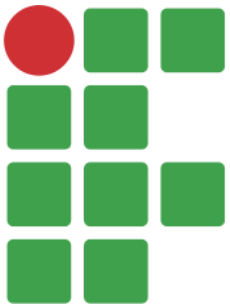


INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
DA PARAÍBA - CAMPUS CAMPINA GRANDE
CURSO SUPERIOR BACHARELADO EM ENGENHARIA DE COMPUTAÇÃO

AMANDA FERREIRA DA SILVA ALVES
SARA AYMÊ MARINHO GASPAR

LICITA ABERTA: INTERFACE AMIGÁVEL PARA VISUALIZAÇÃO DE DADOS DE LICITAÇÕES VIA API DO QUERIDO DIÁRIO

CAMPINA GRANDE - PB
2025

AMANDA FERREIRA DA SILVA ALVES
SARA AYMÊ MARINHO GASPAR

LICITA ABERTA: INTERFACE AMIGÁVEL PARA VISUALIZAÇÃO DE DADOS DE LICITAÇÕES VIA API DO QUERIDO DIÁRIO

Monografia apresentada ao Curso de Bacharelado em Engenharia de Computação, do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba - Campus Campina Grande, como requisito parcial para conclusão do curso de Engenharia de Computação.

ORIENTADOR: PROF. PAULO RIBEIRO LINS JÚNIOR

CAMPINA GRANDE - PB

2025

A474l Alves, Amanda Ferreira da Silva

Licita Aberta: interface amigável para visualização de dados de licitações via API do Querido Diário / Amanda Ferreira da Silva Alves, Sara Aymê Marinho Gaspar. - Campina Grande, 2025.

50 f. : il.

Trabalho de Conclusão de Curso (Curso Superior de Engenharia de Computação) - Instituto Federal da Paraíba, 2025.

Orientador: Prof. Paulo Ribeiro Lins Júnior.

1. Engenharia de computação 2. Transparência pública 3. Licitação pública 4. Desenvolvimento de sistema - Software livre I. Gaspar, Sara Aymê Marinho II. Lins Júnior, Paulo Ribeiro III. Título.

CDU 004.65

AMANDA FERREIRA DA SILVA ALVES
SARA AYMÊ MARINHO GASPAR

LICITA ABERTA: INTERFACE AMIGÁVEL PARA VISUALIZAÇÃO DE DADOS DE LICITAÇÕES VIA API DO QUERIDO DIÁRIO

Monografia apresentada ao Curso de Bacharelado em Engenharia de Computação, do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba - Campus Campina Grande, como requisito parcial para conclusão do curso de Engenharia de Computação.

Prof. Paulo Ribeiro Lins Júnior
Orientador

Prof. Iana Daya Cavalcante Facundo Passo
Membro da Banca

Prof. Katyusco de Farias Santos
Membro da banca

CAMPINA GRANDE - PB
2025

AGRADECIMENTOS

Amanda Ferreira da Silva Alves agradece:

Primeiramente, agradeço a Deus, o meu principal Orientador, por me permitir concluir este trabalho com excelência, concedendo sabedoria, paciência e direcionamento ao longo de todo o curso e desenvolvimento deste trabalho.

Sou grata aos meus pais, Jailda Ferreira e Romildo Alves, ao meu irmão, Samuel Ferreira, e ao meu noivo, Arthur Leandro, que sempre me apoiaram, acreditaram em mim e me incentivaram incondicionalmente a buscar o melhor resultado tanto durante o curso quanto na execução deste trabalho.

Expresso minha gratidão à minha parceira de trabalho, Sara Aymê, cuja dedicação e apoio foram fundamentais para a realização bem-sucedida deste projeto. Agradeço também ao nosso orientador, professor Paulo Ribeiro, por contribuir com a construção da ideia do trabalho e pelas orientações.

Por fim, agradeço aos amigos de turma pelo apoio e motivação ao longo da jornada acadêmica, assim como a todos os professores que contribuíram para meu crescimento e aprendizado.

RESUMO

Este trabalho apresenta o desenvolvimento de uma plataforma para extração, organização e apresentação de informações sobre licitações públicas, com o objetivo de promover maior transparência e acessibilidade aos dados contidos nos diários oficiais municipais. A ideia surgiu a partir da necessidade de disponibilizar informações públicas de forma que o cidadão comum consiga entender, analisar e acompanhar, seja para exercício da cidadania, seja para participar de pregões públicos com suas empresas. Durante a análise de soluções já existentes, foram identificadas plataformas como o **Querido Diário**, que disponibiliza diários oficiais por meio de uma API, e o **Exoonero**, que utiliza essa API para divulgar dados sobre nomeações e exonerações de maneira visual e interativa. No entanto, nenhuma dessas soluções atende especificamente à necessidade de organizar e facilitar o acesso às informações de licitações públicas. Assim, este projeto propõe aprimorar a forma como os dados de contratos e licitações são apresentados, aproveitando as contribuições do Querido Diário e incorporando uma visualização interativa inspirada no Exoonero, mas com uma proposta própria e focada em processos licitatórios nos municípios da Paraíba. A metodologia adotada incluiu o levantamento de requisitos, o desenvolvimento de processos de raspagem de dados e a criação de uma interface amigável, utilizando ferramentas de código aberto e tecnologias modernas de programação. Como resultado, foi desenvolvida uma solução funcional, disponibilizada em repositórios no GitHub. A plataforma demonstra potencial para facilitar o acesso a dados públicos e contribuir para a transparência governamental. Por fim, são apresentadas as limitações do projeto e sugestões para melhorias futuras, como a ampliação da cobertura de municípios e a implementação de funcionalidades analíticas.

Palavras-chave: Transparência pública. Licitações. Diários oficiais. Raspagem de dados. Plataformas digitais.

ABSTRACT

This paper presents the development of a platform for extracting, organizing, and presenting information about public tenders, aiming to promote greater transparency and accessibility to data found in municipal official gazettes. The idea arose from the need to make public information available in a way that is understandable, analyzable, and accessible to the general public, whether for exercising citizenship or participating in public bids with their companies. Existing solutions, such as Querido Diário, which provides official gazettes through an API, and Exoonero, which uses this API to display data on appointments and dismissals in a visual and interactive manner, were analyzed. However, none of these solutions specifically addressed the need to organize and facilitate access to public tender information. This project seeks to improve how data on contracts and tenders are presented, leveraging contributions from Querido Diário and incorporating an interactive visualization inspired by Exoonero, but with its own focus on tender processes in municipalities in Paraíba. The methodology included requirements gathering, data scraping processes, and the creation of a user-friendly interface using open-source tools and modern programming technologies. As a result, a functional solution was developed and made available in repositories on GitHub. The platform demonstrates potential to facilitate access to public data and contribute to government transparency. Finally, the limitations of the project are discussed, along with suggestions for future improvements, such as expanding coverage to more municipalities and implementing analytical features.

Keywords: Public transparency. Tenders. Official gazettes. Data scraping. Digital platforms.

SUMÁRIO

ABSTRACT..... 5
1. INTRODUÇÃO..... 11
2. FUNDAMENTAÇÃO TEÓRICA.....13
2.1 Transparência Pública e Acesso à Informação.....13
2.1.1 Conceito e Importância da Transparência Pública..... 14
2.1.2 Legislação e Normativas no Brasil.....14
Lei de Acesso à Informação (LAI).....14
Lei de Responsabilidade Fiscal (LRF)..... 15
Lei Complementar n.º 131/2009..... 15
Política de Dados Abertos..... 15
Portal da Transparência..... 15
Normativas Específicas no Brasil..... 16
2.2 Diários Oficiais e a Publicidade das Informações Municipais.....16
2.2.1 O Papel dos Diários Oficiais:..... 17
2.2.2 Desafios de Acessibilidade e Formato:..... 18
2.3 Tecnologias de Raspagem de Dados para Transparência.....19
2.3.1 Fundamentos da Raspagem de Dados.....19
OCR para Extração de Texto de PDFs..... 20
2.3.2 Aplicações em Transparência Governamental..... 20
Extração de Informações de Licitações e Contratos..... 21
Análise de Dados Orçamentários e Financeiros.....21
Acompanhamento de Políticas Públicas e Programas Governamentais..... 21
Apoio à Cidadania e Pesquisa..... 22
Facilidade e Acesso a Dados Públicos..... 22
Conclusão..... 22
2.4 APIs Governamentais e Dados Abertos..... 22
2.4.1 Conceito de APIs e Dados Abertos..... 23
2.4.2 A Iniciativa Querido Diário e Outras Fontes de Dados.....26
2.5 Design de Interface e Visualização de Dados para Transparência.....27
2.5.1 Importância da Usabilidade na Exibição de Dados Públicos.....27
2.5.2 Princípios de Visualização e Organização de Dados.....28
2.6 Comparativo com Outras Plataformas de Transparência Pública..... 30
2.6.1 Análise de Plataformas Existentes..... 30
2.6.2 Diferenciais e Inovações da Plataforma Proposta.....31
3. METODOLOGIA..... 32
3.1 Tipo de Pesquisa.....33

3.2 Procedimentos Metodológicos..... 33
3.2.1 Levantamento de Requisitos e Seleção das Fontes de Dados..... 33
3.2.2 Análise da Plataforma Exoonero..... 34
3.2.3 Identificação das Licitações nos Diários Oficiais..... 35
3.2.4 Raspagem e Processamento de Dados:..... 37
3.2.5 Desenvolvimento da Plataforma..... 39
Arquitetura Geral da Plataforma:..... 39
Descrição dos Componentes do Back-end:..... 39
Descrição dos Componentes do Front-end..... 41
3.2.6 Design da Interface e Usabilidade..... 41
3.2.7 Validação da Plataforma..... 41
3.3 Ferramentas Utilizadas..... 42
3.4 Limitações da Pesquisa..... 44
3.5 Validação dos Dados e Resultados..... 45
3.6 Divisão de Tarefas..... 45
3.7 Sustentabilidade e Expansão da Plataforma..... 45
3.8 Ética e Privacidade na Coleta de Dados..... 46
4. RESULTADOS E DISCUSSÕES..... 46
4.1 Funcionalidades da Plataforma..... 46
4.2 Telas da Plataforma..... 47
4.3 Repositórios no GitHub..... 49
4.4 Impacto e Perspectivas..... 49
4. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS..... 50
REFERÊNCIAS..... 51

1. INTRODUÇÃO

A transparência pública é um dos pilares fundamentais para a construção de uma sociedade democrática e participativa. Por meio dela, informações governamentais podem ser acessadas e utilizadas pelos cidadãos para monitorar a gestão pública e garantir o uso responsável dos recursos públicos. Segundo a **Organização para a Cooperação e Desenvolvimento Econômico (OCDE)**, a transparência pública é uma ferramenta essencial para combater a corrupção e aumentar a confiança da sociedade nas instituições públicas, contribuindo para uma governança mais responsável (OCDE, 2016).

No contexto brasileiro, a **transparência pública** não se limita à simples disponibilização de informações. Ela envolve a prática de apresentar dados sobre a gestão dos recursos e atividades governamentais de forma clara, estruturada e acessível. De acordo com Silva (2018), a transparência é fundamental para permitir que os cidadãos acompanhem e fiscalizem as ações do Estado, fortalecendo a democracia e promovendo maior participação social.

A **Lei de Acesso à Informação (Lei nº 12.527/2011)** consolidou o direito de acesso a dados públicos, determinando que os órgãos governamentais disponibilizem informações claras e atualizadas, salvo em casos de sigilo justificado. Complementarmente, a **Lei de Responsabilidade Fiscal (Lei Complementar n.º 101/2000)** e a **Política de Dados Abertos (Decreto n.º 8.777/2016)** reforçam a obrigação da transparência ativa, exigindo que informações de gastos, contratos e licitações sejam divulgadas em tempo real e em formatos abertos e reutilizáveis (BRASIL, 2011; BRASIL, 2016).

No entanto, apesar do avanço normativo, a forma como essas informações são divulgadas não atende às necessidades da população. A publicação de dados, especialmente nos diários oficiais municipais, ocorre de forma pouco acessível, frequentemente em formatos PDF extensos e não padronizados, dificultando a consulta e análise. Diante desse cenário, este trabalho tem como objetivo desenvolver uma plataforma online que realiza a extração e organiza os dados de licitações públicas de maneira visual e acessível, facilitando a interpretação dessas informações. O foco são os diários municipais de cidades do estado da Paraíba.

A motivação para essa iniciativa surgiu da nossa análise das dificuldades enfrentadas por cidadãos e pequenas empresas ao tentar acessar informações sobre licitações públicas. Embora existam ferramentas como o **Portal da Transparência**, a **Política de Dados Abertos** e a própria **publicação dos Diários Oficiais**, ainda há barreiras significativas, como a ausência de padronização, a inconsistência das publicaçõe**s** (com lacunas temporais em alguns municípios) e a complexidade dos formatos disponibilizados.

Essa dificuldade de acesso impacta diretamente pequenos empresários locais, que deixam de participar de processos licitatórios municipais por não conseguirem identificar oportunidades de forma simples e em tempo hábil. Além disso, cidadãos comuns que desejam exercer o controle social sobre o uso dos recursos públicos esbarram em barreiras técnicas e operacionais que inviabilizam esse acompanhamento. O foco nos municípios do estado da Paraíba se deve ao fato de residirmos na região, o que facilitou a identificação das dificuldades locais no acesso às informações sobre licitações públicas. Com isso, buscamos contribuir para a transparência e o controle social em uma realidade próxima de nós

Como facilitador para a coleta de dados, adotamos a **plataforma Querido Diário**, que centraliza diários municipais por meio de uma API. Nos inspiramos na proposta do **Exoonero**, que organiza informações sobre exonerações e nomeações de forma simples e intuitiva. A partir dessas experiências, desenvolvemos o **Licita Aberta**, uma plataforma que visa tornar as informações sobre licitações públicas mais organizadas, acessíveis e fáceis de interpretar, permitindo que qualquer cidadão possa acompanhar de forma clara e intuitiva como o dinheiro público está sendo gasto.

A metodologia adotada incluiu o levantamento de requisitos, análise das fontes de dados, desenvolvimento de processos de raspagem e tratamento das informações, além da criação de uma interface amigável. O projeto busca oferecer uma solução de código aberto, que não apenas promova a transparência, mas também **supere desafios relacionados à padronização e ao formato dos dados governamentais**. A plataforma propõe transformar a apresentação das informações, oferecendo **visualizações interativas** e **ferramentas analíticas** que incentivam maior participação social e engajamento com os dados públicos.

O presente trabalho está estruturado da seguinte forma: no Capítulo 2, são apresentados os conceitos teóricos que embasaram o desenvolvimento da plataforma. O

Capítulo 3 aborda os procedimentos metodológicos e os resultados obtidos, enquanto o Capítulo 4 discute as conclusões, limitações e perspectivas para trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

A transparência pública é crucial para garantir a participação cidadã e a fiscalização da gestão pública. Para que seja realmente competente, é necessário que as informações sejam não apenas disponibilizadas, mas também apresentadas de forma acessível, organizada e de fácil interpretação. No Brasil, a Lei de Acesso à Informação (Lei nº 12.527/2011) trouxe avanços significativos, obrigando os órgãos públicos a fornecer dados em formatos acessíveis ao público. No entanto, a disponibilização de dados em formatos pouco amigáveis, como PDFs extensos, ainda representa um desafio para a plena transparência.

A utilização de plataformas digitais que tratem, organizem e apresentem essas informações de forma interativa tem se mostrado uma solução eficaz. Ferramentas como o Querido Diário e o Exoonero já são exemplos de como a tecnologia pode facilitar o acesso a dados públicos. Contudo, ainda há lacunas na transparência de processos licitatórios, o que motiva o desenvolvimento de uma plataforma que centraliza e organiza informações sobre licitações públicas de forma mais acessível e prática para os cidadãos.

2.1 Transparência Pública e Acesso à Informação

A transparência pública é um princípio fundamental para a governança democrática, permitindo que cidadãos tenham acesso às informações governamentais de maneira clara, acessível e compreensível. A disponibilização de dados sobre a administração pública fortalece o controle social, reduz a assimetria informacional entre governo e sociedade e contribui para o combate à corrupção.

No Brasil, diversas iniciativas e marcos legais têm buscado garantir o acesso aos dados governamentais de forma ampla e estruturada. No entanto, a transparência vai além da mera disponibilização de documentos; é necessário que as informações sejam acessíveis e compreensíveis ao público, permitindo sua análise e utilização. A utilização de tecnologias digitais, como plataformas interativas e bases de dados abertas, tem se mostrado essencial para a efetivação da transparência pública, garantindo que informações relevantes estejam ao alcance de todos.

2.1.1 Conceito e Importância da Transparência Pública

A transparência pública pode ser definida como a prática de disponibilizar informações referentes à administração pública de maneira clara, ampla e acessível para a população, permitindo que a sociedade tenha conhecimento sobre as ações e os recursos do governo. Segundo Silva (2019), a transparência vai além da simples publicação de dados, envolvendo o fornecimento de informações claras e organizadas, de modo a possibilitar o controle social e fortalecer a relação entre governo e sociedade.

A importância da transparência está diretamente ligada à construção de uma sociedade mais participativa e informada. Governos que disponibilizam dados de forma acessível criam condições para que cidadãos e organizações possam fiscalizar ações públicas, contribuindo para a prevenção de práticas corruptas e para a proposição de melhorias no governo.

No Brasil, a transparência pública ganhou impulso com a implementação da Lei de Acesso à Informação (LAI), que representou um marco na regulamentação do direito ao acesso à informação. Essa legislação assegura que qualquer cidadão possa solicitar dados sobre a administração pública, salvo em situações excepcionais, reforçando a importância da transparência como base para uma democracia mais participativa e inclusiva.

2.1.2 Legislação e Normativas no Brasil

A transparência pública no Brasil está fundamentada em um conjunto abrangente de legislações e normativas que buscam assegurar o direito dos cidadãos ao acesso à informação e promover a transparência na gestão pública. Esses instrumentos têm como objetivo principal ampliar o controle social, combater a corrupção e fortalecer os pilares democráticos.

Lei de Acesso à Informação (LAI)

A Lei n.º 12.527/2011, conhecida como Lei de Acesso à Informação (LAI), é o principal marco regulatório sobre o tema no país. Ela regulamenta o direito constitucional de acesso às informações públicas, previsto no artigo 5.º, inciso XXXIII, e nos artigos 37 e 216 da Constituição Federal. A LAI estabelece que toda informação sob custódia de órgãos públicos é, por princípio, acessível, salvo exceções claramente justificadas, como aquelas relacionadas à privacidade, segurança nacional ou sigilo industrial (BRASIL, 2011).

Entre as principais inovações trazidas pela LAI estão:

- A obrigação de os órgãos públicos disponibilizarem informações de forma ativa e proativa, por meio de transparência ativa, como portais eletrônicos acessíveis;
- O direito de qualquer cidadão solicitar informações sem a necessidade de justificar o pedido (transparência passiva);
- Prazos claros para o atendimento às solicitações e penalidades para o descumprimento;
- Classificação de documentos sigilosos em níveis (reservado, secreto e ultrassecreto) e prazos máximos de sigilo.

Lei de Responsabilidade Fiscal (LRF)

Outro marco relevante é a Lei Complementar n.º 101/2000, conhecida como Lei de Responsabilidade Fiscal (LRF). Essa legislação estabelece normas voltadas à gestão fiscal responsável, determinando que todos os entes federativos — União, Estados, Municípios e Distrito Federal — publiquem dados sobre receitas, despesas, execução orçamentária e gestão da dívida pública de forma transparente e acessível. A LRF também exige a criação de relatórios periódicos, como o Relatório de Gestão Fiscal (RGF) e o Relatório Resumido da Execução Orçamentária (RREO), que são fundamentais para o monitoramento da administração financeira pública (BRASIL, 2000).

Lei Complementar n.º 131/2009

Essa lei complementa a LRF ao instituir a obrigatoriedade de publicação de informações financeiras em tempo real, preferencialmente em plataformas digitais. Com isso, ampliou-se a transparência na execução orçamentária e financeira, facilitando o controle social e permitindo maior vigilância por parte da sociedade.

Política de Dados Abertos

O Decreto n.º 8.777/2016 instituiu a Política de Dados Abertos no âmbito do governo federal, com o objetivo de promover a abertura e o uso estratégico de dados públicos. Essa política incentiva a disponibilização de informações em formatos abertos, estruturados e reutilizáveis, fomentando a criação de soluções inovadoras e colaborativas por cidadãos, empresas e organizações civis.

Portal da Transparência

A criação do Portal da Transparência, administrado pela Controladoria-Geral da União (CGU), é outro marco importante na transparência pública brasileira. Lançado em 2004, o portal permite que cidadãos acessem informações detalhadas sobre a execução orçamentária e financeira do governo federal, incluindo despesas, repasses a estados e municípios, contratos, licitações e outros dados relevantes.

Normativas Específicas no Brasil

Esses marcos legais e normativos representam um esforço contínuo do Estado brasileiro para garantir que a administração pública seja mais aberta, responsável e acessível à sociedade. Com o fortalecimento dessas legislações, os cidadãos conquistaram maior capacidade de fiscalização e controle sobre a utilização dos recursos públicos, promovendo uma gestão mais eficiente e ética.

Além disso, a criação de plataformas digitais, como o Portal da Transparência, e a implementação de políticas de dados abertos, têm sido fundamentais para facilitar o acesso a informações e aumentar a interação entre o governo e a sociedade. Embora ainda existam desafios na implementação plena da transparência em todos os níveis de governo, as normas brasileiras oferecem uma base sólida para avançar rumo a uma gestão pública cada vez mais participativa e responsiva às necessidades da população.

2.2 Diários Oficiais e a Publicidade das Informações Municipais

Os diários oficiais são instrumentos fundamentais para a publicidade dos atos administrativos, garantindo que informações de interesse público sejam divulgadas de forma oficial e acessível. Eles representam um meio histórico de registro e comunicação entre o poder público e a sociedade, assegurando a transparência das ações governamentais, especialmente em âmbito municipal.

No entanto, apesar de sua relevância para a publicidade das informações, os diários oficiais enfrentam desafios relacionados à acessibilidade e ao formato dos dados publicados. Essas limitações impactam diretamente a capacidade da sociedade de acessar e compreender as informações neles contidas, especialmente em um contexto de crescente demanda por transparência e participação cidadã.

2.2.1 O Papel dos Diários Oficiais:

Os diários oficiais desempenham um papel crucial como veículos de comunicação pública nos níveis municipal, estadual e federal. Eles são os principais meios de divulgação de atos administrativos e decisões governamentais, como leis, decretos, portarias, nomeações, exonerações, licitações e contratos, entre outros documentos oficiais. A principal função dos diários oficiais é garantir a transparência das ações do poder público, promovendo a publicidade dos atos administrativos e possibilitando o controle social.

Através dos diários oficiais, a sociedade tem acesso direto às informações sobre as ações e decisões governamentais, possibilitando que cidadãos, jornalistas, organizações da sociedade civil e demais interessados acompanhem e fiscalizem as atividades dos governos. Esse acesso à informação é um componente essencial para o exercício da cidadania e para o fortalecimento da democracia, pois permite que a população acompanhe a execução de políticas públicas e exerça o direito de propor melhorias, sugerir e questionar.

Além de promover a transparência, os diários oficiais asseguram o cumprimento do princípio da publicidade, garantido pela Constituição Federal. A publicidade dos atos administrativos não é apenas um dever ético e legal, mas também uma condição para a efetividade da gestão pública, permitindo que a população tenha pleno conhecimento das ações do Estado e possibilite a atuação como um agente de fiscalização. Dessa forma, os diários oficiais funcionam como um canal direto entre o governo e a população, mantendo os cidadãos informados e engajados na administração pública.

Por fim, é importante destacar que, além de sua função informativa, os diários oficiais também desempenham um papel na formalização e validade dos atos administrativos. A publicação desses documentos nos diários oficiais é necessária para que eles tenham efeitos legais, especialmente no caso de contratos e licitações, que precisam ser formalmente divulgados para garantir a transparência e o cumprimento das normativas legais. A publicação no diário oficial confere publicidade e autenticidade aos processos licitatórios e aos contratos firmados, permitindo que a sociedade acompanhe o uso dos recursos públicos e participe ativamente da fiscalização das ações governamentais.

2.2.2 Desafios de Acessibilidade e Formato:

A disponibilização de informações públicas em formatos acessíveis é um dos maiores desafios para a promoção de uma transparência efetiva. Embora a publicação de dados em diários oficiais seja uma prática importante para a transparência, o formato utilizado frequentemente representa um obstáculo para grande parte da população. Muitos órgãos públicos optam por publicar informações, como licitações, contratos e atos administrativos, em documentos como PDF, planilhas ou até mesmo imagens, o que muitas vezes, pode afetar a facilidade com que o cidadão acessa, consulta e utiliza essas informações.

Podemos abordar em especial por um formato específico, o PDF, por ser um formato amplamente utilizado para a publicação de documentos oficiais, apresenta algumas vantagens, como a preservação da formatação e a segurança no compartilhamento de informações. No entanto, quando se trata de disponibilizar dados para consulta pública, o formato PDF se torna um desafio significativo, especialmente quando o conteúdo do documento é extenso e estruturado de maneira complexa. O usuário precisa abrir cada arquivo individualmente, o que pode tornar o processo de busca de informações demorado e ineficiente.

Um dos principais problemas relacionados ao formato PDF é a falta de interatividade e a dificuldade de realizar buscas dentro do documento. Enquanto plataformas digitais modernas permitem a pesquisa de palavras-chave e a visualização dinâmica de informações, os PDFs limitam a experiência do usuário a uma consulta estática e linear. Além disso, documentos longos e não estruturados de maneira amigável podem gerar confusão e dificultar a localização de informações relevantes.

Outro aspecto crítico é a acessibilidade para pessoas com deficiência. Embora existam tecnologias assistivas, como leitores de tela, que permitem que pessoas com deficiência visual acessem documentos digitais, muitos PDFs não são otimizados para essa funcionalidade. Isso ocorre quando os documentos não incluem descrições alternativas para imagens, ou quando a estrutura do texto não está organizada de forma que permita uma leitura fluída pelos softwares de leitura de tela. Quando documentos governamentais são disponibilizados sem essa otimização, as informações ficam fora do alcance de uma parte significativa da população.

Além do PDF, outros formatos de arquivo, como imagens escaneadas ou planilhas não estruturadas, também geram problemas de acessibilidade. Imagens escaneadas, por exemplo, podem ser ilegíveis para leitores de tela, uma vez que não contêm texto editável. As planilhas, por sua vez, podem ser difíceis de interpretar sem uma organização clara das informações, o que limita a capacidade dos cidadãos de realizar análises mais complexas.

Outro desafio importante é a falta de padronização na publicação dos documentos. Cada órgão pode adotar diferentes modelos de relatórios, contratos e licitações, o que gera inconsistências na forma como os dados são apresentados. Isso torna ainda mais difícil a comparação entre informações de diferentes órgãos ou até mesmo a integração dessas informações em sistemas de monitoramento.

Portanto, a disponibilização de dados em formatos como PDF apresenta desafios consideráveis para a transparência pública. A falta de interatividade, a dificuldade de pesquisa, a falta de acessibilidade e a falta de padronização no formato dos documentos são questões que precisam ser resolvidas para garantir que a sociedade tenha acesso fácil e eficiente às informações governamentais. A busca por soluções que envolvam a adaptação dos formatos de dados e a criação de plataformas mais acessíveis é essencial para garantir que a transparência seja efetiva e inclusiva.

2.3 Tecnologias de Raspagem de Dados para Transparência

A raspagem de dados é um conjunto de técnicas e ferramentas utilizadas para extrair informações de fontes de dados não estruturadas ou sem uma interface de programação de aplicativos (API) que permita acesso direto. Essas tecnologias desempenham um papel essencial na promoção da transparência governamental, permitindo que cidadãos, pesquisadores e organizações acessem informações que, de outra forma, estariam difíceis de serem consultadas ou manipuladas.

2.3.1 Fundamentos da Raspagem de Dados

A raspagem de dados, também conhecida como *web scraping*, é uma técnica automatizada que permite a extração de informações de websites, arquivos e documentos que não estão prontamente disponíveis para análise, como páginas web HTML, documentos PDF, imagens e outros formatos. Essas informações podem ser organizadas e transformadas em dados estruturados, prontos para análise.

O processo de raspagem de dados envolve várias etapas: primeiro, é necessário acessar o conteúdo de um site ou documento digital, em seguida, identificar e extrair as partes relevantes dos dados e, finalmente, organizar esses dados de forma estruturada, como em tabelas ou planilhas.

Existem diferentes abordagens e tecnologias usadas para a raspagem de dados. A mais comum é a raspagem de conteúdo HTML de páginas web, onde scripts automatizados navegam nas páginas, identificam os elementos HTML relevantes (como títulos, textos, links e tabelas) e extraem esses dados. No entanto, quando os dados estão armazenados em formatos mais complexos, como PDFs ou imagens, são necessárias técnicas adicionais, como a *extração de texto com OCR* (Reconhecimento Óptico de Caracteres).

OCR para Extração de Texto de PDFs

Um dos maiores desafios na raspagem de dados governamentais está no acesso a documentos em formato PDF. Embora o PDF seja um formato amplamente utilizado para publicação de documentos, sua estrutura pode dificultar a extração automática de informações. Em documentos PDF escaneados, por exemplo, o texto não é "legível" por ferramentas de raspagem convencionais.

A tecnologia OCR é uma solução eficaz para esse problema. OCR é um processo que converte imagens de texto (como aquelas em PDFs escaneados ou imagens de documentos) em texto digital que pode ser editado e manipulado. Usando OCR, é possível extrair informações de documentos governamentais, como atas de licitação, contratos e outros documentos oficiais que foram digitalizados. Ferramentas como Tesseract, Google Vision e Adobe OCR são amplamente utilizadas para realizar esse tipo de extração.

Combinando OCR com raspagem de dados, é possível criar fluxos de trabalho automatizados para capturar dados de uma variedade de fontes e convertê-los em informações acessíveis e analisáveis.

2.3.2 Aplicações em Transparência Governamental

A raspagem de dados desempenha um papel fundamental na promoção da transparência governamental, pois permite que informações que antes estavam em formatos difíceis de acessar sejam transformadas em dados estruturados e acessíveis. Com o uso de

técnicas de raspagem, qualquer pessoa pode obter dados de interesse público de forma eficiente, sem a necessidade de recorrer a solicitações formais ou processos burocráticos.

Extração de Informações de Licitações e Contratos

Uma das principais aplicações da raspagem de dados na transparência governamental é na extração de informações relacionadas a licitações, contratos e outros processos administrativos. No Brasil, por exemplo, essas informações são frequentemente publicadas em diversos diários oficiais ou em páginas de órgãos públicos em formatos como PDF ou HTML. A raspagem de dados pode ser utilizada para coletar informações sobre os participantes, valores, objetos e status das licitações, ajudando cidadãos, empresas e pesquisadores a monitorarem os processos licitatórios de maneira mais eficaz.

Essas informações são essenciais para o controle social e a prevenção de práticas corruptas. A raspagem de dados permite que qualquer pessoa tenha acesso a detalhes sobre contratos firmados entre o governo e empresas privadas, possibilitando uma vigilância constante sobre o uso de recursos públicos.

Análise de Dados Orçamentários e Financeiros

Outra área importante da raspagem de dados para transparência é a análise de dados orçamentários e financeiros. Informações sobre a execução orçamentária do governo, como despesas públicas, receitas, repasses e execução de projetos, são frequentemente divulgadas em sites de transparência ou em documentos de difícil leitura, como PDFs e planilhas. Usando técnicas de raspagem, essas informações podem ser extraídas automaticamente e processadas para gerar relatórios e análises mais fáceis de entender. A visualização de dados, como gráficos e tabelas interativas, pode ser criada para tornar os dados orçamentários mais acessíveis ao público geral.

Acompanhamento de Políticas Públicas e Programas Governamentais

Além disso, a raspagem de dados pode ser usada para monitorar a implementação de políticas públicas e programas governamentais. Por exemplo, dados sobre saúde, educação, infraestrutura e assistência social podem ser extraídos de diversos sites de ministérios e secretarias estaduais e municipais. Essas informações podem ser analisadas para avaliar a

eficácia de programas públicos e a alocação de recursos, além de identificar áreas que necessitam de mais atenção.

Apoio à Cidadania e Pesquisa

Para jornalistas, pesquisadores e organizações não governamentais (ONGs) que monitoram a atuação do governo, a raspagem de dados também é uma ferramenta poderosa. Organizações que promovem a transparência podem usar essas técnicas para coletar grandes volumes de dados e realizar análises aprofundadas sobre o governo e suas práticas. Isso facilita a produção de relatórios e a identificação de possíveis irregularidades ou áreas que precisam de melhorias.

Facilidade e Acesso a Dados Públicos

Com o uso da raspagem de dados, cidadãos podem acessar mais facilmente as informações públicas que, de outra forma, seriam difíceis de encontrar ou estariam inacessíveis devido à falta de uma plataforma unificada de acesso. Por exemplo, em vez de procurar em múltiplos diários oficiais ou sites de diferentes órgãos governamentais, uma raspagem automatizada pode consolidar essas informações em um único repositório acessível e facilmente pesquisável. Isso promove uma maior transparência e um maior engajamento da sociedade no acompanhamento das ações do governo.

Conclusão

A raspagem de dados tem se mostrado uma ferramenta essencial para aumentar a transparência na gestão pública, principalmente quando se lida com grandes volumes de dados dispersos e difíceis de acessar. Ao utilizar tecnologias como OCR e scraping de páginas web, é possível extrair informações de diversas fontes e apresentá-las de forma acessível, o que facilita a análise e a fiscalização por parte da sociedade. Dessa forma, as técnicas de raspagem contribuem diretamente para a promoção da transparência governamental e para o fortalecimento do controle social.

2.4 APIs Governamentais e Dados Abertos

A utilização de tecnologias para facilitar o acesso a informações públicas tem se mostrado essencial para garantir maior transparência nas gestões governamentais. Uma dessas tecnologias, as APIs (Interfaces de Programação de Aplicações), desempenham um

papel crucial na democratização da informação. Por meio delas, é possível tornar dados públicos acessíveis de maneira rápida, estruturada e em tempo real, permitindo que cidadãos, organizações e empresas interajam com as informações de maneira eficiente e prática. Quando combinadas com os conceitos de "dados abertos", essas ferramentas promovem a ampliação da transparência e fortalecem a participação cidadã.

Dados abertos são informações disponibilizadas de forma livre, sem restrições de acesso, permitindo que qualquer pessoa possa usá-las para diversos fins, como pesquisa, fiscalização e desenvolvimento de novos serviços. Essa prática tem se disseminado em diversas esferas governamentais, criando uma oportunidade para que a sociedade exerça o controle social de maneira mais efetiva. No Brasil, iniciativas como a Lei de Acesso à Informação (LAI) e o Portal da Transparência têm sido fundamentais para garantir o acesso aos dados públicos, mas a adoção de APIs e dados abertos elevou ainda mais o impacto dessa legislação, permitindo a criação de plataformas interativas e acessíveis.

2.4.1 Conceito de APIs e Dados Abertos

APIs, ou Interfaces de Programação de Aplicações, são conjuntos de regras e protocolos que permitem que diferentes sistemas e plataformas se comuniquem e compartilhem dados de forma automatizada e eficiente. Elas funcionam como "ponte" entre sistemas, facilitando a troca de informações sem a necessidade de intervenção manual. No contexto da transparência pública, as APIs são utilizadas para tornar dados governamentais acessíveis em tempo real, possibilitando que qualquer pessoa, por meio de simples requisições de software, obtenha informações detalhadas sobre a gestão pública, como gastos, contratos e dados de licitações, entre outros.

As APIs podem fornecer dados de forma estruturada, geralmente em formatos como JSON ou XML, permitindo que sejam facilmente lidos, processados e analisados. Isso facilita a criação de ferramentas de visualização, geração de relatórios e até mesmo de novos serviços baseados nesses dados, como aplicativos móveis ou sites informativos. A utilização de APIs na administração pública é uma estratégia para garantir a transparência das ações governamentais, já que qualquer pessoa pode acessar informações sobre a execução orçamentária, a execução de políticas públicas e outros dados relevantes.

Os dados abertos são informações disponibilizadas sem restrições de acesso, podendo ser utilizados, modificados e compartilhados por qualquer pessoa. Esse conceito de dados

abertos é importante para garantir que a informação pública esteja acessível a todos, de forma livre, e seja possível sua reutilização para diferentes fins. Ao disponibilizar dados abertos, os governos permitem que cidadãos, jornalistas e pesquisadores possam explorar essas informações, realizar análises e promover uma fiscalização mais eficaz das ações governamentais. Isso fortalece a democracia, pois proporciona aos cidadãos o poder de questionar e influenciar as decisões públicas com base em dados reais e atualizados.

A relevância dos dados abertos na transparência pública vai além do simples acesso à informação. Ao disponibilizar dados em formatos abertos e estruturados, os governos incentivam a inovação e o desenvolvimento de soluções tecnológicas por parte de terceiros, como startups e organizações não governamentais, que podem utilizar essas informações para criar novos serviços, como aplicativos que ajudam a monitorar gastos públicos ou sistemas de alerta para práticas ilegais.

Níveis de Classificação dos Dados Abertos (ITL)

Para garantir a transparência e a usabilidade dos dados públicos, a Infraestrutura Nacional de Dados Abertos (ITL) classifica os dados abertos em diferentes níveis, de acordo com sua estruturação e acessibilidade. Esses níveis ajudam a definir o grau de aproveitamento das informações disponibilizadas e sua aplicabilidade para análise, monitoramento e inovação.

- **Nível 1 – Dados Não Estruturados:** Informações disponibilizadas em formatos que não permitem processamento automatizado, como arquivos PDF escaneados ou imagens. Esses dados exigem técnicas avançadas, como OCR (Reconhecimento Óptico de Caracteres), para serem extraídos e analisados.
- **Nível 2 – Dados Estruturados em Formatos Proprietários:** Dados organizados em planilhas ou bancos de dados, mas disponibilizados em formatos fechados e proprietários, como .xls (Excel), dificultando sua interoperabilidade com outras plataformas.
- **Nível 3 – Dados Estruturados em Formatos Abertos:** Informações fornecidas em formatos não proprietários e estruturados, como CSV, JSON ou XML, permitindo fácil leitura e integração com diferentes sistemas e softwares.

- **Nível 4 – Dados Vinculados:** Dados que, além de estruturados, possuem conexões semânticas com outras bases de dados por meio de identificadores padronizados, permitindo maior interconectividade e análise cruzada.
- **Nível 5 – Dados Vinculados e Contextualizados:** Além de serem vinculados, esses dados contam com metadados detalhados e documentação completa, possibilitando uma compreensão mais aprofundada e facilitando seu uso em pesquisas e aplicações avançadas.

A Figura 1, logo abaixo, ilustra esses níveis de classificação:

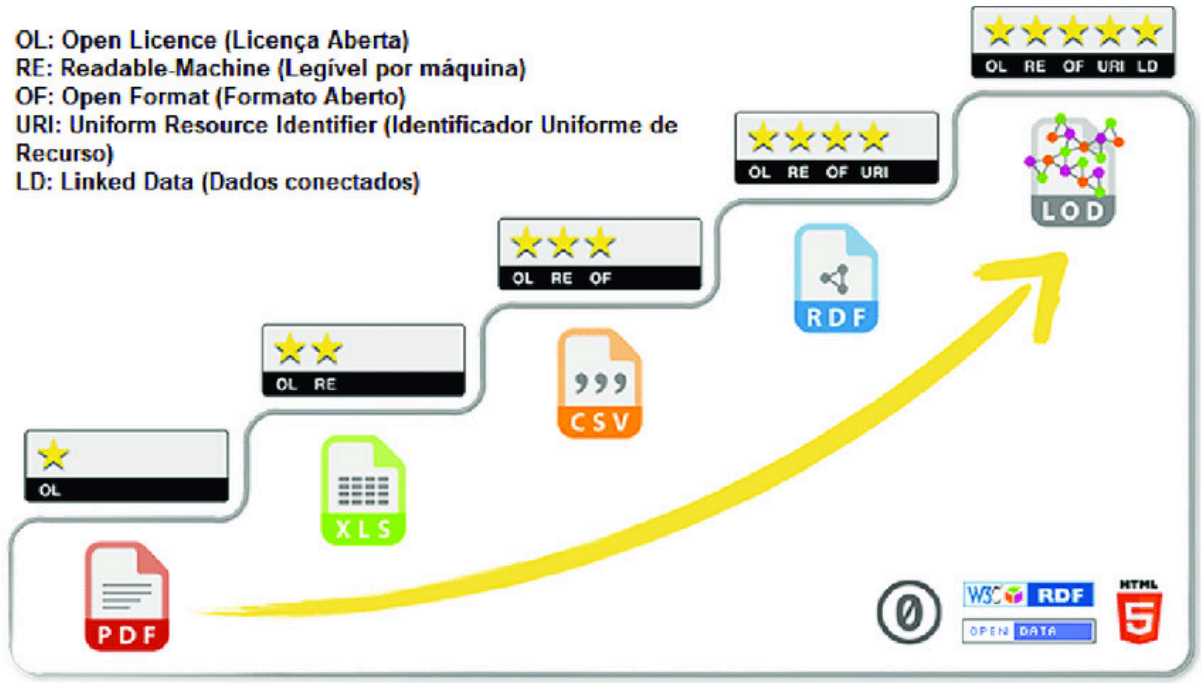


Figura 1 – Níveis de Classificação dos Dados Abertos
(Fonte: Adaptado de 5-STAR (2020). Disponível em: https://www.researchgate.net/figure/Figura-1-As-5-estrelas-dos-Dados-Abertos-Conectados-5-STAR-2020_fig7_349477731. Acesso em: 11 mar. 2025.)

A categorização dos dados abertos é essencial para garantir que as informações públicas sejam organizadas e compreensíveis, facilitando seu uso tanto por cidadãos quanto por desenvolvedores que criam soluções baseadas nesses dados. A transparência efetiva requer não apenas a publicação das informações, mas também sua organização em formatos acessíveis e interoperáveis, promovendo uma maior participação cidadã e inovação no uso de dados públicos

2.4.2 A Iniciativa Querido Diário e Outras Fontes de Dados

Uma das iniciativas mais notáveis no Brasil para promover a transparência por meio de APIs é o Querido Diário, uma plataforma que visa simplificar o acesso aos diários oficiais de diversas esferas governamentais, com destaque para os municipais. O Querido Diário oferece uma API que permite que qualquer usuário acesse informações publicadas nos diários oficiais de forma estruturada e automatizada. Isso inclui dados sobre processos licitatórios, contratos, nomeações, exonerações e outras informações de interesse público que, tradicionalmente, estavam disponíveis apenas em formatos de difícil acesso, como PDFs e imagens.

Antes da criação do Querido Diário, os cidadãos enfrentavam dificuldades para acessar e interpretar os dados nos diários oficiais, já que muitas vezes essas informações eram publicadas em formatos não amigáveis e de difícil leitura. A plataforma resolve esse problema ao oferecer uma interface programática simples, permitindo que desenvolvedores, jornalistas e cidadãos interessados possam acessar, processar e utilizar os dados de maneira eficiente.

A API do Querido Diário é um exemplo perfeito de como as tecnologias de dados abertos podem ser utilizadas para aumentar a transparência e melhorar o acesso à informação pública. Ao disponibilizar os dados dos diários oficiais de forma estruturada, a plataforma facilita a realização de análises e o acompanhamento de ações governamentais em tempo real, promovendo um controle social mais eficaz. A partir dessa API, é possível, por exemplo, monitorar licitações e contratos, acompanhar o andamento de processos administrativos e verificar nomeações e exonerações de servidores públicos, o que contribui para um maior engajamento e fiscalização da população.

Além disso, o Querido Diário se destaca por sua capacidade de abranger diários de diferentes municípios, não se limitando apenas aos diários oficiais da União ou dos Estados. Isso torna a plataforma uma ferramenta valiosa para o acompanhamento da gestão pública em nível local, onde muitas vezes há uma escassez de dados acessíveis e estruturados. Dessa forma, o Querido Diário contribui diretamente para a ampliação da transparência e o fortalecimento do controle social, não apenas em nível federal, mas também nos níveis estadual e municipal.

Outras APIs governamentais, como o Portal da Transparência, complementam essa iniciativa, oferecendo acesso a dados orçamentários e financeiros do governo federal. Essas plataformas são fundamentais para a criação de uma rede de dados públicos interconectada, permitindo que cidadãos e organizações possam acessar informações detalhadas sobre como o dinheiro público está sendo gasto, além de possibilitar o desenvolvimento de novas soluções tecnológicas para o monitoramento e análise da gestão pública.

Essas iniciativas reforçam o compromisso do governo com a transparência e facilitam a participação ativa da sociedade nas questões de interesse público. Ao oferecer informações de maneira estruturada e acessível, o Querido Diário e outras APIs relacionadas contribuem para um ambiente mais transparente e colaborativo, no qual os cidadãos podem exercer seu direito de fiscalizar e influenciar a administração pública com base em dados reais e verificáveis.

2.5 Design de Interface e Visualização de Dados para Transparência

A crescente demanda por transparência na gestão pública tem gerado a necessidade de ferramentas eficazes para garantir o acesso fácil e claro às informações governamentais. O design de interface e a visualização de dados desempenham um papel fundamental nesse processo, uma vez que facilitam a compreensão dos dados públicos e promovem o engajamento dos cidadãos na fiscalização e participação nos processos governamentais. O uso de interfaces intuitivas e a apresentação adequada dos dados permitem que qualquer usuário, independentemente de seu nível de conhecimento técnico, consiga acessar e compreender as informações de maneira eficiente e assertiva.

2.5.1 Importância da Usabilidade na Exibição de Dados Públicos

A usabilidade, no contexto de sistemas de transparência pública, refere-se à capacidade de uma interface ser fácil de usar, acessível e eficiente, de forma a atender às necessidades dos cidadãos sem que eles encontrem obstáculos técnicos ou complexidade desnecessária. A facilidade de navegação em plataformas que disponibilizam dados públicos é crucial para que qualquer usuário consiga encontrar as informações de maneira rápida e intuitiva.

Para garantir a usabilidade, é importante que os portais de transparência sejam projetados com foco na experiência do usuário, considerando aspectos como simplicidade,

clareza e organização dos elementos. A estrutura da página deve ser lógica e coerente, facilitando a navegação e a busca por informações específicas. A visualização deve permitir que dados complexos, como relatórios financeiros, licitações ou gastos públicos, sejam apresentados de maneira clara, com a utilização de gráficos e tabelas interativas que ajudem o usuário a compreender os números e as estatísticas de forma simples.

Além disso, a usabilidade deve abranger a acessibilidade, atendendo a pessoas com diferentes habilidades, incluindo aquelas com deficiências visuais, auditivas ou motoras. O uso de contraste adequado, textos legíveis, compatibilidade com leitores de tela e recursos de navegação por teclado são exemplos de boas práticas que devem ser incorporadas nos projetos de design de interfaces públicas. A implementação de versões responsivas também é essencial, considerando que muitos cidadãos acessam esses dados através de dispositivos móveis. Portanto, o design deve ser adaptável e garantir uma experiência de usuário de alta qualidade em diferentes plataformas, como smartphones, tablets e computadores.

A transparência só será plenamente alcançada quando os cidadãos conseguirem acessar os dados públicos com facilidade e sem barreiras, seja pela complexidade das informações, seja por dificuldades na navegação. Para isso, é fundamental que as plataformas de dados abertos adotem boas práticas de usabilidade, criando um ambiente em que a busca por informações seja simples e direta.

2.5.2 Princípios de Visualização e Organização de Dados

A apresentação de dados públicos, como informações de licitações, despesas públicas ou processos administrativos, deve seguir princípios de visualização que facilitem a análise e interpretação rápida e precisa. A boa visualização de dados é um aspecto essencial na construção de uma interface de transparência pública eficaz, pois transforma grandes volumes de dados em informações claras e compreensíveis, permitindo que o cidadão tenha uma visão holística e detalhada dos processos governamentais.

Entre os princípios fundamentais para a visualização de dados está a clareza, que se refere à capacidade de apresentar as informações de maneira que o usuário compreenda sem esforço. Isso significa que gráficos, tabelas e mapas devem ser simples, mas ricos em informações relevantes. A hierarquia visual também é importante: a disposição dos elementos na página deve refletir a importância das informações, destacando os dados mais significativos e permitindo que o usuário os identifique rapidamente.

Além disso, o uso de gráficos interativos é uma prática recomendada para a apresentação de dados em plataformas de transparência pública. Ferramentas como gráficos de barras, linhas, pizzas e mapas interativos podem ser usadas para representar dados financeiros, contratos, licitações e outros elementos de forma visualmente atraente e facilmente compreensível. Esses gráficos permitem que os cidadãos explorem os dados de maneira dinâmica, filtrando informações por período, categoria ou tipo de gasto, por exemplo. O uso de tabelas interativas também é fundamental, pois permite que o usuário organize e filtre os dados de acordo com sua necessidade, facilitando a análise personalizada.

O layout da página e a disposição dos gráficos e dados devem seguir princípios de organização visual, de modo que os dados sejam agrupados de forma lógica. Por exemplo, informações sobre licitações e contratos podem ser agrupadas em seções específicas, com filtros de pesquisa que permitam ao usuário refinar os resultados por data, valor ou tipo de contrato. Essa organização facilita a navegação e assegura que o usuário tenha uma visão completa e clara dos dados.

Outro ponto importante na visualização de dados é a escolha adequada de cores. As cores devem ser usadas de forma estratégica para transmitir informações, como destacar valores mais altos ou indicar variações em uma tendência. No entanto, é essencial garantir que as escolhas de cores não prejudiquem a legibilidade, especialmente para pessoas com deficiências visuais, como daltonismo. O contraste entre o texto e o fundo deve ser suficiente para garantir a legibilidade e facilitar a compreensão dos dados apresentados.

A contextualização dos dados também é um princípio importante. A apresentação isolada de números pode ser difícil de interpretar, por isso é fundamental que os dados sejam acompanhados de explicações ou comparações que ajudem o usuário a entender o que está sendo representado. Por exemplo, ao exibir dados sobre licitações, pode ser interessante incluir informações sobre o número de participantes ou o valor médio de contratos, oferecendo um contexto adicional que ajude na interpretação.

Por fim, a atualização constante dos dados é um aspecto crucial para garantir que a plataforma de transparência pública mantenha sua relevância e utilidade. A publicação de informações deve ser realizada em tempo real ou, pelo menos, de forma periódica, permitindo que os cidadãos acompanhem as mudanças nos dados e possam analisar a evolução das ações do governo.

A aplicação desses princípios de visualização e organização de dados pode melhorar significativamente a transparência pública, tornando a informação acessível e compreensível. Ao criar interfaces intuitivas e bem projetadas, os governos podem facilitar o controle social e promover a participação ativa da sociedade na fiscalização e no desenvolvimento de políticas públicas.

Esses pontos abordam a importância da usabilidade e a boa visualização de dados como elementos essenciais para a transparência governamental. Ao adotar essas práticas, os portais de transparência pública podem se tornar mais eficazes na disseminação de informações e no fomento à participação cidadã.

2.6 Comparativo com Outras Plataformas de Transparência Pública

A transparência pública no Brasil tem evoluído ao longo dos anos com a implementação de diversas plataformas governamentais, como o Portal da Transparência, que visa garantir o direito do cidadão de acessar informações sobre a administração pública. No entanto, há uma necessidade crescente de soluções que não só disponibilizem dados, mas também proporcionem formas de análise e visualização acessíveis e intuitivas. Nesse contexto, a plataforma *Licita Aberta* se destaca, ao oferecer não apenas a divulgação de informações, mas também ferramentas avançadas para análise e visualização de dados públicos, como gráficos sobre gastos públicos, contratos e licitações. Esta seção analisará plataformas existentes, como o Portal da Transparência, e destacará as inovações da *Licita Aberta*.

2.6.1 Análise de Plataformas Existentes

O Portal da Transparência é a plataforma oficial do governo federal, com o objetivo de garantir que os cidadãos possam acompanhar a execução orçamentária e financeira do governo, incluindo despesas, repasses a estados e municípios, contratos e licitações. Lançado em 2004, o portal representa um marco importante na busca pela transparência pública no Brasil, disponibilizando informações detalhadas sobre a execução financeira, como gastos com pessoal, compras governamentais, e transferências de recursos. Embora a plataforma seja extensa e contenha uma grande quantidade de dados, seu formato de apresentação nem sempre é intuitivo, e a análise desses dados exige um certo conhecimento técnico.

Outras plataformas de transparência pública, como o *Portal da Transparência Estadual* e o *Portal da Transparência Municipal*, replicam a estrutura e a lógica do portal federal, permitindo que os cidadãos tenham acesso a informações sobre os gastos e licitações realizadas no nível estadual e municipal. No entanto, essas plataformas variam em termos de usabilidade, profundidade das informações disponibilizadas e atualizações, com algumas sendo mais completas e outras com dados mais restritos. Um dos principais desafios dessas plataformas é a dificuldade de navegar entre as diferentes seções e a complexidade na interpretação dos dados apresentados. Além disso, a falta de recursos interativos e visuais, como gráficos e relatórios dinâmicos, torna o processo de compreensão das informações um desafio para muitos cidadãos.

A comparação entre essas plataformas revela uma lacuna significativa em termos de usabilidade e de recursos que facilitem a análise dos dados de forma acessível ao público geral. Isso demonstra a necessidade de uma solução que, além de disponibilizar as informações de maneira mais simples, também ofereça funcionalidades que permitam aos usuários realizar análises dinâmicas de forma autônoma.

2.6.2 Diferenciais e Inovações da Plataforma Proposta

A plataforma *Licita Aberta* propõe uma abordagem inovadora para a transparência pública, indo além da simples disponibilização de dados. Ao focar na análise e visualização dos dados governamentais, a plataforma busca não apenas fornecer informações sobre licitações, contratos e gastos públicos, mas também facilitar a compreensão dessas informações por parte do público em geral. Os principais diferenciais da *Licita Aberta* estão na sua interface amigável e na capacidade de gerar análises dinâmicas por meio de gráficos interativos.

Uma das principais inovações da *Licita Aberta* é o uso de gráficos e visualizações que permitem aos usuários entender rapidamente os padrões de gastos públicos, licitações realizadas, e contratos firmados com o governo. Diferentemente de outras plataformas, que apresentam os dados em formatos estáticos, a *Licita Aberta* oferece uma maneira de interagir com os dados, possibilitando a visualização de informações por diferentes perspectivas, como por ano, tipo de licitação, órgão responsável, entre outros. Isso permite que o usuário obtenha uma visão mais clara e abrangente das práticas de contratação pública.

Além disso, a plataforma se destaca por sua usabilidade. A interface foi projetada para ser simples e intuitiva, permitindo que até mesmo usuários sem experiência prévia em análise de dados possam navegar com facilidade. Funcionalidades como filtros dinâmicos, relatórios personalizáveis e dashboards interativos tornam a *Licita Aberta* uma ferramenta poderosa e acessível, capaz de atender desde o cidadão comum até especialistas em controle social e pesquisadores.

Outro diferencial importante da plataforma é sua abordagem focada em licitações. A *Licita Aberta* permite que o usuário visualize, analise e compare licitações e contratos em um único local, com a possibilidade de identificar padrões de gastos, irregularidades e outros aspectos que possam indicar falhas na gestão pública. Ao combinar dados de diferentes fontes e apresentá-los de forma estruturada e visual, a plataforma possibilita um novo nível de análise de transparência pública.

A utilização de tecnologia de raspagem de dados e APIs governamentais para coletar e integrar informações também representa um avanço significativo. Isso permite que a plataforma esteja sempre atualizada e ofereça dados em tempo real, o que é um grande benefício em relação às plataformas existentes, que frequentemente têm problemas com a atualização de informações. Além disso, a *Licita Aberta* oferece funcionalidades de relatórios automatizados, economizando tempo para cidadãos e órgãos de fiscalização.

Em resumo, a plataforma *Licita Aberta* se diferencia por seu compromisso com a usabilidade, a visualização dinâmica de dados e a análise aprofundada das informações sobre licitações e contratos públicos. Ao integrar essas funcionalidades em uma interface acessível, ela tem o potencial de transformar a forma como os cidadãos interagem com as informações públicas e de promover uma maior transparência nas práticas governamentais.

3. METODOLOGIA

Este capítulo apresenta as metodologias utilizadas. Cada etapa é discutida em detalhe, evidenciando os principais achados, os desafios enfrentados e os métodos empregados para alcançar os objetivos propostos. Com o suporte de diagramas, tabelas e análises, são abordados desde o levantamento de requisitos até a implementação e validação da plataforma proposta.

3.1 Tipo de Pesquisa

Este trabalho caracteriza-se como uma pesquisa aplicada, pois busca solucionar o problema real da falta de acessibilidade e transparência nas informações sobre licitações públicas. Além disso, a abordagem adotada foi exploratória, pois envolveu investigação das práticas de publicação em diários oficiais, a análise de plataformas existentes e a criação de soluções inovadoras.

A pesquisa também teve um caráter de desenvolvimento tecnológico, utilizando ferramentas computacionais para implementação do sistema e validação dos resultados.

3.2 Procedimentos Metodológicos

Os procedimentos metodológicos adotados foram organizados em etapas sequenciais, com o objetivo de garantir a sistematização do desenvolvimento e a apresentação clara dos resultados obtidos em cada fase. A seguir, são descritas essas etapas, acompanhadas pelos principais resultados e discussões.

3.2.1 Levantamento de Requisitos e Seleção das Fontes de Dados

O primeiro passo no desenvolvimento da plataforma foi o levantamento dos requisitos, com foco na identificação das necessidades funcionais e não funcionais. O objetivo principal era criar uma ferramenta capaz de visualizar o valor total gasto por ano em licitações públicas, com a possibilidade de detalhamento por município e mês.

Para atingir esse objetivo, a plataforma foi estruturada para extrair dados de diários oficiais municipais, especificamente os publicados pela Federação das Associações de Municípios da Paraíba (FAMUP). Optamos por utilizar as publicações da FAMUP devido à sua abrangência, que contempla a maior parte dos municípios da Paraíba. Embora alguns municípios importantes, como João Pessoa, estejam ausentes, a escolha por essa fonte foi fundamentada na decisão de abranger uma maior quantidade de municípios, permitindo uma análise mais representativa dos gastos públicos no estado.

A escolha da API do Querido Diário para o acesso aos diários oficiais foi motivada pela sua facilidade de uso e pela agilidade no acesso aos dados. Diferentemente de outras fontes de dados, como os próprios sites das prefeituras ou sistemas municipais independentes, que podem exigir métodos mais complexos de scraping ou acessar informações

desestruturadas, a API do Querido Diário oferece uma integração direta com os diários oficiais em formato PDF, simplificando o processo de coleta e garantindo a consistência dos dados. A uniformidade na estrutura desses diários facilitou a implementação de um sistema organizado e transparente, essencial para o monitoramento dos gastos públicos.

Os requisitos foram definidos em duas categorias principais:

Funcionais:

- Exibir o valor total gasto em licitações por ano.
- Permitir a seleção de ano e cidade para exibição dos gastos por mês.
- Oferecer gráficos interativos para uma visualização clara dos dados.

Não Funcionais:

- Garantir uma interface amigável e intuitiva.
- Escalabilidade do sistema para lidar com maiores volumes de dados.
- Prover eficiência e alta disponibilidade na API de coleta de dados.

3.2.2 Análise da Plataforma Exoonero

A plataforma Exoonero.org foi utilizada como ponto de partida para o desenvolvimento do sistema, servindo de inspiração para a lógica de extração e apresentação de dados. O Exoonero trabalha com diários oficiais para exibir informações sobre exonerações e nomeações, e sua estrutura lógica foi analisada como referência inicial.

No entanto, o projeto foi adaptado para atender às especificidades da análise de licitações públicas. A principal diferença está no foco: enquanto o Exoonero trabalha com exonerações e nomeações, este projeto direciona seus esforços para gastos públicos, exigindo ajustes na estrutura de código e nos métodos de extração de dados. Aproveitamos a mesma abordagem utilizada pelo Exoonero para a extração dos diários oficiais e a conversão dos arquivos de PDF para TXT, garantindo um fluxo de processamento eficiente. Entretanto, modificamos a lógica de extração para capturar especificamente informações sobre licitações, como valores, empresas envolvidas e objetos dos contratos.

Essa análise destacou oportunidades de inovação e adaptação, como a necessidade de lidar com diferentes padrões textuais presentes nos diários e o desenvolvimento de uma

lógica específica para extrair informações monetárias. Embora tenha servido como base, o sistema desenvolvido neste projeto é único em seu objetivo e abordagem.

Além da análise conceitual, elaboramos uma tabela e **diagrama de casos de uso** comparando as funcionalidades do Exoonero e da nossa plataforma. O quadro e a figura 2 abaixo ilustra as diferenças entre os dois sistemas:

Licita Aberta	**Exoonero**
Dados de licitações	Dados de nomeações e exonerações
Gráficos com quantidades e valores das licitações	Gráficos focados em quantidades de exonerações
Tabela agrupando mensalmente contratos e respectivas empresas associados às licitações	-

Quadro 1 - Diferenças entre plataformas Licita Aberta e Exoonero

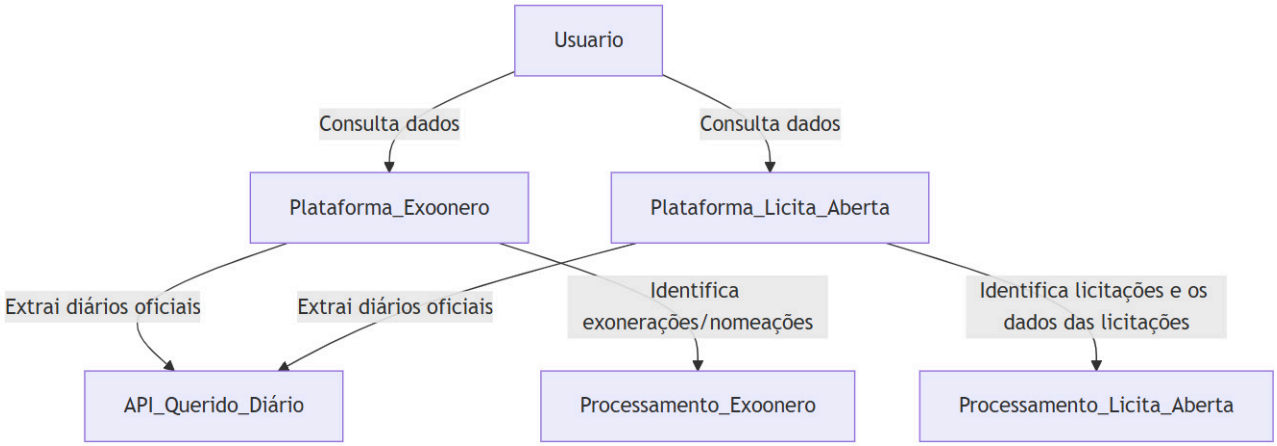


Figura 2 – Diagrama de Casos de Uso: Comparação entre Exoonero e Licita Aberta
Fonte: Autoria Própria.

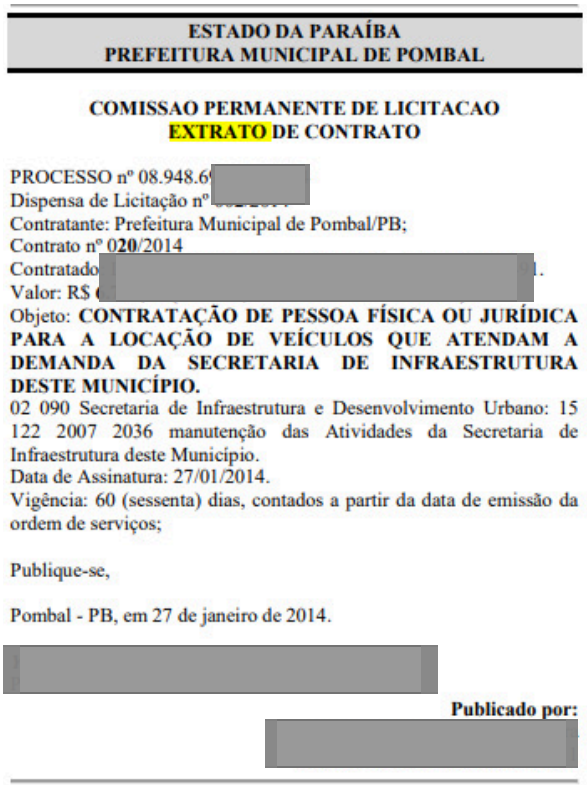
3.2.3 Identificação das Licitações nos Diários Oficiais

Com as fontes de dados e os requisitos definidos, iniciou-se a análise dos padrões textuais nos diários oficiais da Federação das Associações de Municípios da Paraíba (FAMUP) para identificar licitações e contratos. Essa etapa foi essencial para garantir que o sistema fosse capaz de localizar as informações de interesse de forma eficiente.

A análise incluiu a identificação de termos-chave como “extrato de contrato” e “termo aditivo do contrato”, além de suas variações linguísticas. Durante essa fase, foi necessário realizar uma análise inicial manual, realizando busca pelos termos nos arquivos pdf e txt nos documentos publicados entre 2014 e 2024, abrangendo meses e anos distintos. Esse processo foi crucial para entender como os registros de licitações e contratos estavam organizados. Descobriu-se que, apesar da união dos arquivos publicados pelas cidades, havia inconsistências na forma como as informações eram apresentadas nos textos. Por exemplo, termos como “extrato de contrato” ou “extrato do contrato”, assim como valores monetários descritos como “valor total” e “valor global de R$”, exigiam adaptações no processamento automatizado.

Os padrões foram identificados por meio de análises manuais de diversos documentos e testados utilizando expressões regulares (regex). Esses padrões foram continuamente comparados com os resultados esperados para validar a eficiência e precisão na extração das informações.

Logo veremos, 2 exemplos do formato que os contratos são apresentados:



ESTADO DA PARAÍBA
PREFEITURA MUNICIPAL DE POMBAL

COMISSAO PERMANENTE DE LICITACAO
EXTRATO DE CONTRATO

PROCESSO nº 08.948.6[redacted]
Dispensa de Licitação nº [redacted]
Contratante: Prefeitura Municipal de Pombal/PB;
Contrato nº 020/2014
Contratado [redacted].
Valor: R$ [redacted]
Objeto: **CONTRATAÇÃO DE PESSOA FÍSICA OU JURÍDICA PARA A LOCAÇÃO DE VEÍCULOS QUE ATENDAM A DEMANDA DA SECRETARIA DE INFRAESTRUTURA DESTE MUNICÍPIO.**
02 090 Secretaria de Infraestrutura e Desenvolvimento Urbano: 15 122 2007 2036 manutenção das Atividades da Secretaria de Infraestrutura deste Município.
Data de Assinatura: 27/01/2014.
Vigência: 60 (sessenta) dias, contados a partir da data de emissão da ordem de serviços;

Publique-se,

Pombal - PB, em 27 de janeiro de 2014.

[redacted]

Publicado por:
[redacted]

Figura 3 – Extrato de contrato do município de Pombal (2014).
(Fonte: Diário Oficial dos Municípios da Paraíba, publicado pela Federação das Associações dos Municípios da Paraíba (2014))

ESTADO DA PARAÍBA
PREFEITURA MUNICIPAL DE CABACEIRAS

SECRETÁRIA DE ADMINISTRAÇÃO
EXTRATO DO CONTRATO Nº 0064/2017.

OBJETO: Contratação de empresa de construção civil, visando à pavimentação e drenagem na Rua José Osmario de Lima na Zona Urbana do município de Cabaceiras PB, conforme detalhamento no Anexo I do Edital.
FUNDAMENTO LEGAL: Processo nº 0023.0023/2017 - Tomada de Preços nº 0001/2017.
DOTAÇÕES: Os recursos financeiros para execução dos serviços correrão por conta do convênio com o MINISTÉRIO DAS CIDADES - MCIDADES - Termo de Compromisso nº 787036/2013 - Processo nº 2641.1006379-43/2013 e Recursos Próprios do 02.801 - DEPARTAMENTO DE OBRAS E SERVIÇOS PÚBLICOS - 02.801. - 15.451.1003.1.015 - IMPLANTAÇÃO DE PAVIMENTAÇÃO EM PARALELEPIPEDO E ASFÁLTICA.
ELEMENTO: Elemento de Despesas - 4.4.90.51.99.000 - Obras e Instalações no orçamento operativo de 2017 e nos exercício

Figura 4 – Extrato de contrato do município de Cabaceiras (2017).
(Fonte: Diário Oficial dos Municípios da Paraíba, publicado pela Federação das Associações dos Municípios da Paraíba (2017))

Conforme observado nas Figuras 3 e 4, o padrão dos extratos de contrato se diferenciam dependendo do município, mês ou ano. Além disso, foram mapeados outros padrões importantes, como campos relacionados a valores monetários, objetos das licitações e partes contratadas. Essa análise manual e iterativa foi determinante para o desenvolvimento de um sistema de extração eficiente e adaptado aos dados disponíveis.

3.2.4 Raspagem e Processamento de Dados:

A etapa de raspagem de dados deste trabalho foi realizada por meio da **API do Querido Diário**, uma plataforma que visa tornar as informações dos diários oficiais do poder executivo municipal brasileiro acessíveis de forma aberta e de fácil interpretação. O projeto tem como objetivo principal facilitar o acompanhamento, a análise e a conexão entre os dados divulgados, promovendo a transparência e possibilitando o controle social, com foco na melhoria da gestão pública. Ele se destaca pela disponibilização dos dados por meio de uma plataforma de visualização intuitiva e de uma **API pública**, que possibilita a integração com outros sistemas e serviços.

Por meio desta API, é possível realizar buscas específicas por termos nos diários oficiais e acessar documentos de diferentes períodos. Esses documentos são disponibilizados

em formatos como **TXT** e **PDF**, permitindo que sejam analisados e utilizados em diversos contextos de forma flexível.

No contexto deste trabalho, a API foi utilizada para a coleta dos diários oficiais da Federação das associações de municípios da Paraíba (**FAMUP)**. Os arquivos obtidos em formato **PDF** foram processados com o auxílio do **Apache Tika**, ferramenta responsável pela extração do texto desses arquivos para o formato **TXT**. A transformação para texto puro foi essencial para que o conteúdo fosse uniformizado e pudesse ser manipulado de forma eficiente durante a análise.

Após a extração do conteúdo, os arquivos **TXT** passaram por um processo de **limpeza e uniformização**. Esse processo garantiu que as informações relevantes fossem extraídas de forma consistente, permitindo a identificação precisa de dados como contratos, valores, datas e as partes envolvidas nos atos administrativos. A padronização do texto foi crucial para eliminar ruídos e garantir a qualidade dos dados extraídos.

A extração das informações foi realizada utilizando **expressões regulares (regex)**, que possibilitaram a identificação de padrões específicos, como valores monetários, datas e detalhes sobre as partes envolvidas nos contratos. Além disso, um **tratamento de erros** foi implementado para lidar com possíveis inconsistências nos textos, assegurando que o processo de análise fosse robusto e sem falhas.

Para ilustrar o processo de raspagem e extração de dados, um diagrama foi desenvolvido, descrevendo o fluxo de dados desde a requisição à API até a extração e armazenamento dos dados relevantes. Esse diagrama oferece uma visão clara e detalhada das etapas envolvidas, facilitando a compreensão da integração entre a API, o processamento dos documentos e os métodos utilizados para o tratamento e análise dos dados.

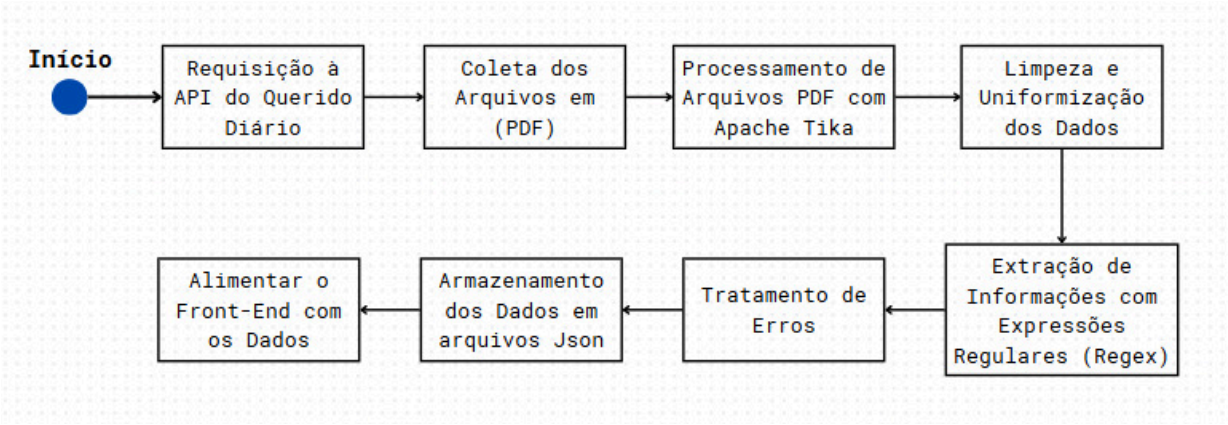


Figura 5 – Fluxo de processamento da plataforma
(Fonte: Autoria própria)

3.2.5 Desenvolvimento da Plataforma

A plataforma foi dividida em duas partes principais, **Back-end** e **Front-end**, cada uma com responsabilidades específicas no fluxo de processamento, organização e apresentação dos dados extraídos. Abaixo, descreve-se detalhadamente cada componente, suas funcionalidades e integrações. Um diagrama foi adicionado para ilustrar a relação entre as partes.

Arquitetura Geral da Plataforma:

A plataforma é composta por três camadas principais:

1. **Coleta de Dados**: Scripts responsáveis por obter diários oficiais a partir da API do Querido Diário.
2. **Processamento de Dados**: Extração e estruturação das informações relevantes dos arquivos coletados.
3. **Apresentação de Dados**: Interface web que permite a visualização e análise interativa dos resultados extraídos.

Descrição dos Componentes do Back-end:

O back-end foi desenvolvido para lidar com a coleta, extração, organização e armazenamento dos dados dos diários oficiais. Ele se divide em três módulos principais:

1. Script de Coleta de Dados (coleta_diarios.sh)

- **Função**: Realiza a integração com a API do Querido Diário e faz o download dos diários oficiais dentro de um intervalo de tempo especificado (01/01/2014) a (31/12/2024).
- **Descrição do Processo**:
  - Configura o ambiente, incluindo a instalação de dependências necessárias.
  - Executa o scrapy para baixar os diários oficiais.
  - Organiza os arquivos em diretórios específicos e nomeia-os de forma padronizada para facilitar o processamento posterior.
  - Utiliza o Apache Tika (via Docker) para converter arquivos PDF em texto, possibilitando sua análise textual.

2.Script de Extração de Dados (extrair_diarios.py)

- **Função**: Processa os arquivos de texto extraídos, identifica os municípios e seus respectivos diários, e estrutura as informações em objetos manipuláveis pelo sistema.
- **Descrição do Processo**:
  - Carrega os textos extraídos dos diários e identifica padrões relevantes (como nomes de municípios e datas de publicação) usando expressões regulares.
  - Cria objetos representativos, como Diario (diário oficial) e Municipio (entidade responsável pela publicação).
  - Gera arquivos estruturados contendo os dados processados e um resumo em formato JSON para facilitar a análise futura.

3.Módulo de Processamento (diario_municipal.py e diario_pb.py)

- **Função**: Contém as classes e funções que implementam a lógica de identificação e extração de informações específicas dos diários oficiais.
- **Descrição das Classes**:
  - Municipio: Representa os municípios identificados no texto dos diários. Realiza a normalização dos nomes e cria um identificador único.
  - Diario: Representa um diário oficial, incluindo informações como município, cabeçalho, texto completo e data de publicação. Possui

métodos para extrair dados específicos, como a data de publicação a partir do cabeçalho.

Descrição dos Componentes do Front-end

O front-end foi projetado para exibir os dados extraídos de maneira clara, interativa e organizada, permitindo que os usuários explorem os resultados de forma intuitiva. Foi implementado no repositório “licita.github.io” e utiliza tecnologias web modernas.

Estrutura da Interface:

- **Página Inicial**: Apresenta uma visão geral dos dados extraídos, com gráficos e tabelas interativas que destacam informações relevantes, como contratos e termos aditivos.
- **Filtros e Pesquisa**: Permite a navegação por ano, município ou tipo de documento, possibilitando a análise segmentada dos dados.
- **Visualização Detalhada**: Cada registro pode ser expandido para exibir informações detalhadas, como datas, valores e empresas envolvidas.

Tecnologias Utilizadas:

- **HTML/CSS/TypeScript**: Para estruturar e estilizar a interface.
- **Bibliotecas de Gráficos**: Ferramentas como Chart.js foram utilizadas para gerar visualizações interativas e dinâmicas.
- **Hospedagem no GitHub Pages**: Permite que a interface seja acessada publicamente, facilitando a disseminação das informações.

3.2.6 Design da Interface e Usabilidade

O design da interface priorizou a usabilidade e a acessibilidade. Foram utilizadas boas práticas de design, como contraste de cores, tipografia legível e organização hierárquica das informações. O foco foi tornar a navegação simples, permitindo que qualquer usuário localize rapidamente os dados desejados, com destaque para os valores apresentados.

3.2.7 Validação da Plataforma

Foram realizados testes tanto manuais quanto automatizados para garantir a precisão da extração de dados dos diários oficiais. A suíte de testes automatizados foi implementada

através do arquivo integracao_test.py utilizando o framework `unittest`, com foco principal na função test_extrai_diarios. Nesse teste, arquivos JSON representando diários municipais são carregados e comparados com os resultados extraídos pela função extrair_diarios_municipais. A validação é feita com base em municípios, cabeçalhos e atos extraídos, verificando se os dados extraídos coincidem com os valores esperados.

Além dos testes automatizados, também foram realizados testes manuais para garantir que situações específicas fossem corretamente tratadas, principalmente em relação à variabilidade dos dados. Esses testes manuais ajudaram a verificar o comportamento do sistema em cenários não contemplados pelos testes automatizados.

3.3 Ferramentas Utilizadas

As ferramentas principais empregadas no projeto foram:

- **Scrapy**: O projeto utiliza o framework Scrapy para coletar os diários oficiais, reaproveitando a spider pb_federacao_municipios do repositório Querido Diário no GitHub.

Figura 6 – Logo do Scrapy
Fonte: Scrapy. Disponível em: https://scrapy.org. Acesso em: *14/03/2025*

- **Apache Tika**: Framework foi utilizado para extrair texto dos diários oficiais em PDF, permitindo a conversão automática para texto simples(txt).

Figura 7 – Logo do Apache Tika

Fonte: Apache Tika. Disponível em:https://en.wikipedia.org/wiki/Apache_Tika. Acesso em: *14/03/2025*

- **Python:** O Python foi utilizado para processar e estruturar os dados extraídos dos diários oficiais. Scripts em Python aplicaram expressões regulares (Regex) para identificar e organizar o conteúdo em arquivos.

Figura 8 – Logo da linguagem Python
Fonte: Python. Disponível em:https:https://www.python.org/community/logos/. Acesso em: *14/03/2025*

- **Bibliotecas python:** Pandas, Matplotlib e Requests, para manipulação de dados e criação de visualizações.
- **API e repositório do Querido Diário:** Fonte de dados dos diários oficiais.

Figura 9 – Logo da plataforma do Querido Diário
Fonte: Querido Diário. Disponível em:https:**https://queridodiario.ok.org.br/**.Acesso em: *14/03/2025*

- **HTML, CSS e TypeScript:** Tecnologias para desenvolvimento do front-end.

Figura 10 – Logo do Apache Tika
Fonte: Apex ensino. Disponível em:**https://apexensino.com.br/base-da-programacao-front-end/.** Acesso em: *14/03/2025*

- **GitHub:** Repositório de códigos e controle de versão.

Figura 11 – Logo do GitHub
Fonte: GitHub, Inc. Disponível em: https://github.com/logos. Acesso em: *14/03/2025*

3.4 Limitações da Pesquisa

A pesquisa enfrentou limitações relacionadas à ausência de um padrão consistente nos registros de contratos publicados nos diários oficiais, casos em que não há registro de diários publicados para uma determinada data ou periodo. Essa falta de uniformidade dificultou a identificação automatizada das informações, tendo em vista que a identificação das licitações é feita por meio da busca por termos específicos que remetem à esses processos, o que acabou servindo como critério na seleção das cidades que que disponibilizam diários oficiais em um formato padronizado e contém os termos-chave para designação dos processos.

Apesar desses desafios, as soluções desenvolvidas acabaram por apresentar mais cidades do que inicialmente foi planejado, assim como garantiu a confiabilidade dos resultados.

3.5 Validação dos Dados e Resultados

A validação dos dados extraídos foi realizada por meio de verificações cruzadas entre os arquivos capturados e as fontes originais disponíveis nos diários oficiais. Essa etapa permitiu identificar a qualidade das informações processadas e detectar inconsistências no processo de raspagem. No entanto, ainda foram observadas falhas em algumas extrações, indicando a necessidade de refinamento nos métodos utilizados. Além disso, foram conduzidos testes manuais em subconjuntos de dados para avaliar a completude e a confiabilidade das informações.

A validação dos resultados também incluiu a análise de gráficos gerados pela plataforma, evidenciando padrões alinhados aos objetivos do projeto. Embora os resultados tenham apresentado coerência com as expectativas iniciais, ainda há margem para aprimorar a precisão das informações extraídas, garantindo maior fidelidade aos dados originais.

3.6 Divisão de Tarefas

A implementação do Licita Aberta foi dividida entre as duas integrantes do projeto da seguinte forma:

- **Amanda Ferreira da Silva Alves:** Responsável pelo desenvolvimento do backend, incluindo a coleta, extração e processamento dos dados.
- **Sara Aymê Marinho Gaspar:** Responsável pela implementação do frontend, incluindo a interface de visualização e organização dos dados extraídos.

Embora cada integrante tivesse atribuições definidas, a colaboração foi essencial para o sucesso do projeto. Sempre que surgiam dúvidas ou desafios técnicos, ambas contribuíam mutuamente, garantindo a integração eficiente entre o backend e o frontend.

3.7 Sustentabilidade e Expansão da Plataforma

A sustentabilidade da plataforma foi planejada com base na modularidade do código, facilitando futuras atualizações e manutenções. Além disso, o uso de ferramentas amplamente documentadas, como Python e bibliotecas populares, garante que o projeto possa ser expandido por outros desenvolvedores.

A plataforma pode ser adaptada para abranger outros tipos de dados governamentais, como contratos de prestação de serviços ou execuções orçamentárias. Essa possibilidade

amplia o impacto do projeto, promovendo maior transparência em outras áreas da administração pública**.**

3.8 Ética e Privacidade na Coleta de Dados

O projeto foi desenvolvido em conformidade com as normas éticas de coleta e uso de dados públicos. Todos os dados utilizados eram de acesso público, conforme garantido pela Lei de Acesso à Informação (Lei n.º 12.527/2011).

Adicionalmente, medidas foram tomadas para garantir que os dados fossem processados de forma anônima, evitando a exposição de informações sensíveis. Essa abordagem reforça o compromisso com a transparência responsável e a proteção dos direitos dos cidadãos.

4. RESULTADOS E DISCUSSÕES

O principal resultado deste trabalho foi o desenvolvimento de uma plataforma funcional para a extração e apresentação de informações sobre licitações públicas a partir de diários oficiais municipais. Este resultado reflete a integração dos processos metodológicos descritos anteriormente e a aplicação das ferramentas e técnicas propostas. A seguir, são apresentados os principais aspectos do projeto concluído.

4.1 Funcionalidades da Plataforma

A plataforma desenvolvida conta com as seguintes funcionalidades principais:

- Raspagem de Dados: Capacidade de extrair informações relevantes de diários oficiais em formato PDF, com foco em licitações públicas.
- Organização dos Dados: Os dados raspados são estruturados e categorizados para facilitar o entendimento e a navegação pelos usuários.
- Interface Amigável: A plataforma apresenta uma interface simples e intuitiva, que permite aos usuários buscar e visualizar os valores das licitações de forma clara e acessível.

Atualmente, a plataforma expõe dados de **109 municípios**, cerca de aproximadamente 48,88% dos municípios da Paraíba, permitindo uma análise abrangente das licitações

municipais e facilitando o acesso às informações por parte da população e de empresas interessadas.

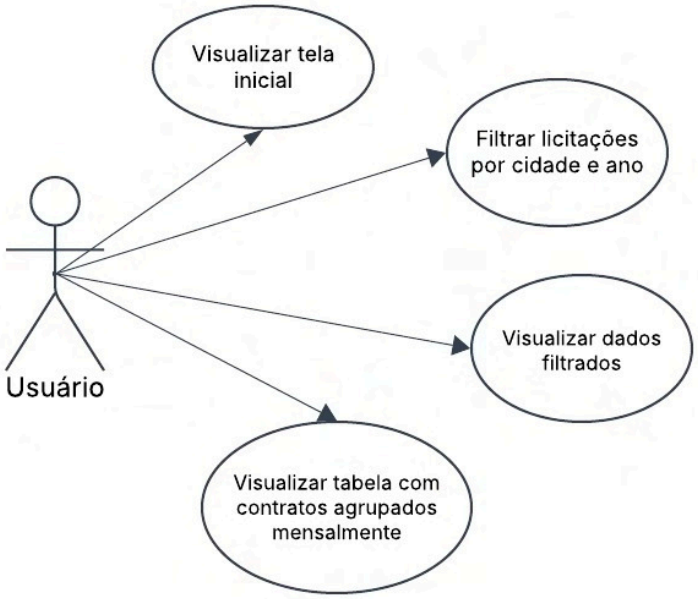


Figura 12 - Diagrama de caso de uso do sistema
Fonte: Autoria Própria (2025).

4.2 Telas da Plataforma

A tela inicial (Figura 13) é a primeira parte do sistema acessada pelo usuário, com o intuito de ser uma plataforma intuitiva já apresentamos a logo, um breve texto explicativo sobre a problemática tratada e a principal funcionalidade do serviço: gráficos de licitações somando as informações gerais que são disponibilizadas por todas as cidades.

Nas figuras 14 e 15 podemos ver quais são os filtros disponíveis para busca específica por períodos nas respectivas cidades. Fechamos a apresentação do principal foco da plataforma com as figura 16 e 17 que traz os dados já filtrados através das opções selecionadas, na primeira seguimos com os gráficos apresentando quantidades e valores gastos, enquanto na segunda há uma tabela agrupando mensalmente os serviços e/ou ferramentais contratados identificados como objetos das licitações e suas respectivas empresas associadas.

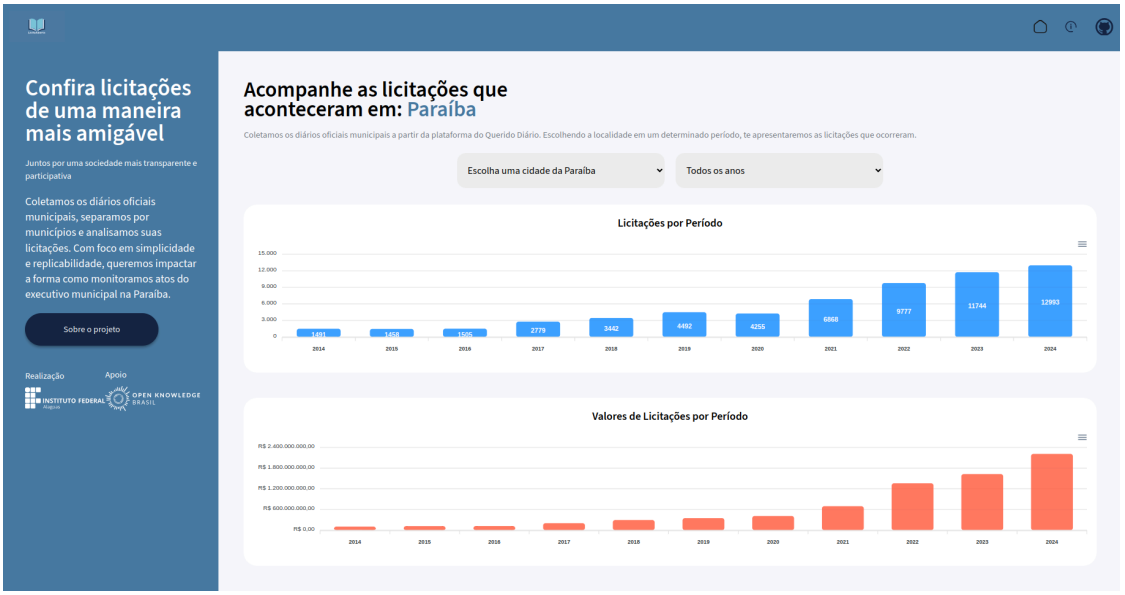


Figura 13 – Tela inicial da plataforma

(Fonte: Autoria própria)

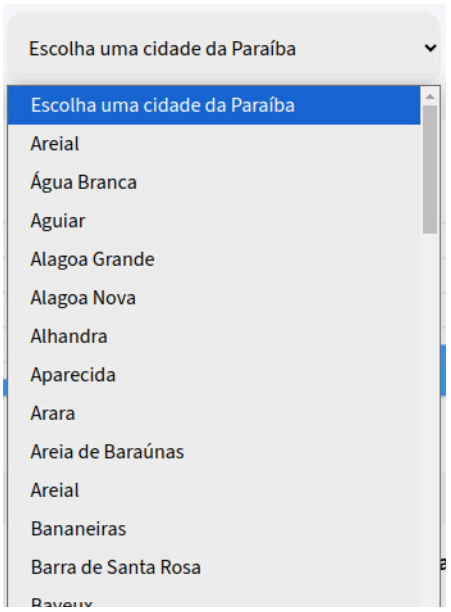


Figura 14 - Filtro por cidades

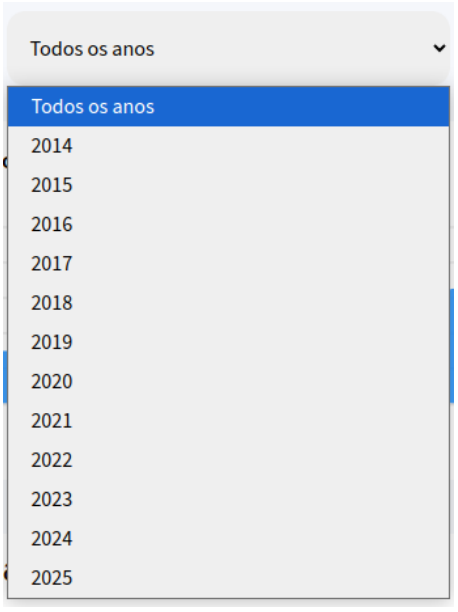
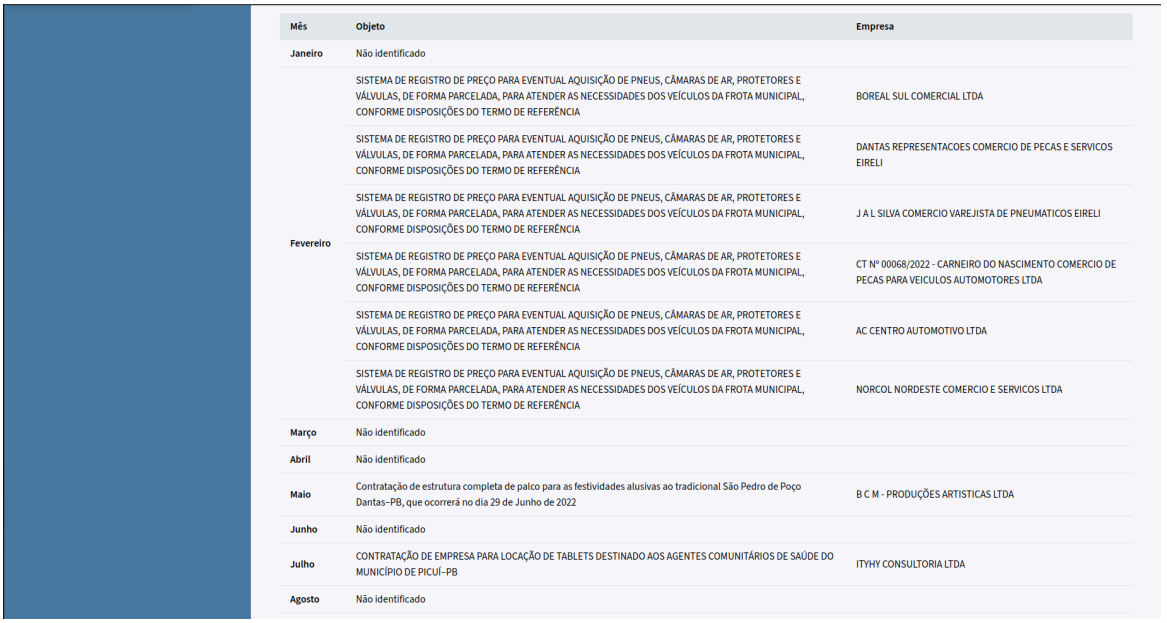


Figura 15 - Filtro por anos

(Fonte: Autoria própria)

Figura 16 - Tela com exposição de valores filtrados para cidade Picuí no ano de 2022

(Fonte: Autoria própria)



Mês	Objeto	Empresa
Janeiro	Não identificado	
Fevereiro	SISTEMA DE REGISTRO DE PREÇO PARA EVENTUAL AQUISIÇÃO DE PNEUS, CÂMARAS DE AR, PROTETORES E VÁLVULAS, DE FORMA PARCELADA, PARA ATENDER AS NECESSIDADES DOS VEÍCULOS DA FROTA MUNICIPAL, CONFORME DISPOSIÇÕES DO TERMO DE REFERÊNCIA	BOREAL SUL COMERCIAL LTDA
	SISTEMA DE REGISTRO DE PREÇO PARA EVENTUAL AQUISIÇÃO DE PNEUS, CÂMARAS DE AR, PROTETORES E VÁLVULAS, DE FORMA PARCELADA, PARA ATENDER AS NECESSIDADES DOS VEÍCULOS DA FROTA MUNICIPAL, CONFORME DISPOSIÇÕES DO TERMO DE REFERÊNCIA	DANTAS REPRESENTACOES COMERCIO DE PECAS E SERVICOS EIRELI
	SISTEMA DE REGISTRO DE PREÇO PARA EVENTUAL AQUISIÇÃO DE PNEUS, CÂMARAS DE AR, PROTETORES E VÁLVULAS, DE FORMA PARCELADA, PARA ATENDER AS NECESSIDADES DOS VEÍCULOS DA FROTA MUNICIPAL, CONFORME DISPOSIÇÕES DO TERMO DE REFERÊNCIA	J A L SILVA COMERCIO VAREJISTA DE PNEUMATICOS EIRELI
	SISTEMA DE REGISTRO DE PREÇO PARA EVENTUAL AQUISIÇÃO DE PNEUS, CÂMARAS DE AR, PROTETORES E VÁLVULAS, DE FORMA PARCELADA, PARA ATENDER AS NECESSIDADES DOS VEÍCULOS DA FROTA MUNICIPAL, CONFORME DISPOSIÇÕES DO TERMO DE REFERÊNCIA	CT Nº 00068/2022 - CARNEIRO DO NASCIMENTO COMERCIO DE PECAS PARA VEICULOS AUTOMOTORES LTDA
	SISTEMA DE REGISTRO DE PREÇO PARA EVENTUAL AQUISIÇÃO DE PNEUS, CÂMARAS DE AR, PROTETORES E VÁLVULAS, DE FORMA PARCELADA, PARA ATENDER AS NECESSIDADES DOS VEÍCULOS DA FROTA MUNICIPAL, CONFORME DISPOSIÇÕES DO TERMO DE REFERÊNCIA	AC CENTRO AUTOMOTIVO LTDA
	SISTEMA DE REGISTRO DE PREÇO PARA EVENTUAL AQUISIÇÃO DE PNEUS, CÂMARAS DE AR, PROTETORES E VÁLVULAS, DE FORMA PARCELADA, PARA ATENDER AS NECESSIDADES DOS VEÍCULOS DA FROTA MUNICIPAL, CONFORME DISPOSIÇÕES DO TERMO DE REFERÊNCIA	NORCOL NORDESTE COMERCIO E SERVICOS LTDA
Março	Não identificado	
Abril	Não identificado	
Maio	Contratação de estrutura completa de palco para as festividades alusivas ao tradicional São Pedro de Poço Dantas-PB, que ocorrerá no dia 29 de Junho de 2022	B C M - PRODUÇÕES ARTISTICAS LTDA
Junho	Não identificado	
Julho	CONTRATAÇÃO DE EMPRESA PARA LOCAÇÃO DE TABLETS DESTINADO AOS AGENTES COMUNITÁRIOS DE SAÚDE DO MUNICÍPIO DE PICUÍ-PB	ITYHY CONSULTORIA LTDA
Agosto	Não identificado	

Figura 17 - Tela de tabela com objetos identificados referentes à cidade Picuí no ano de 2022

(Fonte: Autoria própria)

4.3 Repositórios no GitHub

Para garantir a acessibilidade e a continuidade do projeto, foram criados dois repositórios no GitHub:

- licita.github.io: Scripts de raspagem e processamento de dados.
- extrator_licita: Código do front-end e integração com a interface.

Os repositórios foram disponibilizados em formato público e contêm a documentação necessária para que outros desenvolvedores possam compreender e, eventualmente, contribuir com o projeto.

4.4 Impacto e Perspectivas

Os resultados obtidos com a plataforma demonstram seu potencial para tornar as informações sobre licitações públicas mais acessíveis e transparentes. Além disso, a utilização de ferramentas de código aberto, aliada à disponibilização dos repositórios no GitHub, contribui para a replicabilidade e o aprimoramento contínuo do projeto.

4. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Este trabalho apresentou o desenvolvimento de uma plataforma voltada para a extração, organização e apresentação de informações sobre licitações públicas a partir de diários oficiais municipais. O objetivo principal foi facilitar o acesso a esses dados, promovendo maior transparência e acessibilidade às informações governamentais.

A partir da análise inicial das plataformas Querido Diário e Exoonero, foi possível identificar lacunas na acessibilidade e na organização das informações relacionadas a licitações. Com base nisso, o projeto focou em aprimorar a apresentação desses dados por meio de uma interface amigável e intuitiva, bem como em otimizar o processo de extração e processamento das informações, utilizando técnicas de raspagem de dados.

Os resultados obtidos, como o desenvolvimento funcional da plataforma e a disponibilização dos repositórios no GitHub, demonstram o sucesso na implementação do projeto. A plataforma tem potencial para contribuir significativamente para o aumento da transparência pública e para a conscientização da sociedade sobre os processos licitatórios municipais.

Entretanto, algumas limitações foram observadas, como as dificuldades em lidar com variações no formato dos diários oficiais e a necessidade de um esforço contínuo para manter os dados atualizados. Esses desafios representam oportunidades para melhorias futuras.

Como trabalhos futuros, sugere-se:

- Expandir a base de municípios atendidos pela plataforma;
- Implementar ferramentas de análise e visualização de dados mais avançadas;
- Estudar formas de automatizar ainda mais a identificação de licitações nos diários;
- Explorar parcerias com instituições públicas ou privadas para aprimorar a abrangência e a eficiência da plataforma.

Conclui-se que o projeto alcançou seus objetivos iniciais e oferece uma solução prática para um problema real, promovendo maior transparência e acessibilidade às informações públicas. Espera-se que este trabalho sirva de base para futuros estudos e evoluções no campo da tecnologia aplicada à transparência governamental.

REFERÊNCIAS

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. Open government: the global context and the way forward. Paris: OECD Publishing, 2016.

TRANSPARÊNCIA BRASIL. Guia de boas práticas para portais de transparência. São Paulo: Transparência Brasil, 2022.

BRASIL. Lei n.º 12.527, de 18 de novembro de 2011. **Regula o acesso a informações previsto na Constituição Federal.** Diário Oficial da União, Brasília, DF, 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 1 nov. 2024.

DIAS, M. F.; GOMES, R. T. Desafios na acessibilidade e na organização das informações públicas em diários oficiais. **Revista de Gestão Pública**, v. 7, n. 3, p. 45-60, 2019.

OLIVEIRA, P. R.; SOUZA, J. A. Avanços e limitações da Lei de Acesso à Informação no Brasil. **Cadernos de Administração Pública**, v. 14, n. 2, p. 87-102, 2020.

SILVA, A. M. Transparência pública: fundamentos e práticas na administração pública contemporânea. **Revista Brasileira de Políticas Públicas**, v. 6, n. 1, p. 12-29, 2018.

TRANSPARÊNCIA BRASIL. Para que a transparência seja efetiva, as informações devem ser apresentadas de maneira clara e acessível, de preferência em plataformas digitais que facilitem o acesso aos dados. **Relatório Anual Transparência Brasil**, 2022. Disponível em: https://www.transparenciabrasil.org.br. Acesso em: 1 nov. 2024.

BRASIL. **Lei nº 12.527/2011, de 18 de novembro de 2011**. Lei de Acesso à Informação. Diário Oficial da União, Brasília, 18 nov. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 2 dez. 2024.

BRASIL. **Lei Complementar nº 101/2000, de 4 de maio de 2000**. Lei de Responsabilidade Fiscal. Diário Oficial da União, Brasília, 4 mai. 2000. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm. Acesso em: 2 dez. 2024.

BRASIL. **Decreto nº 8.777/2016, de 11 de maio de 2016**. Política de Dados Abertos. Diário Oficial da União, Brasília, 11 maio 2016. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8777.htm. Acesso em: 2 dez. 2024.

GOMES, S. F. et al. **Web Scraping: Fundamentos e Aplicações**. 1. ed. São Paulo: Atlas, 2019.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. **Introduction to Information Retrieval**. Cambridge: Cambridge University Press, 2008.

NIELSEN, J. **Designing Web Usability: The Practice of Simplicity**. Indianapolis: New Riders, 2000.

TESSERACT. **Tesseract OCR Documentation**. 2024. Disponível em: https://github.com/tesseract-ocr/tesseract. Acesso em: 2 dez. 2024.

WILSON, A. **Understanding OCR and Its Applications in Data Extraction**. Journal of Data Science, v. 34, n. 1, p. 35-50, 2020. Disponível em: https://doi.org/10.1234/jds.2020.0034. Acesso em: 2 dez. 2024

BRASIL. **Lei Complementar nº 131, de 27 de maio de 2009**. Altera a Lei Complementar nº 101, de 4 de maio de 2000, para estabelecer normas para a publicação de informações fiscais e orçamentárias em tempo real. Diário Oficial da União, Brasília, 2009. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp131.htm. Acesso em: 1 dez. 2024.

BRASIL. **Lei Complementar nº 101, de 4 de maio de 2000**. Lei de Responsabilidade Fiscal. Diário Oficial da União, Brasília, 2000. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm. Acesso em: 1 dez. 2024.

BRASIL. **Decreto nº 8.777, de 11 de maio de 2016**. Institui a Política de Dados Abertos no âmbito do Governo Federal e dá outras providências. Diário Oficial da União, Brasília, 2016. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2016/Decreto/D8777.htm. Acesso em: 1 dez. 2024.

CONTROLADORIA-GERAL DA UNIÃO (CGU). **Portal da Transparência**. Disponível em: https://www.portaltransparencia.gov.br/. Acesso em: 1 dez. 2024.

SANTOS, L. B. **A Importância das APIs e Dados Abertos na Transparência Governamental**. Rio de Janeiro: Editora Transparente, 2020.

FERNANDES, J. R. **Tecnologias de Raspagem de Dados e Sua Aplicação no Controle Social**. São Paulo: Editora Pública, 2019.

Querido Diário. **API Querido Diário**. Disponível em: https://www.queridodiario.org.br/. Acesso em: 1 dez. 2024.

SHNEIDERMAN, Ben; PLEMMONS, Donald L. ***Designing the User Interface: Strategies for Effective Human-Computer Interaction*.** 5. ed. Boston: Pearson, 2010.

TUFTE, Edward R. ***The Visual Display of Quantitative Information***. 2. ed. Cheshire: Graphics Press, 2001.

NORMAN, Donald A. ***The Design of Everyday Things*.** 1. ed. New York: Basic Books, 2013.

BERNHOFER, Stephan. ***Data Visualization: A Guide to Visual Storytelling for Libraries***. Chicago: ALA Editions, 2020.

CHI, Ed H. ***Usability Engineering*.** San Francisco: Morgan Kaufmann, 1997.

EVANS, James A. ***Data Visualization: A Practical Introduction*.** 1. ed. Princeton: Princeton University Press, 2020.

RICO, Lucia et al. **"Designing Transparent Government: Data Representation and Public Policy."** *Journal of Digital Government*, v. 23, n. 1, p. 45-61, 2019.

MUNZNER, Tamara. ***Visualization Analysis and Design***. Boca Raton: CRC Press, 2014.

WICKHAM, Hadley. ***ggplot2: Elegant Graphics for Data Analysis*.** 2. ed. New York: Springer, 2016.

BARONI, Hugo et al. **"Usability and Accessibility in Data-driven Websites."** *International Journal of Human-Computer Studies*, v. 77, p. 63-74, 2015.

5-STAR. As 5 estrelas dos dados abertos conectados. 2020. Disponível em: https://www.researchgate.net/figure/Figura-1-As-5-estrelas-dos-Dados-Abertos-Conectados-5-STAR-2020_fig7_349477731. Acesso em: 11 mar. 2025.

Cópia de documento digital impresso por GUSTAVO NOGUEIRA (1653598) em 08/04/2025 16:42.

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA PARAÍBA
Campus Campina Grande - Código INEP: 25137409
R. Tranqüilino Coelho Lemos, 671, Dinamérica, CEP 58432-300, Campina Grande (PB)
CNPJ: 10.783.898/0003-37 - Telefone: (83) 2102.6200

Documento Digitalizado Ostensivo (Público)

VERSÃO FINAL TCC

Assunto:	VERSÃO FINAL TCC
Assinado por:	Amanda Ferreira
Tipo do Documento:	Dissertação
Situação:	Finalizado
Nível de Acesso:	Ostensivo (Público)
Tipo do Conferência:	Cópia Simples

Documento assinado eletronicamente por:

- **Amanda Ferreira da Silva Alves**, **ALUNO (202011250006) DE BACHARELADO EM ENGENHARIA DE COMPUTAÇÃO - CAMPINA GRANDE**, em 18/03/2025 22:35:52.

Este documento foi armazenado no SUAP em 18/03/2025. Para comprovar sua integridade, faça a leitura do QRCode ao lado ou acesse https://suap.ifpb.edu.br/verificar-documento-externo/ e forneça os dados abaixo:

Código Verificador: 1424788
Código de Autenticação: c1a8ca24ae