



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

EMILLY DE ALBUQUERQUE OLIVEIRA

**SENADO ABERTO:
DADOS ABERTOS DAS VOTAÇÕES NOMINAIS**

CAMPINA GRANDE - PB

2024

EMILLY DE ALBUQUERQUE OLIVEIRA

**SENADO ABERTO:
DADOS ABERTOS DAS VOTAÇÕES NOMINAIS**

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador : Professor Dr. Fábio Jorge Almeida Morais

CAMPINA GRANDE - PB

2024

EMILLY DE ALBUQUERQUE OLIVEIRA

SENADO ABERTO:

DADOS ABERTOS DAS VOTAÇÕES NOMINAIS

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

BANCA EXAMINADORA:

Professor Dr. Fábio Jorge Almeida Morais
Orientador – UASC/CEEI/UFCG

Professor Dr. Maxwell Guimarães de Oliveira
Examinador – UASC/CEEI/UFCG

Professor Dr. Francisco Vilar Brasileiro
Professor da Disciplina TCC – UASC/CEEI/UFCG

Trabalho aprovado em: 15 de Maio de 2024.

CAMPINA GRANDE - PB

RESUMO

Diante da crescente demanda por transparência política e participação pública, surge a necessidade de informações acessíveis sobre as votações no Senado brasileiro. No entanto, os registros dispersos em PDF ou páginas da web dificultam o acompanhamento. Este projeto visa preencher essa lacuna ao criar uma aplicação com uma API que coleta, processa e disponibiliza os dados das votações dos representantes no Senado. A metodologia abrange coleta de dados confiáveis, desenvolvimento de algoritmos para estruturação, criação da API, implementação de um *watcher* para novas votações e um front-end interativo. Os resultados esperados incluem uma aplicação amigável que permite aos cidadãos acompanhar as atividades de seus representantes, promovendo a transparência política ao facilitar o acesso às informações sobre as votações no Senado.

OPEN SENATE: OPEN DATA OF BRAZILIAN SENATORS VOTES

ABSTRACT

Given the growing demand for political transparency and public participation, there is a need for accessible information about votes in the Brazilian Senate. However, scattered records in PDFs or web pages make it difficult to follow up. This project aims to fill this gap by creating an application with an API that collects, processes and makes available voting data from representatives in the Senate. The methodology covers reliable data collection, development of algorithms for structuring, creation of the API, implementation of a watcher for new votes and an interactive front-end. The expected results include a user-friendly application that allows citizens to monitor the activities of their representatives, promoting political transparency by facilitating access to information about votes in the Senate.

Senado Aberto: Dados Abertos das Votações Nominais no Senado

Emilly de Albuquerque Oliveira*
emilly.oliveira@ccc.ufcg.edu.br
Universidade Federal de Campina Grande
Campina Grande, Paraíba

Fábio Jorge Almeida Morais
fabio@computacao.ufcg.edu.br
Universidade Federal de Campina Grande
Campina Grande, Paraíba

RESUMO

Diante da crescente demanda por transparência política e participação pública, surge a necessidade de informações acessíveis sobre as votações no Senado brasileiro. No entanto, os registros dispersos em PDF ou páginas da web dificultam o acompanhamento. O projeto almeja reduzir essa lacuna ao criar uma aplicação com uma API Rest que coleta, processa e disponibiliza os dados das votações dos representantes no Senado. A metodologia abrange coleta de dados confiáveis, desenvolvimento de algoritmos para estruturação, criação da API, implementação de um *scheduler* para novas votações e um front-end interativo. Os resultados alcançados incluem uma aplicação simples que permite aos cidadãos acompanhar as atividades de seus representantes, promovendo a transparência política ao facilitar o acesso às informações sobre as votações no Senado.

PALAVRAS-CHAVE

Senado, Política, Transparência, Libertação contínua de dados.

REPOSITÓRIOS

- <https://github.com/emys-alb/votos-senado>
- <https://github.com/emys-alb/votos-senado-api>
- <https://github.com/emys-alb/votos-senado-frontend>

1 INTRODUÇÃO

Os países ocidentais têm vivenciado uma crescente demanda de *accountability*¹ como elemento fundamental à democratização do Estado [3], e o Brasil não é exceção a essa tendência. Essa demanda por transparência política tem ressaltado a necessidade de um acesso às informações que outrora foram consideradas irrelevantes ou complexas demais como, por exemplo, as votações no Senado brasileiro.

Dentro dessa vertente em computação existem os dados abertos. Segundo a definição da Open Knowledge Foundation – OKF, os “dados são abertos quando qualquer pessoa pode livremente usá-los, reutilizá-los e redistribuí-los, estando sujeito a, no máximo, a exigência de creditar a sua autoria e compartilhar pela mesma licença” (Tradução Livre)(OKF; BRASIL).

O problema que nosso sistema se propõe a compreender e resolver é a dificuldade enfrentada pelo público em geral, jornalistas e pesquisadores em obter dados abertos e estruturados sobre as votações dos representantes no Senado. Sabe-se que essas informações

atualmente estão dispersas em páginas web com diferentes formatos como PDF, XML e HTML. Essa falta de acesso claro e organizado aos dados torna desafiador o acompanhamento das atividades políticas e dificulta a avaliação do desempenho dos representantes eleitos.

Já existem algumas soluções para esse problema, incluindo algumas desenvolvidas pelo governo federal, como o projeto Dados Abertos[2], um projeto em construção que disponibiliza diversas informações do Congresso Nacional. Em relação às votações, em específico, os dados são disponibilizados em XML e contemplam dados entre os anos 1991 e 2020. Além disso, um projeto similar foi desenvolvido na UFCG com o uso da linguagem R: o Rcongresso[1]. Este foi criado pelo laboratório Analytics, contudo deixou de ser mantido e pode conter inconsistências.

Dito isso, este trabalho propõe o desenvolvimento de uma aplicação com uma API que coleta, processa e disponibiliza os dados qualitativos e estruturados das votações dos representantes no Senado brasileiro. A metodologia adotada para atingir esse objetivo abrange uma sequência de etapas, que vão desde a coleta de dados confiáveis e a criação de algoritmos para estruturá-los até a implementação de um *scheduler*² que permitirá o acompanhamento de novas votações, além da criação de um *front-end* interativo para apresentar essas informações de forma acessível aos cidadãos.

Os resultados obtidos deste projeto são significativos. Ao permitir que os cidadãos acompanhem as atividades de seus representantes de forma mais eficaz, esta aplicação tem o potencial de fortalecer a democracia brasileira e promover um maior engajamento cívico. Adicionalmente, conseguimos facilitar o desenvolvimento de novas ferramentas com propósitos similares ao disponibilizar dados estruturados e atualizados por nossa API Senado Aberto.

2 ARQUITETURA E PROJETO DA SOLUÇÃO

2.1 Visão Geral do Projeto

A transparência e a participação cívica são pilares fundamentais da democracia pois promovem a responsabilidade dos governantes perante os governados e fortalecem a legitimidade das instituições democráticas. A transparência garante que as ações do governo sejam abertas e acessíveis ao público, permitindo que os cidadãos entendam as decisões tomadas em seu nome e avaliem se estão alinhadas com seus interesses e valores. Por sua vez, a participação cívica engaja os cidadãos no processo político, capacitando-os a expressar suas opiniões, demandas e preocupações, e a contribuir ativamente para a formulação e implementação de políticas públicas. Quando a transparência e a participação cívica são promovidas e respeitadas, cria-se um ambiente democrático mais saudável, no qual os cidadãos se sentem parte integrante do sistema político e

*Os autores retêm os direitos, ao abrigo de uma licença Creative Commons Atribuição CC BY, sobre todo o conteúdo deste artigo (incluindo todos os elementos que possam estar contidos, tais como figuras, desenhos, tabelas), bem como sobre todos os materiais produzidos pelos autores que estejam relacionados ao trabalho relatado e que estejam referenciados no artigo (tais como códigos-fonte e bases de dados). Essa licença permite que outros distribuam, adaptem e evoluam seu trabalho, mesmo comercialmente, desde que os autores sejam creditados pela criação original.

¹Termo utilizado para indicar que empresas privadas e órgãos públicos têm responsabilidade perante a sociedade.

²Agendador de tarefas que executa uma tarefa após um tempo pré-determinado.

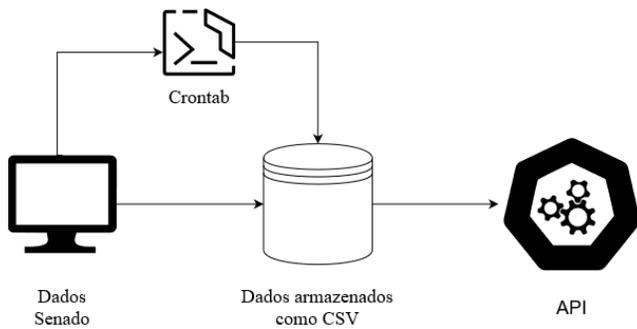


Figura 1: Arquitetura do projeto

têm confiança de que seus direitos serão protegidos e seus interesses serão representados de forma adequada.

O projeto Senado Aberto visa facilitar o acompanhamento das votações realizadas pelos representantes do Senado, oferecendo uma plataforma mais acessível para os cidadãos fiscalizarem as matérias propostas e discutidas pelos seus representantes.

Para abordar essa questão, desenvolvemos uma ferramenta de extração que coleta todas as votações nominais usando o formato de páginas web disponibilizado pelo Senado. Ao serem processados pelo Senado Aberto, os dados serão disponibilizados no formato JSON, comumente usado para análises, por meio de uma API criada neste projeto, com o objetivo de tornar essas informações mais acessíveis para a população interessada em informações detalhadas e de fácil acesso, capazes de serem facilmente processadas para análises e criação de novas ferramentas³.

Para garantir a atualização constante dos dados, implementamos um programa para executar o *script* de raspagem de dados periodicamente, utilizando o Crontab com frequência semanal. A Figura 1 ilustra a arquitetura atual do projeto, destacando a integração entre os diferentes componentes.

Para garantir acessibilidade a usuários com diversos níveis de habilidade técnica, foi desenvolvido um portal que apresenta uma interface amigável com filtros assertivos, evitando distrações de interface e facilitando a tarefa de buscar informações sobre as votações mais recentes ocorridas no Senado Federal. O portal organiza as informações, exibindo os eventos mais recentes primeiro e retrocedendo até o ano de 2019.

2.2 Senado Aberto: A API

A API do Senado Aberto foi desenvolvida para suprir a necessidade de acesso a informações que até o presente momento não foram disponibilizadas pelo governo federal. Ela opera tanto utilizando os dados fornecidos em formato XML (pelo site do governo federal), quanto capturando as informações de votações nominais publicadas no site do Senado, como datas, descrição e votantes.

A partir desta API, ilustrada na Figura 2, os usuários podem acessar uma variedade de informações, incluindo o nome de todos os senadores, votações realizadas por data, todos os votos de um senador específico, todos os votos em uma votação específica e todos os votos de um senador em uma votação particular.

³Os cidadãos podem acessar os dados pelo site <https://app-70aqyqce6q-uw.a.run.app/docs#>

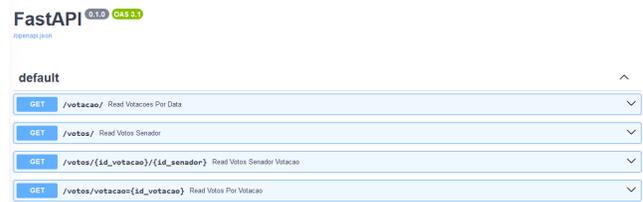


Figura 2: API Senado Aberto

Votações Nominais

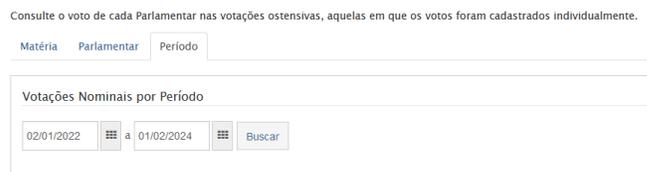


Figura 3: Sistema de Busca no Senado Federal

Para a construção dessa ferramenta foi utilizada FastAPI e Swagger devido a familiaridade com as mesmas, e principalmente pela facilidade que a documentação gerada automaticamente pelo Swagger é capaz de fornecer para o uso da API.

O objetivo principal da API é facilitar a busca e a análise de informações, uma vez que no site do Senado Federal é necessário realizar buscas específicas por matéria, parlamentar ou período de tempo (Figura 3). Através dela, outros sistemas conseguem utilizar esses dados de forma estruturada sem a necessidade de realizar raspagens ou reunir informações de diversas fontes. Essa API está disponível para acesso e hospedada no Google Cloud⁴.

2.3 Senado Aberto: Portal Online

Inicialmente, planejamos manter o front-end disponibilizado pelo Swagger. Contudo, percebemos que a interface não era agradável para o público em geral, não atendendo ao propósito de ser uma ferramenta acessível para todos os cidadãos. Portanto, ao longo do desenvolvimento do projeto, concentramos nossos esforços na construção de uma interface mais amigável, que inclui ferramentas de pesquisa mais abrangentes e destaque para as votações mais recentes.

Para esse fim, optamos por utilizar o Next.js⁵ para o desenvolvimento do front-end e publicamos o projeto na plataforma Vercel⁶. Essas ferramentas foram utilizadas por serem complementares uma a outra no desenvolvimento, terem alta curva de aprendizagem e serem capazes de gerar modelos de páginas estáticas que otimizam a resposta da página criada ao cliente.

Atualmente, temos uma interface livre de distrações é composta por todas as votações coletadas, organizadas do mais recente ao mais antigo, conforme ilustrado na Figura 4. Os usuários podem

⁴<https://cloud.google.com/>

⁵<https://nextjs.org/>

⁶<https://vercel.com/>

Últimas votações

As votações nominais mais recentes do Senado Federal.

13/12/2023	Aprovado	13/12/2023	Aprovado	13/12/2023	Aprovado
OFS 00018 de 2023		OFS 00020 de 2023		MSF 00085 de 2023	
Votação nominal do Ofício nº 18, de 2023 - Eduardo Nilo de Almeida (C/PMB)		Votação nominal do Ofício nº 20, de 2023 - Cirina Menezes Burelato (C/PMB)		Votação nominal da Mensagem nº 85, de 2023 - Caio Mário Trivelato Seabra Filho (PMDB)	
OFS 00019 de 2023		RQS 01103 de 2023		MSF 00090 de 2023	
Votação nominal do Ofício nº 19, de 2023 - Alexandre Augusto Sobrinho de Andrade (PT)		Solicita o encaminhamento da discussão do PL nº 3.629/2023.		Votação nominal da Mensagem nº 90, de 2023 - Daniel Walter Macedo Bernardes (C/PMB)	

Figura 4: Sistema com votações dos parlamentares

```
data_votacao,descricao,link,materia,obs,parlamentar,status_votacao,voto
```

Figura 5: Estrutura dos dados em CSV

acessar as votações através do site <https://votos-senado.vercel.app/votacao/1>.

2.4 Dados publicados

Para a raspagem dos dados, foi utilizado a ferramenta Scrapy. Esta foi escolhida por ter sido construída em cima da linguagem Python, amplamente utilizada para ciência de dados, oferecendo uma ferramenta sólida e ampla comunidade técnica. Os dados capturados podem ser encontrados no repositório de raspagem do GitHub, organizados como arquivos *comma-separated values* (CSVs) na pasta *out*, estruturados como mostra a Figura 5. Para aqueles usuários que desejam utilizar o script para extrair apenas uma parte dos dados, é possível configurar os parâmetros de data de início e fim através de variáveis de ambiente. Uma vez configurado, basta executar o script para que todo o fluxo de execução, desde a coleta até a extração e geração do arquivo, seja realizado automaticamente.

O arquivo final será nomeado de acordo com a variável definida como "FILENAME" e conterá informações como a data da votação, sua descrição e um link para acesso direto. Ademais, o arquivo incluirá o nome da matéria em votação, o nome do parlamentar que votou, seu voto e, caso não tenha votado, uma observação correspondente.

3 AVALIAÇÃO DO PROJETO

A fim de testar a qualidade do sistema quanto à sua usabilidade e à satisfação dos usuários, foi feito um levantamento utilizando o Computer System Usability Questionnaire (CSUQ)[4], no qual os usuários respondem a 19 afirmações de boa usabilidade do sistema com um valor de 1 a 7. Nesse caso, 1 significa completa discordância e 7 significa completa concordância com a afirmação. A moda dos valores das respostas de cada pergunta indica o quão boa é a usabilidade do sistema quanto àquele aspecto, sendo considerado um resultado melhor quanto mais o valor se aproxima de 7.

O levantamento contou com respostas de 12 usuários, de diferentes formações, e teve como resultado um valor próximo a 7 em diversos aspectos. Os pontos que exigem melhorias foram indicados nas Figuras 6 e 7. Esses quesitos falam sobre eficácia da informação e se o sistema possui todas as funções e capacidades esperadas.

É fácil encontrar a informação que eu precisava.

12 respostas

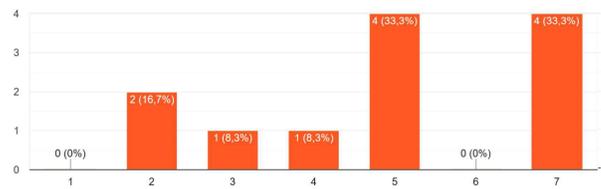


Figura 6: Respostas no tópico: A informação é eficaz para me ajudar a completar as tarefas

Este sistema tem todas as funções e capacidades que eu espero que ele tenha.

12 respostas

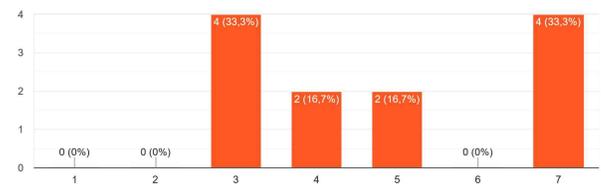


Figura 7: Respostas no tópico: Este sistema tem todas as funções e capacidades que eu espero que ele tenha

Como é possível observar na Figura 6, no quesito de facilidade para encontrar as informações no portal, os votos ficaram dispersos entre 2 e 7, a maioria ficou dividida entre as notas 5 e 7, indicando que há espaço para melhorias. Na seção de pontos positivos e negativos do sistema, foi apontado a necessidade de incluir um sistema de busca, um usuário também apontou que "seria importante ter explicações sobre as nomenclaturas e outras informações dispostas nas telas".

Além disso, podemos observar na Figura 7 que com a maioria dos votos concentrados entre 3 e 5, existe a necessidade do usuário final de mais funcionalidades. Dentre os *feedbacks* coletados, recebemos um comentário pertinente sobre o sistema de filtros: "Falta uma forma de pesquisa/filtragem. Dependendo da informação que se procura, demora muito para encontrar". Buscando melhorar nosso projeto, mapeamos as versões futuras para contemplar filtros para senadores, datas de votações e situações da votação. Essa melhoria estará também refletida nas versões futuras da API.

Para os quesitos de facilidade de uso e auxílio no trabalho as respostas se mantiveram acima de 5, tendo como sua maioria nota 7. Entretanto, quando ocorreram erros no Portal, os votos ficaram divididos entre 7 e valores próximos de 4, como mostra a Figura 8. Esse resultado indica que falta clareza nas mensagens de erro e dificuldade para se recuperar dos mesmos.

Para a API, o principal quesito ressaltado foi a importância de uma documentação complementar e uma interface mais agradável. A Figura 9 mostra que essa foi uma das maiores dificuldades encontradas, tendo os votos divididos de 4 a 7 mas atingindo o mínimo máximo de 1 (Discordo totalmente). Através dos comentários dos usuários, notamos que a documentação automática gerada pelo FastAPI não informa corretamente informações como o que é obrigatório e esperado de cada função e oferece uma visualização dos

O sistema dá mensagens de erro que me dizem claramente como corrigir problemas.

12 respostas

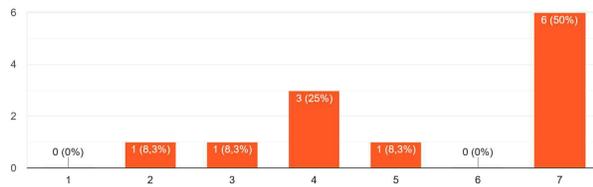


Figura 8: Respostas no tópico: O sistema dá mensagens de erro que me dizem claramente como corrigir problema

As informações fornecidas para o sistema são fáceis de entender.

11 respostas

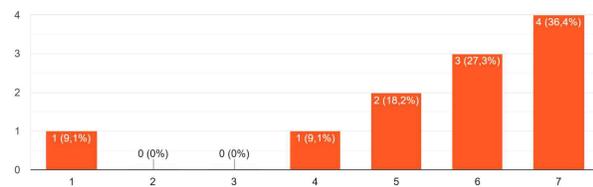


Figura 9: Respostas no tópico: As informações fornecidas para o sistema são fáceis de entender.

dados limitada. O portal busca contornar a necessidade do uso da API através da interface do Swagger. Contudo, pensando em melhorar essa documentação é possível customizá-la com a utilização de PyDoc⁷, permitindo a agregação de mais informações para o usuário final.

4 EXPERIÊNCIAS E LIÇÕES APRENDIDAS

Nesta seção, é discutido sobre como foi realizado o processo de desenvolvimento do projeto, bem como os principais desafios encontrados ao longo do processo.

4.1 Processo de Desenvolvimento

O projeto foi desenvolvido utilizando um processo incremental, no qual novas funcionalidades foram sendo implementadas e integradas à base de código principal ao longo do tempo. Por ter sido desenvolvido por uma única pessoa, não houve revisão de código, e levou aproximadamente 5 meses para alcançar o estado atual de desenvolvimento, no qual os cidadãos podem acessar a plataforma e os dados estão sempre atualizados.

Cada tarefa foi cuidadosamente avaliada quanto ao seu risco, estimada em termos de tempo necessário para conclusão e atribuída a mim. O projeto foi dividido em três partes distintas: Estudo de viabilidade, Desenvolvimento e Melhorias do Sistema. No estudo de viabilidade, as principais tarefas incluíram a preparação do arcabouço do projeto, definindo as ferramentas a serem utilizadas e delineando o fluxo do sistema, além de estudar o sistema de votações do Senado para identificar seus benefícios e limitações.

⁷<https://docs.python.org/pt-br/3/library/pydoc.html>

Quanto às melhorias realizadas, incluíram-se a adição de variáveis de ambiente e a implementação do Crontab, ambas já concluídas. Todas as demais tarefas estavam relacionadas ao desenvolvimento do sistema, incluindo a raspagem de dados, a estruturação dos dados em uma API, a implementação do *watcher*, o cálculo de desempenho da API e o deploy do sistema.

4.2 Principais desafios

As principais dificuldades enfrentadas durante a realização deste projeto residiram na definição do escopo, uma vez que a melhoria da usabilidade de um sistema não é uma métrica facilmente mensurável. Ainda que existam ferramentas que ajudam a monitorar a usabilidade de sistemas web, tais como Hotjar⁸ e Mixpanel⁹ mas que ambas necessitam de licença paga. Além disso, o desafio de criar uma pipeline complexa, envolvendo a extração, manipulação e visualização dos dados dentro do prazo estabelecido para este Trabalho de Conclusão, representou um aspecto complicado de gerenciar.

Um requisito especialmente desafiador foi a necessidade de manter o projeto operacional sem incorrer em custos de infraestrutura. Dada a nossa intenção de garantir uma alta disponibilidade para os usuários, juntamente com a exigência de armazenar grandes volumes de dados e realizar operações de extração de forma constante, nos deparamos com a dificuldade de encontrar uma ferramenta que atendesse a esses requisitos sem gerar custos. Apesar de várias ferramentas terem sido consideradas, até o momento presente não encontramos uma solução que nos permitisse manter a estrutura operacional sem incorrer em custos adicionais.

5 CONCLUSÕES E TRABALHOS FUTUROS

No projeto Senado Aberto, conseguimos criar um sistema complexo e integrado contendo extração, divulgação e visualização de dados do Senado. A partir dele, começamos o processo de liberação de dados para um público maior e disponibilizamos os mesmos para que sejam analisados por terceiros sem maiores dificuldades. Entretanto, ainda temos espaço para melhorias.

Como já mencionado o Senado, no processo de aumentar a transparência, também possui uma API de dados¹⁰ dos próprios parlamentares, ementas, vetos e outras informações. Acredito que seria de grande valia incorporar essas informações no portal online e cumprir nosso propósito de auxiliar na busca por informação. Além da criação de uma documentação complementar e implementação de um sistema de interface mais agradável para a população geral.

Além disso, há oportunidades para aprimoramento no Portal tais como a inclusão de recursos de filtragem por senadores e por ano, bem como a possibilidade de visualizar detalhes das votações em cada *card* como porcentagem de participação, abstenções e explicação das nomenclaturas usadas. Essas melhorias visam enriquecer a experiência do usuário e fornecer um acesso mais eficiente e personalizado às informações sobre as votações no Senado Federal. Além disso, o projeto facilita a implementação de novos sistemas que possam utilizar ele de apoio, uma vez que a preparação dos dados já foi feita.

⁸<https://www.hotjar.com/pt-BR/>

⁹<https://mixpanel.com/>

¹⁰<https://legis.senado.leg.br/dadosabertos/docs/ui/index.html>

AGRADECIMENTOS

Quero expressar minha sincera gratidão a todos que estiveram ao meu lado durante esta jornada. Meus colegas de curso, amigos e, especialmente, minha família, vocês foram minha rocha e minha inspiração. Sem o apoio de vocês, este marco não seria possível.

Também quero agradecer às comunidades do curso, como Guardians¹¹, Elas Computação¹² e OpenDevUFCG¹³, por me acolherem calorosamente e me proporcionarem momentos de muito crescimento e profissional.

A todos que desempenharam um papel importante nesta jornada, meu mais profundo obrigado. Estou verdadeiramente grata por fazer parte de uma comunidade tão solidária e inspiradora.

REFERÊNCIAS

- [1] 2019. Rcongresso. <https://github.com/analytics-ufcg/rcongresso> [Online; acessado em 24 de Setembro de 2023].
- [2] 2021. Senado Federal. <https://www12.senado.leg.br/dados-abertos/conjuntos?portal=Legislativo&grupo=plenario> [Online; acessado em 14 de Janeiro de 2023].
- [3] Fernando Filgueiras. 2011. Além da transparência : accountability e política da publicidade. *Lua Nova Revista de Cultura e Política* 84 (2011).
- [4] James R. Lewis. 1995. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction* (1995). <https://doi.org/10.1080/10447319509526110>

¹¹<https://github.com/Guardians-DSC>

¹²<https://elas.computacao.ufcg.edu.br/>

¹³<https://opendevufcg.org/>