiar essa importante tarefa, o Conselho Nacional de Secretarias Estaduais de Saúde (CONASS) ofereceu todos os recursos

⁴ <https://www.conass.org.br/>

técnicos e práticos necessários para a realização deste projeto. Elaborado em convergência aos objetivos deste programa de mestrado em gestão de políticas públicas para saúde, cujo objetivo é o desenvolvimento de soluções para problemas relevantes para a área de atuação profissional.

Portanto, o presente estudo constitui-se como uma pesquisa aplicada ao desenvolvimento de um protótipo que automatize a base de dados da produção legislativa em saúde na Câmara dos Deputados, integrando, processando e sistematizando as informações relevantes. Além disso, será desenvolvido um painel de visualização dos dados, permitindo a realização de cruzamentos e análises importantes para a gestão estadual do SUS.

Sua execução pode ser compreendida em imprescindíveis quatro etapas: (1) criação de um mecanismo de extração dos dados da CD, para um Banco de Dados sob guarda do Conass. (2) Utilização de técnicas pré-processamento e modelagem. (3) Disponibilização de um modelo de visualização, (4) com acesso pelo público alvo.

2 OBJETIVOS

O objetivo geral deste estudo é desenvolver um protótipo capaz de automatizar o processo de monitoramento da produção legislativa em saúde no Congresso Nacional Brasileiro, por meio da integração e sistematização de dados abertos.

2.1 OBJETIVOS ESPECÍFICOS

- Integrar, processar e sistematizar os dados de matérias legislativas em saúde a fim de torná-los acessíveis e manipuláveis pelo protótipo em desenvolvimento.
- Desenvolver um painel de visualização dos dados que permita o cruzamento e análise de informações de interesse da Gestão Estadual do SUS, tais como os temas prioritários em discussão no Congresso Nacional e as ações legislativas que impactam diretamente o funcionamento do sistema de saúde brasileiro.

3 REVISÃO DA LITERATURA

A revisão de literatura é uma etapa fundamental na produção de conhecimento, pois permite ao pesquisador conhecer a produção científica já existente e estabelecer as bases teóricas para sua pesquisa. Também permite avaliar a evolução do conhecimento e as tendências na área de sistematização de dados abertos e análise de informações legislativas em saúde.

Revisar a literatura existente é uma etapa mais aprofundada da pesquisa. Envolve uma análise mais detalhada da literatura existente, incluindo a identificação de pontos fortes e fracos, tendências e lacunas na área de estudo.¹⁵ Baseia-se na leitura de artigos, livros, teses, dissertações, entre outras fontes de informações.

O processo consiste na identificação, seleção, leitura e análise dos documentos relacionados ao assunto, para produzir um documento que apresente um resumo das informações encontradas. Reunindo informações que sejam relevantes para o objeto da pesquisa, permitindo compreender o estado atual da literatura sobre o tema em questão, bem como identificar lacunas e oportunidades para novas pesquisas.^{16,15,17}

É utilizada em diversas áreas do conhecimento como ciências sociais, humanas e exatas, e considerada um passo importante na produção de conhecimento científico. Ela também é ferramenta auxiliar na tomada de decisões, na elaboração de projetos e planejamento, além de servir como fonte de informação para a produção de artigos acadêmicos^{18, 19,17}

Esta revisão, foi realizada em bases científicas indexadas, norteadas pela seguinte pergunta de pesquisa: “Qual é o escopo da literatura publicada sobre a base de dados das matérias legislativas em saúde na Câmara dos Deputados?”

As bases de dados utilizadas, por serem abrangentes e interdisciplinares, foram *LILACS*, *ERIC*, *SCIENCE DIRECT*, *SCIELO*, *PUBMED*, *SCOPUS*, *WEB OF SCIENCE*, *EMBASE*, *GOOGLE ACADÊMICO*, *DIGITAL BRASILEIRA DE TESES e DISSERTAÇÕES (BDTD)* sendo concentrada em artigos científicos. Diante da questão de pesquisa, definiu-se a utilização combinada dos descritores *Medical Subject Headings (MeSH)* e *Descritores em Ciências da Saúde (DeCS)*.^{18,20}

A eleição de palavras-chave e suas combinações contaram com a utilização de booleanos. A sintaxe da estratégia de busca foi testada, através da adaptação de termos e operadores booleanos em cada base de dados, o resultado com maior concordância da revisão foi incluído para avaliação dos critérios de inclusão e exclusão contidos no protocolo de busca, segue os termos utilizados:¹⁶

((health legislation); OR (legislation as topic); OR (legislative analysis), OR (Legislative Production), AND (open data); OR (open government); OR (open parliament); OR (Chamber of Deputies); AND (Health Policy); AND (Linked Open Data); (Open Government). (Legislative Data), (Dirichlet Allocation), (Natural Language Processing); (Politics); (database modeling); (monitoring of Legislative Production)). (“câmara dos deputados” OR câmara legislativa“ AND ”dados abertos“ OR dados OR dataset OR ”data lake” OR datawarehouse “câmara dos deputados”

AND “business intelligence”)

Não houve definição de intervalo de tempo como filtro de pesquisa tendo em vista o número reduzido de produções e a atualidade das reflexões sobre a temática

Conforme buscas realizadas em maio de 2022, foram avaliadas 230 publicações e suas respectivas referências bibliográficas, a partir da leitura de títulos e resumos, resultando na seleção de 19 artigos disponibilizados em anexo. Para leitura em profundidade e sujeitas à incorporação no *corpus* de balizamento teórico deste estudo.

4 DISCUSSÃO

4.1 Política de Dados Abertos

A Resolução 59 da ONU, adotada pelo Conselho de Segurança em 1946²¹, é considerada o primeiro documento internacional que reconhece o acesso à informação como um direito humano fundamental. Essencial para a promoção da paz e da segurança internacional.^{22,23,24}

No Brasil o acesso à informação é direito de todos os cidadãos garantido pela Constituição Federal de 1988²⁵. Com vistas a contribuição para o fortalecimento da democracia, permitindo que os cidadãos acompanhem os trabalhos do governo e fiscalize seu funcionamento.^{24,26}

Apesar de ser um direito fundamental, acessá-lo ainda é um desafio no Brasil e também em outros países. Pois, existem diversas barreiras para o acesso à informação, como o analfabetismo e o custo da internet, ainda impedem que muitas pessoas se beneficiem deste direito. Muitas vezes, as leis de acesso à informação são insuficientes ou não são cumpridas.^{22,26,24}

Em 2009, o Ministério do Planejamento, Orçamento e Gestão (MPOG), publicou uma cartilha técnica, sobre: “Dados Abertos no Brasil“. Com orientações para facilitar o acesso à informação pública e promover a transparência. A cartilha também destaca alguns princípios que deveriam ser seguidos para implementação de Dados Abertos, incluindo disponibilização de dados em plataformas online e em formato padronizado. Evidenciando a adoção de medidas de segurança e criação de mecanismos para mensurar o impacto gerado pela iniciativa.²⁷

Em 2011, o Brasil foi um dos primeiros países a aderir ao Governo Aberto. Passou a integrar a Open Government Partnership (OGP), uma iniciativa global que reúne países que se comprometem a melhorar o formato de relacionamento do governo com a sociedade. Através da promoção da transparência, da participação e da responsabilização. Desde então, o país tem se empenhado em honrar os compromissos assumidos na Declaração de Princípios do Governo Aberto^{28 29}

A Lei de Acesso à Informação (LAI), n.º 12.527 é uma lei federal brasileira que regula o acesso a informações públicas pelos cidadãos. A lei foi sancionada em 2011 pelo então presidente da República, Dilma Rousseff, e entrou em vigor no ano seguinte.²³

A LAI é considerada uma das mais avançadas leis de acesso à informação do mundo, garantindo o direito de todos os cidadãos brasileiros de solicitar e obter informações detalhadas sobre o funcionamento da administração pública. A lei estabelece que todas as informações consideradas públicas devem ser disponibilizadas pela entidade de origem. Salvo nos casos em que a divulgação possa causar algum tipo de dano à segurança nacional, à ordem pública ou às investigações criminais.^{27,23,24}

A LAI também criou o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), órgão que tem como principal função garantir o cumprimento da lei. Desde a sua entrada em vigor, a Lei de Acesso à Informação já contribuiu para a melhoria do funcionamento da administração pública brasileira, tornando-a mais transparente e acessível à população.^{29,30,26}

Ao reconhecer a importância do direito à saúde, entregue à população por meio do acesso

às ações e serviços públicos de saúde, reconhece-se a necessidade de aperfeiçoamento constante do SUS. E o forte impacto da tecnologia, na sistematização de dados e facilitação de análises. Contudo, para se ter informação e análise, o primeiro passo está no acesso aos dados.^{27,30}

A Lei direciona a União, Estados, Distrito Federal e Municípios passaram a observar os procedimentos a serem seguidos, que garantam o acesso à informação previsto nos artigos 5º, 37º e 216º da Constituição Federal (CF). Na mesma esteira estão as entidades privadas sem fins lucrativos que recebam recursos públicos. Art. 4º Para os efeitos desta Lei, considera-se:^{4,31}

Informação	Dados, processados ou não, que podem ser utilizados para produção e transmissão de conhecimento, contidos em qualquer meio, suporte ou formato;
Documento:	Unidade de registro de informações, independentemente do suporte ou formato;
Informação sigilosa:	Aquela submetida temporariamente à restrição de acesso público em razão de sua imprescindibilidade para a segurança da sociedade e do Estado;
Informação pessoal:	Aquela relacionada à pessoa natural identificada ou identificável;
Tratamento da informação:	Tratamento da informação: conjunto de ações referentes à produção, recepção, classificação, utilização, acesso, reprodução, transporte, transmissão, distribuição, arquivamento, armazenamento, eliminação, avaliação, destinação ou controle da informação;

Informação	Dados, processados ou não, que podem ser utilizados para produção e transmissão de conhecimento, contidos em qualquer meio, suporte ou formato;
Disponibilidade:	Qualidade da informação que pode ser conhecida e utilizada por indivíduos, equipamentos ou sistemas autorizados;
Autenticidade:	Qualidade da informação que tenha sido produzida, expedida, recebida ou modificada por determinado indivíduo, equipamento ou sistema;
Integridade:	Qualidade da informação não modificada, inclusive quanto à origem, trânsito e destino;
Primariedade:	Qualidade da informação coletada na fonte, com o máximo de detalhamento possível, sem modificações.

Fonte: ADAPTADO (BRASIL,2012)

A Fundação do Conhecimento Aberto (OKFN) é uma organização internacional sem fins lucrativos que promove acesso livre ao conhecimento e dados de todo o tipo. A OKFN lista uma série de características para que o dado seja considerado aberto:³²

- Completos - Para os dados serem completos, todos os conjuntos de dados públicos devem ser disponibilizados, e não somente uma parte deles.
- Atuais - Os dados devem ser disponibilizados ainda “quentes”, tão rapidamente quanto seja necessário para serem úteis
- Acesso não discriminatório - Os dados devem ser disponíveis a todos os cidadãos interessados, sem ser necessária identificação, registro ou cadastro.
- Formatos não-proprietários - Os dados devem ser disponíveis em pelo menos um formato sobre o qual nenhum ente tenha controle exclusivo.
- Processáveis por máquina - Os dados devem ser razoavelmente estruturados para possibilitar o seu processamento automatizado.
- Livres de licenças - Os dados não devem ser sujeitos a regulações de direitos autorais, marcas, patentes ou segredo industrial (exceto, claro, os dados que exigem sigilo e respeito à privacidade).

As informações disponibilizadas por consequência da execução da LAI podem ser obtidas por um pedido de acesso a informações diretamente aos órgãos com procedimentos dispostos nos artigos 10 ao 14 dessa lei⁴, porém em alguns casos também pode ser obtida de forma automatizada por portais de divulgação ou até mesmo por uma API (*Application Programming Interface*).

4.2 Arquitetura da Informação

A Arquitetura da Informação (AI) é uma área da ciência da computação que se concentra na análise, na organização, na estruturação e na apresentação de informações em sistemas de computador. É usada para melhorar o acesso a dados e informações, tornando-os mais acessíveis para os usuários.^{33,34}

Consiste na análise de métodos de armazenamento e recuperação de informações, na criação de sistemas de navegação, na criação de interfaces de usuário e na implementação de sistemas de acesso a dados. Por meio de sua análise, ela fornece uma estrutura de organização para a informação, que inclui métodos de recuperação de informações, como busca, classificação, organização e navegação. Incluindo a criação de mapas mentais, diagramas de fluxo, etc.^{34,35,36}

Além disso, a Arquitetura da Informação fornece um conjunto de princípios para a criação de interfaces de usuário intuitivas e fáceis de usar, que consigam fornecer acesso rápido e eficiente às informações.³⁵ O termo “Arquitetura da Informação” foi utilizado pela primeira vez em 1975, quando Richard Saul Wurman criou o termo para descrever o seu trabalho de organização do conteúdo de livros. Desde então, a AI tornou-se um campo cada vez mais popular de estudo, sendo aplicado em vários contextos, desde sites web até softwares e sistemas de informação.³⁷

Esta prática é frequentemente associada com design de interação, taxonomia, pesquisa e descobrimento. Além disso, a AI também envolve a utilização de técnicas e ferramentas específicas para ajudar a criar estruturas de informação bem organizadas, representadas e recuperadas, com acesso e disseminação de informação advindos da Biblioteconomia.^{38,36}

Taxonomia: é uma ferramenta para criar classificações hierárquicas para a organização de informações. Uma taxonomia é geralmente representada por hierarquias de categorias, subcategorias e termos que auxiliam na classificação dos conteúdos.^{39,33} Também pode ser usada para melhorar a indexação de conteúdos por mecanismos de busca, aumentando assim sua visibilidade na *web*.⁴⁰

Mapas de conteúdo, são ferramentas importantes de AI. Eles fornecem uma representação visual do conteúdo dentro de um *site*, mostrando como as seções e páginas relacionam-se entre si. Permitem que os usuários naveguem facilmente por um site, ajudando-os a encontrar informações relevantes e a localizar o conteúdo que precisam. Também podem contribuir os desenvolvedores a criar sites mais eficientes e intuitivos.^{39,40,41}

Wireframes, são esboços básicos que mostram a estrutura e o *layout* de uma página *web*.

Eles permitem que os designers vejam o projeto em seu estágio inicial, sem se preocupar com a aparência visual. É um termo que descreve a criação de esboços de telas que se relacionam entre si. Estes esboços são essenciais para o desenvolvimento de aplicações *web* e para garantir a qualidade na experiência do usuário.^{40,42}

Prototipagem, é uma abordagem que se concentra na criação de um modelo funcional e interativo para testar e validar as ideias e requisitos de design. Antes de entrar na fase de desenvolvimento formal. É um processo iterativo para coletar feedback do usuário e melhorar a usabilidade do produto.^{34,41}

A prototipagem é iterativa, exige que os designers alterem seu projeto à medida que o feedback do usuário é obtido. Os designers também devem avaliar a usabilidade do produto e considerar as restrições de tempo e orçamento.^{33,34,35}

A relação da informação com a gestão estratégica das organizações vem sendo destacada por vários autores. Considerada fundamental para a tomada de decisões estratégicas e para o sucesso das organizações. Entendida como um recurso estratégico que permite as organizações aproveitarem oportunidades e minimizar riscos.^{35,43,8}

Podem ser usada para acompanhar o desempenho organizacional, para aprimorar processos e para melhorar a eficiência e a eficácia das suas operações. Assim, a informação é fundamental para que as organizações alcancem seus objetivos estratégicos e aproveitem as oportunidades existentes.^{34,43}

Além disso, o gerenciamento da informação também envolve a criação, a utilização e a avaliação de sistemas de informação, bem como estratégias de comunicação eficaz para atingir os objetivos da organização.^{44,37}

1. Captação: Coleta de informações, tanto internas quanto externas.
2. Processamento: Transformação das informações em um formato que possa ser acessado e compreendido.
3. Armazenamento: Armazenamento de informações.
4. Recuperação: Busca e recuperação de informações armazenadas.
5. Distribuição: Compartilhamento das informações com outras pessoas.
6. Utilização: Uso das informações para a tomada de decisões.

4.3 Processo Legislativo Brasileiro

O processo legislativo é o mecanismo responsável pela criação, alteração ou extinção de leis. O Legislativo (Congresso Nacional, Assembleias Legislativas e Câmaras de Vereadores) exerce sua atribuição de legislar, ou seja, produzir leis. Esse processo é iniciado com a iniciativa de algum parlamentar, que apresenta um Projeto de Lei ao Congresso.⁴⁵

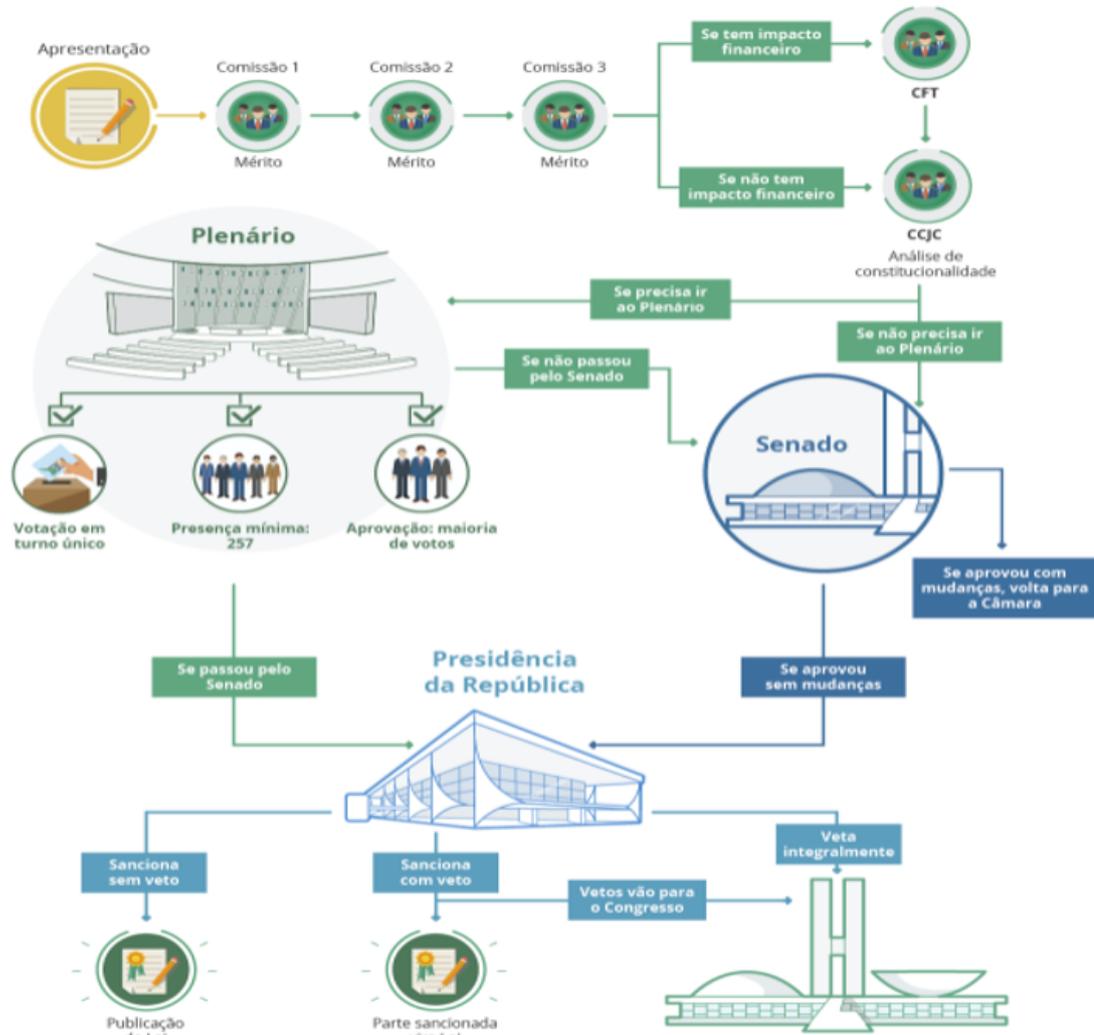
Os principais tipos são as Propostas de Emenda à Constituição (PEC); Projetos de Lei Complementar (PLP); Projetos de Lei Ordinária (PL); Projetos de Decreto Legislativo (PDC); Projetos de Resolução (PRC) e Medidas Provisórias (MPV).^{46,47,48}

As proposições responsáveis por iniciar o processo legislativo são a Proposta de emenda à Constituição e os projetos, que podem ser de lei, de decreto legislativo e de resolução.^{45,47} Podendo ser apresentadas por agentes políticos, como parlamentares, governadores ou o próprio Presidente da República, ou por membros da sociedade civil, como associações e organizações não governamentais. Por órgãos públicos (Tribunais superiores, Comissões da Câmara e do Senado, Procuradoria Geral da República) e por qualquer cidadão, respeitando as diferentes regras de tramitação para cada ato e ator explicitadas na Constituição Federal.^{46,47,45}

A proposição é distribuída para as comissões das Casas Legislativas por meio de processo administrativo, a fim de que sejam realizadas análises e discussões sobre as possíveis implicações da proposta. As comissões, ao final de sua análise, podem acatar, rejeitar ou modificar a proposta, submetendo assim a proposta para votação no Plenário, onde cada parlamentar vota a favor ou contra a proposta.^{49,47,46}

Caso seja vetada, uma nova votação pode ser feita, onde para que a proposta seja aprovada, é necessário o voto favorável de 3/5 dos votos dos parlamentares presentes. Se a maioria dos votos for a favor da proposta, a mesma é aprovada e segue para a sanção do Executivo, com a decisão de sancionar ou vetar a proposta. Mantida a aprovação segue para a promulgação, a etapa final da tramitação legislativa, onde a proposta passa a ter efeitos legais.^{13,45,46}

Figura 1 – Fluxograma de tramitação.



Fonte: Câmara Deputados (BRASIL,2022)

Dado que, o processo legislativo é organizado por várias etapas, desde a formulação de um projeto, passando pela votação e aprovação, até a sua regulamentação. Estas fases são essenciais para a criação de leis que regem a vida de todos os cidadãos e que são necessárias para garantir a boa funcionalidade e a eficiência do Estado. No entanto, não é objetivo desse trabalho aprofundar as nuances do seu funcionamento, mas focar na importância de obter as informações e na transformação para uso institucional e social.

Por meio de ferramentas como o Portal da Transparência, as pessoas podem acompanhar, em tempo real, os gastos do governo e tomar decisões mais informadas sobre a gestão dos recursos. Além disso, desde a aprovação da Lei de Acesso à Informação, é possível solicitar informações oficiais de órgãos públicos e acompanhar o andamento dos processos.^{48,47,46}

Em síntese, o acesso público à informação governamental é fundamental para garantir a transparência na administração pública. Para isso, as informações relevantes devem ser disponibilizadas publicamente com utilização de ferramentas tecnológicas para facilitar o acesso e

a disseminação desses dados. Só assim é possível assegurar que a população e outras instituições possa acompanhar de perto o trabalho desenvolvido pelo governo e exercer seu direito de fiscalizá-lo.^{50,51,52}

4.4 A Informação Legislativa

A Constituição Federal de 1988,²⁵ atribui ao Congresso Nacional a função de elaborar e fiscalizar a legislação. Dessa forma, a informação legislativa é oriunda do exercício das funções de elaboração legislativa e de fiscalização advindas da competência constitucional do Congresso Nacional. Sua principal atribuição a elaboração das leis, são propostas pelo Poder Executivo e discutidas e votadas pelos parlamentares. Além disso, o Congresso também é o órgão responsável por fiscalizar o Poder Executivo, incluindo o controle da execução orçamentária e a fiscalização da atuação dos órgãos públicos.^{53,6,2}

A informação legislativa é, em sentido amplo, toda a informação produzida ou recebida pelo órgão legislativo e tratada de forma corporativa (de acordo com políticas, regras, diretrizes).⁵⁰ Inclui, portanto, a informação sobre o processo legislativo (emenda à Constituição, projetos de lei, etc.), bem como a informação sobre a atuação do órgão legislativo (votações, comissões, etc.). Também pode ser entendida um conjunto de normas, princípios e procedimentos que regem a produção, o armazenamento, a disseminação.^{50,51,53}

É um instrumento fundamental para o desenvolvimento das atividades legislativas, uma vez que permite o acesso às informações necessárias para o exercício do mandato parlamentar. A informação também dá ao cidadão a oportunidade de participar ativamente da política, pois ele pode acompanhar e discutir as propostas e ações dos governantes.^{52,51}

Assim, a informação do processo legislativo ajuda a fortalecer a democracia, pois permite que o cidadão fiscalize o poder público e verifique se as leis e decisões estão sendo cumpridas. Isso ajuda a assegurar que as leis sejam justas e que a democracia seja mantida.^{25,50,54,51}

4.5 Matérias legislativas no Contexto do Sistema Único de Saúde Brasileiro

A construção do sistema de saúde pública no Brasil deve ser analisada como um processo político e histórico.⁵⁵ A luta pela democratização da saúde como parte da vida social foi protagonizada por diversos setores organizados, especialmente acadêmicos e trabalhadores em saúde, que se opunham à mercantilização capitalista e à cooptação permanente da saúde liderada pela medicina.

Essa conquista social do Brasil, se deu a partir da idealização de um modelo universal e gratuito, em contraposição ao modelo fragmentado e excludente existente até então. Esse processo teve início ainda no contexto do regime militar, quando a luta pela democratização da saúde era uma das reivindicações do movimento sanitário.⁵⁶

Assim, compreender a saúde como um processo político e histórico é fundamental para entendermos a importância do Sistema Único de Saúde (SUS) e a necessidade de continuarmos

lutando por sua manutenção e aprimoramento. A saúde não é apenas uma questão técnica ou econômica, mas sim uma questão de justiça social e cidadania.

A Constituição Federal de 1988 representou um avanço na saúde pública do Brasil, ao institucionalizar o Sistema Único de Saúde (SUS) e reconhecer a saúde como direito fundamental do cidadão e dever do Estado. O SUS trouxe mudanças significativas ao romper com a lógica de um sistema fragmentado e excludente, garantindo acesso universal e igualitário às ações e serviços públicos de saúde.^{55,56}

O papel do Poder Legislativo federal é fundamental para o fortalecimento do SUS. É por meio de leis e proposições legislativas que se pode promover ações e programas que visem a qualidade do sistema de saúde, garantindo a ampliação do acesso e a melhoria dos serviços prestados. Por meio de uma série de instrumentos normativos utilizados para sistematizar a estrutura e as práticas institucionais do sistema de saúde, incluindo emendas constitucionais, leis ordinárias, medidas provisórias, decretos e portarias.^{57,58}

A Lei n. 8.080/90 foi uma importante conquista legislativa na área da saúde no Brasil. Surgiu a partir de um projeto de lei que tratava da participação da sociedade na gestão do Sistema Único de Saúde (SUS). Esse projeto de lei foi amplamente discutido e debatido antes de ser aprovado e se tornar lei, e seu objetivo principal era garantir a participação popular na formulação das políticas de saúde e na fiscalização da sua implementação. Com essa lei, o SUS foi consolidado como um sistema de saúde universal, equânime e integral, com a participação da sociedade como um dos seus princípios fundamentais.⁵⁹

Entretanto, para que o SUS possa ser consolidado e aprimorado, é necessário que a produção legislativa em saúde esteja em consonância com os princípios de democratização reivindicados pela sociedade.⁵⁸

Em um estudo realizado por Marques e Delduque (2012)⁵⁷ pontua que, o comprometimento do Poder Legislativo federal com o SUS tem sido insuficiente. É comum que projetos de leis e emendas constitucionais relacionados à saúde fiquem parados no Congresso Nacional, sem avanço ou discussão efetiva. Além disso, muitas vezes, a produção legislativa em saúde é influenciada por interesses corporativos e políticos, prejudicando a efetividade das políticas públicas.

O Congresso Nacional, como detentor do monopólio da produção legal no Brasil, é responsável por legislar sobre questões relacionadas à saúde. No entanto, há uma falta de acompanhamento e controle da eficiência e eficácia das leis aprovadas nessa área, além da inexistência de mecanismos de avaliação de impacto dessas leis. Isso pode levar a problemas no sistema de saúde e agravar a desigualdade social.^{59,57,58}

O Sistema Único de Saúde (SUS) é influenciado pela esfera das políticas públicas e sua interconexão com o Poder Legislativo. Isso implica em diversas consequências para os gestores e pesquisadores que necessitam monitorar e avaliar a produção legislativa em saúde. É importante buscar informações atualizadas e confiáveis para esses fins, para compreender como as decisões políticas afetam a saúde da população e contribuir para aprimorar as políticas de saúde.^{57,60,58}

A análise da produção legislativa em saúde é uma importante ferramenta para avaliar o desempenho do Sistema Único de Saúde (SUS) e identificar a efetividade ou lacunas das políticas públicas nessa área. No entanto, essa análise pode ser metodologicamente desafiadora devido à sobreposição de fontes de dados numerosos e complexos, como a extensa produção científica sobre o tema. Além disso, o escopo da produção legislativa em saúde é vasto e inclui uma grande volume de proposições sob o filtro temático da saúde. Nesse sentido, é fundamental compreender que a origem e dependência do SUS na esfera das políticas públicas são marcadas pela interconexão do legislativo na totalidade. Por isso, gestores e pesquisadores precisam buscar informações atualizadas para fins de monitoramento e avaliação da produção legislativa em saúde, a fim de entender como as políticas públicas têm sido implementadas e quais são seus impactos na saúde da população.^{57,58}

Isso acontece, entre outros motivos, porque esse tipo de sistema gera uma massa de dados significantes, mas predominantemente heterogêneos, motivo pelo qual aproximações da natureza de análise de dados são cotadas. Não só para o aumento da eficiência e espectro de tarefas anteriormente realizadas, mas para a inauguração de tantas outras que seriam improváveis sem a computação desses insumos. Nominalmente a realização de cruzamento e classificação de informação obtida através de uma alimentação automática de dados oriundos, primariamente, de portais públicos desembaraçados.^{58,60}

Em linhas gerais, podem ser traçados indicadores para a mensura da transparência, ou seja, do nível de abertura de um determinado portal ou banco de dados. Estes indicadores podem ser ilustrados, mas não se limitam às questões tais quais: a) qual a natureza, tanto em nível geral quanto em detalhe, de temas e categorias subtemáticas sobre quais os dados versam (saúde, segurança, finanças governamentais, poluição etc.); b) os parâmetros técnicos que garantem a abertura dos dados para leitura automática por máquinas (se estão em arquivos com formatos condizentes); c) existência de barreiras para o livre uso e análise (licenças livres; não restrições ao acesso); d) existência de incentivos estatais para a apropriação social de dados abertos; e) legislação condizente; f) produção de dados com qualidade para o livre uso e análise; f) desenvolvimento de ações de participação, entre outros.^{57,60,58}

Esse conjunto de indicadores vigentes tem duas incompletudes ou fragilidades relevantes: (i) carece de uma efetiva análise qualitativa dos conjuntos de dados em sua estrutura interna e; (ii) carece de análises específicas em relação aos ‘Dados Abertos do Poder Legislativo’. Para tentar avançar sobre essas duas lacunas, estudos de análises envoltos em tomada de ação são bastante desejáveis e indispensáveis.^{60,58}

Até o momento, a Câmara dos Deputados utiliza um portal¹ de pesquisa com dados das Matérias Legislativas apresentadas, ainda que seja possível pesquisar por tipo de proposições conforme os filtros disponíveis. O portal não disponibiliza nenhum tipo de visualização dessas informações aos usuários.⁴⁶

Essa aplicação apenas libera resultados em formato de uma planilha, Durante testes

¹ <https://www.camara.leg.br/buscaProposicoesWeb/pesquisaSimplificada>

feitos, a capacidade encontrada para extração, foi limitada a 1500 linhas por busca. Dificultando a análise e processamento das variáveis pelo usuário. Que também são reduzidas, quando comparadas a disponibilização dos dados brutos via API.

Figura 2 – Portal de buscas por proposições

Pesquisa simplificada
Informe pelo menos um dos argumentos de pesquisa abaixo:

Assunto

Tipo da Proposição

Tipos mais pesquisados (Selecionar todos, Limpar seleção)

- PEC - Proposta de Emenda à Constituição
- PLP - Projeto de Lei Complementar
- PL - Projeto de Lei
- MPV - Medida Provisória
- PLV - Projeto de Lei de Conversão
- PDC - Projeto de Decreto Legislativo
- PRC - Projeto de Resolução
- REQ - Requerimento
- RIC - Requerimento de Informação
- RCP - Requerimento de Instituição de CPI
- MSC - Mensagem
- INC - Indicação

Outros Tipos (Selecionar todos, Limpar seleção)

- AVN - Aviso (CN)
- CAC - Comunicado de Alteração do Controle Societário
- CAE - Relatório do CAE
- CCN - Consulta do Congresso Nacional
- CMC - Comunicação de Medida Cautelar

Número Ano

Formato 9999. A partir de 1946.

Autor

Em tramitação

Todas Sim Não

PESQUISAR **LIMPAR**

Fonte: Câmara Deputados (BRASIL, 2022)

A qualidade e a completude dos dados são essenciais para uma análise precisa e adequada de qualquer tema. No campo da saúde, a falta de informações completas pode prejudicar a compreensão e avaliação adequada do sistema de saúde, sobretudo para os usuários que muitas vezes não possuem as habilidades necessárias para trabalhar com grandes volumes de dados.³

A análise comparativa e a identificação de tendências são cruciais para a tomada de decisão em saúde, e a falta de completude dos dados pode levar a conclusões errôneas ou imprecisas. É importante, portanto, garantir a disponibilidade e a qualidade dos dados para permitir uma análise adequada e uma tomada de decisão embasada.^{11,61,8}

A adoção de sistemas de informação automatizados pode ser uma estratégia organizacional eficiente, isso se torna especialmente importante diante do grande volume de dados

relacionáveis, mas não diretamente relacionados, coletados e analisados atualmente.^{43,10} No entanto, a falta de dados completos pode prejudicar a análise e dificultar a comparação de variáveis e identificação de tendências, levando a conclusões imprecisas (Silva et al., 2020). Assim, a adoção de sistemas de informação eficientes pode facilitar a tomada de decisões em políticas públicas, ao invés de causar desnorteamento diante das grandes proporções de dados disponíveis.^{43, 46}

Tal fato pode apontar para a facilitação da tomada de decisões em políticas públicas, em oposição a um desnorteamento causado pelas grandes proporções que tomaram os dados coletados no mundo contemporâneo. Que não se traduzem, necessariamente, em avaliação.⁸

A ciência de dados, se consolidou como um campo de convergência interdisciplinar, científica, acadêmica, filosófica e prática, principalmente por cientistas da matemática, matemática do problema estatístico e acadêmico com básico em análise. Como os médicos e sanitaristas no caso da saúde, no entanto, é possível incluir grupos de ciência de dados nas equipes de especialistas como biólogos, geneticistas, economistas, financistas, geógrafos, advogados e historiadores.⁴³

Um dos seus principais objetivos são descobrir novas percepções a partir da análise de dados, modelando e interpretando-os. São usadas técnicas de estatística, ciência da computação, inteligência artificial e aprendizado de máquina para exploração e análise de grandes conjuntos de dados. Auxiliando na tomada de decisões estratégicas e na criação de soluções para problemas complexos.⁴³

5 MÉTODO, TÉCNICA E PROJETO APLICADO

5.1 Identificação de Método e Técnicas

Tendo em vista a natureza de conhecimentos interdisciplinares e de projeto aplicado desse trabalho, O ponto de partida para automatização dos dados em saúde da CD. O ponto de partida foi reconhecer a complexidade da combinação dos léxicos diferenciados, das práticas e da combinação de métodos empírico-teóricos.^{47,2,46}

Nesse sentido, o objeto da pesquisa-intervenção ora tratado é transversal e assume dimensões predominantemente de conciliação de práticas. Em geral, ainda que se perceba que o projeto lida com quantidades vultosas de dados, seu âmago é de abordagem qualitativa, que exigiu no decorrer da implantação, técnicas diferenciadas para alcance de resultados.^{62,63}

As subseções a seguir detalharão as técnicas aplicadas, resgatando enunciados e princípios sobre o processo de modelagem/ mineração de dados. Cabe destacar, que a linha mestra da modelagem concebida foi feita com base no repositório de dados disponibilizado pela API da CD.

Tradicionalmente, a mineração de dados e a descoberta de conhecimento eram realizadas manualmente. Com o passar do tempo, a quantidade de dados – disponíveis em diferentes softwares - cresceu para mais de um *terabyte* inviabilizando qualquer manutenção ou controle manual.⁶⁴

Knowledge Discovery in Databases (Descoberta de conhecimentos em base de dados) ou KDD, é uma área interdisciplinar com foco no desenvolvimento de métodos e técnicas para a extração de padrões e conhecimento a partir de dados.^{65,66} É um processo de descoberta de padrões, relacionamentos e tendências úteis a partir de grandes volumes de dados. É um processo iterativo que normalmente envolve limpeza de dados, integração de dados, seleção de dados, transformação de dados, descoberta de padrão e avaliação de padrão.^{67,68} O KDD também inclui o uso de mineração de dados e algoritmos de aprendizado de máquina para descobrir padrões e relacionamentos nos dados. Permitindo que as organizações tomem melhores decisões descobrindo conhecimento útil e acionável a partir dos dados.

Entre tais etapas, a que se faz mais importante é a mineração de dados, que comprova o pressuposto de que dados se transformam em informação, e posteriormente em conhecimento, levando à sabedoria, tornando a mineração de dados um processo imprescindível para as organizações, este processo é composto por cinco etapas, sendo: (1) Seleção: (2) Pré-processamento: (3) Transformação: (4) Mineração de dados: (5) Análise dos resultados:³⁴Interessa, então, que a decisão para a execução do projeto aplicado, ora apresentado, contou com KDD e todas as suas etapas de formação, descrito a seguir.^{66,65,64,68,67}

1. Etapa de seleção de Dados: Nessa fase foi preciso desenvolver uma compreensão do domínio do Projeto Aplicado, como passo preparatório inicial que interliga um cenário com decisões que refletem no resultado alcançado (sobre transformação, algoritmos, desempenho). As pessoas responsáveis por um KDD projeto precisam entender e definir os objetivos do

usuário final e o ambiente em que o processo de descoberta de conhecimento ocorrerá (incluindo conhecimento prévio relevante). Essa fase permite a sua revisão à medida que o processo de KDD avança.^{64,65}

Os objetivos do KDD, inicia-se com pré-processamento dos dados, definido nas próximas três etapas. Alguns dos métodos são semelhantes aos algoritmos de mineração de dados usados no contexto de pré-processamento.⁶⁷

Selecionando e fabricando um conjunto de dados onde a descoberta será realizada. Definidos os objetivos, os dados que serão utilizados para a descoberta do conhecimento devem ser escolhidos. Incluindo descobrir quais fontes estão disponíveis, obter dados complementares quando for necessário. A seguir, integrar todos os dados para a descoberta de conhecimento em um conjunto de dados, incluindo os atributos que serão considerados para o processo.^{67,66}

Este processo é muito importante porque a mineração dos dados aprende e descobre a partir dos dados disponíveis. Forma-se a base de evidências para a construção dos modelos. Se alguns atributos importantes estiverem faltando, o trabalho pode ser comprometido.^{68,66}

Assim, quanto mais características forem consideradas, maiores são as chances de um resultado relevante.^{65,64} Em compensação, coletar, organizar e fazer o manejo de repositórios de dados complexos é oneroso e há uma janela de oportunidade de melhor entendimento dos fenômenos. Essa contrapartida representa um aspecto em que o aspecto interativo e iterativo do KDD está ocorrendo. Nesse início, importa ter o melhor conjunto de dados acessíveis, expandindo e observando o impacto referente à descoberta e modelagem do conhecimento.⁶⁸

2. Pré-processamento e limpeza. Nesse momento, a confiabilidade dos dados passa por refinamento, integra a limpeza de dados, remoção de ruídos ou discrepâncias, retirada de células vazias. Pode envolver métodos estatísticos complexos ou usar algoritmos de mineração de dados especificamente desenvolvidos para essa situação.⁶²

Exemplificando, se alguém reear que um determinado atributo é de baixa confiabilidade ou tem muitos dados insignificantes (lixo) ou ausentes. Esse atributo pode se tornar o objetivo de um algoritmo supervisionado de mineração de dados. Um modelo de previsão para este atributo será desenvolvido e, em seguida, os dados ausentes podem ser previstos. A aplicação responsável por essa ação tem prudência e se concentra em níveis que dependem de diversas condições estabelecidas. De todo modo, se mostra imprescindível estudar os aspectos que por vezes, revelam por si só.^{65,64}

3. Transformação de dados: Nesta etapa, é preparada e desenvolvida a geração de melhores dados para a mineração de dados. Os métodos aqui incluem redução de dimensão (como seleção e extração de recursos e amostragem de registros) e transformação de atributos (como discretização de atributos numéricos e transformação funcional). Essa etapa pode ser crucial para o sucesso de todo o projeto KDD e está descrita na íntegra no próximo capítulo.^{65,64}

Essa fase é muito específica no projeto. Por exemplo, em Proposições legislativas, o quociente de características pode se tornar o fator mais relevante, e não somente a categoria. Em análises dos temas, talvez precisemos considerar efeitos além de nosso controle, bem

como esforços de outras questões, como similaridade das palavras-chave para criação de novo agrupamento. No entanto, mesmo que não usemos a transformação no início, podemos obter um efeito surpreendente que nos indica a transformação necessária (na próxima iteração).^{67,68 7}

Assim, o processo KDD reflete sobre si mesmo e leva a uma compreensão da transformação necessária. Essa modelagem permite, por exemplo, mecanismo do *Business Intelligence*¹, demonstra que a disponibilidade do dado ficou mais rápida e efetiva.⁶⁴

4. Na fase de Mineração de dados: É importante utilizar o tipo mais adequado. Por exemplo, regressão, classificação. Isso depende principalmente dos objetivos do KDD e também das etapas passadas. Dois objetivos são principais na Mineração de Dados: previsão e descrição. A previsão é muitas vezes referida como mineração de dados supervisionada, enquanto a mineração de dados descritiva inclui os aspectos não supervisionados e de visualização da mineração de dados.⁶⁵

A maioria das técnicas de mineração de dados é baseada em aprendizado indutivo, onde um modelo é construído explícita ou implicitamente pela generalização de um número suficiente de exemplos de treinamento. A suposição subjacente da abordagem indutiva é que o modelo treinado é aplicável a casos futuros. A estratégia também considera o nível de meta-aprendizagem para o conjunto específico de dados disponíveis.^{68,67}

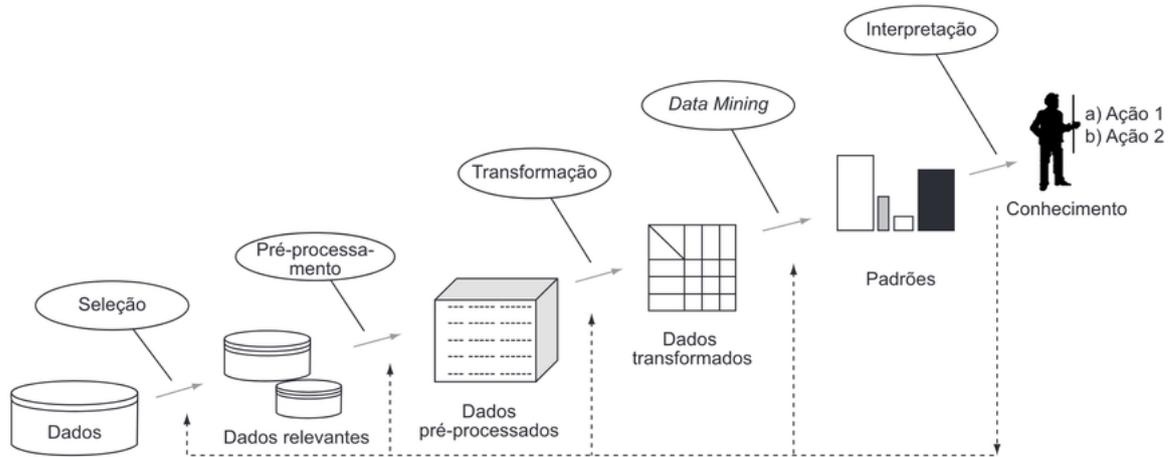
5. Análise dos resultados: Assim, é preciso escolher o algoritmo de mineração de dados. Esta etapa inclui a seleção do método específico a ser usado para a busca de padrões (incluindo múltiplos indutores). Por exemplo, considerando precisão em relação à compreensibilidade, a primeira é melhor com redes neurais, enquanto a segunda é melhor com árvores de decisão. Para cada estratégia de meta-aprendizagem diversas são as possibilidades de como isso pode ser executado.^{65,66}

A meta-aprendizagem consiste em explicar o que faz com que um algoritmo de mineração de dados seja bem-sucedido ou não em um problema específico. Assim, esta abordagem tenta compreender as condições sob as quais um algoritmo de Data Mining é mais adequado. Cada algoritmo tem parâmetros e táticas de aprendizado (como validação cruzada de dez vezes ou outra divisão para treinamento e teste).^{67,68,66}

Nesta etapa, é provável que se tenha que rodar e testar várias vezes até que um resultado satisfatório seja obtido. Por exemplo, ajustando os parâmetros de controle do algoritmo, como o número mínimo de atributos em uma árvore de decisão. Feito isso, avaliamos as interpretações dos padrões minerados (regras, confiabilidade etc.), em relação aos objetivos definidos na etapa.^{64,67}

¹ O business intelligence (BI) combina análise empresarial, mineração de dados, visualização de dados, ferramentas/infraestrutura de dados e práticas recomendadas para ajudar as organizações a tomar decisões impulsionadas por dados.⁶⁹

Figura 3 – Etapas do processo KDD



Fonte: Fayyad et al. (1996).

5.2 Integração dos Dados da Câmara dos Deputados API (Application Programming Interface)

Nesta etapa é provável que se tenha que rodar e testar várias vezes até que um resultado satisfatório seja obtido. Por exemplo, ajustando os parâmetros de controle do algoritmo, como o número mínimo de atributos em uma árvore de decisão. Uma API permite a comunicação entre soluções e serviços sem a necessidade de entendimento de ambos os lados sobre a implementação. A partir de uma forma de comunicação padrão, cada solução possui autonomia para definir seu formato de desenvolvimento.^{70,71}

Ao desenvolver novas ferramentas e soluções (ou ao gerenciar aquelas já existentes), as APIs oferecem a flexibilidade necessária para simplificar o design, a administração e o uso, além de fornecer oportunidades de inovação. Com a utilização de API, a integração de novos componentes de aplicações a uma arquitetura já existe sendo feito de forma simplificada. Muitas vezes, as necessidades organizacionais mudam rapidamente para responder aos mercados digitais em transformação.^{5,71}

Nesse ambiente, novas tecnologias podem redefinir o setor inteiro com uma nova aplicação. Por isso, é importante oferecer suporte à implantação e desenvolvimento de serviços inovadores.⁵ Utilizando como exemplo um departamento que permite a consulta da disponibilidade de um produto diretamente com a distribuidora, essa aplicação poderia ter um alto custo, com funcionalidades limitadas e exigiria longos períodos de desenvolvimento além de manutenções.^{70,5}

Como alternativa, esse mesmo departamento poderia fornecer uma API para consultar a disponibilidade de um produto diretamente no estoque. Dessa forma, o acesso aos dados por meio de uma API permite aos usuários consolidarem informações sobre seu inventário em um único local, além de possibilitar a contribuição de desenvolvedores que trabalham para a distribuição em questão. Independentemente de alterações nos sistemas, contanto que o comportamento da

API não mude o cliente não ao ir a ter nenhum impacto.^{72,73,74}

A Câmara dos Deputados é um exemplo de disponibilização de informações legislativas por API. No website do órgão é possível ter acesso às instruções necessárias para extrair os dados sobre normas jurídicas utilizando essa tecnologia. A tramitação da proposição é o que indica uma mudança na mesma. Dessa forma, por outro recurso também fornecido pela API, é possível listar todas as tramitações de uma proposição por meio de seu identificador (id). Para cada tramitação temos as seguintes informações:^{72,75 74}

5.3 OpenAPI

A iniciativa Open API foi desenvolvida por um grupo de especialistas da indústria que reconhecem a significância da padronização dos dados e da utilização criação de APIs.^{76,77} A OpenAPI Specification (OAS) define uma interface padrão independente de linguagem para APIs, permitindo que humanos e computadores descubram e entendam os recursos de uma aplicação sem exigir acesso a código-fonte, documentações adicionais ou inspeção de tráfego de rede.⁷⁶

Quando implementado seguindo as orientações da OAS a aplicação consumidora poderá utilizar o serviço disponibilizado com esforços mínimos de implementação. Também poderá ser utilizada por ferramentas de geração de documentação para exibição da API, ferramentas de geração de códigos e ferramentas de teste.^{77,76,70}

Um arquivo que implementa o padrão Open API permite que seja documento é descrito toda a API, incluindo:^{76,77}

- *Endpoints*² disponíveis (usuários) e operações em cada *Endpoints* (GET usuários, POST /usuários);
- Parâmetros de operação de entrada e saída para cada operação
- Métodos de autenticação;
- Informações de contato;
- Licenças;
- Termos de uso e outras informações.

² Um endpoint de comunicação é um tipo de nó de rede de comunicação. É uma interface exposta por uma parte comunicante ou por um canal de comunicação.⁵

Figura 4 – Estrutura Básica OpenAPI

```

1. openapi: 3.0.0
2. info:
3.   title: Sample API
4.   description: Optional multiline or single-line description in [CommonMark](http://commonmark.org/help/) or HTML.
5.   version: 0.1.9
6.
7. servers:
8.   - url: http://api.example.com/v1
9.     description: Optional server description, e.g. Main (production) server
10.  - url: http://staging-api.example.com
11.    description: Optional server description, e.g. Internal staging server for testing
12.
13. paths:
14.   /users:
15.     get:
16.       summary: Returns a list of users.
17.       description: Optional extended description in CommonMark or HTML.
18.       responses:
19.         '200': # status code
20.           description: A JSON array of user names
21.           content:
22.             application/json:
23.               schema:
24.                 type: array
25.                 items:
26.                   type: string

```

Fonte: Autores

A estrutura base de uma escrita seguindo as definições *OpenAPI* se dá pela seguinte forma, na estrutura base são identificadas algumas sessões. São elas: A seção de informações contém as seguintes características de uma API: Título, descrição (opcional) e versão. No campo *Title* (título) deve ser especificado nome da API e o campo *description* (*descrição*) deve ser utilizado para descrever sobre sua API. Sendo utilizada mais de uma linha, são suportados comandos de uma linguagem simples de marcação.^{76,77}

5.4 Json

A formatação *Json*, é específica em que as informações são enviadas e enviadas recebidas entre duas aplicações que se comunicam e tem como base um subconjunto da linguagem de programação *Javascript*. Com alta legibilidade tanto por ser humano como por máquinas, o formato *Json* é completamente independente, a linguagem é utilizada em aplicações desenvolvidas nas mais diversas linguagens de programação.⁷⁸

Figura 5 – Exemplo de Arquivo Json

```

{
  "orders": [
    {
      "orderno": "748745375",
      "date": "June 30, 2088 1:54:23 AM",
      "trackingno": "TN0039291",
      "custid": "11045",
      "customer": [
        {
          "custid": "11045",
          "fname": "Sue",
          "lname": "Hatfield",
          "address": "1409 Silver Street",
          "city": "Ashland",
          "state": "NE",
          "zip": "68003"
        }
      ]
    }
  ]
}

```

Fonte: Autores

5.5 Processo de Desenvolvimento

5.5.1 Dados Disponíveis

As relações e acontecimentos da Câmara dos Deputados geram várias informações que são de extrema importância no contexto sociopolítico do Brasil. Boa parte dessas informações estão disponíveis seguindo a lei de dados abertos. Buscando facilitar a busca dessas informações, foi disponibilizada uma API pela Câmara dos Deputados.⁷⁹

Esse serviço foi disponibilizado no formato Swagger, que tem sua especificação definida no OpenAPI. Essa forma de apresentação, permite consultar os dados necessários para interagir com o serviço, bem como os resultados gerados a cada consulta. As informações disponíveis na API da Câmara dos Deputados são divididas nas sessões e composta de diversos Endpoints para as mais diversas finalidades.⁸⁰

5.5.2 Estrutura de disposição dos dados

A solução disponibilizada pelo órgão federativo fornece três formatos de resposta para as buscas em (*text, xml e json*). Destes, o mais utilizado atualmente é o JSON, devido à facilidade de manipulação.^{78,81}

A listagem de proposições é feita em períodos, que podem ser informados durante a requisição. Caso nenhum período seja informado, serão retornadas pela API as proposições que tiveram alterações nos últimos trinta dias. As informações de cada proposição são:

- **Id:** Número identificador da proposição;
- **Uri:** Caminho de requisição para detalhamento da proposição;
- **Sigla Tipo:** Sigla do tipo relacionado a proposição;
- **Cod Tipo:** Identificador do tipo relacionado a proposição;
- **Número:** Número da proposição;
- **Ano:** Ano em que a proposição foi publicada;
- **Ementa:** Ementa relacionada à proposição.

Devido ao grande volume de tramitações diárias na CD, a quantidade de registros nesse *endpoint* pode se tornar muito grande a depender do período de busca, então, além dos dados de resposta são retornadas informações adicionais para auxiliarem em uma busca paginada.

O *endpoint* mencionado retorna informações de identificação da proposição sem entrar em detalhes sobre seu conteúdo e disposições uma forma de se obter esse tipo de detalhamento e por uma requisição para o *endpoint* que recebe o identificador (*id*) de uma proposição como parte dos parâmetros e como resultado de requisição, são retornados os seguintes campos:⁸⁰

- **Data Apresentação:** Data de apresentação da proposição
- **Uri Órgão Numerador:** Caminho de requisição para detalhamento do Órgão numerador;
- **Status Proposição:** Objeto com informações sobre o status da proposição;
- **Uri Autores:** Caminho de requisição para listagem dos autores da proposição;
- **Descrição Tipo:** Descrição do tipo relacionado a proposição;
- **Ementa Detalhada:** Ementa associada a proposição descrita de forma detalhada;
- **Keywords:** Lista de palavras-chave, relacionadas a proposição;
- **Uri Prop Principal:** Caminho de requisição para detalhamento da proposição inicial;
- **Uri Prop Anterior:** Caminho de requisição para detalhamento da proposição anterior;
- **Uri Prop Posterior:** Caminho de requisição para detalhamento da proposição posterior;
- **Url Inteiro Teor:** Caminho de requisição para o texto completo da proposição
- **Uri Final:** Caminho de requisição para detalhamento da proposição final;
- **Texto:** Caminho de requisição para detalhamento da proposição inicial;
- **Justificativa:** Caminho de requisição para detalhamento da proposição inicial.

Além disso, é possível obter as informações sobre o proponente de uma respectiva proposição através de outro Endpoints que recebe como *parâmetro* o identificador (id) da proposição e como resultado da requisição, são retornados os seguintes campos:

- **Uri:** Caminho de requisição para detalhamento do deputado;
- **Nome:** Nome do deputado;
- **Cod Tipo:** Código do tipo relacionado ao deputado;
- **Tpo:** Tipo do deputado;
- **Ordem Assinatura:** Número relativo à ordem de assinatura da proposição;
- **proponente:** Número de proponente do deputado para a proposição.

Utilizando o código de deputado disponível como retorno da chamada explicitada no parágrafo anterior, podemos utilizar mais um recurso da API e listar as seguintes informações do deputado:

- **Id:** Código identificador do deputado;
- **Uri:** Caminho de requisição para detalhamento do deputado;
- **Nome civil:** Nome civil relacionado ao deputado;
- **Data Hora:** Data e hora da tramitação;
- **Sequência:** Número identificando a ordem do registro;
- **Sigla Órgão:** Sigla do órgão associado a tramitação;
- **Uri Órgão:** Caminho de requisição para detalhamento da tramitação; órgão relacionado a
- **Uri Último Relator:** Caminho de requisição para detalhamento do parlamentar vinculado a tramitação como relator;
- **Regime:** Texto descrevendo o regime da tramitação;
- **Descrição Tramitação:** Descrição da tramitação;
- **Cod Tipo Tramitação:** Código do tipo relacionado a tramitação;
- **Descrição Situação:** Situação atual da proposição;
- **Cod Situação:** Código da situação atual da proposição;
- **Despacho:** Texto de despacho da tramitação;

- **Url:** Caminho para documento completo da tramitação;
- **Âmbito:** Âmbito da tramitação

A tramitação da proposição é o que indica uma mudança na mesma. Dessa forma, outro recurso também fornecido pela API, é possível listar todas as tramitações de uma proposição por meio de seu identificador (id). Para cada tramitação temos as seguintes informações:^{75,82}

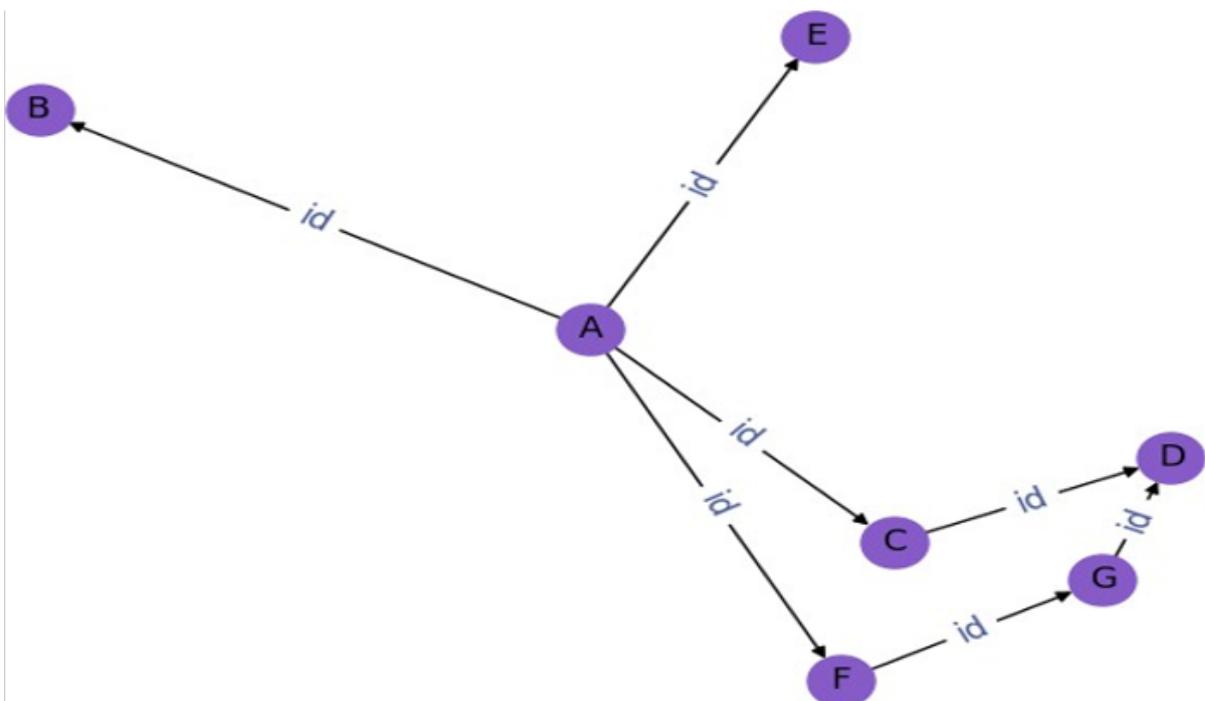
Da mesma forma que se faz importante o conhecimento de quem é o autor da proposição, é necessária a aferição dos dados do deputado. Essas informações podem ser recuperadas consultando o endpoint de busca do deputado já mencionado.⁷⁴

Já em relação aos temas das proposições, pode-se observar que a API da Câmara dos Deputados não fornece nenhum recurso para a listagem de proposições de acordo com um tema específico. Visando a recuperação de proposições relacionadas apenas a saúde, faz-se necessário para cada proposição, requisitarmos outro endpoint responsável por retornar o seu respectivo tema.^{75,83}

Outra informação que não está detalhada ao buscar uma proposição e que impacta diretamente suas tramitações são as votações a qual tal proposição é submetida. Dito isso, é preciso realizar uma requisição que retorna as informações.

Assim, é necessário utilizar o identificador do proponente de uma proposição para termos seu detalhamento, a votação segue o mesmo fluxo. É observado depois da descrição dos recursos nos parágrafos anteriores uma forte relação entre os mesmos. Essa relação é evidenciada na Figura 6, onde é possível ver um grafo direcional apontando a relação de cada recurso.

Figura 6 – Grafo da Relação dos Endpoints



Fonte: Autores

Para a compreensão completa da imagem anterior é importante a observação da Tabela que indica qual serviço o nó do grafo se refere. É relevante ressaltar que o nome atribuído a aresta é relacionado ao atributo do nó de saída.

Tabela 1 – Relação no Endpoint

Serviço	Endpoint
Listagem de Proposições	/proposições
Busca de Proposição	/proposições/id
Autores da Proposição	/proposições/id/autores
Buscar Deputado	/deputados/id
Tramitação da Proposição	/proposições/em/tramitação
Votações da Proposição	/proposições/id/votações
Detalhamento da Votação	/votações/id

Fonte: Autores

Considerando todas as informações explanadas nos parágrafos e o grafo que relaciona todos os recursos. É possível realizar a busca dos dados começando no 'A' e indo até um nó qualquer que não tenha nos adjacentes, repetindo o processo até que todos os nós sejam visitados.

5.5.3 Ferramenta de integração

Todas as informações são extraídas a partir do site via protocolo *HTTP* com método *GET*, gerando *JSONs* para serem interpretados. As principais bibliotecas para realizar essas buscas são: *Requests (Python)*, *URL (PHP)*, *Retrofit (Java)* e *http (Go)*.^{76,72}

Alguns campos de interesse ainda são descritos como identificadores ou siglas. É disponibilizado um recurso na API da Câmara dos Deputados para solucionar essas demandas. Nesse serviço, os dados são *JSON* estáticos e as buscas devem ser feitas para ser possível chegar à informação desejada.^{82,83}

Devido ao alto grau de relação entre os recursos disponíveis, as informações relevantes dos mesmos e as buscas em *JSONs* estáticos, é proposto que a informação seja extraída e armazenada em um banco relacional a fim de facilitar a gestão das informações e as consultas das relações entre estas que não estão disponíveis no serviço oferecido pela Câmara dos Deputados. Ao realizar a modelagem desse armazenamento é possível avaliar informações nulas no retorno dos serviços.⁸¹

A partir da API da Câmara é possível identificar a ausência de alguns dados, tornando necessário a criação de estratégias para buscar as informações faltantes em outros repositórios.

Além de evidenciar situações que podem representar fragilidades nos fatos que essas informações representam, podemos utilizar uma estrutura de percorrimento em grafos para realizar a busca nos *endpoints* da Câmara dos Deputados. Tendo em vista que existe um grande

volume de informações a serem recuperadas e que essas informações se relacionam de alguma forma.^{74,84,8}

Apesar disso, a Câmara dos Deputados ao permitir a recuperação dessas informações, se enquadra ao referenciado na Lei n.º 12.527. Garantindo assim, o acesso aos dados públicos por pessoas físicas ou empresas, com interesse ou necessidade de portá-los.²⁹ Tal como é a proposta desse documento, o uso dessas informações para uma futura categorização temática das matérias legislativas em saúde.

5.6 Técnicas de Pré-Processamento e Modelagem Aplicadas ao Banco de dados da API da CD.

O número de projetos de leis relacionados com a área da saúde, publicados diariamente. Inviabiliza a extração, análise, pré-processamento e categorização desses projetos de forma humana/manual devido ao tempo necessário para que todas essas etapas aconteçam.⁶⁸

No entanto, por automações torna-se viável a realização dos procedimentos listados anteriormente de forma mais assertiva e rápida, uma vez que, os seres humanos estão sujeitos a distrações.⁶⁷ Podem levar a falhas na execução das etapas, prejudicando assim, o resultado da análise. No entanto, isso não significa que uma inteligência artificial não tenha falhas, mas é bem menos suscetível a falhas quando comparado com o de um ser humano.

Por isso, o Swagger que utiliza o OpenAPI como diretriz aliado às linguagens Python e o formato Json permitem que as disponibilizações das informações consolidadas possam ser feitas de forma padronizada. Com suas respectivas documentações disponibilizadas para os usuários que forem utilizá-la.⁷⁶

5.6.1 *Python*

O Python é uma linguagem de programação de alto nível, interpretada, imperativa, orientada a objetos, funcional, de tipagem dinâmica e de *script*. Foi criada em 1989 por Guido Van Rossum no Instituto de Pesquisa Nacional para Matemática e Ciência da Computação (CWI). No dia 13 de maio de 2020, uma nova versão foi disponibilizada pela comunidade, a versão 3.8.3.⁸⁵

Um pouco divergente das demais linguagens de programação como Java, *GO* e *Dart*,³ propriedades de grandes empresas como Oracle e Google, o *python* possui um modelo de desenvolvimento comunitário. Aberto e gerenciado pela organização sem fins lucrativos, Python Software Foundation.⁸⁷

3

Java é uma linguagem multiplataforma, orientada a objetos e centrada em rede que pode ser usada como uma plataforma em si.

Go é uma linguagem de programação concorrente e procedural.

Dart é uma linguagem de programação orientada a objetos.

86

A linguagem *python* possui uma característica singular, a tipagem dinâmica, no entanto, ainda assim é possível utilizar tipos de dados bem definidos pela linguagem, bem como, realizar a criação de novos tipos de dados através de classes.⁸⁷

Tabela 2 – Tipos de Dado

Tipo de dado	Descrição
Str, unicode	Uma cadeia de caracteres imutável
List	Lista heterogênea mutável
Tuple	Tupla imutável
Set, frozenset	Conjunto não ordenado, não contém elementos duplicados
Dict	Conjunto associativo
Int	Número de precisão fixa, e ´ transparentemente convertido para <i>long</i> caso não caiba em um <i>int</i>
Float	Ponto flutuante
Complex	Número complexo
Bool	Booleando

Fonte: Autores

A proposta de projetar uma linguagem de propósito geral de alto nível e multi paradigma. Capaz de suportar paradigmas, orientado a objetos, imperativo, funcional e processual. Deixou como uma das principais características a simplicidade de leitura do código, uma vez que, em poucas linhas é possível elaborar um programa.

Com a filosofia de enfatizar a importância do esforço do programador sobre o esforço computacional, priorizando a legibilidade do código sobre a velocidade ou expressividade.^{85,87}

Por ser de fácil leitura e visualmente agradável, a separação de blocos de código é feita por espaços em branco (indentação) ao invés de delimitadores visuais como chaves, palavras ou caracteres especiais.

Nesse aspecto, é de suma importância que o desenvolvedor esteja atento. Uma vez que, a indentação, se faz obrigatória, ou seja, o aumento da indentação, indica o início de um novo bloco que termina a diminuição da indentação como exemplificado no Código 1.⁸⁵

Outra característica interessante do *python* é a sua capacidade de interoperar com diversas outras linguagens, principalmente código nativo. Através da documentação é possível encontrar exemplos de como usar o *Python C-API* para escrever funções em C que podem ser chamadas

diretamente do código *python*, no entanto, existem outras alternativas mais robustas e adequadas como o *Cython*, *Swig* ou até no mesmo o *CFFI*.⁸⁷

Por fim, a linguagem e seu interpretador estão disponíveis em diversas plataformas, como o *Unix (Linux, FreeBSD, Solaris, MacOS X, etc)*, *Windows*. Para os sistemas operacionais que eventualmente não suportem o *python*, basta que exista um compilador *C* disponível para ser possível gerar o código *python* a partir da fonte.^{87,85}

5.6.1.1 Biblioteca Pandas

Ainda em relação ao Data Frame é possível selecionar elementos segundo os critérios condicionais criados, segundo a documentação da biblioteca, chama-se essa operação de *Boolean Indexing*.⁸⁸

Essa indexação também possibilita a utilização de múltiplas colunas para checagem de condição, no entanto, para ser possível utilizar esse recurso, é necessário a utilização dos operadores *bitwise*, são utilizados quando precisamos realizar operações ao nível de bits com números inteiros, ou seja, trabalhar com sua representação binária. ao invés de *and*, *or* ou *not*.⁷⁹

Por fim, a biblioteca ainda oferece inúmeros recursos para visualização dos dados, leitura de arquivos, manipulação dos dados, dentre outros recursos que podem ser acessados e vistos através da documentação oficial .⁷⁹

5.6.1.2 Preparação dos Dados

Uma das principais funções do profissional que atua com dados com foco na análise ou na utilização de inteligência artificial é a atividade de pré-processamento de dados.⁷⁹

O pré-processamento/preparação é um conjunto de atividades que envolvem a preparação, organização e, claro, a estruturação dos dados. É uma etapa fundamental que precede as etapas de análises e previsões, uma vez que, será determinante para a qualidade final dos dados que serão utilizados, podendo, inclusive, impactar na predição.⁸⁷

Alguns dos principais problemas tratados pela fase de pré-processamento, são: elencar atributos faltantes, *outliers* (desvios) e escalas diferentes para valores iguais. Problemas estes sendo encontrados e tratados a partir de um conjunto de dados, também chamado de *dataset*.⁸⁷

5.6.1.3 Limpeza de dados

Quando analisado o conjunto de dados original (*dataset*) é possível identificar diversas partes irrelevantes, ou seja, que não tenham relevância para o seu uso futuro. Dessa forma, a limpeza envolve o manuseio e preenchimento de dados ausentes, redução de ruídos, remoção de valores e resolução de possíveis inconsistências.^{80,64}

Em relação aos dados faltantes/ausentes, essa situação é notória quando alguns dados não estão presentes no *dataset*. Existem diversas técnicas que podem ser aplicadas para resolver o problema, são exemplos:

- Remoção de registros nulos;
- Realizar uma média com os valores de mesmo atributo;
- Realizar uma mediana com os valores de mesmo atributo;
- Preencher o atributo ausente com os valores que mais aparecem no dataset

Uma importante observação, é que para cada técnica acima, é preciso a aplicação de uma estratégia diferente para lidar com o problema e nenhuma por si só resolve o problema por completo.^{65,64}

Já em relação aos dados ruidosos, são aqueles que não fazem sentido e que não podem ser interpretados pelas máquinas. Existem diversas causas para que esses dados estejam ruidosos, por exemplo: falha na coleta dos dados e erro na entrada dos dados.^{75,74,65}

Para tratamento do problema mencionado anteriormente, é possível a utilização do método de Binning, um processo de suavização de dados utilizado para minimizar os efeitos de pequenos erros de observação. Dessa forma, os valores originais são divididos em pequenos intervalos conhecidos como compartimentos e, na sequência, são substituídos por um valor geral calculado para esse compartimento.^{85,71,64}

Outra técnica que pode ser aplicada é a da regressão, onde os dados são suavizados ajustando-os a uma função de regressão. Podendo ser linear (variável independente) ou múltipla (várias variáveis independentes).

Por fim, também é possível citar a técnica do agrupamento, cuja abordagem agrupa os dados semelhantes em um cluster (grupos). Onde desvios podem ser tratados separadamente ou deixados de fora desses clusters.^{64,84}

5.6.1.4 Redução dos dados

Por fim, a redução de dados é utilizada para aumentar a eficiência e reduzir os custos, buscando tratar o volume de dados.

Computacionalmente processar toda essa quantidade pode se tornar um problema. Dentre as diversas atividades/técnicas utilizadas, podemos citar: agregação de cubo de dados, seleção de subconjunto de atributos e a redução da numerosidade.⁸⁸

A atividade de agregação de cubo de dados é utilizada na construção de um cubo de dados em um formato multidimensional que, apesar de gerar maior necessidade de armazenamento, permite um processamento mais rápido, uma vez que não necessita percorrer toda a base em busca de um valor específico.⁷⁹

Em relação à técnica de seleção de subconjunto de atributos, é a opção de utilizar os atributos altamente relevantes em detrimento dos menos relevantes. Para ser possível realizar a seleção de atributos, pode-se usar o nível de significância e o valor "p" do atributo, aqueles atributos com valor maior que o nível de significância podem ser descartados.^{85,87}

Por fim, a atividade da redução da numerosidade permite que os dados sejam substituídos ou estimados por alternativas de representação de dados menores. Tais como, modelos paramétricos (armazenam apenas os parâmetros do modelo em vez de dados reais) ou métodos não paramétricos, tais como, amostragem e o uso de histogramas.^{79,85}

5.6.2 Modelagem de banco de dados

Um Banco de Dados representa um aspecto real, nomeado de "mini mundo" ou "universo de discurso". Baseado no aspecto real, as mudanças decorrentes do "mundo real" são refletidas no "mini mundo" que, implicando em um banco de dados. O Banco de Dados pode ser definido como um conjunto de dados sobre o qual uma comunidade de usuários realiza operações de recuperação, atualização e remoção⁸⁹

Um Banco de Dados computadorizado pode ser mantido e criado por um grupo de aplicações escritas estritamente para aquela tarefa ou por um SGBD (Sistema de Gerência de Banco de Dados).

O SGBD é uma coleção de programas de propósito geral que facilita o processo de definição, construção e manipulação de banco de dados para várias aplicações. Onde a combinação do banco de dados e do SGBD pode-se chamar genericamente de sistema de banco de dados. Entre as suas funções estão:^{89,90}

- Definir: um banco de dados significa especificar os tipos de dados a serem armazenados;
- Construir: processo de armazenamento de dados em algum meio de armazenamento controlado pelo SGBD;
- Manipular: o banco de dados compreende certas funções como consulta para recuperar dados específicos, atualização de dado para refletir as mudanças percebidas no "mini mundo" e produzir relatórios.

5.6.3 Sistema de Banco de Dados

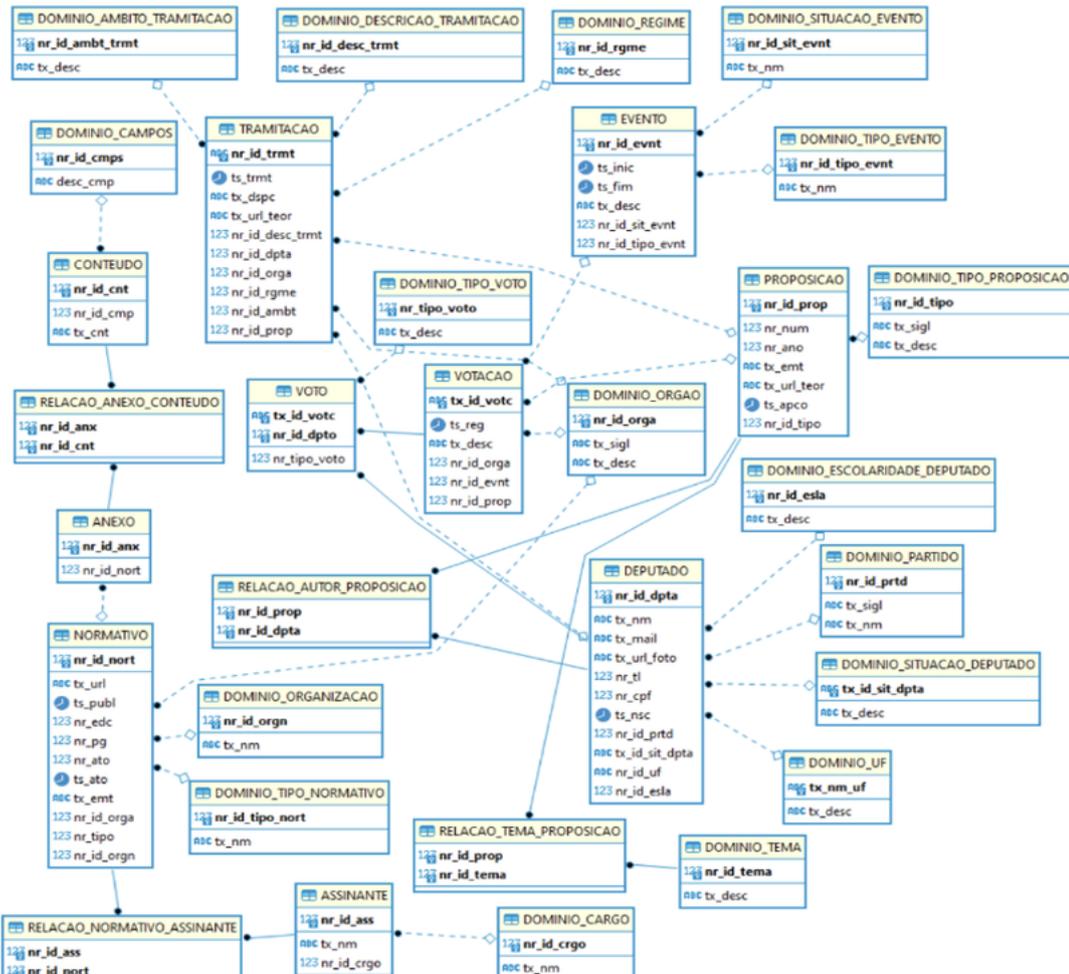
O Sistema de Banco de Dados pode ser definido como um sistema computadorizado de manutenção de registros, ou seja, é um sistema computadorizado cuja finalidade geral é o armazenamento de informações. Bem como, permitir que os usuários busquem e atualizem essas informações quando solicitado. Essas informações podem ser qualquer coisa que tenham algum significado ao indivíduo (representação do mini mundo) ou a organização a que o sistema deve servir.⁸⁹

5.7 Desenvolvimento do modelo

É importante observar que ao longo do artigo, serão utilizados os termos informação e dado, cujo significado serão interpretados e usados como sinônimos. Na literatura, é possível

encontrar autores que tratam o dado como algo que é realmente armazenado no banco de dados, enquanto” informação para se referir ao significado desses dados para determinado usuário.⁸⁹

Figura 7 – Modelagem de Dados CD



Fonte: Autores

5.7.1 Carregando o Dataset

A primeira etapa do pré-processamento e a criação de um alfabeto de preposições, conjunções e acentos que em momento oportuno serão utilizados como filtros. Considerando o processamento final apenas palavras que possuem relevância semântica no contexto geral.

Após uma seleção de palavras-chave que serão desconsideradas do texto geral, é realizada uma limpeza de tubulações incorretas e espaçamentos entre as palavras que em alguns casos estão em excesso.

Como primeira pasta podemos utilizar os dados de proposições legislativas do período de janeiro de 2022 a agosto de recuperados através da API da Câmara dos Deputados. Após a

recuperação São dispostos em um arquivo CSV (*common separated values*) e transformados em uma data frame através da seguinte instrução:

Figura 8 – Seleção de Registros das Tabelas Presentes no Banco de Dados da Câmara.

```
query="select 'tramitacao' as tabela, count(*) as registro from bd_camara.tramitacao T
union select 'proposicao', count(*) from bd_camara.proposicao P
union select 'proposicao_tema', count(*) from bd_camara.proposicao_tema PT
union select 'proposicao_votacao', count(*) from bd_camara.proposicao_votacao PV
union select 'autoria', count(*) from bd_camara.autoria A
union select 'deputado', count(*) from bd_camara.deputado D
union select 'deputado_profissao', count(*) from bd_camara.deputado_profissao DP
union select 'profissao', count(*) from bd_camara.profissao P
union select 'dominio_tema', count(*) from bd_camara.dominio_tema DT
union select 'orgao', count(*) from bd_camara.orgao O
union select 'votacao', count(*) from bd_camara.votacao V
union select 'voto', count(*) from bd_camara.voto Vt
ORDER BY 1"

kb=pg_roda_query(query)

kb$registro=format(kb$registro,0)
colnames(kb)=set_colnames(colnames(kb))
library(kableExtra)
kbl(kb,
  align = c("l", "r"),
  caption = "Registros das tabelas presentes no schema bd_camara" %>%
  kable_classic() %>%
  add_footnote("Fonte - CIEGES", notation="alphabet" , threeparttable = TRUE)
```

Fonte: Autores

Figura 9 – Número de Registros das Tabelas Presentes no Banco de Dados da Câmara.

tabela	registro
autoria	9.785
deputado	513
deputado_profissao	688
dominio_tema	30
orgao	3.226
profissao	172
proposicao	4.772
proposicao_tema	4.938
proposicao_votacao	2.388
tramitacao	110.871
votacao	2.366
voto	83.856

Fonte: Autores

Não existe um caminho único para se aplicar o pré-processamento em todos os dados, pois a depender do dado a ser analisado e da perspectiva de cada cientista de dados é possível optar por uma técnica ou outra. O recomendado é experimentar diversas técnicas, analisar os resultados, e ver qual terá a melhor relevância para o seu conjunto de dados. Apesar dos dados sobre proposições estarem disponibilizados a partir de uma API.⁸⁸

É possível identificar algumas inconsistências na disponibilização desses dados ou até mesmo na caracterização do fato representado pela informação. Proposições retornadas sem informações da ementa, temas e URL são exemplos disso.

Podemos propor soluções alternativas para complementar as informações faltantes como uma implementação de uma *web scraping* para buscar essas informações em outra fonte, porém, esse procedimento não é parte e nem responsabilidade do pré-processamento, além disso, a extração de conteúdo de um PDF, como as proposições, não é uma tarefa fácil, uma vez que, em alguns casos será preciso a elaboração de algoritmos de reconhecimento de texto ou imagem.^{88,79}

Com o grande volume de informações dispostas a necessidade evidente de meios computacionais para os aferimentos, cálculos e análises computacionais desses dados. Sua disponibilização da informação de forma íntegra e correta é essencial para que esses objetivos sejam atingidos.⁷⁹

Adicionado ao fato de que os dados recuperados precisam ser organizados para serem disponibilizados para um modelo de inteligência artificial, a etapa de limpeza de dados no pré-processamento se confirma essencial quando aplicado ao contexto da API da Câmara dos Deputados.

A linguagem *Python* aliada às bibliotecas *Pandas* e *Matplotlib*⁸⁸ fornecem os recursos necessários para ser realizado o pré-processamento de informações que serão utilizadas em contextos de inteligência artificial e análise de dados, além de permitir a visualização e disponibilização dos dados de forma visualmente clara e com sumarização de relevância semântica para o contexto da análise.

6 PROTÓTIPO DO PAINEL DE MONITORAMENTO DE PROPOSIÇÕES LEGISLATIVAS

As ferramentas de Business Intelligence (BI) apresentam uma combinação de técnicas, ferramentas, processos e sistemas usados para transformar dados brutos e complexos em informações úteis e percepções. Que podem ser usados para melhorar a tomada de decisão das instituições.⁹¹

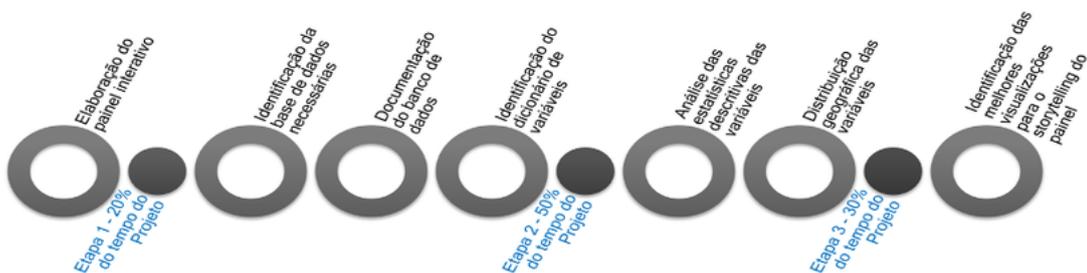
Estas ferramentas podem ajudar os usuários a identificar padrões, descobrir relações e criar previsões usando dados estruturados e não estruturados. Apoiando as instituições a melhorar os resultados, aumentar a eficiência, reduzir custos, melhorar a qualidade das informações. Algumas das principais ferramentas de Business Intelligence incluem dashboards, relatórios, análises, mineração de dados, otimização de processos, análise preditiva, análise de sentimento e outras.⁶⁹

O BI apresenta uma excelente solução para as etapas de visualização, organização, tratamento e análise dos dados, com a vantagem de terem pequena curva de aprendizagem para o usuário final. Além disso, essas ferramentas de BI também oferecem recursos de automação, como a capacidade de agendar relatórios para serem gerados e enviados automaticamente. Isso permite que os interessados na informação acessem os dados mais recentes sempre que precisarem.^{69,92}

A plataforma escolhida para o desenvolvimento desse trabalho foi a Tableau¹, software licenciado e utilizado pelo CONASS, para produção de diversos painéis dinâmicos.

É uma plataforma de análise visual de dados robusta que ajuda os usuários a explorar, visualizar e compartilhar os achados de seus dados. Construído para conectar-se a uma variedade de fontes de dados, permitindo que os usuários criem visualizações interativas para apresentar e compartilhar suas descobertas.⁶⁹

Figura 10 – Etapas de elaboração do protótipo



Fonte: Autores

As etapas anteriores a criação do painel, foram a de coleta de dado. Isso pode ser executado utilizando as ferramentas de ETL (extração, transformação, carregamento) usado

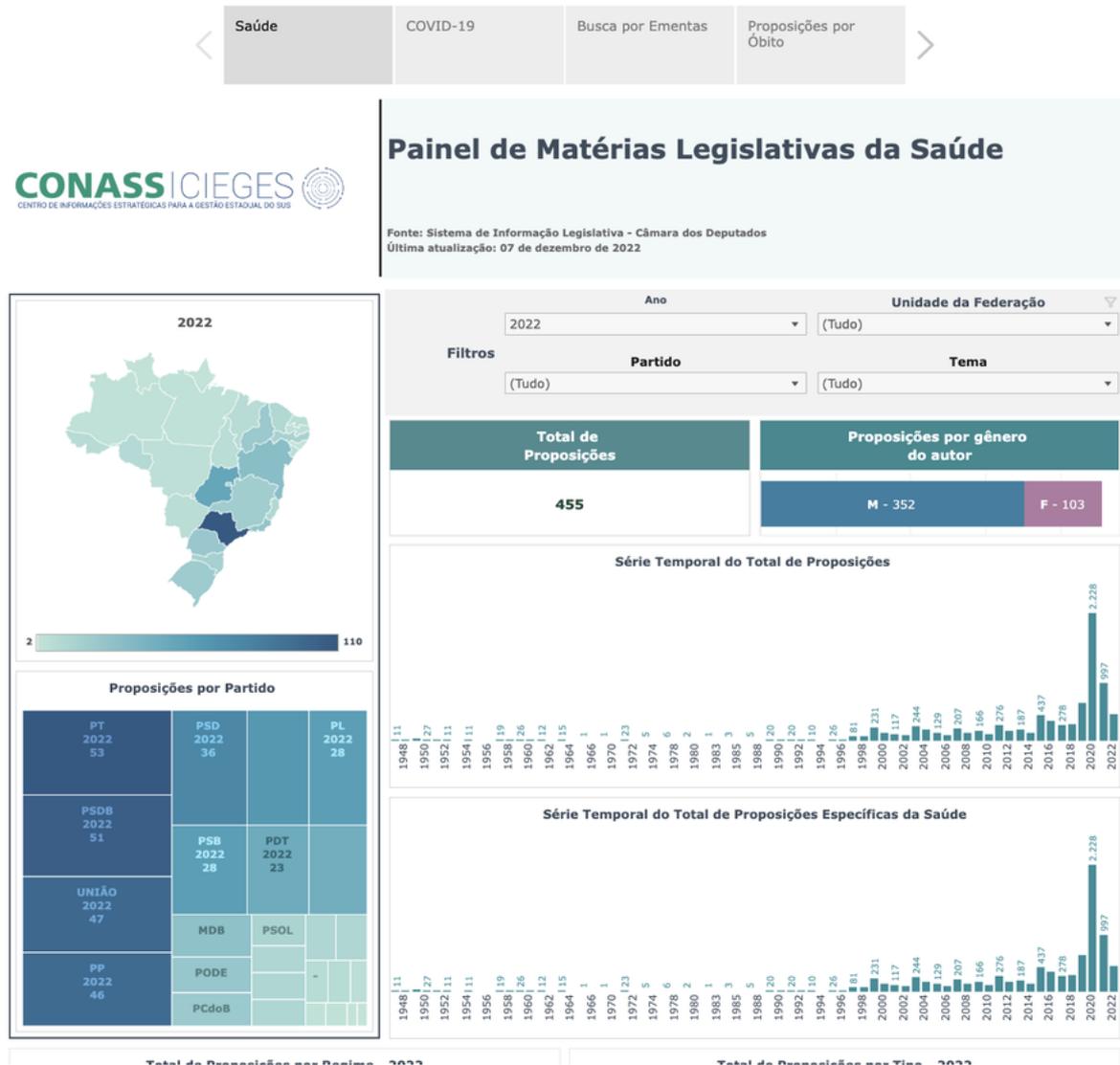
¹ Tableau é um software de análise de dados de visualização de informações. Os usuários podem utilizá-lo para explorar e analisar seus dados, criar dashboards e contar histórias em seus dados. O Tableau também oferece recursos de colaboração, como compartilhamento de relatórios e trabalho colaborativo.⁹²

combinando dados de diversas fontes. Extraíndo, limpando e carregando os dados em um banco de dados relacional da CD.

Em sequência a Modelagem de dados que foi feita acordo com as especificidades em apresentadas na metodologia do trabalho. Por último, usando ferramentas Business Intelligence (BI) a visualização dos dados da CD, a partir de um Painel dinâmico, contendo gráficos e visualizações para análises.

O painel foi elaborado e dividido em três sessões, saúde geral, covid-19, e uma tela de busca por emendas. Os filtros possíveis são: proposições por ano; por unidade da federação do parlamentar; proposições por partido; proposições por área temática; proposições por regime e por tipo. Na visão geral é possível verificar os números totais e algumas análises baseada em séries temporais.

Figura 11 – Painel de Matérias legislativas em Saúde no Congresso Nacional.



Fonte: Autores

A visualização em gráfico demonstra a possibilidade de filtrar por temas agregados à

saúde, segundo tipos de proposições em tramitação na Câmara dos Deputados. Esta informação é essencial para análise, pois, uma proposição pode estar vinculada a mais de tema, podendo ser acrescentado ou alterado de acordo com seguimento da tramitação. Também é um exemplo de informação que não está disponível no portal da CD, podendo ser obtido apenas no repositório de dados abertos. Ao selecionar um tema ou mais, todas as informações do painel são alteradas para demonstrar essas informações por filtro específicos.

Figura 12 – Pesquisa por tema

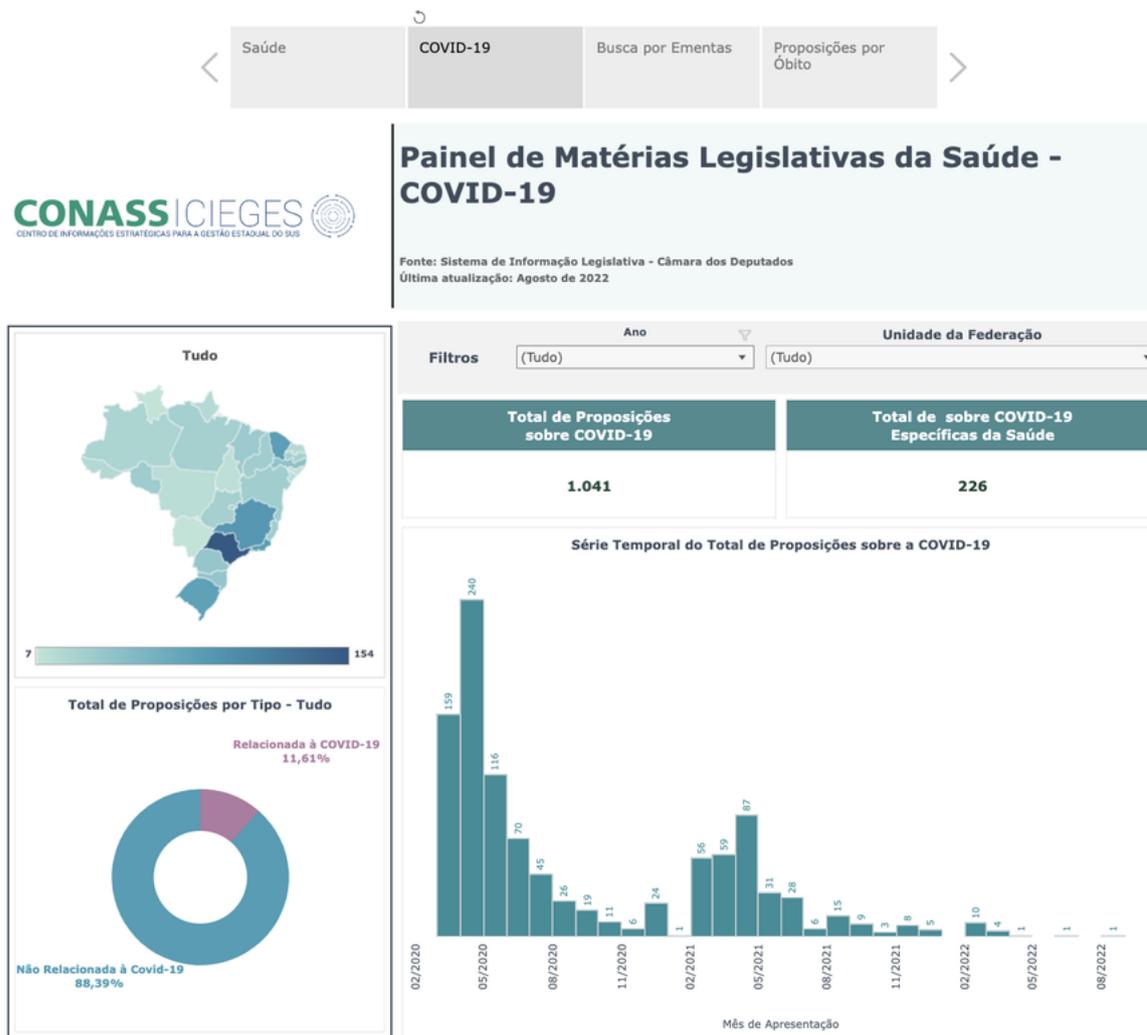
Tema	
▼	(Tudo) ▼
<input checked="" type="checkbox"/>	Administração Pública
<input checked="" type="checkbox"/>	Agricultura, Pecuária, Pesca e Extrativismo
<input checked="" type="checkbox"/>	Arte, Cultura e Religião
<input checked="" type="checkbox"/>	Cidades e Desenvolvimento Urbano
<input checked="" type="checkbox"/>	Ciência, Tecnologia e Inovação
<input checked="" type="checkbox"/>	Ciências Exatas e da Terra
<input checked="" type="checkbox"/>	Ciências Sociais e Humanas
<input checked="" type="checkbox"/>	Comunicações
<input checked="" type="checkbox"/>	Defesa e Segurança
<input checked="" type="checkbox"/>	Direito Civil e Processual Civil
<input checked="" type="checkbox"/>	Direito Constitucional
<input checked="" type="checkbox"/>	Direito e Defesa do Consumidor
<input checked="" type="checkbox"/>	Direito e Justiça
<input checked="" type="checkbox"/>	Direito Penal e Processual Penal
<input checked="" type="checkbox"/>	Direitos Humanos e Minorias
<input checked="" type="checkbox"/>	Economia
<input checked="" type="checkbox"/>	Educação
<input checked="" type="checkbox"/>	Energia, Recursos Hídricos e Minerais
<input checked="" type="checkbox"/>	Esporte e Lazer
<input checked="" type="checkbox"/>	Estrutura Fundiária
<input checked="" type="checkbox"/>	Finanças Públicas e Orçamento
<input checked="" type="checkbox"/>	Homenagens e Datas Comemorativas
<input checked="" type="checkbox"/>	Indústria, Comércio e Serviços
<input checked="" type="checkbox"/>	Meio Ambiente e Desenvolvimento Sustentá...
<input checked="" type="checkbox"/>	Política, Partidos e Eleições
<input checked="" type="checkbox"/>	Previdência e Assistência Social
<input checked="" type="checkbox"/>	Processo Legislativo e Atuação Parlamentar
<input checked="" type="checkbox"/>	Relações Internacionais e Comércio Exterior
<input checked="" type="checkbox"/>	Saúde

Fonte: Autores

A partir de 2020, houve um intenso processo no CONASS de qualificar a informação em saúde para a gestão estadual no Brasil. O que levou ao desenvolvimento de diversos painéis interativos sobre variados temas, mas sobretudo aqueles vinculados à pandemia de COVID-19.

Por isso, esse trabalho se dedicou a uma sessão com plano de visualização do desempenho da câmara dos deputados, na proposição de matérias legislativa sobre o tema.

Figura 13 – Painel de Matérias legislativas em Saúde: Sessão Covid-19



Fonte: Autores

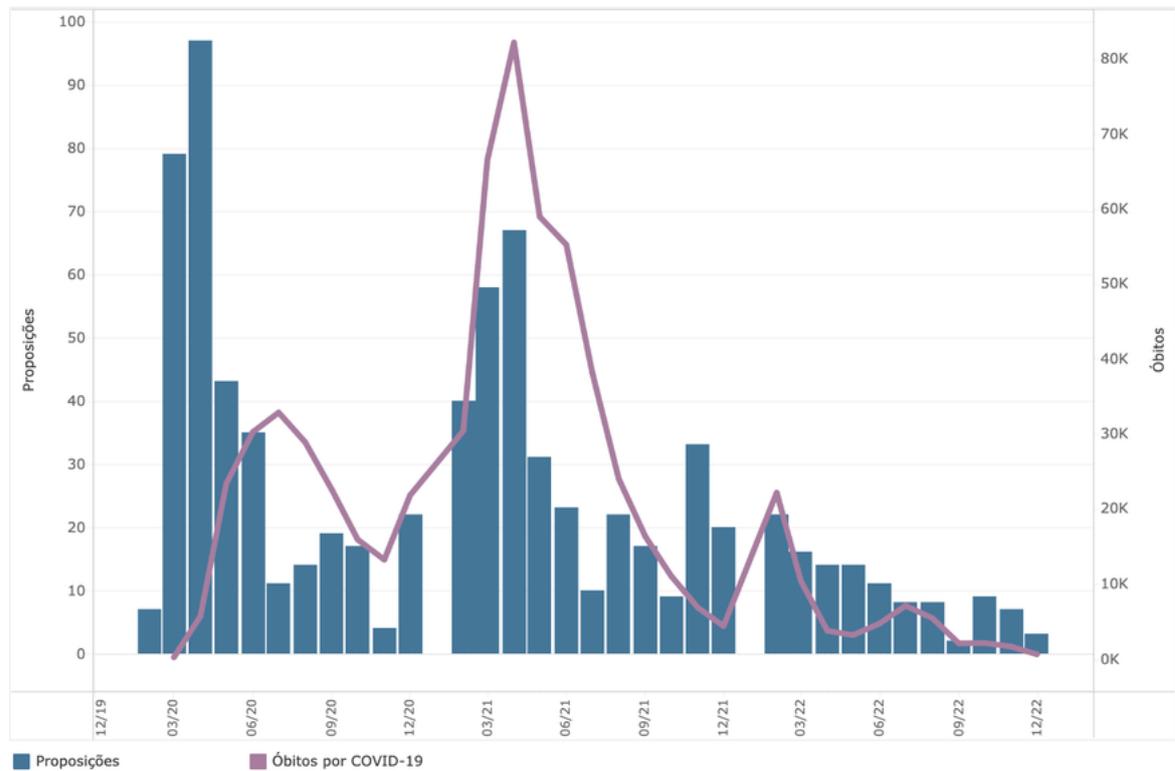
Os mesmos filtros da sessão geral podem ser aplicados, no entanto, todas as informações estão segmentadas ao tema específico “Covid-19”. Essa forma de visualização pode ser aplicada a qualquer temática específica, facilitando a compreensão de técnicos e gestores. Com isso, cria a possibilidade de monitoramento oportuno, de outros temas através da modelagem desta base de dados.

O cruzamento de bases de dados é uma ferramenta extremamente útil para a compreensão de dados complexos, pois permite que as partes interessadas obtenham percepções mais profundas e informações mais detalhadas sobre os temas em questão. Nesse sentido, apresentamos no painel, o cruzamento a partir da base dados, modelada para esse trabalho, com outra base pública, de casos e óbitos por covid-19, no Brasil.² No gráfico a seguir, a curva de óbitos, sobreposta aos

² <https://www.conass.org.br/painelconasscovid19/>

dados das proposições publicadas no período pandêmico.

Figura 14 – Painel de Matérias legislativas em Saúde: Proposições e Óbitos por COVID-19



Fonte: Autores

É importante lembrar que esses dados apresentados são mutáveis e atualizados diariamente, podendo os projetos serem arquivados, retirados pelo autor, apensados em outros projetos ou mesmo arquivados. Por isso a importância de uma ferramenta que acompanhe esse nível de dinâmica da CD.

A terceira sessão, do painel foi dedicada ao desenvolvimento de um ambiente de busca. Com seguintes filtros disponíveis: proposições por ano; por unidade da federação do parlamentar; proposições por partido; proposições por área temática; proposições por regime e por tipo. Além de ser possível pesquisar pela emenda utilizando palavras-chave.

<https://covid.saude.gov.br/>

Figura 15 – Painel de Matérias legislativas em Saúde: Sessão Covid-19: Busca por Ementas

↻

< Saúde COVID-19 **Busca por Ementas** Proposições por Óbito >

Filtros

	Ano	Partido Autor	Regime	Tipo	Relacionada à COVID
	2022	(Tudo)	(Tudo)	(Valores múltiplos)	(Tudo)
	Tema	Proposicao	Ementa		
	Saúde	(Tudo)			

Proposicao Nenhum valor de medida

MPV 1126/2022	Revoga a Lei nº 14.125, de 10 de março de 2021, que dispõe sobre a responsabilidade civil relativa a eventos adversos pós-vacinação contra a co..
PDL 10/2022	Susta as Portarias nºs 5, 6, 7 e 8 da SCTIE/MS, de 25 de janeiro de 2022 e a NOTA TÉCNICA Nº 3/2022-SCTIE/MS, de 25 de janeiro de 2022 e os ..
PDL 11/2022	Susta as Portarias nºs 1, 2, 3 e 4 da SCTIE/MS, de 20 de janeiro de 2022 e a NOTA TÉCNICA Nº 2/2022-SCTIE/MS, de 20 de janeiro de 2022 e os ..
PDL 351/2022	Susta os efeitos da alteração ao art. 21-A do Anexo LXXVII - DO PROGRAMA FARMÁCIA POPULAR DO BRASIL (PFPPB), promovida pela PORTARIA G..
PDL 359/2022	Susta a RESOLUÇÃO Nº 2324/2022 do Conselho Federal de Medicina - CFM, que Aprova o uso do canabidiol para o tratamento de epilepsias da cri..
PDL 360/2022	Susta a Resolução nº 2.324 de 2022 do Conselho Federal de Medicina - CFM, que altera as normas referente ao uso do canabidiol para o tratamen..
PDL 362/2022	Susta a Resolução nº 2.324, de 11 de outubro de 2022, que "Aprova o uso do canabidiol para o tratamento de epilepsias da criança e do adolesce..
PDL 363/2022	Susta os efeitos da Resolução Nº 2.324, de 11 de outubro de 2022, do Conselho Federal de Medicina, que aprova o uso do canabidiol para o trata..
PDL 364/2022	Susta os efeitos da Resolução CFM nº 2.324, de 11 de outubro de 2022, que "aprova o uso do canabidiol para o tratamento de epilepsias da criang..
PDL 366/2022	Susta os efeitos da Resolução 2.324/2022, do Conselho Federal de Medicina.
PDL 384/2022	Susta a Resolução RDC 754, de 22 de novembro de 2022, da Agência Nacional de Vigilância Sanitária (ANVISA), que aprova a obrigatoriedade do ..
PDL 397/2022	Susta os efeitos da Resolução da Diretoria Colegiada - RDC nº 761, de 23 de novembro de 2022, da Agência Nacional de Vigilância Sanitária - Anvi..

Fonte: Autores

CONSIDERAÇÕES FINAIS

Diante da complexidade da produção legislativa em saúde no Congresso Nacional Brasileiro e da importância do monitoramento dessas informações para a tomada de decisões na Gestão Estadual do SUS. Este trabalho desenvolveu uma pesquisa aplicada a construção de um painel com dados de matérias legislativas, com foco na categoria 'saúde'.

O painel facilita o cruzamento e análise de informações relevantes para a Gestão Estadual do SUS, como temas prioritários em discussão no Congresso Nacional e as ações legislativas que impactam diretamente o funcionamento do sistema de saúde brasileiro. A partir de dados abertos disponibilizados pelo governo, fundamentada em princípios e técnicas de Ciência da Informação e Ciência da Computação e Ciência de dados.

Para alcançar esse objetivo, foi necessário integrar, processar e sistematizar os dados de matérias legislativas em saúde, tornando-os acessíveis e manipuláveis pelo protótipo.

O desafio na obtenção dos dados foi considerável, devido aos diversos formatos em volumes significativos, exigindo computadores de alto desempenho e conhecimento multidisciplinares para criação e aplicação do projeto.

É vital ter em mente que as proposições legislativas estão em constante evolução devido à natureza dinâmica do processo legislativo. A digitalização do banco por meio da API, proporciona a atualização diária, resultado alcançado por esse trabalho, torna possível o acompanhamento diário das propostas relacionadas à saúde pública. Isso possibilita a criação de uma estratégia para monitorar e avaliar essas informações, a fim de que atores capazes de influenciar o Processo Legislativo possam usá-las.

Outro ponto a ser destacado, está na etapa de modelagem da base de dados, e seu armazenamento em um repositório institucional, potencializando a possibilidade desses dados serem mantidos e disseminados.

Pretende-se, em projetos futuros, a utilização do modelo como fonte para alimentar algoritmos preditivos, que por exemplo: avaliem o discurso dos deputados sobre uma questão específica, permitindo prever as intenções de voto dos legisladores. Nesse sentido, criar diversas possibilidades de automação com funções analíticas.

Para avaliar os resultados de uso do painel, é preciso aplicar a solução de forma efetiva e mensurar seu uso. É necessário manter o trabalho de aprimoramento, seguindo as novidades e adaptações na coleta de dados abertos e disponibilizando novos tipos de visualizações. Também é importante estudar formas de correlacionar outras bases de dados que sejam significativas.

Espera-se que esta etapa inicial possa ser a base para novas investigações sobre o uso de dados abertos para acompanhar as propostas legislativas em saúde. Os resultados alcançados contribuam para o fortalecimento da democracia e para a melhoria da gestão pública da saúde no Brasil.

REFERÊNCIAS

- [1] Silva PN, Pinheiro MMK. Execução da política de dados abertos no Brasil. *Tendências da Pesquisa Brasileira em Ciência da Informação*. 2019 dez;12(2). Número: 2. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/495>.
- [2] Vargas DA, Frozza AA. Como vai a Câmara? Uma proposta de uso de dados abertos.
- [3] Cavalcante GV, de Sousa FR, Vaz RCR, Araújo CHG. Dados abertos legislativos: O parlamento e o cidadão.
- [4] BRASIL. Lei nº 12.527, de 18 de novembro de 2011. *Diário Oficial da República Federativa do Brasil, Poder Legislativo*. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm.
- [5] Ofoeda J, Boateng R, Effah J. Application Programming Interface (API) Research: A Review of the Past to Inform the Future. *International Journal of Enterprise Information Systems (IJEIS)*. 2019 jul;15(3):76–95. Editora: IGI Global. Disponível em: <https://www.igiglobal.com/article/application-programming-interface-api-research/www.igiglobal.com/article/application-programming-interface-api-research/232166>.
- [6] de Brito RD. Análise e predição nas votações de leis federais na Câmara dos Deputados. *Brasil*; 2022. Aceito: 2022-08-29T16:15:00Z. Disponível em: <http://repository.ufrpe.br/handle/123456789/3162>.
- [7] Campos LDA. Interface Visual Interativa para Dados Abertos sobre Proposições na Câmara de Deputados. Disponível em: <https://lume.ufrgs.br/handle/10183/223627>.
- [8] da Silva SP, dos Santos ER, Rabelo LEM, Luciano MSF. Avaliando a política de Dados abertos no Legislativo brasileiro. *Compólitica*. 2020 maio;10(1):137–160. Disponível em: <http://compolitica.org/revista/index.php/revista/article/view/351>.
- [9] Batista AH. Infraestrutura nacional de dados abertos.
- [10] Cardoso CHR. O uso de dados abertos para maior transparência do setor governamental: um estudo sobre o panorama global da cultura do accountability, controle social e de análise de dados. 2020 fev; Aceito: 2020-03-23T15:38:12Z. Disponível em: <http://bibliotecadigital.fgv.br:80/dspace/handle/10438/28927>.
- [11] Cavalcante D. Organização do conhecimento e representação da informação na curadoria de dados. *VI Congresso Internacional em Tecnologia e Organização da Informação*. 2020;6(1). Número: 1. Disponível em: <https://toi.eca.usp.br/vitoi/article/view/33>.
- [12] Marques AIA. Desenvolvimento de API para aplicação cloud. 2018. Aceito: 2018-06-04T14:09:43Z. Disponível em: <https://iconline.ipleiria.pt/handle/10400.8/3263>.
- [13] Miranda E. Os desafios do processo legislativo brasileiro em tempos de pandemia. *Direito Izabela Hendrix*. 2020;25(25). Número: 25. Disponível em: <https://www.metodista.br/revistas-izabela/index.php/dih/article/view/2318>.
- [14] de Oliveira Santos A, Alves JA. Volume 3 – Competências e Regras; COVID-19: Desempenho da Câmara dos Deputados do Brasil e Ação Intersetorial na Saúde. 2021. Disponível em: <https://www.conass.org.br/biblioteca/volume-3-competencias-e-regras/>.
- [15] SciELO - Brasil. Uma revisão bibliográfica sobre a interdisciplinaridade no ensino das ciências da natureza. Disponível em:

<https://www.scielo.br/j/epec/a/CgpBrMQzDYPqkHZ7yKKdqGk/abstract/?lang=pt>.

[16] Rojas MAR. Introducción a la teoría de conjuntos, los operadores booleanos y la teoría del concepto para profesionales de la información documental. Universidad Nacional Autónoma de México. Instituto de Investigaciones Bibliotecológicas y de la Información. 2017. Aceito: 2018-09-06T13:41:20Z. Disponível em: https://ru.iibi.unam.mx/jspui/handle/IIBI_UNAM/L140.

[17] Botelho LLR, de Almeida Cunha CC, Macedo M. O método da revisão integrativa nos estudos organizacionais. *Gestão e Sociedade*. 2011 dez;5(11). Disponível em: <http://www.gestaoesociedade.org/gestaoesociedade/article/view/1220>.

[18] SciELO - Brasil. A distribuição do conhecimento científico público em informação, comunicação e informática em saúde indexado nas bases de dados MEDLINE e LILACS. Disponível em: <https://www.scielo.br/j/csc/a/QhFhWcPJHZB6zrdLvNmRPQ/abstract/?lang=pt>.

[19] Anderson S, Allen P, Peckham S, Goodwin N. Asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. *Health Research Policy and Systems*. 2008 jul;6.

[20] Leta J. Indicadores de desempenho, ciência brasileira e a cobertura das bases informacionais. *Revista USP*. 2011 maio;(89):62–77. Número: 89. Disponível em: <https://www.revistas.usp.br/revusp/article/view/13869>.

[21] Declaração sobre os princípios fundamentais relativos à contribuição dos meios de comunicação de massa DHnet Direitos Humanos ONU. Disponível em: <http://www.dhnet.org.br/direitos/sip/onu/paz/dec78.htm>.

[22] SciELO - Brasil. Observância da lei de acesso à informação pelas autarquias federais do Brasil. Disponível em: <https://www.scielo.br/j/pci/a/f4X6mHwf83S46YHhWr3SbkM/abstract/?lang=pt>.

[23] Avaliação da Lei de Acesso à Informação Brasileira – Uma abordagem metodológica interdisciplinar. *Revista de Estudos Empíricos em Direito*. 2019 out. Disponível em: <https://revistareed.emnuvens.com.br/reed/article/view/128>.

[24] Michener G, Contreras E, Niskier I. Da opacidade à transparência? Avaliando a Lei de Acesso à Informação no Brasil cinco anos depois. *Revista de Administração Pública*. 2018 ago;52:610-29. Rio de Janeiro: Fundação Getulio Vargas. Disponível em: <http://www.scielo.br/j/rap/a/xJVxcSMSQpQ5qvjBsV7z7ph/abstract/?lang=pt>.

[25] BRASIL. Constituição da República Federativa do Brasil 1988. Senado Federal. 2016; p. 496. Disponível em: https://www2.senado.leg.br/bdsf/bitstream/handle/id/518231/CF88_Livro_EC91_2016.pdf.

[26] SciELO - Brasil. O papel das controladorias locais no cumprimento da Lei de Acesso à Informação pelos municípios brasileiros. Disponível em: <https://www.scielo.br/j/cebape/a/yJ4SxCzCWkvLRjTv7YjCKDL/?lang=pt&format=html>.

[27] Ministério do Planejamento Orçamento e Gestão (MPOG). Cartilha Técnica para Publicação de Dados Abertos no Brasil. BRASIL; 2012. Disponível em: <https://wiki-dados-h.cgu.gov.br/CartilhaTecnicaParaPublica%C3%A7%C3%A3odeDadosAbertosnoBrasil.aspx?Code=1>.

[28] SciELO - Brasil. Open Government Partnership in São Paulo City and the São Paulo Aberta program: challenges in the diffusion and institutionalization of a global policy. Disponível em:

<https://www.scielo.br/j/cebape/a/dGTchmmd9HWWyhxwg8XRZMH/?lang=pt>.

[29] Angélico F. Lei de acesso à informação pública e seus possíveis desdobramentos para a accountability democrática no Brasil. 2012 jun. Disponível em: <http://bibliotecadigital.fgv.br:80/dspace/handle/10438/9905>.

[30] Jardim JM. A implantação da lei de acesso à informação pública e a gestão da informação arquivística governamental Access to Public Information Act: implementation and the management of government archival information. Liinc em Revista. 2013 nov;9(2). Rio de Janeiro: Instituto Brasileiro de Informação em Ciência e Tecnologia. Disponível em: <https://revista.ibict.br/liinc/article/view/3495>.

[31] Braga JV, Gouveia FAR. Iniciativas brasileiras para Dados Governamentais Abertos Conectados: Uma Análise do nível de abertura dos dados nas plataformas governamentais brasileiras. Revista Ibero-Americana de Ciência da Informação. 2022 ago;15(2):495 – 515. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/42801>.

[32] Brelaz GD, Crantschaninov TI, Bellix L. Open Government Partnership na cidade de São Paulo e o programa São Paulo Aberta: desafios na difusão e institucionalização de uma política global. Cadernos EBAPEBR. 2021 mar;19:123 – 137. Rio de Janeiro: Fundação Getúlio Vargas, Escola Brasileira de Administração Pública e de Empresas. Disponível em: <http://www.scielo.br/j/cebape/a/dGTchmmd9HWWyhxwg8XRZMH/?lang=pt>.

[33] Silva LC, Segundo JES. COMPONENTES DE REPRESENTAÇÃO DA INFORMAÇÃO EM AMBIENTES DE INFORMAÇÃO DIGITAL: ESTUDO DO SISTEMA DE ORGANIZAÇÃO DO SOFTWARE TAINACAN. In: ENANCIB 2019; 2019. Disponível em: <https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/1333>.

[34] Arquitetura da informação para ambientes informacionais digitais. Disponível em: <https://brapci.inf.br/index.php/res/v/181958>.

[35] SciELO - Brasil. Os Sistemas de Recomendação, Arquitetura da Informação e a Encontrabilidade da Informação. Disponível em: <https://www.scielo.br/j/tinf/a/YsgLRc86K3WZfcbXPQHq7Vg/abstract/?lang=pt>.

[36] de Araújo de UNESP Camargo LS. Metodologia de desenvolvimento de ambientes informacionais digitais a partir dos princípios da arquitetura da informação. Aleph. 2010 fev;p. 289. São Paulo: Universidade Estadual Paulista (Unesp). Disponível em: <https://repositorio.unesp.br/handle/11449/103357>.

[37] Fidelis JRF, Cândido CM. A administração da informação integrada às estratégias empresariais. Perspectivas em Ciência da Informação. 2006 dec;11:424 – 432. Publisher: Escola de Ciência da Informação da UFMG. Available from: <http://www.scielo.br/j/pci/a/qhxrPnFxf5dYlWmZScyhQQR/?lang=pt>.

[38] da Costa de Oliveira BB. A participação popular no processo legislativo: o exercício da cidadania ativa e o discurso do Estado democrático de direito no Brasil. Universidade de São Paulo; 2010. Available from: <http://www.teses.usp.br/teses/disponiveis/2/2134/tde-16022011-154849/>.

[39] Brandt MB, Vechiato FL, Vidotti SABG. Encontrabilidade da Informação na Câmara dos Deputados. Em Questão. 2018 jan;24(1). Available from: <http://seer.ufrgs.br/index.php/EmQuestao/article/view/71734>.

[40] Moreira D. Com a Palavra os Nobres Deputados: Ênfase Temática dos Discursos dos Parlamentares Brasileiros. Dados. 2020 may;63. Publisher: Instituto de Estudos Sociais e Políticos (IESP) da Universidade do Estado do Rio de Janeiro (UERJ). Available from:

<http://www.scielo.br/j/dados/a/qRgPLv6XqGxBtrTDZ9JH5Xx/?lang=pt>.

- [41] Polato R, Balaniuk R, Ferneda E. Coesão partidária na câmara dos deputados: uma abordagem com ciência de dados. *Revista Eletrônica de Ciência Política*. 2022 mar;12(2). Available from: <https://revistas.ufpr.br/politica/article/view/83334>.
- [42] Guimarães RG. O poder legislativo e a criação da lei: uma análise do processo legislativo brasileiro sob a perspectiva do princípio da tripartição do poder. 2008 may; Accepted: 2016-04-26T20:27:22Z Publisher: Pontifícia Universidade Católica de São Paulo. Available from: <https://repositorio.pucsp.br/xmlui/handle/handle/8135>
- [43] Informação legislativa sobre inovação tecnológica;. Available from: <https://brapci.inf.br/index.php/res/v/200237>.
- [44] da Rocha Miranda RC, de João Braga R. Informação legislativa e correlatas: como conceituar? *Revista de Informação Legislativa*. 2021;58(230):85 – 109. Publisher: Senado Federal. Available from: https://www12.senado.leg.br/ril/edicoes/58/230/ril_v58_n230_p85.
- [45] Feitosa ALG. A integração entre sistemas legislativos, terminologia e web semântica na organização e representação da informação legislativa. 2005 jan; Accepted: 2019-05-13T11:49:37Z. Available from: <https://repositorio.unb.br/handle/10482/34606>.
- [46] Informação legislativa, acesso e cidadania - Dialnet;. Available from: <https://dialnet.unirioja.es/servlet/articulo?codigo=6142011>.
- [47] Torres S, de Almeida MB, da Graça Simões M, Torres S, de Almeida MB, da Graça Simões M. Relações semânticas em sistemas de organização do conhecimento: uma investigação no domínio da informação legislativa; 2015. Publisher: Universidad Complutense de Madrid. Available from: <https://eprints.ucm.es/id/eprint/34618/>.
- [48] Paim JS. O que É o SUS. Rio de Janeiro: Fiocruz; 2015.
- [49] Cordeiro H. O Instituto de Medicina Social e a luta pela reforma sanitária: contribuição à história do SUS. *Physis: Revista de Saúde Coletiva*. 2004 07;14:343 – 362. Available from: <http://www.scielo.br/scieloOrg/php/articleXML.php?lang=pt&pid=S0103-73312004000200009>.
- [50] Delduque MC, Romero LC. Produção normativa em saúde : políticas setoriais e regulação. Senado Federal, Subsecretaria de Edições Técnicas; 2012. Accepted: 2021-01-28T00:00:31Z. Available from: <https://www2.senado.gov.br/bdsf/handle/id/584058>.
- [51] Bem IPD, Delduque MC. Análise da Produção Legislativa em Saúde na 54ª e 55ª Legislaturas do Congresso Nacional do Brasil: o que os parlamentares produzem em saúde? *Physis: Revista de Saúde Coletiva*. 2019 feb;28. Publisher: PHYSIS- Revista de Saúde Coletiva. Available from: <http://www.scielo.br/j/physis/a/xqCYJCGx39G4JfLVD9PyFJc/?lang=pt>.
- [52] Santos ADO. TESES DA SAÚDE NO RELATÓRIO FINAL DA VIII CONFERÊNCIA NACIONAL DE SAÚDE E NA LEGISLAÇÃO FEDERAL NO PERÍODO COMPREEN- DIDO ENTRE 1986 E 2016: uma análise comparada. 2019;.
- [53] Barros IS, Romero LC. A produção legislativa em saúde da Câmara Legislativa do Distrito Federal na Quinta Legislatura (2007-2010). *Cadernos Ibero-Americanos de Direito Sanitário*. 2015 sep;4(3):25 – 38. Number: 3. Available from: <https://www.cadernos.prodisa.fiocruz.br/index.php/cadernos/article/view/162>.
- [54] da Rocha CAJ. PROCESSO DE EXTRAÇÃO DE CONHECIMENTO DE BASES DE DADOS - ELEMENTOS DE APOIO E PRINCIPAIS PROBLEMAS. *Revista Traços*.

2017 nov;3(5). Number: 5. Available from: <http://revistas.unama.br/index.php/revistatracos/article/view/833>.

[55] SciELO - Brasil - A análise de implicação com ferramenta na pesquisa-intervenção A análise de implicação com ferramenta na pesquisa-intervenção;. Available from: <https://www.scielo.br/j/psoc/a/YWZKDkyF5zBjQvhjJZkdK7m/abstract/?lang=pt>.

[//www.scielo.br/j/psoc/a/YWZKDkyF5zBjQvhjJZkdK7m/abstract/?lang=pt](https://www.scielo.br/j/psoc/a/YWZKDkyF5zBjQvhjJZkdK7m/abstract/?lang=pt).

[56] Fernandes AMD, Rozenowicz A, Ferreira JP. Avaliação qualitativa e a construção de indicadores sociais: caminhos de uma pesquisa/intervenção em um projeto educacional & A NAME="n1"& & A& &. Psicologia em Estudo. 2004 aug;9:243 – 253. Publisher: Universidade Estadual de Maringá. Available from: <http://www.scielo.br/j/pe/a/kcRnQ8GP7rPChyJ7gTYJFHd/abstract/?lang=pt>.

[57] Data Preparation Process for Construction Knowledge Generation through Knowledge Discovery in Databases \textbar Journal of Computing in Civil Engineering \textbar Vol 16, No 1;. Available from: [https://ascelibrary.org/doi/abs/10.1061/\(ASCE\)0887-3801\(2002\)16:1\(39\)](https://ascelibrary.org/doi/abs/10.1061/(ASCE)0887-3801(2002)16:1(39))

[58] Fayyad U, Piatetsky-Shapiro G, Smyth P. From Data Mining to Knowledge Discovery in Databases. AI Magazine. 1996 mar;17(3):37 – 37. Number: 3. Available from: <https://ojs.aaai.org/index.php/aimagazine/article/view/1230>.

[59] Matheus CJ, Chan PK, Piatetsky-Shapiro G. Systems for knowledge discovery in databases. IEEE Transactions on Knowledge and Data Engineering. 1993 dec;5(6):903 – 913. Conference Name: IEEE Transactions on Knowledge and Data Engineering.

[60] Fayyad U. Knowledge Discovery in Databases: An Overview. In: Džeroski S, Lavrac N, editors. Relational Data Mining. Berlin, Heidelberg: Springer; 2001. p. 28 – 47. Available from: https://doi.org/10.1007/978-3-662-04599-2_2.

[61] IMPLEMENTASI BUSINESS INTELLIGENCE UNTUK MENENTUKAN TINGKAT KEPOPULERAN JURUSAN PADA UNIVERSITAS \textbar Jurnal Ilmiah Informatika;. Available from: <https://journal.ibrahimy.ac.id/index.php/JIMI/article/view/465>.

[62] revisões por pares API \textbar Anais da conferência ACM 2010 sobre trabalho cooperativo suportado por computador;. Available from: <https://dl.acm.org/doi/abs/10.1145/1718918.1718957>.

[63] The Materials Application Programming Interface (API): Uma API simples, flexível e eficiente para dados de materiais com base nos princípios REpresentational State Transfer (REST) - ScienceDirect;. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S0927025614007113>.

[64] de Sousa Baptista VMP. Um modelo para a detecção das mudanças de posicionamento dos deputados federais; 2015. Accepted: 2016-02-17T11:30:52Z Company: Brasil Distributor: Brasil Institution: Brasil Label: Brasil Publisher: Universidade Federal da Paraíba. Available from: <https://repositorio.ufpb.br>.

[65] Baptista V, Brito F, Brasileiro J, Duarte AN, Bezerra EP, Almeida F, et al. Uma ferramenta para analisar mudanças na coesão entre parlamentares em votações nominais. In: Anais do Brazilian Workshop on Social Network Analysis and Mining (BraSNAM). SBC; 2014.

p. 219 – 224. ISSN: 2595-6094. Available from: <https://sol.sbc.org.br/index.php/brasnam/article/view/6817>.

- [66] Brandt MB. Modelagem da informação legislativa: arquitetura da informação para o processo legislativo brasileiro. 2020;Publisher: Universidade Estadual Paulista (Unesp).
- [67] Passos NRS, Rodrigues AF, Macedo HT, Prado BOP, Silva GJFD, Matos LN. Extração, Transformação e Carga de Dados Abertos como Ferramenta de Apoio ao Voto nas Eleições de 2018. In: Anais do Simpósio Brasileiro de Sistemas de Informação (SBSI). SBC; 2019. p. 175 – 182. ISSN: 0000-0000. Available from: <https://sol.sbc.org.br/index.php/sbsi/article/view/13904>.
- [68] Schwichtenberg S, Gerth C, Engels G. From Open API to Semantic Specifications and Code Adapters. In: 2017 IEEE International Conference on Web Services (ICWS); 2017. p. 484 – 491.
- [69] Koren I, Klamma R. The Exploitation of OpenAPI Documentation for the Generation of Web Frontends. In: Companion Proceedings of the The Web Conference 2018. WWW '18. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee; 2018. p. 781 – 787. Available from: <https://doi.org/10.1145/3184558.3188740>.
- [70] Fundamentos do Esquema JSON \textbar Anais da 25ª Conferência Internacional sobre World Wide Web;. Available from: <https://dl.acm.org/doi/abs/10.1145/2872427.2883029>.
- [71] Berti A, van Zelst SJ, van der Aalst W. Process Mining for Python (PM4Py): Bridging the Gap Between Process- and Data Science. arXiv; 2019. ArXiv:1905.06169 [cs]. Available from: <http://arxiv.org/abs/1905.06169>.
- [72] Documentação da API \textbar SpringerLink;. Available from: https://link.springer.com/chapter/10.1007/978-1-4842-1305-6_4.
- [73] Ooms J. The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects. arXiv; 2014. ArXiv:1403.2805 [cs, stat]. Available from: <http://arxiv.org/abs/1403.2805>.
- [74] Farias DP. Ética hacker na comunicação de dados públicos : o caso do Laboratório Hacker da Câmara dos Deputados. 2015;Accepted: 2015-12-02T19:27:32Z. Available from: <https://bdm.unb.br/handle/10483/11514>.
- [75] Araújo J. Desenvolvimento de uma aplicação para acompanhamento da atividade parlamentar dos deputados federais. 2021 dec;Accepted: 2021-12-19T02:15:28Z Publisher: Pontifícia Universidade Católica de Goiás. Available from: <https://repositorio.pucgoias.edu.br/jspui/handle/123456789/3531>.
- [76] Passos NRS. Transformando dados em conhecimento: LADDRES, uma aplicação prática;.
- [77] Dobesova Z. Programming language Python for data processing. In: 2011 International Conference on Electrical and Control Engineering; 2011. p. 4866 – 4869.
- [78] Oehlman D, Blanc S. Pro Android Web Apps. Develop for Android using HTML5, CSS3, & Java script, Apress. 2011;.
- [79] IDM, A VE, M SA, V ME. Creation of a chemical-technological system digital twin using the Python language. Journal of Applied Informatics. 2021;16(1 (91)). Available from: <https://elibrary.ru/item.asp?id=44792496>.
- [80] Análise de Dados com Pandas \textbar SpringerLink;. Available from: <https://>

link.springer.com/chapter/10.1007/978-1-4842-7189-6_3.

[81] Murari MA. Desenvolvimento de um software para migração de um banco de dados relacional Firebird para o não relacional MongoDB. 2014 dec;Accepted: 2018- 04-10T11:36:42Z Publisher: Universidade Federal de Santa Maria. Available from: <http://repositorio.ufsm.br/handle/1/12904>.

[82] ESTUDO COMPARATIVO DA APLICAÇÃO DOS PROGRAMAS PYTHON E ORANGE PARA A ANÁLISE APROFUNDADA DE BANCOS DE DADOS

\textbar Colloquium Exactarum. ISSN: 2178-8332;. Available from: <https://revistas.unoeste.br/index.php/ce/article/view/3827>.

[83] A survey on business intelligence tools for university dashboard development \textbar IEEE Conference Publication \textbar IEEE Xplore;. Available from: <https://ieeexplore.ieee.org/abstract/document/7460347>.

[84] Implementasi Business Intelligence untuk Menganalisis Data Persalinan Anak di Klinik Ani Padang dengan Menggunakan Aplikasi Tableau Public \textbar Zikri \textbar Jurnal Online Informatika;. Available from: <http://join.if.uinsgd.ac.id/index.php/join/article/view/v2i14>.

ANEXOS

Ano da publicação	Autor	Título
2019	McDonnell, Robert Myles; Duarte, Guilherme Jardim; Freire, Danilo	Congress: An R Package for Analyzing Data from Brazil's Chamber of Deputies and Federal Senate
2020	Moreira, Davi	Com a Palavra os Nobres Deputados: Ênfase Temática dos Discursos dos Parlamentares Brasileiros
2020	Brito, Ana Caroline Medeiros; Silva, Filipi Nascimento; Amancio, Diego Raphael	A complex network approach to political analysis: Application to the Brazilian Chamber of Deputies
	Vargas, Dimi Antoni; Frozza, Angelo Augusto	COMO VAI A CÂMARA? Uma proposta de uso de dados abertos
2017	Victorino, Marcio de Carvalho; Shiessl, Marcelo; Oliveira, Edgard Costa; Ishikawa, Edson; Holanda, Maristela Terto de; Hokama, Marçal de Lima	Uma proposta de ecossistema de big data para a análise de dados abertos governamentais conectados
2018	Santos, Marlene Cristina dos; Marin, Heimar de Fátima	Análise do uso de um sistema informatizado por gestores hospitalares
2018	Brandt, Mariana Baptista; Vechiato, Fernando Luiz; Vidotti, Silvana Aparecida Borsetti Gregorio	Encontrabilidade da Informação na Câmara dos Deputados
2019	Oliveira, Danilo Amaral de	Compreendendo e prevendo o processo legislativo via ciência de dados
	Campos, Lisiane De Aguiar	Interface Visual Interativa para Dados Abertos sobre Proposições na Câmara de Deputados
2018	Formigoni, Vanessa Lourenço Mattede; Honorato, Icaro Henrique	MINERAÇÃO DE DADOS NA BASE DE DADOS ABERTA DA CÂMARA LEGISLATIVA FEDERAL BRASILEIRA: ÊNFASE NA ANÁLISE DOS DADOS DA LEGISLATURA 54 (2011-2013)
2021	Cannavale, Victor	ANÁLISE DE RISCO PARA PROJETOS LEGISLATIVOS QUE IMPACTAM O MODELO ZONA FRANCA DE MANAUS: UMA PROPOSTA DE ANÁLISE DE DADOS ABERTOS
2022	Polato, Ricardo; Balaniuk, Remis; Ferneda, Edilson	Coesão partidária na câmara dos deputados: uma abordagem com ciência de dados
	Passos, Néelson Rangel Santos	Transformando dados em conhecimento: LADDRES, uma aplicação prática
2021	Moeira, Felipe Lélis	Programa de Pós-Graduação em Direito
2020	Brandt, Mariana Baptista	Modelagem da informação legislativa: arquitetura da informação para o processo legislativo brasileiro
	Cavalcante, Gustavo Vasconcellos; de Sousa, Fabricio Rocha; Vaz, Rodolfo Cezar Ranulfo; Araújo, Carlos Henrique Gouveia	DADOS ABERTOS LEGISLATIVOS: O PARLAMENTO E O CIDADÃO
2021	Bartholomay, Eduardo Luís	A VOZ DO POVO É A VOZ DE DEUS? A INFLUÊNCIA DAS MANIFESTAÇÕES NA COMUNICAÇÃO POLÍTICA

	Cardoso, Paulo Henrique	Ciência de dados aplicada a dados governamentais abertos sob a ótica da Ciência da Informação
2016	Cruvinel, Gustavo Warzocha Fernandes	CÂMARA DOS DEPUTADOS CENTRO DE FORMAÇÃO, TREINAMENTO E APERFEIÇOAMENTO PROGRAMA DE PÓS-GRADUAÇÃO MESTRADO PROFISSIONAL EM PODER LEGISLATIVO

Códigos Python

```
import requests
from flask import Flask, jsonify

app = Flask(__name__)

@app.route('/api/camara-deputados')
def buscar_dados_camara():
    # código para buscar os dados da Câmara dos Deputados
    return jsonify(dados)
```

```
@app.route('/api/camara-partidos')
def buscar_partidos():
    partidos = dados['siglaPartido'].unique()
    return jsonify(list(partidos))
```

```
import requests
import pandas as pd
from flask import Flask, jsonify

app = Flask(__name__)

response = requests.get('https://dadosabertos.camara.leg.br/api/v2/deputados')
dados_json = response.json()
dados = pd.DataFrame.from_records(dados_json['dados'])

@app.route('/api/camara-deputados/<int:id>')
def buscar_deputado(id):
    deputado = dados.loc[dados['id'] == id].to_dict('records')
    if len(deputado) > 0:
        return jsonify(deputado[0])
    else:
        return jsonify({'erro': 'Deputado não encontrado.'})

@app.route('/api/camara-partidos')
def buscar_partidos():
    partidos = dados['siglaPartido'].unique()
    return jsonify(list(partidos))

if __name__ == '__main__':
    app.run()
```

Códigos Python

```
import requests
import pandas as pd
from pymongo import MongoClient

client = MongoClient('mongodb://localhost:27017/')
db = client['camara-deputados']
deputados = db['deputados']
```

```
response = requests.get('https://dadosabertos.camara.leg.br/api/v2/deputados')
dados_json = response.json()
dados = pd.DataFrame.from_records(dados_json['dados'])
```

```
for i, row in dados.iterrows():
    deputados.insert_one(row.to_dict())
```