

UNIVERSIDADE FEDERAL DE SÃO PAULO

DEPARTAMENTO DE CIÊNCIA E TECNOLOGIA

PROGRAMA DE MESTRADO PROFISSIONAL EM INOVAÇÃO TECNOLÓGICA

**VISUALIZAÇÃO E COMPARTILHAMENTO DE
DADOS UTILIZANDO GIS COM PARTICIPAÇÃO
PÚBLICA**

TIAGO EDUARDO

ORIENTADOR: PROF. DR. WALTER TEIXEIRA LIMA JUNIOR

São José dos Campos - SP
Março/2024

UNIVERSIDADE FEDERAL DE SÃO PAULO

DEPARTAMENTO DE CIÊNCIA E TECNOLOGIA

PROGRAMA DE MESTRADO PROFISSIONAL EM INOVAÇÃO TECNOLÓGICA

**VISUALIZAÇÃO E COMPARTILHAMENTO DE
DADOS UTILIZANDO GIS COM PARTICIPAÇÃO
PÚBLICA**

TIAGO EDUARDO

Dissertação apresentada ao Programa de Pós-Graduação Profissional em Inovação Tecnológica da Universidade Federal de São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Inovação Tecnológica, área de concentração: Inovação Tecnológica
Orientador(a): Dr. Walter Teixeira Lima Junior

São José dos Campos - SP
Março/2024

Elaborado por sistema de geração automática com os dados fornecidos pelo(a) autor(a).

Eduardo, Tiago

Visualização e compartilhamento de dados utilizando GIS com participação pública / Tiago Eduardo
Orientador(a) Walter Teixeira Lima Junior. - São José dos Campos, 2024.

112 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Mestrado Profissional Interdisciplinar em Inovação Tecnológica) - Universidade Federal de São Paulo - Instituto de Ciência e Tecnologia, 2024.

1. GIS. 2. Hiperlocal. 3. Participação Pública. 4. Tecnologia Móvel. 5. Lupa NH. I. Lima Junior, Walter Teixeira, orientador(a). II. , , coorientador(a), coorientador(a). III. Título.

UNIVERSIDADE FEDERAL DE SÃO PAULO

DEPARTAMENTO DE CIÊNCIA E TECNOLOGIA

PROGRAMA DE MESTRADO PROFISSIONAL EM INOVAÇÃO TECNOLÓGICA

VISUALIZAÇÃO E COMPARTILHAMENTO DE DADOS UTILIZANDO GIS COM PARTICIPAÇÃO PÚBLICA

TIAGO EDUARDO

Dissertação apresentada ao Programa de Pós-Graduação Profissional em Inovação Tecnológica da Universidade Federal de São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Inovação Tecnológica, área de concentração: Inovação Tecnológica.

Aprovado em 06 de março de 2024.

Membros da Banca:

Prof. Dr. Walter Teixeira Lima Junior
(Orientador – DCT-Unifesp)

Prof^a. Dr^a. Denise Stringhini
(Professora - DCT-Unifesp)

Prof^a. Dr^a. Agnes de Souza Arruda
(Professora - UMC)

Prof. Dr. Francisco Gilson Rebouças Pôrto Junior
(Professor - UFT)

São José dos Campos - SP
Março/2024

A minha família e amigos.

AGRADECIMENTO

Em primeiro lugar, a meus pais e meu irmão, que sempre me guiaram e incentivaram a seguir o caminho do estudo, aprendizagem, e a constante busca pelo conhecimento.

À minha querida esposa, por sempre me apoiar, aconselhar e ajudar nas reflexões, medos e anseios que apareceram ao longo desta pesquisa, e por entender todas as minhas ausências e ainda sim zelar por mim.

Ao meu professor e orientador, Prof. Dr. Walter Teixeira Lima Junior, por todos os ensinamentos, pela motivação, e especialmente, pela dedicação em organizar todos os afazeres que uma pesquisa desta magnitude exige. Sou muito grato por sempre estar disponível para sanar dúvidas e auxiliar com ideias.

Agradeço ao grupo de pesquisa Lupa NH e aos profissionais e alunos da escola estadual Raimunda dos Passos Santos, por terem sido parte integrante deste processo, e que dedicaram muitas horas na execução da atividade de captação de dados, possibilitando uma troca mútua entre pesquisador e voluntário. Me sinto privilegiado de ter trabalhado ao lado de grandes pesquisadores e pessoas como vocês.

Agradeço ao grupo de estudo Sistemas Cognitivos Artificiais e Robótica Social, onde tive a oportunidade de dar meus primeiros passos como pesquisador e onde meus colegas colaboraram nas reflexões e dúvidas ao longo do percurso acadêmico.

E por último, a instituição Universidade Federal de São Paulo, polo São José dos Campos, e ao seu corpo docente, pela oportunidade e privilégio me dado.

Precisamos de diversidade de pensamento no mundo para enfrentar os novos problemas

Tim Berners-Lee

RESUMO

Este é um trabalho de cunho interdisciplinar que exige que busquemos conhecimentos em áreas de estudos diversos, como tecnologia, comunicação e ciências sociais. Essa abordagem de conhecimentos especializados é necessária, pois quando trabalhamos com problemas complexos é fundamental um olhar disruptivo e amplo. Torna-se de grande importância a busca de conceitos e métodos de diferentes disciplinas do conhecimento humano para o desenvolvimento de uma solução efetiva para o problema pesquisado. Com o foco em explorar o problema de falta de informação estruturada sobre a situação de infraestrutura em áreas periféricas, e a baixa cobertura de noticiário local sobre, propõe-se utilizar conceitos de Participação Pública (PP) combinada com tecnologia GIS (*Geographic Information System*) na coleta de informações georreferenciadas (longitude e latitude), com o propósito de montar uma base de dados pública, e a partir desta gerar recomendações de criação de notícias hiperlocal utilizando inteligência artificial generativa (IAG). A partir do desenvolvimento do aplicativo móvel Lupa NH que habilita cidadãos do bairro Novo Horizonte em Macapá (AP) a cadastrarem ocorrências sobre o fornecimento de serviços de iluminação pública, calçadas e asfalto, coleta e tratamento de esgoto, coleta de lixo, limpeza urbana e água potável, objetiva-se implementar um algoritmo de priorização e seleção desses dados a partir de sua geolocalização, e com o uso do *ChatGPT* gerar recomendações de narrativas para criação de notícias. Tanto o conteúdo coletado, quanto as recomendações compõem um portal para compartilhamento da informação. Será decorrido aqui a contextualização do problema de pesquisa, seus impactos identificados na literatura, além de demonstrar um caminho possível para minimizar a falta de conteúdo hiperlocal com o uso de tecnologias digitais georreferenciadas com participação pública.

Palavras-chave: GIS, Hiperlocal, Participação Pública, Tecnologia Móvel, Interdisciplinar, Lupa NH

ABSTRACT

This interdisciplinary work requires us to seek knowledge in diverse areas of study, such as technology, communication, and social sciences. The approach of specialized knowledge is necessary because when we work with complex problems, it is essential to have a disruptive, broad look, based on the collaboration of concepts and methods from different disciplines of human knowledge to find an effective solution to the researched problem. With the focus on exploring the problem of lack of structured information about the infrastructure situation in marginal areas, and the low coverage of local news, it is proposed to use concepts of Public Participation (PP) combined with GIS (Geographic Information System) technologies in the collection of georeferenced information (longitude and latitude), with the purpose of creating a public database, and from this use generative artificial intelligence to generate recommendations for creating hyperlocal news. Based on the development of Lupa NH mobile application that enables citizens of the Novo Horizonte neighborhood in Macapá (AP) to register occurrences about the provision of public lighting services, sidewalks and asphalt, sewage collection and treatment, garbage collection, urban cleaning and drinking water, the objective is to implement an algorithm for prioritizing and selecting this data based on its geolocation, and using ChatGPT integration, generates narrative recommendations for news creation. Both the collected data and the news recommendations compose a portal for sharing information. The contextualization of the problem, its impacts identified in the literature, and a possible viable way to minimize the lack of hyperlocal content with the use of georeferenced digital technologies with public participation will be covered here.

Keywords: GIS, Hyperlocal, Public Participation, Education, Mobile Technology, Interdisciplinary, Lupa NH

LISTA DE FIGURAS

Figura 1 - Mapa do estado do Amapá, cidade de Macapá e bairro Novo Horizonte.	19
Figura 2 - Resultado totalizado por base	35
Figura 3 - Número de pesquisas realizadas por ano	35
Figura 4 - Número de pesquisas importadas X selecionadas na primeira análise por base.....	36
Figura 5 - Número de pesquisas importadas X selecionadas na primeira análise e segunda fase por base	38
Figura 6 - Mapa do bairro Novo Horizonte	46
Figura 7 - Diagrama de arquitetura.....	62
Figura 8 - Caso de uso do aplicativo móvel	63
Figura 9 - Caso de uso do <i>site</i>	66
Figura 10 - Diagrama de fluxo do cadastro de ocorrência e processo assíncrono ...	68
Figura 11 - Probabilidade de distribuição utilizando <i>Alias Method</i>	69
Figura 12 - Diagrama de fluxo do algoritmo de sorteio e geração de texto	71
Figura 13 - Aplicativo disponível na loja da <i>Google</i>	72
Figura 14 - Rota de captura de dado	74
Figura 15 – Enquete do aplicativo móvel	75
Figura 16 – Enquete do <i>site</i>	76
Figura 17 - Gráfico do desempenho da API de asfalto e calçadas.....	78
Figura 18 - Gráfico do desempenho da API de coleta de lixo	79
Figura 19 - Gráfico do desempenho da API de iluminação pública	79
Figura 20 - Gráfico do desempenho da API de coleta de esgoto	80
Figura 21 - Gráfico do desempenho da API de limpeza urbana.....	80
Figura 22 - Gráfico do desempenho da API de água potável.....	81
Figura 23 - Gráfico do desempenho da API de recomendação de criação notícia ...	81
Figura 26 - Fotos da captação de dados	85
Figura 27 - Total de ocorrências captadas	87
Figura 28 - Cadastro de trecho.....	88
Figura 29 - Respostas da enquete do aplicativo	91
Figura 30 - <i>Post</i> Instagram sobre o projeto	92

Figura 31 - Acesso ao <i>site</i>	93
Figura 32 - Visualização de dados em mapa	94
Figura 33 - Visualização do detalhe de uma ocorrência no mapa	95
Figura 34 – Visualização de dados em tabela	95
Figura 35 - Texto de recomendação de criação de notícia	96
Figura 36 - Respostas da enquete do <i>site</i>	98

LISTA DE TABELAS

Tabela 1 - Resultado do desempenho da API de asfalto e calçadas.....	78
Tabela 2 - Resultado do desempenho da API de coleta de lixo	78
Tabela 3 - Resultado do desempenho da API de iluminação pública	79
Tabela 4 - Resultado do desempenho da API de coleta de esgoto	79
Tabela 5 - Resultado do desempenho da API de limpeza urbana.....	80
Tabela 6 - Resultado do desempenho da API de água potável	80
Tabela 7 - Resultado do desempenho da API de recomendação de criação notícia	81
Tabela 8 - Distribuição de alunos presentes por dia e período	83
Tabela 9 - Total de ocorrência por classificação	89
Tabela 10 - Dados de instalação do aplicativo	90

LISTA DE QUADROS

Quadro 1 - Palavras-chave e sinônimos	33
Quadro 2 - <i>Strings</i> de busca	34
Quadro 3 - Critério de inclusão.....	34
Quadro 4 - Questionário de critérios para segunda análise	37
Quadro 5 - Pesquisas categorizadas	39
Quadro 6 - Descrição das atividades e ferramentas.....	48
Quadro 7 - Requisitos do aplicativo móvel.....	55
Quadro 8 - Requisitos do <i>site</i>	57
Quadro 9 - Requisitos do algoritmo de recomendação de criação de notícia	58
Quadro 10 - Estrutura de montagem do texto para envio ao <i>ChatGPT</i>	70
Quadro 11 - Configuração de carga	78
Quadro 12 - Lista de melhorias	85

LISTA DE ABREVIATURAS E SIGLAS

AI – *Artificial intelligence*

ABRAJI - Associação Brasileira de Jornalismo Independente

GIS - *Geographic Information System*

GPS – *Global Positioning System*

HTML - *Hypertext Markup Language*

IDE - *Integrated Development Environment*

NH - Novo Horizonte

PP - Participação Pública

PPGIS - *Public Participation Geographic Information System*

RSL - Revisão Sistemática da Literatura

TA – Tecnologia Apropriada

TC – Tecnologia Convencional

TS – Tecnologia Social

SIG - Sistema de Informação Geográfica

VGI - *Volunteered Geographic Information*

SUMÁRIO

CAPÍTULO 1 - INTRODUÇÃO.....	17
1.1 Introdução	17
1.2 Contextualização	20
1.3 Objetivo geral	22
1.4 Objetivos específicos	23
1.5 Justificativa	23
1.6 Organização do texto	24
CAPÍTULO 2 - REVISÃO BIBLIOGRÁFICA	26
2.1 Fundamentação teórica.....	26
2.1.1 Geographic Information System (GIS)	26
2.1.2 Participação pública	27
2.1.3 Hiperlocal	28
2.1.4 Inteligência Artificial Generativa.....	30
2.1.5 Tecnologia Social.....	31
2.2 Revisão sistemática da literatura (RSL)	32
2.2.1 Preparação	32
2.2.2 Execução.....	34
2.2.3 Primeira análise	36
2.2.4 Segunda análise	36
2.2.5 Resultados.....	38
CAPÍTULO 3 - MATERIAIS E MÉTODOS	45
3.1 Metodologia	45
3.1.1 Alcance da pesquisa	45
3.1.2 Público-alvo e amostra	46
3.1.3 Desenho da pesquisa.....	47
3.1.4 Inovação esperada.....	47
3.1.5 Análise dos resultados	48
3.1.6 Comunicação.....	49
3.2 Linguagem de programação e ferramentas.....	49
3.3 Banco de dados.....	51

3.4 Integrações sistêmicas	52
3.5 Recursos não funcionais	52
3.5.1 Escalabilidade.....	53
3.5.2 Confiabilidade	54
3.5.3 Segurança	54
3.5.4 Manutenção	54
3.6 Recursos funcionais.....	55
3.6.1 Aplicativo	55
3.6.2 <i>Site</i>	57
3.6.3 Algoritmo de recomendação de criação de notícia	58
3.7 Segurança e privacidade.....	58
3.8 Termo de uso, política de privacidade, TCLE e TALE	60
CAPÍTULO 4 - DESENVOLVIMENTO	61
4.1 Aplicativo móvel	62
4.2 <i>Site</i>	65
4.3 Algoritmo de recomendação de criação de notícia	67
4.4 Publicação e divulgação.....	71
4.5 Captação de dados	72
4.6 Enquetes	75
CAPÍTULO 5 - RESULTADOS E DISCUSSÃO.....	77
5.1 Teste de desempenho	77
5.2 Atividade de captação de dados.....	82
5.3 Métricas da atividade de captação de dados	87
5.4 Enquete do aplicativo.....	90
5.5 <i>Site</i> e algoritmo de recomendação de geração de notícia	91
5.6 Enquete do <i>site</i>	97
5.7 Discussão.....	98
CAPÍTULO 6 - CONCLUSÃO.....	102
REFERÊNCIAS.....	105
APÊNDICE A	110
TABELAS DO BANCO DE DADOS	110

Capítulo 1

INTRODUÇÃO

1.1 Introdução

Apesar da existência de mídias tradicionais a nível local como TVs, rádios comunitárias e jornais impressos, o foco desses veículos muitas vezes é em comunicar notícias do cotidiano, e não estabelecer debates estruturados sobre questões e temas de interesse coletivo (Lima Junior, 2018). Para impulsionar este tipo de geração de informação é que o jornalismo hiperlocal se estabelece.

Apesar de não haver uma definição padrão (Radcliffe, 2013), segundo a análise de (Trasel *et al.*, 2021), o jornalismo hiperlocal é a cobertura de notícias jornalísticas com foco em áreas geográficas restritas utilizando sistemas informativos como *blogs*, *sites*, entre outros, por cidadãos locais jornalistas ou amadores. Ainda segundo o autor, o objetivo desta ação é compartilhar as dificuldades e necessidades da comunidade visando proporcionar um melhor entendimento da situação local e resolução dos problemas com a participação coletiva.

Este modelo de comunicação pode ser beneficiado através do uso de tecnologias móveis digitais (Pavlik, 2011; Radcliffe, 2013), que são compostas por plataformas sociais, as quais permitem micro publicação, e celulares/*tablets* com GPS (*Global Positioning System*), pois esses habilitam o uso de dados georreferenciados. Através deste tipo de dado pode-se alavancar outra tecnologia espacial que é o acrônimo GIS (*Geographic Information System*), também traduzido como SIG (Sistema de Informação Geográfica). Este define-se como um sistema que manipula e processa dados de georreferenciamento, combinando localização

com descrições e características de um determinado local. Esses dados podem ser vinculados e visualizados em mapas, criando assim um relacionamento do dado espacial com o informativo. Esta tecnologia possibilita o entendimento de padrões e comportamento, promovendo melhor comunicação e apoio na tomada de decisão (ESRI, 2023).

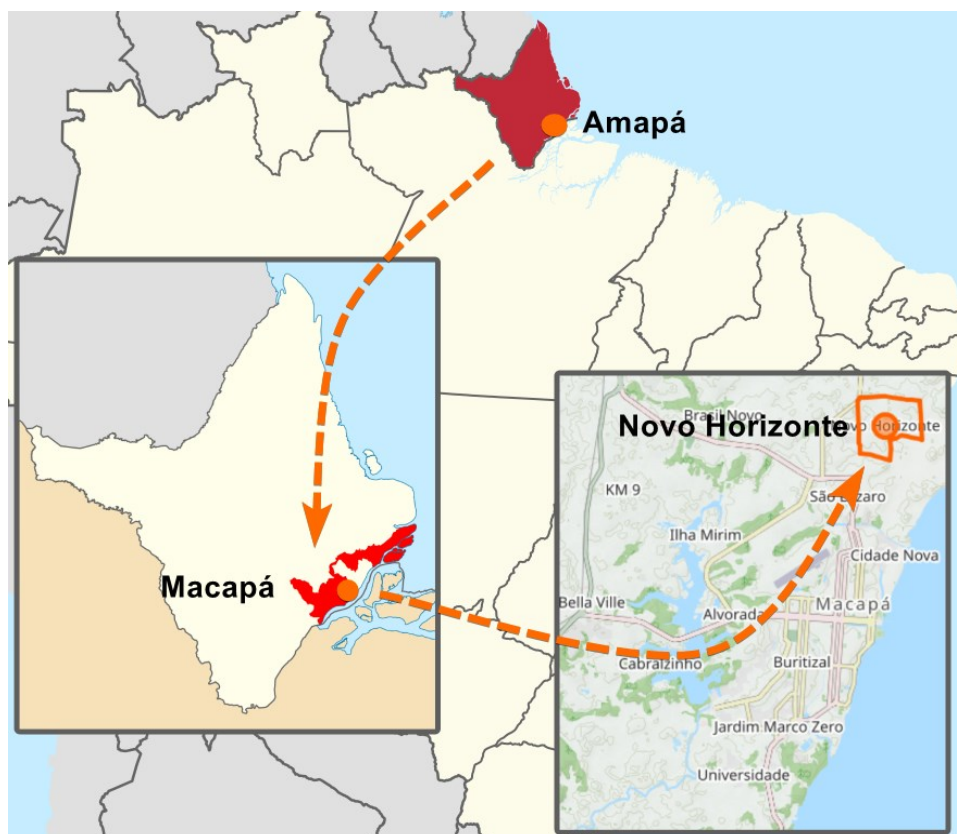
A partir do dado capturado, pode-se utilizar algoritmos de inteligência artificial para integrar essas bases e gerar de forma responsiva novas possibilidades de uso. Nesta linha, ferramentas como *ChatGPT*, que utilizam inteligência artificial generativa (IA Generativa) possibilitam uma ação de cooperação durante o processo de criação de conteúdo. A IA Generativa possui a capacidade de prever uma sequência de palavras que um ser humano utilizaria para montar uma sentença. Esses algoritmos aprendem a partir de interações com o usuário, geralmente via *chat*. Utilizando-se de um modelo não supervisionado onde o algoritmo calcula a probabilidade de uso de uma sentença baseado nas integrações passadas e um segundo modelo pré-treinado para refinar o resultado, cria-se uma forma de interação onde possibilita a colaboração entre homem e máquina (Radford *et al.*, 2018).

Combinando esses conceitos e tecnologias, tem-se a oportunidade de desenvolver soluções tecnológicas inovadoras que englobem informações georreferenciadas com participação pública e inteligência artificial, visando a geração de conteúdo informativo hiperlocal. Essas soluções, quando criadas e mantidas pelos cidadãos de um determinado local, possuem o potencial de criar um senso de pertencimento e responsabilidade social (Johnson *et al.*, 2017).

Objetivou-se nessa pesquisa a criação de uma solução tecnológica integrada (aplicativo móvel e *site*) que sirva como instrumento para criação de base de dados e auxilie na geração de notícias hiperlocal. O aplicativo móvel é responsável pela captura, armazenamento e processamento dos dados georreferenciados sobre infraestrutura deficitária. Estas são coleta e tratamento de esgoto, água potável, iluminação pública, calçamento, limpeza pública e coleta de resíduos. O *site* encarrega-se de gerar uma visualização completa dos dados capturados e, a partir desta, gerar recomendações de criação de notícias utilizando IAG, para publicações em diversos meios como *blogs* ou redes sociais. Para validação desta solução, foi executado um estudo de caso em que o público-alvo foram estudantes do ensino

médio integral que cursam a disciplina eletiva de comunicação jornalística da Escola Estadual Raimunda dos Passos Santos, localizada no bairro Novo Horizonte em Macapá (AP) (Figura 1).

Figura 1 - Mapa do estado do Amapá, cidade de Macapá e bairro Novo Horizonte.



Fonte: Autor utilizando OpenStreetMap

O *site* conta também com diversos recursos de visualização e exportação dos dados. Essa transparência possui uma relevância grande nesta pesquisa, pois deve-se considerar a tendência do setor público em aderir ao "dado aberto" incentivado pelas políticas públicas de transparência vigentes. Nota-se uma pressão do setor privado em adquirir esses dados para uso comercial e ações de marketing. Como consequência, áreas com pouco ou nenhum dado são pouco contemplados com o avanço dessas ações (Johnson *et al.*, 2017). Um dos fatores que impede o governo de realizar capturas e disponibilização de dados de georreferenciamento de forma igualitária nas regiões é o seu custo. Segundo Johnson *et al.* (2017), esses setores teriam que arcar com os recursos necessários para toda a etapa do processo,

incluindo não apenas a coleta, processamento e disponibilização, mas também a constante manutenção e atualização desses dados.

Em contrapartida, no setor civil existe uma iniciativa de captação de dados de georreferenciamento de forma voluntária com disponibilização pública, esse movimento é conhecido como VGI (*Volunteered Geographic Information*), onde a captura é realizada pelos próprios cidadãos utilizando-se da rede social ou de outra tecnologia móvel (Kleinmans; Ham; Evans-Cowley, 2015). Essas soluções ganham uma importância ainda maior pois, não apenas promovem visibilidade, mas também possibilitam que a região exista espacialmente nas bases de dados públicas e sejam contempladas por ações da área privada.

O Instituto de Tecnologia Social (2004) apresenta o conceito de Tecnologia Social (TS) para soluções que abordam essa relação entre conhecimento local, processos existentes, baixo custo e escalável para resolver algum problema existente de forma a diminuir a dependência de setores públicos e privados. Também pode-se traçar similaridades com o conceito de Inovação Social (IS), onde busca-se desenvolver inovações para suprir necessidades e amenizar problemas sociais (*Stanford Social Innovation Review*, 2003). Entretanto, este último ainda possui o lucro como foco.

Esta pesquisa seguiu o viés da TS, com inovação tecnológica criada de forma participativa, com uso voluntário (não governamental ou privado) e voltado a amenizar lacunas sociais através de soluções alternativas baratas. Construir conhecimento, ciência e tecnologia para uma região e público específico interessados em jornalismo, a fim de jogar uma luz sobre um problema social, neste caso a falta de notícia de conteúdo informativo em regiões periféricas.

1.2 Contextualização

O bairro Novo Horizonte é uma região periférica que segundo o Censo 2010¹ contém cerca de 25 mil moradores, distribuídos em 272 quadras, numa área de 15,3 quilômetros quadrados (IBGE, 2010). Atualmente é o segundo maior bairro em

¹ IBGE – Instituto Brasileiro de Geografia e Estatística. Censo 2010. Disponível em <https://censo2010.ibge.gov.br/resultados.html> Acessado em 29 de janeiro de 2024.

números de habitantes de Macapá. Os moradores da região alegam existir uma carência relacionado a prestação de serviço público de infraestrutura básica, além da falta de um poder executivo comprometido e ativo². Contudo o bairro apresenta um potencial em comunicação comunitária, pois possui em seu passado iniciativas de rádio comunitária e jornal impresso realizados pela própria população.

Devido a esse contexto, este local foi escolhido como alvo do Projeto Experimental e de Extensão Lupa NH³ do Programa de Mestrado Profissional Interdisciplinar em Inovação Tecnológica do Instituto de Ciência e Tecnologia da Universidade Federal de São Paulo (UNIFESP), liderado pelo Prof. Dr. Walter Teixeira Lima Junior. O objetivo do projeto era promover o engajamento cívico dos moradores e frequentadores do bairro através da participação pública e por intermédio de tecnologias digitais conectadas móveis. Das diversas frentes deste projeto surgiu o aplicativo móvel Lupa NH onde os moradores podiam coletar informações georreferenciadas sobre o bairro, gerando uma base de dados com informações sobre a infraestrutura deficitária nas áreas de água potável, coleta e tratamento de esgoto, iluminação pública, calçadas e asfalto, limpeza urbana e coleta de lixo (Lima Junior, 2018).

Com o início dos trabalhos em 2017, uma parceria foi formada com a Escola Estadual Raimundo dos Santos Passos situada na região, onde professores e alunos participaram de *workshops* sobre técnicas jornalísticas e debateram sobre necessidades e desafios da região. Em um dos encontros foi realizado um *Hackaton*⁴ onde um protótipo para o aplicativo móvel Lupa NH foi desenvolvido. Em seguida, foi executado uma etapa piloto onde membros do projeto e voluntários caminharam pelas ruas da região e testaram as telas de registro de ocorrências e captura via georreferenciamento. Devido ao seu propósito de teste, nenhum dos dados capturados foram mantidos, porém foi possível validar se o modelo de trabalho proposto é efetivo.

² Jornal do Amapá. Disponível em <https://globoplay.globo.com/v/11179993> Acessado em 29 de janeiro de 2024.

³ Projeto de Extensão cadastrado na Pró-reitoria de Extensão e Cultura (Proec/Unifesp), sob o nº 22197, denominado "Lupa NH 2.0: formação de inteligência social hiperlocal". Disponível em <https://www.ppgcom.ufpa.br/index.php/projetos-de-pesquisa> e <https://www.unifesp.br/campus/sjc/ppgit/projetos/tic>. Acessado em 29 de janeiro de 2024.

⁴ Hackaton: É uma combinação entre os termos hack (programar) e marathon (maratona), são eventos que reúnem programadores com o objetivo de criar uma solução por completo em um determinado período pré-estipulado.

A partir do aprendizado adquirido na etapa piloto, e dos *feedbacks* coletados com os voluntários, um documento com novos requisitos de melhorias foi proposto para o aplicativo móvel final. Porém devido a questões de financiamento para o projeto e o acontecimento da pandemia de Covid-19, o projeto foi interrompido e permaneceu parado até o ano de 2022. A partir do seu retorno, e um novo incentivo adquirido através do CNPq⁵, foram traçadas novas metas, onde uma delas foi implementar uma solução integrada que se apoie no aprendizado e requisitos do projeto de extensão existente. Esta deve conter o desenvolvimento de uma inovação tecnológica que possibilite a criação de uma base de dados sobre infraestrutura e incentive a geração de notícias de conteúdo informativo no bairro Novo Horizonte.

Para este fim, foi necessária uma abordagem interdisciplinar para compreender o impacto deste tipo de solução em uma comunidade periférica e ao mesmo tempo utilizar boas práticas de desenvolvimento de *software* para garantir uma solução aderente, confiável e escalável para outras regiões. Adicionalmente, também foi necessário que todos os dados captados estivessem disponíveis para os cidadãos, pois estes eram os reais donos da informação.

Dado as características do bairro, busca-se com essas tecnologias incentivar os estudantes da Escola Estadual Raimunda dos Passos Santos a participarem do processo de criação e divulgação dessas informações que relatam problemas de infraestruturas, consequentemente possibilitar o empoderamento da população e expor a situação da região através de dados coletados por ela mesma, pois, segundo (Sieber, 2006, p.503), a participação pública e o GIS oferecem uma abordagem única para envolver o público na tomada de decisões, onde os mesmos interagem dinamicamente com as informações, analisam alternativas e capacitam indivíduos e grupos.

1.3 Objetivo geral

Pretende-se demonstrar que tecnologias digitais móveis que utilizam GIS com participação pública podem gerar inovações tecnológicas que promovem a

⁵ CNPq: Conselho Nacional de Desenvolvimento Científico e Tecnológico. Agência governamental de fomento a ciência. Chamada CNPq/MCTI/FNDCT nº 18/2021-Universal.

estruturação e visualização de dados georreferenciados sobre as lacunas existentes no serviço público periférico, além de possuir potencial para fomentar a geração de notícias hiperlocal.

1.4 Objetivos específicos

- a) Gerar uma base de dados da região alvo da pesquisa com informações sobre infraestrutura pública precária;
- b) Promover o acesso a dados através da criação de um portal com visualização em mapa dos dados estruturados da região;
- c) Gerar recomendação de criação de notícia através do uso de dados georreferenciados e IA (Inteligência Artificial) generativa;

1.5 Justificativa

O site ABRAJI (Associação Brasileira de Jornalismo Independente) traz uma matéria onde pontua que "mais da metade dos municípios brasileiros são desertos de notícia"(Oliveira, 2018)⁶. A informação foi obtida através de um levantamento chamado Atlas de Notícia realizado pelo Instituto para o Desenvolvimento do Jornalismo e agência Volt Data Lab. Nos Estados Unidos da América um estudo similar intitulado *America's Growing News Deserts*⁷ realizado pelo *Columbia Journalism Review* desenvolveu um mapa interativo que apresenta as áreas onde a desertificação de notícia tem aumentado (Bucay *et al.*, 2017).

No Brasil, o estudo aponta que 51% dos municípios não possuem um veículo de comunicação. Isso representa cerca de 30 milhões de pessoas sem cobertura

⁶ ABRAJI: 30 milhões de pessoas vivem em desertos de notícias. Disponível em <https://www.abraji.org.br/noticias/30-milhoes-de-pessoas-vivem-em-desertos-de-noticia-mostra-atlas-da-noticia-2-0>. Acessado em 29 de janeiro de 2024.

⁷ CJR: America's Growing News Deserts. Disponível em https://www.cjr.org/local_news/american-news-deserts-donuts-local.php. Acessado em 29 de janeiro de 2024.

mediática local. Outras 34 milhões (30%) possuem entre um ou dois veículos, sendo caracterizados como em risco de se tornarem desertos (Oliveira, 2018).

As áreas mais afetadas tendem a ser as mais isoladas e com alta vulnerabilidade social, onde esses veículos ocupam um lugar importante na formação do cidadão cívico.

Essa pesquisa justificou-se pela necessidade e importância que a informação estruturada sistemicamente possui a nível local e a forma como sua veiculação se dispõe no cotidiano de cidadãos de áreas periféricas.

1.6 Organização do texto

- Capítulo 1: Apresenta-se uma introdução da base teórica e técnica que fundamenta essa pesquisa. Decorre-se sobre GIS, jornalismo hiperlocal e inteligência artificial generativa, e como esses conceitos se inter-relacionam para criação de novas tecnologias que possuem foco na geração de notícia. Na seção de contextualização, traz-se o histórico em que esse projeto se enquadra e sua colaboração almejada. Por último, apresenta-se o problema de pesquisa e o objetivo proposto.
- Capítulo 2: Objetiva-se nessa seção aprofundar na base conceitual e técnica trazendo sua evolução ao longo do tempo. Divide-se especialmente nos tópicos de GIS, Participação Pública, Jornalismo Hiperlocal e Inteligência Artificial Generativa, pois são o alicerce que permeiam a solução aqui proposta. Na seção de trabalhos correlatos traz-se alguns exemplos práticos de aplicação de soluções que utilizam GIS para resolver questões na área de comunicação e infraestrutura. Na revisão sistemática da literatura (RSL) apresenta-se o resultado da análise em bases científica sobre projetos que seguem a mesma linha de pesquisa.
- Capítulo 3: Apresenta-se os recursos tecnológicos escolhidos para a construção da solução e seus propósitos. Também é apresentado a metodologia utilizada durante o processo de pesquisa, desenvolvimento e validação.

- Capítulo 4: Detalha-se as etapas do desenvolvimento do aplicativo móvel, *site* e execução da captação de dados. Neste capítulo utiliza-se termos técnicos e relevantes para a reprodução e validação da pesquisa. Também é discutido requisitos funcionais, não funcionais e de desempenho.
- Capítulo 5: Apresentação dos resultados obtidos.
- Referências bibliográficas.
- Apêndice: Estão organizados na seguinte ordem: A) Artigos selecionados da RSL. B) Tabelas do Banco de Dados com detalhamento dos campos e tipo de dados utilizado. Este auxilia projetos futuro de integração e possibilita transparência de dados.

Capítulo 2

REVISÃO BIBLIOGRÁFICA

2.1 Fundamentação teórica

2.1.1 Geographic Information System (GIS)

GIS não é uma tecnologia nova, historicamente ela foi incorporada de outras disciplinas como estatística social, econometria, botânica e zoologia, e adaptada para o contexto espacial (Longley, 2000). Porém, sua aplicabilidade se inova a cada momento impulsionado pelo avanço da tecnologia e mudanças sociais (Goodchild, 1991). Na última década, o "*boom*" ocasionado pelo movimento de abertura e transparência de dados demandou que as novas soluções tecnológicas considerassem também dados georreferenciados (Johnson et al., 2017). Isso se deve ao fato de que GIS muda a forma como os dados são visualizados, possibilitando através do geoprocessamento a combinação de informações de diferentes fontes e sua exibição em camadas, resultando em uma melhor compreensão e auxiliando na tomada de decisão (Wright; Duncan; Lach, 2009; Sieber, 2006).

Algumas áreas se destacam na realização de pesquisa e desenvolvimento tecnológico utilizando GIS, entre elas planejamento urbano, mapeamento de relevos, preservação de áreas verdes e prevenção de desastres naturais. Apesar de seu potencial, o uso de dado georreferenciado sempre demandou recursos técnicos e especialistas (Sieber, 2006). Entretanto com o advento das tecnologias móveis, esse cenário mudou e possibilitou a diversificação ainda mais das áreas que embarcaram

essa tecnologia. Cada vez mais, ambos os aspectos técnicos e sociais têm sido superados e possibilitado resultados mais prolíficos (Wright; Duncan; Lach, 2009).

Ainda na década de noventa (Goodchild, 1991) chegou a listar 9 áreas de pesquisa distintas para GIS, com suas próprias questões e conceitos. São elas:

- Coleta e medição de dados;
- Estatísticas espaciais;
- Modelagem de dados;
- Estruturas de dados e índices;
- Algoritmos e processos;
- Exibição de dados espaciais;
- Ferramentas analíticas;
- Teoria da decisão e análise de risco;
- Raciocínio e cognição.

Três décadas depois essas disciplinas se ramificaram e proliferou-se um número irastreável de soluções. Ainda que muitas com sucesso comprovado, Longley (2000, p.2:40) ressalta uma preocupação: "A base de habilidades de análise espacial dentro da geografia é estreita, e há perigos se o software proprietário for rapidamente assimilado em uma disciplina que carece de uma direção clara sobre como se deve 'agregar valor' a esses desenvolvimentos". Contudo, o número de pessoas se especializando e capacitando em GIS é crescente e não está restrito apenas a área de geografia (Longley, 2000).

2.1.2 Participação pública

Este é um termo utilizado para representar o envolvimento de indivíduos no processo de decisão de alguma ação que os afetam de alguma maneira. No contexto político, onde este é mais difundido, tem-se o Projeto de Lei nº 678, de 2021, que regula o artigo 193 da Constituição Federal Brasileira de 1.988 para "estabelecer princípios e diretrizes para a participação da sociedade nos processos de formulação, de monitoramento, de controle e de avaliação das políticas sociais" (BRASIL, 1988). No que tange esta pesquisa, o termo Participação Pública (PP) se

refere ao uso da tecnologia para disponibilizar ferramentas e dados a um nível apropriado para uso das pessoas (ESRI, 2023).

GIS com participação pública, também referenciado como PPGIS (*Public Participation Geographic Information System*), habilita a tecnologia o potencial de engajar a participação de cidadãos nos processos decisórios, e consequentemente fortalecer a democracia (Sieber, 2006). Com a evolução das pesquisas desse tema na área, é necessário não apenas considerar o aspecto tecnológico, mas também fazer uma reflexão do impacto social (Pina; Painho, 2012). Governos têm sido influenciados nos seus processos de tomada de decisão devido o potencial de transparência e engajamento público que essas soluções têm proporcionado (Sieber, 2006).

A globalização do dado tem conectado comunidades ao contexto mundial, possibilitando uma coexistência entre local (geográfico) e global (digital). Esse fator ocasiona um compartilhamento maior de informação descentralizada, reduzindo a relevância da mídia tradicional e definindo uma identidade local desses veículos (Saar; Lima Junior; Ferreira, 2021). A instrumentação desses atores com tecnologias digitais que possibilita a captura e compartilhamento de dados georreferenciado, se torna relevante nesse processo de relação entre comunidade e mundo. GIS com participação pública passa a possibilitar o engajamento da comunidade no processo de planejamento, assim como, na incorporação da cultura e conhecimento local nas hipóteses de mudanças que acontecem ao longo do tempo (Delmelle, 2022).

Em um cenário ideal, cidadãos de um bairro poderiam capturar os próprios dados, analisá-los e processá-los sem a necessidade de algum ator externo com algum conhecimento técnico específico. Tal cenário, embora mais raro, possibilita uma evolução maior no processo de uso de tais ferramentas, resultando assim em uma melhor qualidade de dados e poder de negociação com o poder público (Harris; Weiner, 2002).

2.1.3 Hiperlocal

Este tipo de operação teve seu início nos anos 1980 nos Estados Unidos, onde o propósito foi classificar conteúdo televisivo local. Logo após, este foi associado a conteúdos produzidos para *web* (Lima Junior; Coelho, 2015).

Segundo Pavlik (2011) e Radcliffe (2013), o termo utilizado no jornalismo refere-se a conteúdo informativo online referente a uma região geográfica específica, independente de suas características, podendo ser bairros, vilas, comunidades, condomínios, entre outros. Já Glaser (2010) define como o conjunto de informações relevantes para uma comunidade que tem sido ignorada pelos veículos tradicionais de rádio ou televisão. Contudo, além de vasto, o termo vem modificando-se com o tempo, alterando o entendimento dos conceitos e seus significados (Lima Junior, 2017).

Como apontam Metzgar, Kurpius e Rowley (2011) esse tipo de operação ou mídia é híbrida, pois são muitas vezes instrumentos de movimentos cívicos e comunitários para discussões de temas de interesse coletivo, beneficiando-se também dos meios digitais alavancados com a *web 2.0*. Além do mais, alguns componentes são comuns entre as discussões sobre o assunto e ajudam a formular melhor o entendimento de hiperlocal. Os autores argumentam que o local geográfico, envolvimento dos cidadãos e o meio digital ajudam a diferenciar esse tipo de mídia das demais já existentes a nível comunitário. (Metzgar; Kurpius; Rowley, 2011)

Entretanto, é necessário assumir uma definição para o direcionamento desta pesquisa. Neste caso, o conceito de comunicação hiperlocal se refere a mecanismos tecnológicos digitais que favorecem a comunicação em um espaço geográfico delimitado ou restrito, a fim de proporcionar uma melhor divulgação de conteúdo informativo e apoiar o desenvolvimento social (Lima Junior, 2018).

O uso de tecnologia móvel favorece esse tipo de comunicação, pois, devido a evolução tecnológica, às novas ferramentas sociais e de georreferenciamento, o compartilhamento de informação tornou-se um processo mais democrático. Ainda que diversos desafios precisem ser superados como por exemplo financiamento, engajamento e visibilidade (Radcliffe, 2013), existe um real potencial para geração de notícia local através de cidadãos não jornalistas (Saar; Lima Junior; Ferreira, 2021).

Metzgar, Kurpius e Rowley (2011) comentam sobre quatro categorias sugeridas por Miel e Faris (2008) que descrevem o comportamento de geração de conteúdo em uma mídia tradicional, as quais são incorporados também pelo jornalismo hiperlocal:

- Comentários, críticas e conversas: Esse tipo de conteúdo são os mais populares em *sites* das mídias tradicionais.
- *Reviews*: São comentários onde já existe uma contribuição agregada.
- Opinião e análise: Incorpora um conteúdo mais extenso e analítico.
- Reportar e Denunciar: Este é o nível de comentário que caracteriza mídias hiperlocal, onde não existe apenas uma interação a um conteúdo, mas sim a geração dele.

Enquanto veículos tradicionais reorganizam-se devido ao grande corte de audiência⁸ e diminuição das redações dos últimos anos, este tipo de mídia hiperlocal procura ser uma das opções viáveis para preencher um *gap* de cobertura onde a participação pública é deixada de lado. (Metzgar; Kurpius; Rowley, 2011).

2.1.4 Inteligência Artificial Generativa

O termo inteligência artificial vem sendo utilizado para descrever ou classificar diversos produtos e soluções. Um exemplo bastante difundido na *internet* são os *chatBots*. Devido a sua característica de estabelecer comunicação rápida com textos curtos, que se assemelha com o modo atual de comunicação *online*, refletiu em uma aceitação maior entre os usuários (Lokman; Ameen, 2018). Essas soluções utilizam algoritmos NLP (*Natural Language Process*) que são responsáveis por traduzir a linguagem humana (escrita ou falada) em uma linguagem de máquina, considerando aspectos de sentimento, contexto e intenção do locutor ("*Natural Language Processing (NLP) Solutions | IBM*"). Isso é possível através do uso de uma modelagem baseada em regras da linguagem humana, combinada com modelos estatísticos e uma extensa base de dados (Ibid).

Avanços no processamento de dados colaboraram para a evolução dos *chatBots*, que passaram a responder em tempo real uma gama maior de informações. Isso se dá devido ao uso de *deep learning* que possibilita o

⁸ *The State of News Media*: Disponível em <https://assets.pewresearch.org/wp-content/uploads/sites/13/2016/06/30143308/state-of-the-news-media-report-2016-final.pdf> Acessado em 29 de janeiro de 2024.

mapeamento da informação utilizando redes neural, que simularam, de forma primitiva, o cérebro humano (Dwivedi *et al.*, 2023).

A ferramenta *ChatGPT* utiliza *deep learning* e uma modelagem generativa para processamento de texto e geração de respostas. Na modelagem generativa combina-se o uso do modelo não supervisionado para geração de respostas utilizando bases de dados extensas e um modelo supervisionado para refinamento baseado em interações passadas (Radford *et al.*, 2018, Dwivedi *et al.*, 2023). Esse modelo de método híbrido de geração de respostas tem sido utilizado por outras ferramentas do tipo e se mostrado eficiente.

Ainda há a necessidade de uma grande discussão sobre questões éticas e autorais relacionado ao uso geral de ferramentas de AI generativas, entretanto o resultado apresentado pela ferramenta *ChatGPT* em específico mostra um grande potencial, ainda em curva crescente, para uso como apoio em atividades diversas (Dowling; Lucey, 2023).

2.1.5 Tecnologia Social

O termo Tecnologia Social (TS) estabeleceu-se no Brasil a partir do ano de 2004 quando o Instituto de Tecnologia Social (2004) publicou sua definição e fez com que o termo se popularizasse nas mesas de debates. A TS estabelece uma solução alternativa e possível para aqueles que não usufruem das soluções existentes e disponibilizadas pelas empresas e governos. A partir do conhecimento local e processos existentes, é possível criar soluções inovadoras, de baixo custo e escalável para resolver ou amenizar algum problema em um contexto local. Desta forma, diminui-se a dependência de atores externos responsáveis pela provisão dessas soluções.

Este termo, porém, é a evolução de uma discussão iniciada na década de 70 onde os termos Tecnologia Convencional (TC) e Tecnologia Apropriada (TA) estavam em discussão. Segundo Dagnino (2014), a TC caracteriza-se por ser mais poupadora de mão de obra, intensiva em consumo sintético, possui escalas ótimas de produção sempre crescentes, sua produção é dada por máquinas e possui controles coercitivos que diminuem a produtividade. Podemos associar a essas características empresas de produção em massa por exemplo, onde o foco no

máximo lucro e menor perda guia sua administração. Devido a sua característica de escalas ótimas de produção, a cada novo ciclo é necessário aumentar os ganhos, ou seja, produz-se mais, ou minimiza-se os custos de produção e mão de obra, adquirindo assim o lucro.

Em contrapartida, a TA aparece como uma opção para os que não se beneficiavam da TC. Críticas a países de primeiro mundo foram surgindo reivindicando que soluções tecnológicas inovadoras fossem criadas para resolver os problemas que afetavam países subdesenvolvidos. Fraga (2011) esclarece que soluções para combater a pobreza, que fossem de desenvolvimento barato e de fácil replicação deveriam ser incorporados nas linhas de pesquisas. Mais tarde características da TA ajudariam a formular o que se entende hoje por TS. Foi incorporado o conceito de que os produtos e técnicas desenvolvidas devem ser direcionadas a partir dos problemas existentes na sociedade, não pelas universidades ou governos. Ela é elaborada e gerenciada pela população de forma coletiva e auto gestionário (Dagnino, 2014).

No contexto deste projeto, utiliza-se GIS com participação pública, algoritmos de seleção/clusterização e IA generativa como um arcabouço tecnológico para criação de soluções inovadoras com potencial de geração de valor para transformação em comunidades.

2.2 Revisão sistemática da literatura (RSL)

2.2.1 Preparação

Para melhor delineamento dos objetivos e clara mensuração dos desafios desta pesquisa, fez-se necessário não apenas o entendimento das tecnologias, conceitos e contexto, mas também uma consulta nas bases científicas para compreender o quão profundo esse tópico tem sido explorado por outros pesquisadores. Buscou-se identificar nesta RSL como esta pesquisa se enquadra em relação a originalidade, relevância científica, delimitação e viabilidade. Para este objetivo, definiu-se as seguintes bases para consulta:

- *El Compendex* (<http://www.engineeringvillage.com>)
- *Web of Science* (<http://www.isiknowledge.com>)
- *Scopus* (<http://www.scopus.com>)

A opção por essas três bases foi devido a sua característica de ser motor de busca. Seus resultados tendem a serem mais ecléticos pois é feita a indexação de diversas revistas.

Para iniciar as buscas, foi necessário preparar a *string*⁹ de consulta base. Para este fim utilizou-se as palavras-chave do projeto combinado com alguns sinônimos, como mostrado no quadro 1. A utilização de sinônimos maximiza o alcance da busca, entretanto, os mecanismos de busca geralmente possuem um limite de quantidade de termos para utilização nas *strings*. Devido a isso não é recomendado utilização de muitos termos.

Quadro 1 - Palavras-chave e sinônimos

Palavra-Chave	Sinônimo	Categoria
GIS	<i>PGIS, PPGIS</i>	<i>Intervention</i>
hyperlocal		<i>Population</i>
interdisciplinar		<i>Intervention</i>
<i>public Participation</i>		<i>Intervention</i>
<i>news</i>	<i>knowledge base</i>	<i>Outcome</i>

Fonte: Autor

As palavras-chave foram classificadas seguindo o método PICOC, que define cinco categorias para formulação do objetivo de uma RSL (Claes Wholin, 2012). Cada letra do acrônimo possui um significado:

- *Population*: Quem?
- *Intervention*: O que ou como?
- *Comparison*: Comparado a que?
- *Outcome*: O que se espera atingir / melhorar?
- *Context*: Qual a circunstância ou organização?

⁹ *String*: Texto contendo as palavras-chave e condicionais como “AND” e “OR” para busca em uma determinada base de dados.

Considerando essa estrutura, distribuiu-se as palavras-chave na categoria mais adequada, como visto na última coluna do quadro 1. Utilizou-se as categorias como base para a determinação do uso de *AND* e *OR* na construção da *string*. Palavras da mesma categoria foram separadas por *OR*, enquanto palavras de categoria diferente foi utilizado *AND*. O resultado é apresentado no quadro 2.

Quadro 2 - Strings de busca

<p><i>("Hyperlocal") AND ("GIS" OR "PGIS" OR "PPGIS" OR "Interdisciplinary" OR "Public Participation") AND ("news" OR "knowledge base")</i></p>

Fonte: Autor

2.2.2 Execução

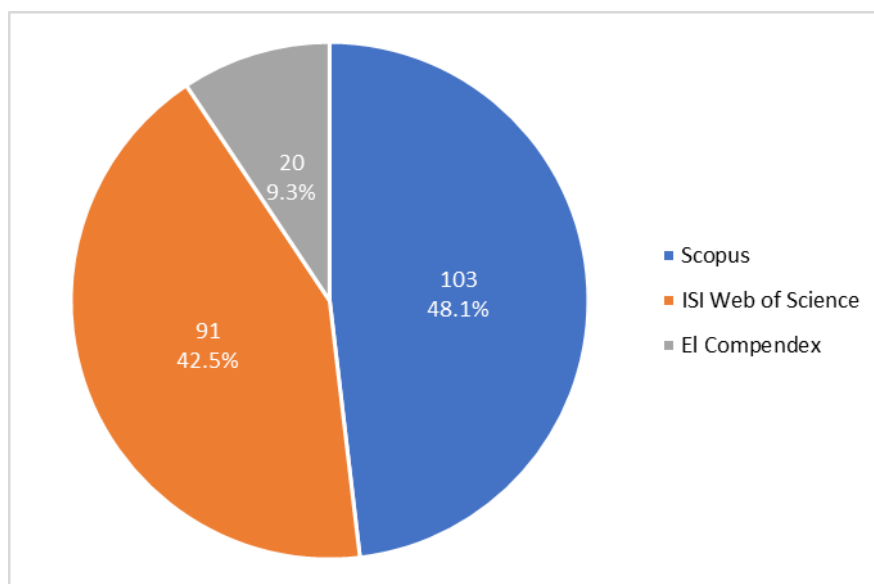
A partir da *string* base, passou-se para a etapa de execução da busca em cada uma das bases. Foram considerados três critérios para inclusão (quadro 3) e estes foram adicionados como filtros nos respectivos *sites*.

Quadro 3 - Critério de inclusão

Critério de inclusão
Artigo
Conferência
Últimos 10 anos

Fonte: Autor

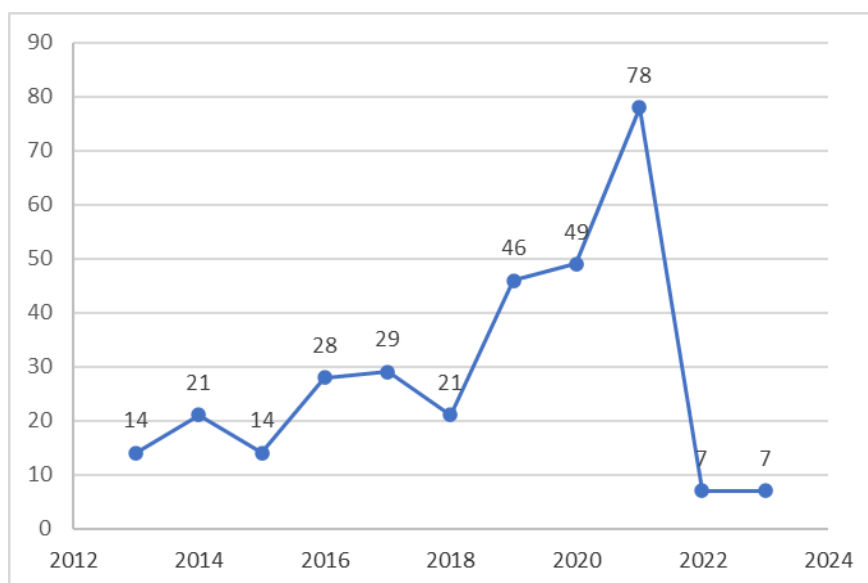
O resultado da busca com o uso dos critérios de inclusão totalizou **214** pesquisas. Na figura 2 podemos ver o número por base e sua proporção em relação ao total.

Figura 2 - Resultado totalizado por base

Fonte: Autor

O resultado foi exportado no formato .csv e logo após importado na ferramenta *Parsif*¹⁰, que é uma plataforma para apoio na execução de RSL.

Ao projetar o resultado em uma linha do tempo pode-se notar um aumento do número de produção de pesquisa a partir do ano de 2018 e uma queda brusca em 2022.

Figura 3 - Número de pesquisas realizadas por ano

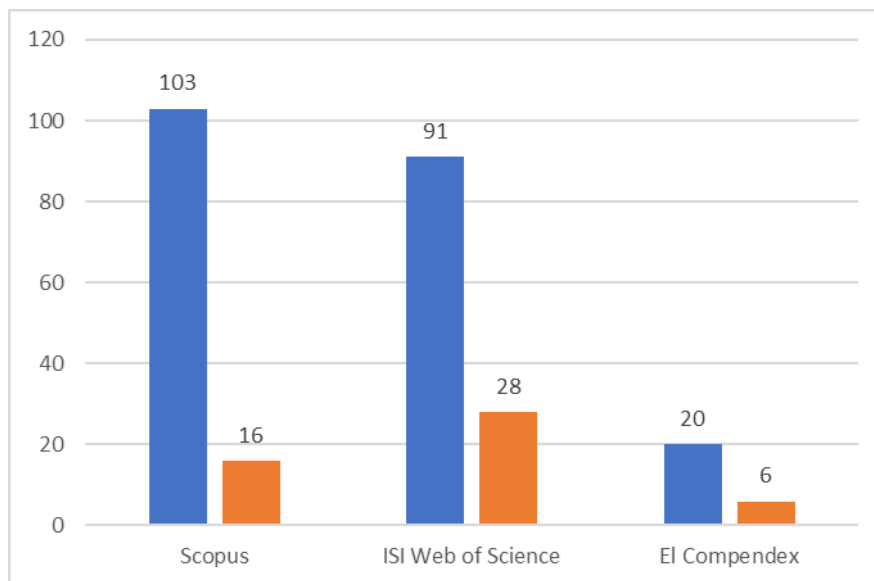
¹⁰ *Parsif* é um *site* para elaboração de RSL. Disponível em <https://parsif.al>. Acessado em 29 de janeiro de 2024.

Fonte: Autor

2.2.3 Primeira análise

Para realizar a análise das 214 pesquisas importadas foram executadas algumas validações. A primeira foi eliminar os resultados duplicados. A segunda foi eliminar pesquisas que não estavam relacionadas com a grande área de estudo e área específica deste projeto. As grandes áreas consideradas foram Ciências Sociais, Multidisciplinar e Ciência da Computação. Para as áreas específicas foram consideradas comunicação, urbanismo, administração pública, e todas as subáreas que esses contemplam. Adicionalmente, foram consideradas também itens que algumas bases classificavam como “*general*” ou “*miscellaneous*”. O resultado desta primeira análise reduziu o número de pesquisas para 50. Na figura 4 pode-se ver sua distribuição por base.

Figura 4 - Número de pesquisas importadas X selecionadas na primeira análise por base



Fonte: Autor

2.2.4 Segunda análise

Nesta etapa realizou-se a leitura do resumo (*abstract*) de cada uma das 50 pesquisas selecionadas. Para guiar esta validação utilizou-se um questionário de critérios a serem considerados durante a leitura (quadro 4). Em alguns casos,

quando o resumo não era suficiente para determinar uma resposta para algum dos itens do questionário, uma leitura dinâmica nos capítulos de introdução, metodologia e conclusão era realizado.

Quadro 4 - Questionário de critérios para segunda análise

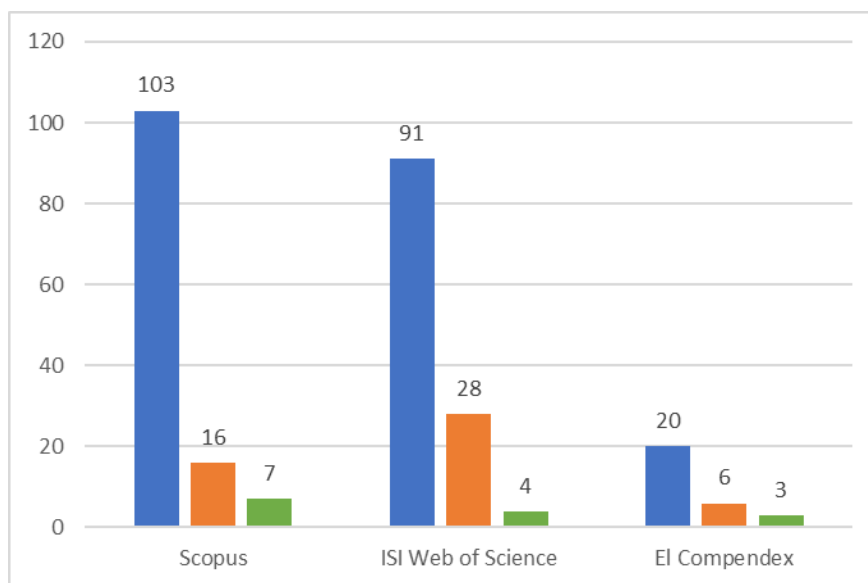
1	A pesquisa se relaciona com populações vulneráveis e em áreas periféricas?
2	A pesquisa colabora para a discussão sobre empoderamento ou auto-organização social?
3	A pesquisa é aplicada no contexto de infraestrutura pública urbana?
4	A pesquisa contempla desenvolvimento de inovações tecnológicas <i>web</i> , móveis ou IA?
5	A pesquisa utiliza participação pública ou voluntária na coleta ou análise de dados?
6	A pesquisa possui resultados claro, prático e relevante a nível social?
7	A pesquisa discute a geração de conteúdo informativo a nível local?

Fonte: Autor

Cada pergunta no questionário possuía três opções de respostas e um valor em ponto designado. Sim (1.0), Parcialmente (0.5) e Não (0). Tendo como menor e maior valor total o 0 e 7, cada uma das pesquisas foram respondidas e classificadas por nota. Foi definido por conveniência uma nota de corte de igual ou maior a 5 que resultou em uma nova lista com **14** pesquisas, como mostrado na figura 5.

O resultado possibilitou uma separação entre pesquisas que focavam em discussões conceituais e exploratórias de pesquisas aplicadas que consideravam o uso de tecnologia móveis. Essa nova lista foi separada para leitura completa e execução de uma análise crítica. Seus achados estão no item 2.2.5.

Figura 5 - Número de pesquisas importadas X selecionadas na primeira análise e segunda fase por base



Fonte: Autor

2.2.5 Resultados

Esta análise identificou que muitos tipos de problemas podem ser trabalhados utilizando tecnologias digitais móveis que embarcam GIS. Combinado com a participação pública, um impacto positivo nas comunidades pode ser alcançado. Foram selecionadas soluções de diferentes países, culturas e condições socioeconômicas. Alguns setores como turismo, desastres naturais e conservação de áreas verdes, se destacam e lideram a linha de frente no desenvolvimento de pesquisas. Sem uma fronteira aparente, essas tecnologias em conjunto com ações sociais mostraram potencial para ser uma ferramenta para alavancar mudanças a nível hiperlocal.

Após análise crítica foi optado por categorizar as pesquisas por área de atuação, facilitando assim a identificação de tendências e abordagens técnicas que estão sendo seguido quando se trata de tecnologias móveis, jornalismo hiperlocal e participação pública. As categorias foram definidas pós leitura, a partir do entendimento de cada projeto, como mostrado quadro 5.

Quadro 5 - Pesquisas categorizadas

Categoria	Qtd	Referência
Impacto da tecnologia hiperlocal na comunidade	2	Liu (2018) / Robaeyst (2021)
Perfil dos veículos hiperlocal	4	Jangdal (2020) / Tshabangu (2022) / Brannock Cox; Poepel (2020) / Jumisko-Pyykkö (2014)
Perfil das tecnologias sociais	1	Mukhtar <i>et al.</i> (2021)
Tecnologias colaborativas	2	Han; Shih; Carroll (2014) / Leyshon; Rogers (2020)
Tecnologias <i>data-driven</i>	4	Hepp; Loosen (2019) / Carroll <i>et al.</i> (2015) / Han (2014) / Noceta (2021)
Outras RSL	1	Saldivar (2018)

Fonte: Autor

Embora o foco desta RSL ser pesquisas tecnológicas aplicadas, percebe-se que a grande maioria dos trabalhos na área concentram-se no campo conceitual. Há um grande foco em analisar o impacto da mídia hiperlocal nos veículos de comunicação, nos profissionais envolvidos, nos materiais produzidos e no impacto desses materiais nas comunidades.

Na categoria de **impacto da tecnologia hiperlocal na comunidade**, a pesquisa de Liu (2018), nos estados Unidos, comparou o nível de participação civil de residentes de áreas com ou sem veículos de mídia hiperlocal. O resultado foi que existe uma influência recíproca, onde o nível de participação civil dos residentes aumentou, assim como, o veículo moldou-se para absorver esse reflexo, reforçando a ideia de que veículos hiperlocal não são necessariamente uma via de transmissão única.

Na pesquisa de Robaeyst (2021), foi investigado o nível de coesão social nos bairros que utilizam a ferramenta *Hoplr* como rede social hiperlocal. O estudo foi realizado em 597 bairros, totalizando 230.198 usuários, na região de *Flemish* na Bélgica. Nessa ferramenta, o governo local e a população podiam criar perfis e utilizar o espaço para dialogar, promovendo o engajamento e participação pública. Segundo o estudo, duas características colaboram mais para a criação da coesão social. São elas: 1) uma comunicação trivial sobre interesses em comum entre os usuários; 2) e uma rede de apoio onde os residentes possam trocar informações e recomendações sobre serviços em geral como compras e vendas da região.

Mudando o foco para pesquisas que analisam o **perfil dos veículos hiperlocal**, o estudo de Jangdal (2020), realizado na Suécia, faz um levantamento dos tópicos mais abordados pelos veículos hiperlocal. Segundo o estudo, no país, o conteúdo que apresenta maior número de notícias relacionadas é a política. Isso demonstra que esses veículos cumprem com um dos objetivos centrais deste tipo de mídia, que é proporcionar informações que promovam o senso crítico e democratização. Entretanto, esses assuntos tendem a terem menor engajamento comparado a outros (29% e 33%). O autor ressalva que se deve considerar que cerca de metade dos veículos hiperlocal na Suécia permitem interação via comentários nas notícias. Isso pode comprometer a discussão dos residentes sobre os tópicos de seu interesse.

Na mesma linha, o estudo de Tshabangu (2022) traz a mudança no processo de trabalho do grupo *Alpha Media Holding*, no Zimbábue. O jornal mudou seu modelo de trabalho não participativo, onde todo processo, desde a geração de ideia, coleta de dados, escrita, edição e publicação eram produzidos interno, e passou para um modelo participativo com a inclusão do cidadão e processo descentralizados. Essa mudança foi possível através do uso de tecnologias digitais e processos estruturados, comenta o autor. Este modelo possibilitou um novo foco no conteúdo hiperlocal pois muitos materiais passaram a ser enviados pelos colaboradores não jornalistas da própria região onde as notícias eram veiculadas.

Com outro viés, a pesquisa de Brannock Cox e Poepsel (2020) traz uma avaliação no uso do modelo *Hearken*¹¹, que apoia veículos hiperlocal na criação de conteúdo que supra a necessidade e interesses dos cidadãos utilizando técnicas jornalísticas para maximizar a identificação de pauta e promover o engajamento do público. O uso do modelo é interessante para evitar que o jornalismo hiperlocal falhe no propósito de ser um veículo que gere conteúdo relevante aos cidadãos. Isso também ajuda a evitar que alguns cidadãos utilizem esses veículos para práticas antidemocráticas.

Ainda na mesma linha, Jumisko-Pyykkö (2014) em uma pesquisa na Finlândia, analisou os fatores que colaboram para a melhor qualidade de criação de conteúdo por não jornalistas no país. Após uma série de entrevistas com o público-

¹¹ *Hearken* é um modelo de trabalho para aumentar engajamento. Disponível em <https://wearehearken.com/about-us>. Acessado em 29 de janeiro de 2024.

alvo, foi identificado que o *feedback* do veículo é a melhor forma de promover a melhoria da qualidade. O *feedback* dos leitores em relação a aquele conteúdo também mostrou-se ser um item motivador. Outros fatores como bonificação monetária ou outros benefícios também colaboram com a motivação para se desenvolver melhores conteúdos.

As pesquisas acima focaram em analisar os veículos hiperlocal assim como o material produzido e seu impacto. As pesquisas seguintes já utilizam tecnologias móveis como objeto de estudo ou meio para atingir algum objetivo.

O estudo de Mukhtar *et al.* (2021), analisa o **perfil das tecnologias sociais** e discute que tecnologias desenvolvidas nos países de primeiro mundo podem não ser aplicável a países em desenvolvimento devido a diferenças nas características comportamentais da população desses países. Como exemplo, o autor traz, na Índia há a perspectiva de que 60% da população do país deixará de viver nas vilas e passará a viver nas grandes cidades até 2050. Lagos, cidade mais populosa da África, tem um aumento populacional de 3,26% anual. Além disso, as populações em países em desenvolvimento podem possuir problemas latentes relacionado a transporte de qualidade, acesso à *internet* ou mesmo uma tendencia cultural de ter como fonte de informações primária amigos e família. Esse estudo argumenta que soluções tecnológicas social para esses países precisam considerar a cultura, característica e problemas sociais locais. Importar soluções podem não ser o melhor caminho.

Em relação a **tecnologias colaborativas**, Leyshon e Rogers (2020) trazem um estudo sobre a ferramenta CJNN (*Citizen Journalism News Network*) desenvolvida em Cornwall, na Inglaterra. O aplicativo CJNN possibilita a criação de fragmentos de notícia, onde o usuário seleciona qual o tipo de conteúdo quer adicionar, como texto, imagem ou vídeo, e este é publicado em um *site* do aplicativo, ao lado dos conteúdos de outros usuários. Através do uso de *tags* e geolocalização, o conteúdo é *clusterizado* e classificado para melhor consumo do leitor. O autor discute que tecnologias sociais possibilitam uma via de mão dupla para jornalismo hiperlocal. Elas possibilitam o cidadão a tornar-se parte do processo de criação de notícia. Entretanto, o projeto sofreu com o baixo engajamento dos cidadãos e a dificuldade em moderar os conteúdos. Ainda assim, discute-se a importância em disponibilizar meios para comunidades periféricas terem acesso a conteúdo digital.

Han, Shih e Carroll (2014) trazem em sua pesquisa que o uso de aplicativos para necessidades em geral a nível hiperlocal é vasto e popular. A tecnologia desenvolvida em sua pesquisa é o aplicativo *Mobile Timebanking*, que possibilita um morador a disponibilizar-se para fazer uma atividade voluntária para outro morador da mesma área. Cada atividade gera um crédito que pode ser utilizado para requisitar serviços de outro morador da área. É uma forma de troca de serviços onde quem use um serviço pode receber em troca outro serviço de sua necessidade.

Outra característica identificada nas pesquisas selecionadas são as **tecnologias data-driven** que possuem como foco a coleta de dados de diferentes fontes.

O projeto realizado por Hepp e Loosen (2019) na Alemanha, apresenta o aplicativo *Molo.news*, onde o objetivo é veicular conteúdo de diversas fontes em um mesmo local e com moderação automatizada para exibir informações que sejam relevantes para um determinado local. O projeto discute a importância da abordagem focada na relacionalidade, onde o jornalista e o cidadãos estabelecem um papel com mútuas responsabilidades e expectativas. O aplicativo foi especificado e desenvolvido utilizando esse entendimento da relação entre esses dois atores que buscam um resultado comum.

No estudo de Carroll *et al.* (2015), é apresentada a tecnologia *CiVicinity.org* como uma ferramenta criada para agregar notícias de diversos veículos de mídia local. Além das notícias, pode-se utilizar um mapa para localização de eventos na região. Outra funcionalidade é um espaço aberto para publicação e comentários em conteúdo dos próprios usuários. O autor discute que soluções tecnológicas a nível hiperlocal colaboram com a criação da identidade, senso de pertencimento e participação.

O estudo de Han, Shih e Carroll (2014) apresenta um aplicativo que agrega informações sobre um determinado ponto geográfico em formato de *microblog*, possibilitando trazer informações descentralizadas de pessoas que comentaram ou mencionaram o local em alguma postagem em outras redes sociais. O estudo discute que é necessário tecnologias descentralizadas onde as redes se conectam para geração de mais informação relacionada. O aplicativo foca em promover uma melhor visibilidade a locais geográficos que são comentados por cidadãos.

No estudo de Nocera (2021) é apresentada a criação da ferramenta *Crosstown Foundry* para apoiar jornalistas no processo de criação de notícias

hiperlocal. A ferramenta tem a característica de ser um agregador de conteúdo onde conecta-se em fontes diversas como, trânsito, ocorrências criminais e saúde, para a criação de um painel central. Através desses dados é possível fazer uma análise de predição e destacar mudanças em indicadores, possibilitando a identificação de possíveis tópicos para novos conteúdos. A ferramenta foi desenvolvida e está em uso por um veículo de notícias sem fins lucrativo vinculado a *University of Southern California's Annenberg School for Communication and Journalism*.

Por último, Saldivar *et al.* (2018) realizou uma **RSL** que identificou inovações social tecnológicas que colaboram com o processo de engajamento social e participação pública. Em seu estudo que selecionou 35 tecnologias, 20 eram soluções *web*, 6 aplicativos móveis e 2 híbridos. As demais possuíam outros formatos de mídia como painéis digitais públicos. 5 das soluções usavam mapas interativos, e apenas 3 soluções utilizavam GIS ou realidade virtual. No caso de GIS, eram utilizados para colher *feedback* de certos locais. Apoio a planejamento urbano era o foco de 16 das 35 tecnologias. 34 das 35 dependiam da interação do usuário para funcionar, seja ela comentário, opinião ou avaliação.

Em todas as pesquisas selecionadas não foi identificado nenhuma solução para transparência de dados das soluções tecnológicas desenvolvidas. Não parece haver uma preocupação em disponibilizar de forma rastreável exportável todas as interações que os usuários dessas tecnologias realizam. Nota-se também que muitas tecnologias focam em relacionar, classificar e centralizar dados já existentes com o intuito de facilitar o acesso do cidadão as informações. Entretanto não é explorado a possibilidade de geração de conteúdo primário pelo próprio cidadão. Muitas das participações cível ocorrem na forma de comentário, compartilhamento e divulgação de conteúdo.

Nenhuma das soluções utilizam inteligência artificial em algum processo ou funcionalidade. Existem muitas pesquisas que discutem o uso de IA na criação de conteúdo, porém o debate gira em torno dos benefícios e riscos, e não experimentos práticos integrando essas tecnologias a tecnologias moveis hiperlocal.

Conclui-se que esta pesquisa é relevante pela sua originalidade em permitir a criação de conteúdo informativo primário através do uso do aplicativo georreferenciado e *site* com recomendação de criação de notícia através de IA generativa. Como aprendizado, tem-se que tecnologias digitais social de âmbito

hiperlocal deve focar em resolver ou amenizar um problema específico de contexto local. A replicação dessas tecnologias deve ser executada com devida cautela e considerar sempre uma etapa de adequação para novo contexto. Tecnologias facilmente configurais e escaláveis onde não há a necessidade de equipe especializada são as mais adequadas nesse cenário. Por último, deve-se considerar o fator motivação e engajamento para garantir o uso e relevância da tecnologia. Características como gamificação, interação entre usuários e personalização de perfil podem colaborar de forma positiva.

Capítulo 3

MATERIAIS E MÉTODOS

3.1 Metodologia

Este capítulo apresenta o detalhamento metodológico que compõem as etapas para execução dessa pesquisa. São apresentadas as ferramentas, tecnologias e sequência de atividades executadas para alcançar o objetivo definido. Espera-se que cada etapa esclareça de forma sistemática, empírica e crítica os caminhos percorridos, pois como descreve Sampiere e Collado (2013), em uma pesquisa científica os dados devem ser coletados, analisados e criticados para avaliação e aperfeiçoamento constante, não recorrendo assim a causalidades.

3.1.1 Alcance da pesquisa

Esta é uma pesquisa de natureza aplicada onde optou-se por utilizar o método de pesquisa estudo de caso descritivo, onde a partir de análises quantitativas e qualitativa, buscou-se descrever o resultado do uso das ferramentas aqui apresentadas.

Guiado pelos estudos de Metzgar, Kurpius, Rowley (2011), Radcliffe (2013) e Lima Junior (2017), onde é apresentado que tecnologias móveis com dados geográficos e participação pública podem ser utilizados como instrumentos para minimizar o deserto de notícias em áreas periféricas, discute-se a viabilidade dessa abordagem no bairro Novo Horizonte, destacando dificuldades e aprendizados.

Este estudo não proverá informações válidas para uma generalização, porém espera-se que ele colabore com a preparação de hipóteses e construção de teorias.

3.1.2 Público-alvo e amostra

O local de aplicação é o bairro Novo Horizonte situado na cidade do Macapá (AP). Região de zona periférica com cerca de 25 mil habitantes distribuídos em uma área de 15,3 quilômetros e 240 quadras (Figura 6).

Figura 6 - Mapa do bairro Novo Horizonte



Fonte: Google Maps

O público-alvo foi selecionado de forma não probabilístico através da Escola Estadual Raimundo dos Passos Santos, na localidade, onde através de uma matéria eletiva de jornalismo promovida pela escola, alunos dos anos primeiro, segundo e terceiro do ensino médio foram convidados a utilizarem as ferramentas e a fazerem a captação de dado. O projeto está aprovado pelo Comitê de Ética da UNIFESP,

com parecer CAAE: 65307822.6.0000.5505¹², onde essa parceria com a escola foi formalizada.

3.1.3 Desenho da pesquisa

Utilizou-se a seguinte sequência de atividades.

1. RSL: Esta etapa avaliou a originalidade, relevância científica, delimitação e viabilidade da pesquisa.
2. Arquitetura sistêmica: Levantamento e análise de todos os *softwares* necessários para atender o objetivo da pesquisa incluindo o desenho do diagrama de fluxo utilizando ferramenta Miro.
3. Desenvolvimento técnico do aplicativo móvel para captação de dados.
4. Desenvolvimento técnico do *site* para exposição de dados.
5. Desenvolvimento do algoritmo de recomendação e priorização de notícia.
6. Teste de desempenho das APIs de cadastro do aplicativo e do algoritmo de geração e priorização de notícia.
7. Etapa de captação de dados georreferenciados em campo.
8. Aplicação de envelope para avaliação do aplicativo.
9. Aplicação de envelope para avaliação do site.
10. Validação dos dados capturados, aplicativo e *site*.

3.1.4 Inovação esperada

Criação e validação de um algoritmo que utiliza dado georreferenciado a partir de ocorrências cadastradas (latitude e longitude), para seleção, priorização e montagem do texto para recomendação de criação de notícia.

Para montagem do texto de recomendação utilizou-se AI generativa, neste caso o *ChatGPT*. Entretanto a solução é escalável para outras plataformas, devido a seu desacoplamento, exigindo pouco desenvolvimento de código.

¹² Plataforma Brasil: Acessar o endereço <https://plataformabrasil.saude.gov.br> e buscar pelo número CAAE. Acessado em 29 de janeiro de 2024.

O algoritmo de seleção e priorização utiliza regra de atribuição de peso a cada ocorrência dado a sua proximidade com outras ocorrências, gerando uma regra de *clusterização*. Exemplo:

- *Cluster* por proximidade geral: + 20 pontos. Cada ocorrência com 50 metros ou menos de proximidade de outra ocorrência.
- *Cluster* por proximidade por tipo: + 30 pontos. Cada ocorrência com 50 metros ou menos de proximidade de outra ocorrência e ser do mesmo tipo.
- Ocorrências sem *cluster*: +10

Dado o peso de cada ocorrência, utilizou-se o método *Alias Method* (Walkter, 1974) para sorteio da ocorrência. Por último, utilizou-se a integração com *OpenAI* para formulação do texto de recomendação da notícia e envio para visualização no *site*. O algoritmo completo está detalhado no item 4.3.

3.1.5 Análise dos resultados

Para formulação dos resultados desta pesquisa, executou-se uma análise descritiva da atividade e dos dados capturados através do método de estudo de caso (Wazlawick, 2009), além de uma análise de desempenho dos softwares propostos e por fim, uma enquete para coleta de feedback sobre o aplicativo e site. As atividades para cumprir essa análise foram divididas da seguinte forma:

Quadro 6 - Descrição das atividades e ferramentas

Atividade	Ferramenta
Execução do teste de desempenho das APIs de entrada de dados utilizado pelo aplicativo móvel	<i>Postman</i>
Execução do teste de desempenho do algoritmo de geração de recomendação de criação de notícia	<i>Postman</i>
Descrição da atividade de captação de dados	Diário de campo
Métricas da atividade de captação de dados	<i>Site</i>
Enquete para captação da percepção do usuário em relação ao aplicativo	Papel
Validação do algoritmo de recomendação de criação de notícia utilizando dados reais	<i>Site</i>

Enquete para captação da percepção do usuário em relação ao site	Formulário <i>web</i>
--	-----------------------

Fonte: Autor

Relacionado as características dos dados, as tabelas detalhadas no apêndice A descrevem suas identificações e tipagens.

3.1.6 Comunicação

Os entregáveis estão disponíveis publicamente no seguinte formato:

- O código fonte dos *softwares* como aplicativo móvel, *site*, algoritmos de recomendação de narrativas e banco de dados estão disponíveis na plataforma *GitHub* no endereço <https://github.com/lupa-novo-horizonte>
- O aplicativo móvel está hospedado na *PlayStore* no endereço <https://play.google.com/store/apps/details?id=com.lupa20>
- *Site* contendo todos os dados captados com opção de exportação para formato CSV disponível para acesso no endereço <http://lupa.tecccog.net/mapa>
- Publicações relacionadas a essa pesquisa em <http://lupa.tecccog.net/publicacoes>
- Por último, essa pesquisa está disponível na biblioteca digital da Unifesp na seção de dissertações de mestrado.

3.2 Linguagem de programação e ferramentas

O desenvolvimento do aplicativo móvel tem como base seu protótipo entregue a essa pesquisa pelo projeto Lupa NH. Seu desenvolvimento inicial ocorreu em 2018 em um *hackaton* que reuniu estudantes e professores do bairro Novo Horizonte. Em princípio foi considerado reutilizá-lo e trabalhar com uma evolução incremental. Entretanto, as características técnicas do protótipo escolhidas na época consideravam apenas a velocidade de desenvolvimento e entrega para uso. Isso se

justifica devido ao formato do evento onde os participantes tinham apenas 8 horas de trabalho em equipe.

Tendo isto, decidiu-se desenvolver uma nova versão utilizando a tecnologia *React Native*¹³ pertencente a *Meta*. Sua principal característica é ser uma linguagem de programação para desenvolvimento de aplicativos móveis. Com isso, recursos como câmeras, GPS, armazenamento, compatibilidade com aparelhos e desempenho são mais bem aproveitados. Esta linguagem de programação também permite o desenvolvimento multiplataforma. O mesmo código-fonte foi utilizado para a versão *Android* e *iOS*, gerando um ganho de tempo de desenvolvimento e melhor escalabilidade de plataformas. Para a escrita do código-fonte optou-se pelo *Visual Studio Code*¹⁴ disponibilizada pela *Microsoft*. Esta é uma plataforma leve para uso em computadores domésticos e possui uma grande base de conhecimento *online*. Para o armazenamento de dados, optou-se pela tecnologia *MySQL*¹⁵. Isso se deve ao fato de ser um banco de dados relacional, onde os dados são salvos em tabelas e pode-se criar relacionamentos entre elas, aumentando a capacidade de geração de relatórios. Como um dos objetivos dessa pesquisa foi disponibilizar os dados em um *site* para geração de transparência, essa possibilidade de recurso se mostra benéfica.

Para o *site* onde os dados e as sugestões de notícias são apresentados, optou-se pela linguagem de programação *Angular*¹⁶ desenvolvida e mantida pela *Google* e que possui uma vasta base de conhecimento pública. Para edição do código-fonte utiliza-se o mesmo editor do aplicativo móvel, *Visual Studio Code*. Após a conclusão da etapa de desenvolvimento, foi publicado o código compilado em um servidor *linux Ubuntu*¹⁷ que possibilita o acesso dos usuários através de uma URL

¹³ *React Native*: Linguagem de programação desenvolvida pela empresa Meta. Disponível em <https://reactnative.dev>. Acessado em 29 de janeiro de 2024.

¹⁴ *Visual Studio Code*: Editor de código fonte. Disponível em <https://code.visualstudio.com/learn>. Acessado em 29 de janeiro de 2024.

¹⁵ *MySQL*: Banco de dados relacional *open source*. Disponível em <https://www.mysql.com/why-mysql>. Acessado em 29 de janeiro de 2024.

¹⁶ *Angular*: Linguagem de programação mantida pela Google. Disponível em <https://angular.io/docs>. Acessado em 29 de janeiro de 2024.

¹⁷ *Ubuntu*: Versão de sistema operacional *Linux*. Disponível <https://ubuntu.com>. Acessado em 29 de janeiro de 2024.

pública. Para que o meio de acesso seja fácil, disponibilizou-se um endereço DNS¹⁸, desta forma um nome amigável pode ser vinculado ao endereço do servidor e permitir o acesso de uma forma simplificada. Opta-se pela disponibilização anônima do *site*, possibilitando acesso irrestrito a todo conteúdo gerado.

Na camada de API que é responsável por inserir, processar e armazenar o dado do aplicativo móvel, onde o *site* busca as informações para exibição, optou-se pela linguagem C#¹⁹ da *Microsoft*. Esta camada é fundamental para o desenvolvimento da solução pois é onde os algoritmos analisam os dados georreferenciais, criam as regras de priorização e seleção, fazem as montagens dos textos para a recomendação de criação de notícia e disponibiliza para exibição no *site*.

Ressalta-se que o código-fonte da solução integrada (*site* e aplicativo) é aberto e está armazenado e disponibilizado em um repositório público. A opção escolhida foi a plataforma *GitHub*²⁰ que oferece registro grátis e é comumente utilizada entre desenvolvedores de *softwares*.

3.3 Banco de dados

Em relação as estruturas das tabelas e colunas, manteve-se as existentes da versão piloto pois elas mostraram-se suficientes para atenderem os dados previstos. As tabelas e campos utilizados para o aplicativo e seus respectivos formatos de dados estão adicionados no apêndice A.

¹⁸ DNS: *Domain Name System* é o serviço que converte um endereço IP em um nome legível para acesso via URL.

¹⁹ C#: Linguagem de programação desenvolvida pela *Microsoft*. Disponível em <https://learn.microsoft.com/pt-br/dotnet/csharp/tour-of-csharp>. Acessado em 29 de janeiro de 2024.

²⁰ *GitHub*: Repositório de código fonte para projetos de *software*. Disponível em <https://github.com>. Acessado em 29 de janeiro de 2024.

3.4 Integrações sistêmicas

Optou-se por utilizar duas integrações externas para implementação de algumas das funcionalidades que compõem a solução.

A primeira é o *Google Maps API*²¹, que são um conjunto de APIs disponibilizados gratuitamente pela empresa *Google* que oferece o recurso de visualização em mapa e busca de endereço via longitude e latitude.

O segundo recurso é o *OpenAI APIs*²², que também é uma integração via API e possibilita a geração de texto generativo utilizando a ferramenta *ChatGPT*.

3.5 Recursos não funcionais

Para o desenvolvimento de um *software* deve-se considerar também os requisitos não funcionais. Enquanto os requisitos funcionais referem-se as telas, botões e eventos que o usuário realiza, os não funcionais compõem as características de segurança, integridade e escalabilidade. Devido ao grande número de aparelhos móveis no Brasil atualmente (Meirelles, 2023), esses requisitos tornam-se fundamentais para garantir o funcionamento apropriado das soluções e minimizar o risco para aqueles que as utilizam.

A disciplina de arquitetura de software lista diversos requisitos não funcionais que podem ser endereçados para melhor qualidade de uma solução. Para os *softwares* desta pesquisa foca-se nos itens de escalabilidade, confiabilidade, segurança e manutenção (Bass, Clements e Kazman, 2003).

- Escalabilidade refere-se à capacidade de um sistema aumentar e diminuir seus recursos para suportar um número maior de acesso, como também suportar novas funcionalidades.

²¹ *Google Maps API*: Disponível em <https://cloud.google.com/apis/docs/overview>. Acessado em 29 de janeiro de 2024.

²² *OpenAI API*: Disponível em <https://openai.com/blog/chatgpt>. Acessado em 29 de janeiro de 2024.

- Confiabilidade refere-se à habilidade de um sistema executar sua função sem falhas por um período de tempo definido.
- Segurança refere-se a permitir que apenas usuários autorizados possuem acesso a áreas e dados restritas.
- Manutenção refere-se à capacidade do sistema sofrer alterações de forma segura e rápida.

3.5.1 Escalabilidade

Implementou-se no aplicativo móvel e *site* recursos para habilitar a escalabilidade de forma apropriada e garantir o desempenho das funcionalidades sob diferentes volumes de dados. O primeiro recurso é o de micro serviço, onde a solução é dividida em pequenas partes que rodam de forma independente, mas que se complementam (Richardson, 2018, p.11). Outro recurso são os *containers*, que são pequenos ambientes isolados que recebem e executam o código-fonte dos micros serviços (Richardson, 2018, p.393). Cada *container* contém apenas um dos micros serviços (partes) da solução. Ao menos um *container* é necessário com cada parte para uma execução completa. Nesta pesquisa usou-se dois micros serviços sendo um para as APIs utilizadas pelo aplicativo móvel e um para as APIs e páginas utilizadas pelo *site*. São necessários dois *containers* para execução completa. Neste cenário, caso um dos *containers* falhe ou fique indisponível, basta iniciar um novo do mesmo tipo. A criação leva segundos pois utiliza uma imagem do ambiente que contém todo o código-fonte daquela parte e configurações pronto para uso.

Essa tecnologia também possibilita a escalação automática e gradual dos recursos. Quando o aplicativo ou *site* recebe muitos acessos, novos containers podem ser criados de forma automatizada possibilitando suprir a nova demanda. Após o retorno ao tráfego padrão, os mesmos podem ser deletados voltando a escala inicial definida.

3.5.2 Confiabilidade

Com o objetivo de diminuir os impactos causados por novas mudanças implementou-se testes unitários com cobertura de 90%. A cada nova versão, executa-se os testes para garantir que todas as unidades utilizadas pelo aplicativo móvel e *site* estão em funcionamento correto. Testes unitários são rápidos e eficientes para apontar falhas antes que o código-fonte seja entregue (Richardson, 2018, p.307).

3.5.3 Segurança

Tanto o aplicativo quanto o *site* utilizam APIs para integração entre o *frontend* e *backend*. Esta é uma tecnologia que permite que um *software* possa interagir com outros *softwares* através de um contrato que especifica o que pode ser lido, escrito, atualizado e deletado (Higginbotham; Safari, 2015). Para garantir uma comunicação autorizada implementou-se um *token* JWT (*Json Web Token*) para autenticação entre as partes. Quando o *site* é acessado, uma requisição é realizada para *backend* contendo esse token. Ele é avaliado e, caso válido, a requisição retorna com *status* de sucesso e os dados requisitados. Para geração do *token* é necessário um usuário e senha cadastrados no banco de dados. Sem essas informações o acesso não é autorizado. Essa é uma boa prática recomendada por padrões de projetos para garantir que usuários indevidos não acessem os dados da aplicação (Richardson, 2018, p.356).

3.5.4 Manutenção

Implementou-se um código-fonte em camadas que se conectam através de interfaces, possibilitando assim um baixo acoplamento e fácil manutenção. Este modelo também permite uma maior legibilidade o que resulta em uma melhor adaptação e interpretação da solução por novos desenvolvedores que venham a interagir com o código-fonte.

No que tange o armazenamento, utiliza-se *Git* para controle das versões. A cada nova mudança um novo identificador é gerado e atribuído a versão. Este recurso permite identificação de data, autor e linhas alteradas na ação, como

também retornar para uma versão anterior (Loeliger, Mccullough, 2012). O sistema de versão utilizado pelo *Git* é o semântico que é composto por três números (ex. 1.0.0) que representam o nível de complexidade da atualização (Richardson, 2018, p.69–70). Neste projeto estamos utilizando apenas as duas últimas posições.

3.6 Recursos funcionais

3.6.1 Aplicativo

Os requisitos foram definidos a partir do protótipo existente onde podia-se basicamente fazer o registro de usuário e cadastrar ocorrências. Buscando melhorar a experiência do usuário, foi proposto algumas melhorias como *login* automático após primeiro acesso, mapa geral com todos os cadastros, mapa específico com o local exato da ocorrência sendo cadastrada e contador de cadastros na tela de perfil.

Em relação ao *layout* seguiu-se a premissa de que o aplicativo deveria ser fácil e intuitivo, por isto optou-se por deixar todos os acessos na tela inicial utilizando ícones grandes. As cores, fontes e logo foram inspiradas no protótipo e logo oficial do projeto de extensão LupaNH.

Abaixo segue o quadro de requisitos e melhorias definidas durante o processo de levantamento e análise.

Quadro 7 - Requisitos do aplicativo móvel

	Funcionalidade	Campos	Objetivo
1	<i>Login</i>	<i>E-mail</i> e senha	Acessar o sistema
2	Cadastro de usuário	<i>E-mail</i> , senha, confirmar senha novamente e assinalar consentimento	Criar usuário
3	Recuperar senha	<i>E-mail</i>	Enviar e-mail com link para usuário redefinir a senha.
4	Iluminação Pública	Possui poste? As luzes estão funcionando? Há fios elétricos soltos? Cadastro de trecho?	Cadastrar ocorrência. É classificada como problema ou em conformidade dependendo dos valores

			selecionados. Quando optado por cadastro de trecho o GPS captura de um ponto A para B. O usuário deve se deslocar para registrar o trecho.
5	Limpeza Urbana	A prefeitura faz a limpeza? Se sim, qual a frequência semanal? Existe lixo acumulado? A prefeitura faz a capinagem?	Cadastrar ocorrência. É classificada como problema ou em conformidade dependendo dos valores selecionados
6	Calçadas e Asfalto	A via é asfaltada? A via possui buracos ou cratera? Há calçadas pavimentadas? Cadastro de trecho?	Cadastrar ocorrência. É classificada como problema ou em conformidade dependendo dos valores selecionados. Quando optado por cadastro de trecho o GPS captura de um ponto A para B. O usuário deve se deslocar para registrar o trecho.
7	Água Potável	Possui poço amazônico? Há água encanada? Quantos dias faltam água na semana? Alguma obra de saneamento sendo executada?	Cadastrar ocorrência. É classificada como problema ou em conformidade dependendo dos valores selecionados
8	Tratamento de Esgoto	Há coleta ou tratamento de esgoto? Possui fossa? Alguma obra de saneamento sendo executada?	Cadastrar ocorrência. É classificada como problema ou em conformidade dependendo dos valores selecionados
9	Coleta de Lixo	Há coleta de lixo? Qual a frequência semanal?	Cadastrar ocorrência. É classificada como problema ou em conformidade dependendo dos valores selecionados
10	Informação	Texto	Texto informando que o GPS deve estar ligado para possibilitar o cadastro

11	Dúvida	Texto	Texto divulgando links do <i>site</i> do projeto, redes sociais e políticas e termos
12	Perfil	Texto	Texto com nome do usuário, link para redefinição de senha, sair do sistema e contador de ocorrências cadastradas
13	Mapa	Mapa do <i>google maps</i>	Visualização em mapa de todas as ocorrências de todos os usuários. Opção de clicar na ocorrência e ver detalhes do preenchimento

Fonte: Autor

3.6.2 Site

Os requisitos do *site* foram pensados considerando um dos objetivos do projeto que é a disponibilização e compartilhamento dos dados capturados. Toda a disposição dos objetos em tela deveria transmitir a mensagem de que os dados estão naquele lugar para uso de todos. Com isso em mente, decidiu-se pelos seguintes requisitos:

Quadro 8 - Requisitos do *site*

	Requisito
1	Acessar de forma anônimo. Não há necessidade de usuário e senha
2	Utilizar página única que foque diretamente no objetivo
3	Utilizar mapa grande central com barra de filtro superior
4	Mapa deve apresentar todos as ocorrências onde ao clicar apresentar detalhes
5	Filtro por tipo de ocorrência, data início e fim e classificação
6	Opção de exportar para arquivo CSV
7	Opção de visualizar os dados em mapa, tabela e gráficos
8	Área ao lado do mapa para exibição das recomendações de criação de notícia

Fonte: Autor

3.6.3 Algoritmo de recomendação de criação de notícia

Os requisitos propostos para esse algoritmo foram pensados com o objetivo de apoiar o usuário que cadastrou dado via aplicativo e acessa *site* para visualização. O algoritmo proposto precisa combinar as informações do banco de dados e acessar as integrações de *Google Maps* e *OpenAI* de forma rápida para produzir um resultado em tela para o usuário. Para atingir esse objetivo foi proposto:

Quadro 9 - Requisitos do algoritmo de recomendação de criação de notícia

	Requisito
1	O algoritmo deve acessar todos os dados cadastrados no banco de dados que sejam classificados como problema
2	O algoritmo deve criar 2 <i>clusters</i> com os dados de acordo com sua geolocalização
3	O primeiro cluster deve conter ocorrência do mesmo tipo com proximidade menor ou igual a 50 metros
4	O segundo cluster deve conter ocorrência de tipo diferente com proximidade menor ou igual a 50 metros
5	O algoritmo deve utilizar os dados de longitude e latitude para integrar a API do <i>Google Maps</i> e identificar nome da rua e número, quando houver
6	O algoritmo deve utilizar o dado de endereço e tipo de problema para integrar a API do <i>OpenAI</i> e gerar o texto de recomendação de criação de notícia, em português
7	Uma nova recomendação deve ser apresentada na tela do <i>site</i> a cada carregamento utilizando um endereço sorteado no banco de dados.
8	O cluster deve influenciar a probabilidade de cada ocorrência ser sorteada.
9	Limite do tamanho do texto de recomendação deve ser 200 caracteres.

Fonte: Autor

3.7 Segurança e privacidade

Objetivou-se abordar alguns procedimentos de segurança para garantir a integridade dos dados assim como a segurança dos usuários dos sistemas. Alguns desses procedimentos aplicam-se a nível de servidor onde os *softwares* estão instalados, outros ao aplicativo móvel e outros ao *site*. Determinou-se que apenas o

aplicativo móvel exija um cadastro de usuário para acesso. O *site* possui acesso anônimo. Segue abaixo uma listagem dos procedimentos e seu escopo:

- Banco de dados *MySQL*: Utilizou-se uma senha alfanumérica de 8 dígitos gerado por um gerador de senha.
- Acesso ao servidor *linux* onde o *site* é hospedado: Utilizou-se uma senha alfanumérica de 8 dígitos gerado por um gerador de senha.
- Acesso do aplicativo móvel e do *site* ao banco de dados para buscar ou salvar registros: Utilizou-se APIs para esta comunicação. Durante uma transação a API utiliza o *token* JWT. Este *token* é uma chave única gerada para uma transação entre o aplicativo/*site* e o banco de dados. Para geração deste *token*, usa-se um usuário administrador e senha.
- Acesso do usuário: Requer um cadastro de usuário para habilitar acesso ao aplicativo móvel. O campo de *e-mail* e uma senha são requisitados e salvo no banco de dados. Essas informações são utilizadas somente para acesso ao sistema pelo próprio usuário. Todo registro de inclusão é vinculado a um *e-mail*, porém, não se usa ou exibe esta informação em nenhuma tela de consulta ou listagem tanto no aplicativo móvel quanto no *site*.
- Senha de acesso do usuário: Durante o registro de usuário é exigido uma senha de 6 dígitos. Essa informação é criptografada utilizando tecnologia *SymmetricSecurityKey*²³, sendo assim, dificulta-se a utilização da informação caso o dado seja acessado de forma indevida.
- Exclusão dos dados: Todos os registros efetuados por um usuário podem ser excluídos caso solicitado. Esta é a razão do vínculo de um registro e um usuário no banco de dados. Todas as informações acessíveis por usuários são anônimas.

²³ *SymmetricSecurityKey*: É uma tecnologia utilizada na programação para geração de *token* utilizando chave simétrica.

3.8 Termo de uso, política de privacidade, TCLE e TALE

Durante o registro de *login* de um usuário no aplicativo móvel, requer-se a adesão a TCLE (Termo de Consentimento Livre e Esclarecido) ou TALE (Termo de Assentimento Livre e Esclarecido). Executa-se essa ação através de um campo assinalado pelo usuário. A informação é salva no banco de dados vinculado ao *login* e data do aceite. Foi elaborado também um Termo de Uso e Política de Privacidade para esclarecimentos sobre o uso do dado e responsabilidades das partes. Todos esses documentos encontram-se anexados ao protocolo do comitê de ética da Unifesp. Também é possível acessá-los pelo aplicativo e *site* caso o usuário necessite consultá-los. É obrigatório o aceite dos documentos para ter acesso ao aplicativo móvel.

Capítulo 4

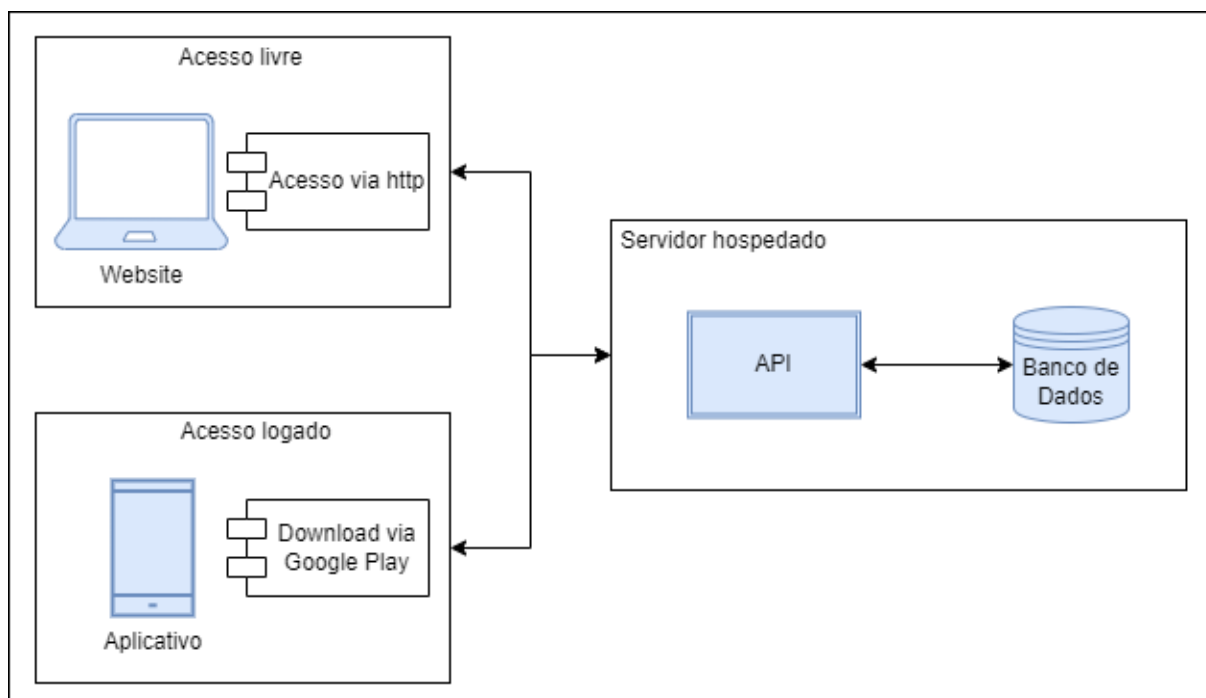
DESENVOLVIMENTO

O processo de desenvolvimento desta pesquisa dividiu-se em algumas etapas como é listado no item 3.1.3. Essas etapas podem ser agrupadas em itens técnicos, que sustentam e permitem que o trabalho seja realizado, e itens de pesquisa, que é a execução conceitual como revisão bibliográfica, estudo de caso em que os dados são captados e analisados.

Para atender o objetivo da pesquisa, onde tomou-se como premissa que soluções tecnológicas digitais que utilizam GIS e participação pública podem fornecer dados estruturados georreferenciados, além de fomentar a geração de notícia hiperlocal, foi proposto o desenvolvimento de uma solução integrada tecnológica que possibilite a criação de uma base de dados sobre infraestrutura, e a partir dela, gerar recomendações de criação de notícias.

Esta solução é composta por dois softwares, sendo o primeiro um aplicativo móvel para cadastro de dado, e o segundo um *site web* para visualização e compartilhamento do dado, como apresentado na figura 7. Para validação da solução, realizou-se um estudo de caso no bairro Novo Horizonte por intermédio da Escola Estadual Raimunda dos Passos Santos e de seus alunos da disciplina eletiva sobre técnicas jornalísticas. Através deles, executou-se a captação dos dados via aplicativo móvel e utilização das recomendações de criação de notícia via *site*.

Figura 7 - Diagrama de arquitetura



Fonte: Autor

As seções a seguir descrevem os itens fundamentais para compreensão e execução dessa pesquisa. Não serão apresentados detalhes técnicos do desenvolvimento das ferramentas utilizadas, a não ser que está se mostre relevante para pesquisa.

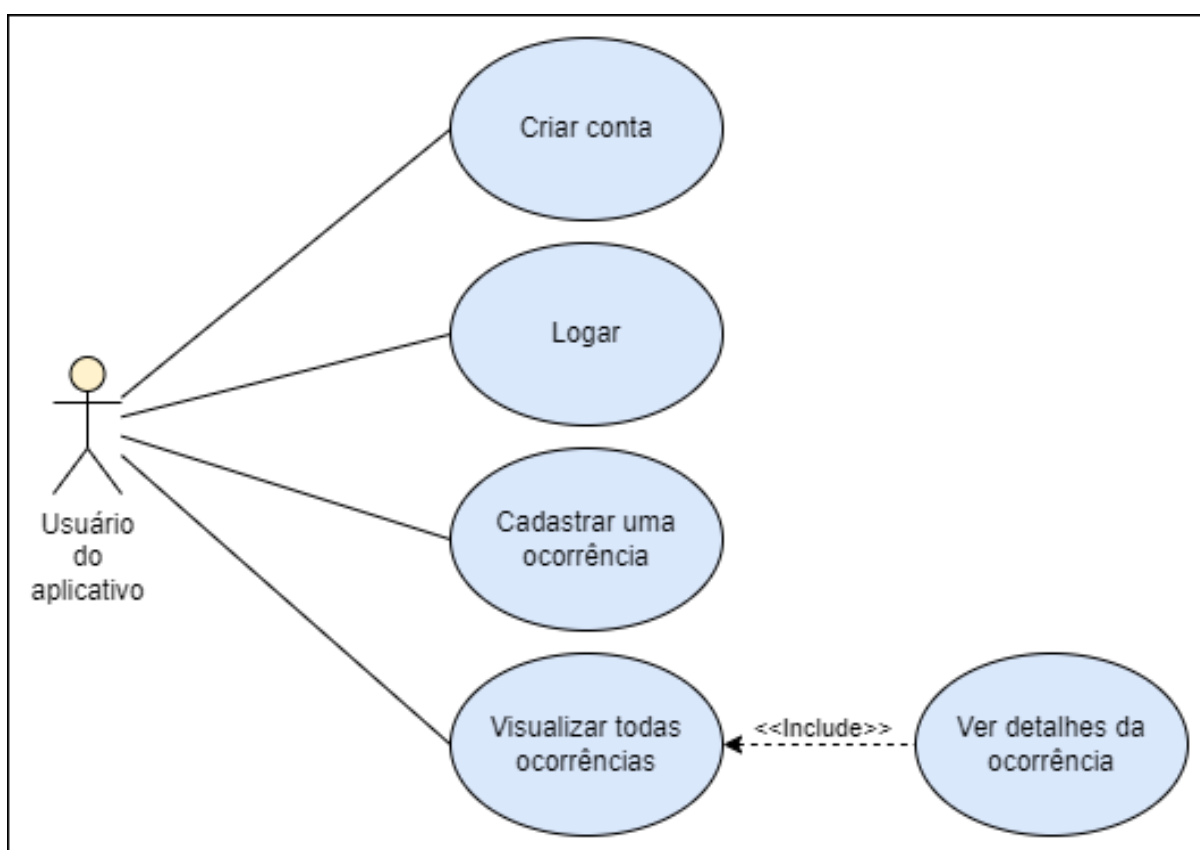
4.1 Aplicativo móvel

Para captação de dados desenvolveu-se um aplicativo móvel. Foram necessários dois meses de codificação para a conclusão desta etapa. Através de um emulador de celular disponibilizado pela ferramenta *Android Studio*²⁴, cada funcionalidade pode ser desenvolvida e testada apropriadamente. Utilizando também o *Android Studio* foi possível simular o aplicativo rodando em diversos modelos de aparelhos móveis, com diferentes resoluções de tela, *hardware* e versão do sistema operacional *Android*, incluindo *tablets*. Na imagem abaixo (Figura 8) temos o caso de uso com o ator “usuário do aplicativo” e as ações disponíveis a ele.

²⁴ *Android Studio*: Editor de código fonte desenvolvido pela Google e JetBrain.

Como pode-se notar o aplicativo contém apenas um perfil de acesso que executa todas as ações. Não foi previsto propositalmente a ação de deletar ocorrências. Este deve ser feita diretamente via banco de dados.

Figura 8 - Caso de uso do aplicativo móvel



Fonte: Autor

Para o primeiro acesso ao aplicativo, pede-se um cadastro de uma conta utilizando um *e-mail* e o aceite dos termos da pesquisa. Essa etapa cobre os requisitos exigidos pelo comitê de ética e detalhado no item 3.7. A partir do segundo acesso a senha não é requisitada novamente pois utilizou-se um sistema de armazenamento na memória do aparelho móvel. Desta forma, o acesso é facilitado, entretanto, é possível utilizar o botão “sair” que é localizado na tela de perfil, fazendo com que a senha volte a ser exigida no próximo acesso. Nesta mesma tela de perfil há o recurso de atualização de senha. O usuário recebe um *link* no *e-mail* cadastrado com o procedimento. Este *link* também é utilizado para o processo de recuperação de senha, que fica na tela de *login*, quando um usuário encontra

difficuldade ao acessar. Cada etapa deste processo foi desenhada para facilitar o acesso de quaisquer perfis de pessoa, familiarizados com tecnologia ou não.

No que tange as ocorrências disponíveis para cadastro, considerou-se as listadas nos requisitos mencionado no capítulo 3.6.1. Destaca-se a melhoria para adicionar um mapa na parte superior das telas de cadastro de ocorrência, possibilitando ao usuário visualizar, e caso necessário, alterar a localização do cadastro. Este recurso visa minimizar erros de precisão do GPS que tendem a acontecer dependendo do sinal do aparelho e da região. Como resultado espera-se aumentar a credibilidade e precisão do dado.

Para as ocorrências calçadas e asfalto e iluminação pública foi adicionado a opção de registro em formato de trecho. Essa opção habilita um botão onde ao acionado, o usuário desloca-se de um ponto A ao B, e o aplicativo captura a longitude e latitude do trecho, criando no mapa uma linha, diferente do registro habitual onde cria-se um ponto. Essa opção de registro reflete melhor a característica de ruas que não possuem calçamentos ou iluminação, pois pode-se adicionar um registro de trecho para toda a extensão da via, ao contrário do registro por ponto, que necessitaria de múltiplos pontos para representar a extensão da rua.

Outro recurso é a disponibilização de uma visualização geral em mapa de todos os cadastros na página inicial do aplicativo. Esse recurso permite saber se uma ocorrência deparada por um usuário já está cadastrado ou não. Caso esteja, ao clicar na ocorrência é exibido seu detalhe e data de criação. Todas as ocorrências cadastradas são salvas no banco de dados utilizando APIs.

Durante o processo de desenvolvimento notou-se que seria possível cadastrar ocorrências onde dependendo da combinação de respostas selecionadas nos campos, significaria que não há problema identificado naquele ponto. Como exemplo podemos ter um cadastro de coleta de lixo onde as respostas para “Há coleta de lixo?” seja sim e “Qual a frequência semanal?” seja 7. Essa informação não poderia ser apresentada no *site* ou nas recomendações de criação de notícia como um problema identificado. Tendo isso, foi definido um novo requisito onde dado a combinação das respostas, uma ocorrência poderia ser classificada como “problema” ou “em conformidade”. Caso problema, o ícone no aplicativo e *site* ficaria em amarelo, caso em conformidade, o ícone ficaria em azul. No *site* foi adicionado um botão para filtrar a classificação.

Na etapa de publicação na *PlayStore* foi necessário um registro prévio com uma conta *Google*. Exigiu-se algumas etapas de configuração onde inicialmente um público-alvo é definido para realização do teste beta. Apenas contas *Google* pré-cadastrados possuem autorização para baixar e instalar o aplicativo. Duas semanas de pré-teste foram executadas pelo próprio desenvolvedor antes da abertura do aplicativo para livre instalação.

Também foi exigido a elaboração do termo de compromisso e da política de segurança, ambas prévia a disponibilização do aplicativo. Essas informações devem estar disponíveis tanto dentro do aplicativo quanto na *PlayStore*, onde os usuários possam consultar e entender quais os possíveis riscos e medidas a serem tomadas em caso de qualquer problema.

A *Google* tem exigido cada vez mais comprometimento com a segurança e integridade dos aplicativos disponíveis em sua loja. Após cada publicação, pode-se levar até 3 dias para a liberação, pois a *Google* executa testes e procedimentos de validação para analisar se o aplicativo está dentro dos padrões exigidos. A não conformidade impossibilita a publicação ou a remoção no caso de um aplicativo já existente.

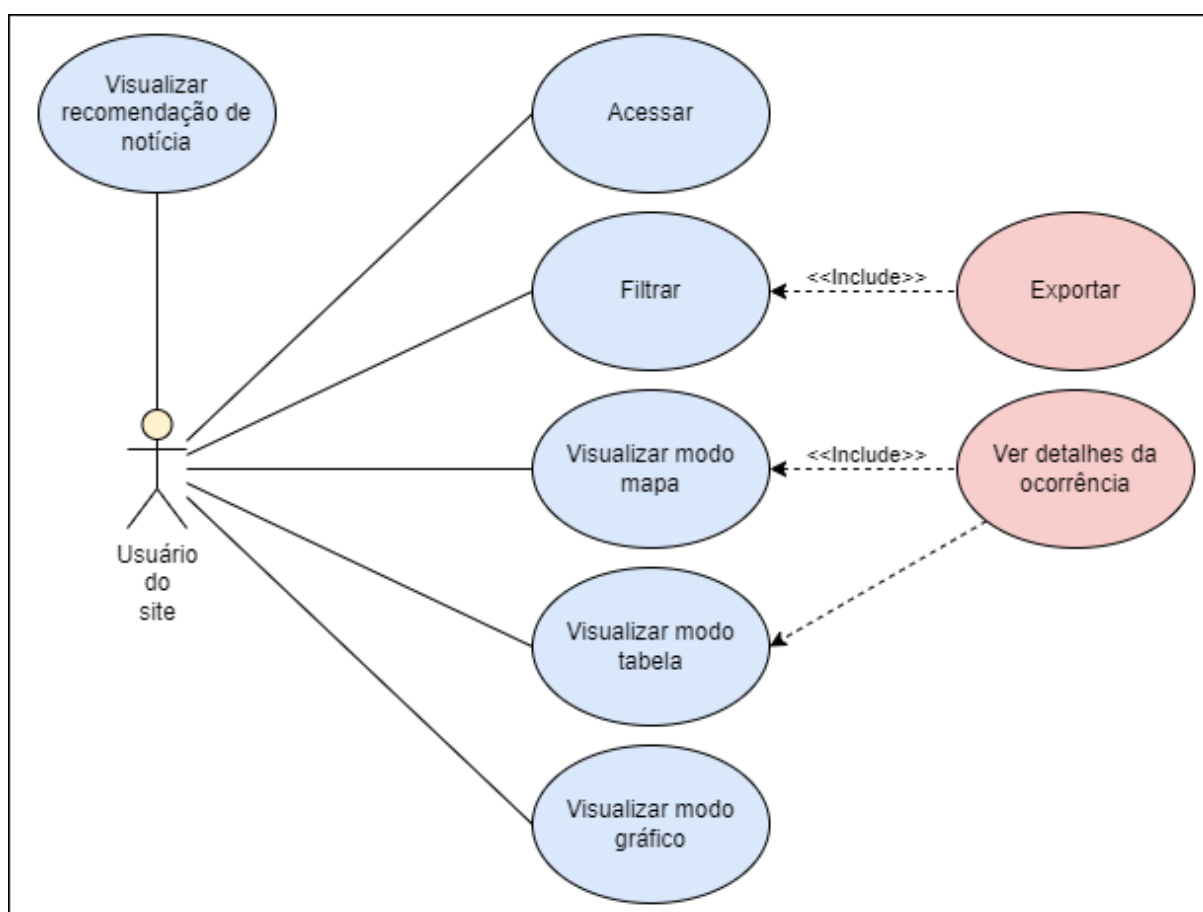
4.2 Site

O desenvolvimento de um *site* utiliza diferentes recursos tecnológicos se comparado a um aplicativo. Como exemplo, em um aplicativo móvel é necessário o uso de um emulador *Android* ou *Apple*. Já no desenvolvimento *web* deve-se lidar com diferentes versões de navegadores além de garantir que a visualização via aparelho móvel esteja disponível também.

Assim como no aplicativo móvel, o *site* contou com a utilização de APIs. Estas são as responsáveis por conectar no banco de dados e trazer os dados de forma estruturada para apresentação na página. Optou-se por desenvolver uma página com visualização única de um grande mapa. Este recurso é possível através do uso do *Google Maps API*. Com ele pôde-se adicionar um mapa em um *site* e criar integrações a ele. Os dados de longitude e latitude armazenados no banco de dados

são integrados ao mapa onde marcadores são criados para sinalizar uma ocorrência. Para exibição dos detalhes do cadastro utilizou-se o recurso de *lightbox*, onde ao clicar no marcador, uma caixa abre-se para a apresentação dos detalhes da ocorrência. Como apresentado na figura 8, o *site* não possui recurso de inserção ou atualização de dados, pois seu acesso é anônimo e permitir o cadastro desta forma poderia comprometer a qualidade dos dados. Assim como o aplicativo, o acesso ao *site* possui apenas um ator, não existindo assim um acesso de administrador. O *layout* utilizado é adaptado também do protótipo móvel existente. O logotipo, cores, fontes e alguns ícones foram reutilizados garantindo assim uma harmonia visual entre os *softwares*.

Figura 9 - Caso de uso do *site*



Fonte: Autor

Dispõe-se também na página inicial alguns recursos para melhor utilização do *site*. O visitante pode gerar uma listagem dos registros com opção de filtro por data,

tipo e classificação, além de exportação em arquivo CSV²⁵ (*Comma-Separated Values*).

Foi adicionado na lateral esquerda do *síte* a funcionalidade de recomendação de criação de notícia onde a cada carregamento uma nova recomendação é gerada. Os critérios para essa exibição são os calculados pelo algoritmo de prioridade e o texto gerado pela AI generativa de texto, neste caso *ChatGPT*. Caso o usuário não queira utilizar a sugestão apresentada, deve-se atualizar a página para uma nova sugestão ser carregada. Durante o desenvolvimento adicionou-se a opção de habilitar um *cache* para essa funcionalidade. O intuito foi evitar que múltiplas atualizações ao mesmo tempo na tela sobrecarregasse o servidor. Desta forma, atualizações dentro do intervalo de 1 minuto receberão a mesma recomendação. Esse tempo é configurável.

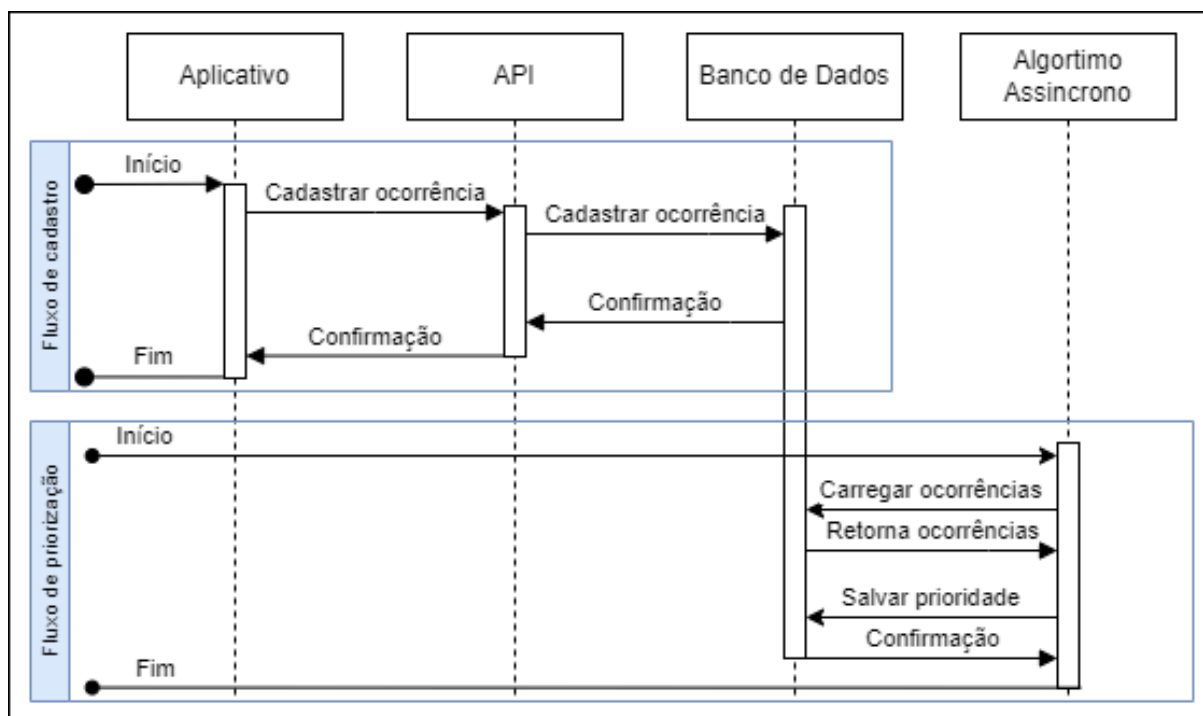
Para a publicação do *síte* e sua disponibilização para acesso foi utilizado uma ferramenta chamada *Nginx*. Esta é responsável por garantir que o *síte* publicado no servidor *linux* Ubuntu possa ser acessado via navegador por qualquer usuário.

4.3 Algoritmo de recomendação de criação de notícia

Dado os requisitos apresentados no capítulo 3.6.3, foi necessário dividir a implementação deste recurso em 4 etapas. A primeira acontece após a inserção de uma ocorrência de qualquer tipo no banco de dado via aplicativo móvel. Após o item ser salvo e a confirmação enviada ao usuário, é iniciado um processo assíncrono. Um processo assíncrono em *software* se refere a uma atividade que é executada no servidor após a conclusão da interação do usuário. Como mostra figura 10.

²⁵ CSV: Formato de arquivo onde os valores são separados por vírgula.

Figura 10 - Diagrama de fluxo do cadastro de ocorrência e processo assíncrono



Fonte: Autor

Neste caso, o processo assíncrono é um algoritmo que acessa o banco de dados a cada 10 minutos (configurável) e verifica a existência de novos cadastros. Caso sim, faz-se uma comparação entre a distância em metros da nova ocorrência e cada uma das ocorrências existentes na base de dados. Caso uma ocorrência esteja a 50 metros ou menos de distância, ambas recebem uma marcação de que são um *cluster* de proximidade. Caso elas sejam do mesmo tipo de ocorrência, ambas recebem uma marcação de que são um *cluster* de proximidade e tipo. Ao final deste processo, esse resultado gera uma tabela nomeada como TBPrioridade contendo o ID da ocorrência, o ID do tipo da ocorrência, uma marcação identificando se é um cluster de proximidade e uma marcação se é um *cluster* de proximidade e tipo.

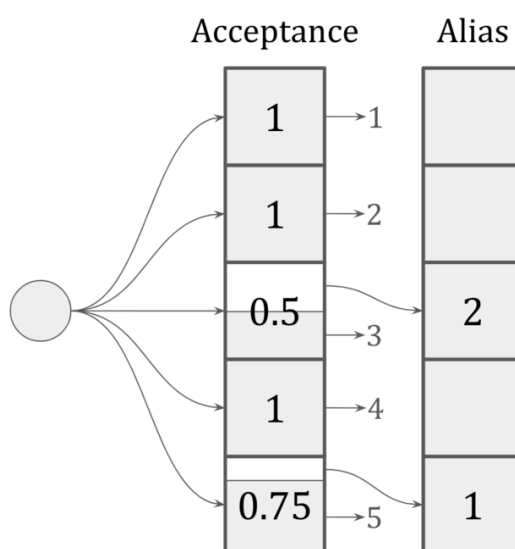
Na segunda etapa é executado o cálculo do peso da ocorrência. Todos os itens da tabela de prioridade são lidos e uma pontuação é adicionada de acordo com o *cluster* pertencente. Caso uma ocorrência pertença a um *cluster* de proximidade, este recebe + 3 de ponto. Caso pertença a um *cluster* de proximidade e tipo, este recebe +5. Caso não pertença a nenhum *cluster*, o item permanece com a pontuação de +1 ponto. O cálculo do peso é realizado através da equação abaixo, onde, faz-se a soma do total de pontos da ocorrência, multiplicado por 100 e dividido

pelo total de pontos de todas as ocorrências combinadas. O valor final representa o percentual de chance daquela ocorrência ser selecionada para a recomendação de criação de notícia. Este total também é salvo na tabela de prioridade.

$$PESO = \frac{(PONTOS * 100)}{TOTALPONTOS}$$

A terceira etapa é responsável pelo sorteio da ocorrência. O algoritmo carrega em uma lista o ID da ocorrência e o peso calculado na etapa anterior. A partir desta lista é utilizado o método *Alias Method* (Figura 11) para realização da seleção. De acordo com Walkter (1974) este método consiste essencialmente na manipulação da amplitude de números sequenciais estatisticamente independentes e uniformemente distribuídos. Na implementação este método executa uma seleção randômica considerando o peso de cada item da lista. Itens com maior peso possuem maior chance de serem selecionados. O método *Alias Method* garante que o desempenho da seleção seja otimizado comparado a outros mecanismos de sorteio randômico por peso, pois mesmo com pesos diferentes cada item ocupa um mesmo tamanho de espaço na lista.

Figura 11 - Probabilidade de distribuição utilizando *Alias Method*



Fonte: Autor

Com o item selecionado, a quarta e última etapa é executada. O algoritmo monta os requisitos do texto a ser enviado a API do *ChatGPT* para geração da recomendação. Como mostrado no quadro 4, é utilizado o tipo de ocorrência (t2) e o endereço, que é buscado através de uma integração com o *Google Maps API* utilizando o dado de latitude e longitude (t4). Com essas duas informações o algoritmo intercala com outros textos base carregados do banco de dados (t1, t3, t5 e t6) e gera a pergunta a ser enviado a integração com o *ChatGPT*. A linha t1 possui uma variação de 10 modelos para requisitar diferentes tipos de escrita. É necessário adicionar na pergunta a informação para escrever em português, pois a linguagem padrão da API é inglês (t6).

Quadro 10 - Estrutura de montagem do texto para envio ao *ChatGPT*

t1	<i>Recommend a person to write a post informing citizen about</i>
t2	<i>the poor garbage collection system</i>
t3	<i>at</i>
t4	<i>Maria das Flores street</i>
t5	<i>in Novo Horizonte neighborhood</i>
t6	<i>Write in brazilian portugues.</i>

Fonte: Autor

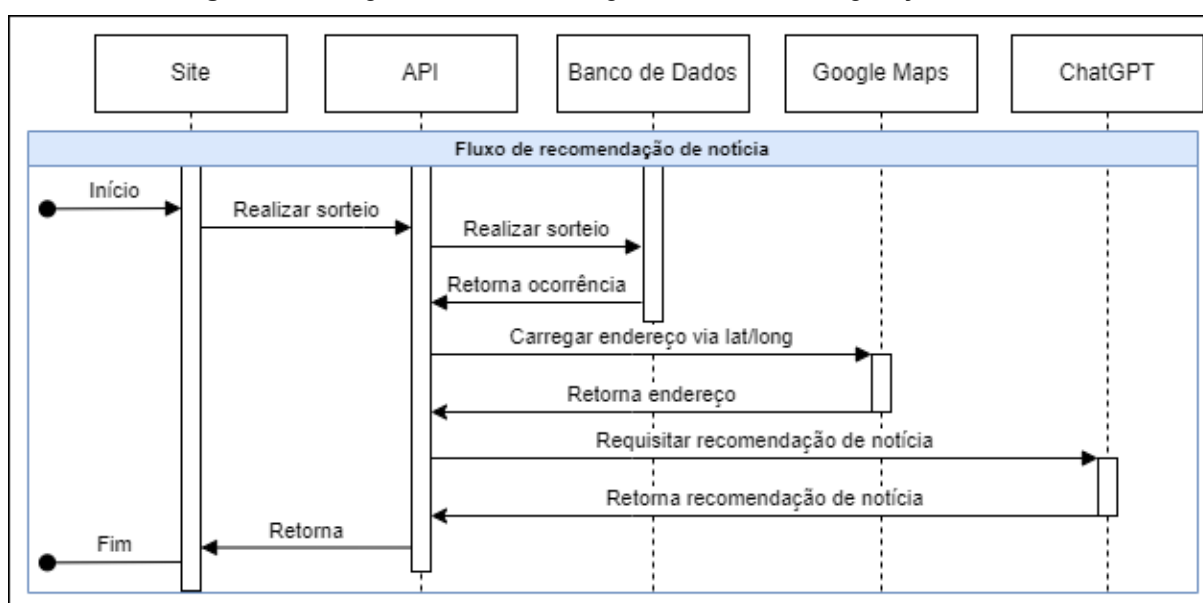
O quadro 10 acima gera o seguinte texto:

Exemplo de envio: *"Recommend a person to write a post informing citizen about the poor garbage collection system at Maria das Flores street in Novo Horizonte neighborhood. Write in portugues"*.

Exemplo de resultado: "Escreva uma postagem para informar os cidadãos sobre o deficiente sistema de coleta de lixo na Rua Maria das Flores, no bairro Novo Horizonte. Destaque os problemas enfrentados pelos moradores, como atrasos na coleta, acúmulo de lixo nas calçadas e os impactos negativos desta situação. Incentive a comunidade a se unir, reportar as irregularidades às autoridades competentes e exigir melhorias imediatas nesse serviço essencial. Juntos, podemos transformar a Rua Maria das Flores em um local mais limpo e saudável! #ColetaDeLixoMariaDasFlores #NovoHorizonteLimpo".

Ao final, o resultado retornado do *ChatGPT* é enviado diretamente para a tela. Caso o usuário recarregue a página, o processo é reexecutado a partir da etapa 3 onde apenas um novo sorteio e a integração com *ChatGPT* é executado. As etapas de geração de *cluster* e cálculo de ponto e peso são executadas apenas após uma inserção de dados via aplicativo. Desta forma ao acessar o *site*, a tabela *TBPrioridade* já está pronta para ser utilizada.

Figura 12 - Diagrama de fluxo do algoritmo de sorteio e geração de texto

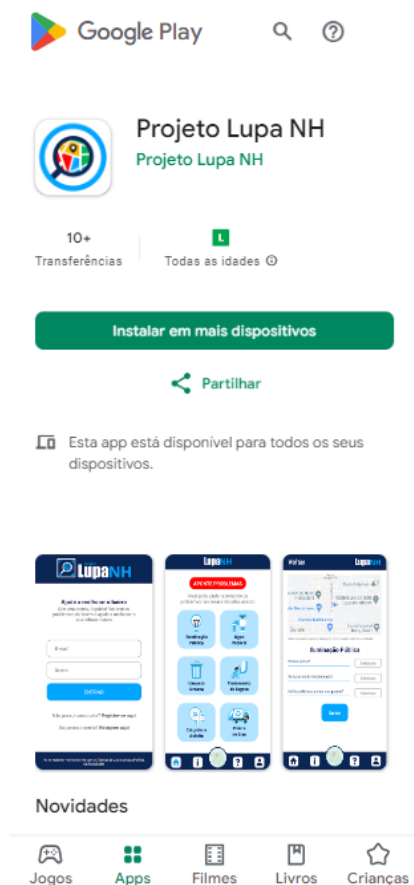


Fonte: Autor

4.4 Publicação e divulgação

Após concluídos os desenvolvimentos, os recursos técnicos devem ser divulgados para uso. Foi optado em disponibilizar o *link* para acesso dentro do *site* do projeto Lupa NH, pois é um lugar já conhecido e divulgado, possibilitando também que um visitante possa conhecer demais ações que são executados por este projeto e acompanhar outras iniciativas. Referente ao aplicativo, ele pode ser encontrado na *PlayStore* ao realizar a busca pelo termo "Lupa NH".

Figura 13 - Aplicativo disponível na loja da *Google*



Fonte: *Google PlayStore*

Com os itens publicados e disponíveis, a execução da etapa de estudo de caso foi planejada para acontecer durante as atividades da disciplina eletiva do projeto Lupa NH na escola estadual Raimunda dos Passos Santos. Os professores desta disciplina foram convidados assim como os alunos a participarem da coleta dos dados e utilização do algoritmo de recomendação de criação de notícia. O convite formal realizou-se através de um *e-mail* aos coordenadores da disciplina eletiva. Após o aceite, o trabalho foi organizado.

4.5 Captação de dados

Para captação de dados planejou-se executar três dias de trabalho com a duração de 2 horas de atividade cada. As datas foram 15, 16 e 17 de novembro no

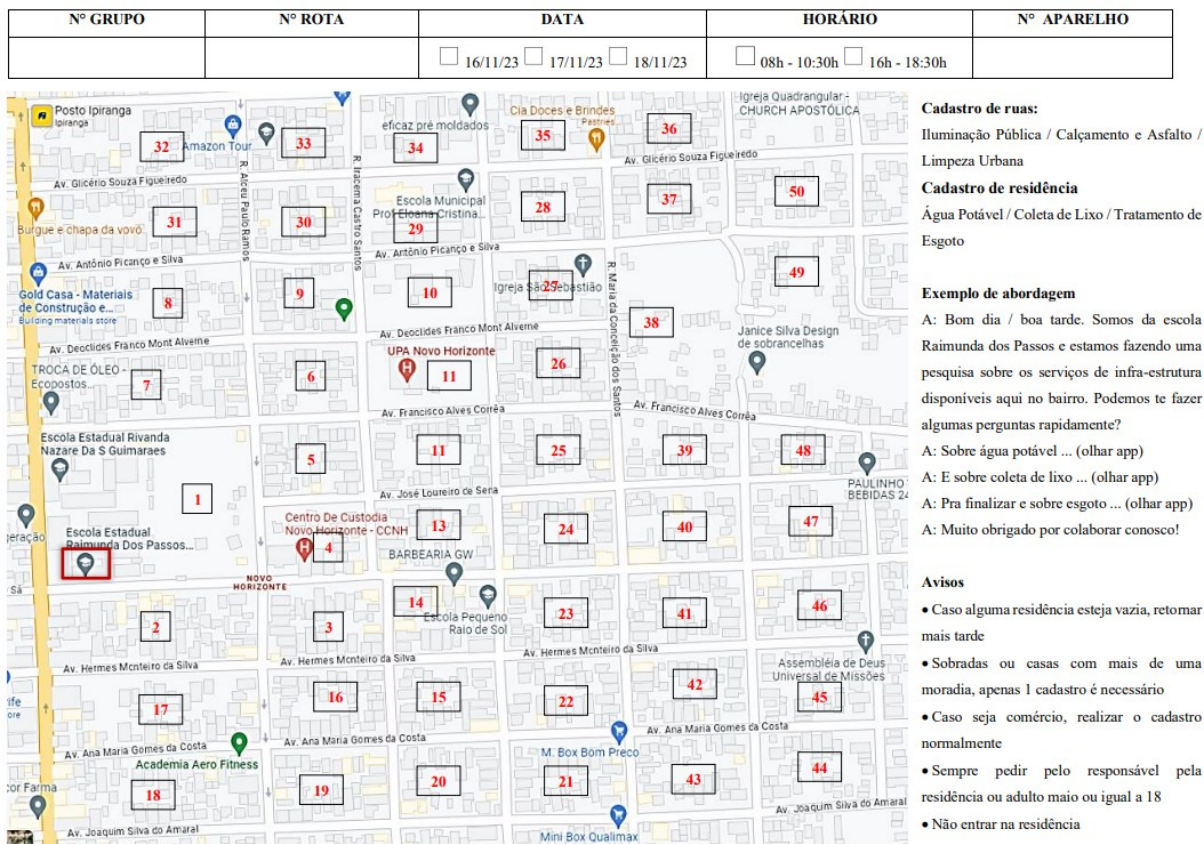
horário das 08hs às 10hs e 16:00hs às 18:00hs. Optou-se pelo período de início da manhã e final da tarde devido as altas temperaturas na região. O local de encontro foi uma sala de aula cedida pela escola.

O primeiro turno da manhã do dia 15 foi utilizado para apresentar o objetivo da atividade, cadência do trabalho, avisos de segurança e divisão dos grupos. Cada grupo foi formado por 5 ou mais pessoas, sendo a maior parte estudantes da escola e ao menos 1 professor ou pesquisador em posição de monitor e facilitador. Cada grupo foi designado a uma rota pré-estipulada e um período de 2:00hs para percorrê-la. O monitor do grupo foi responsável por gerenciar a captação, auxiliar nos diálogos com moradores e garantir que o trajeto fosse cumprido.

A separação das rotas foi elaborada através de uma imagem do bairro capturada via *Google Maps* onde cada quadra foi demarcada com um número sequencial de identificação da rota. O documento também possuía um cabeçalho para identificação dos grupos, data, horário e aparelho utilizado. Na lateral foi adicionado alguns informativos e textos de ajuda para consulta rápida. O documento em formato PDF foi impresso e entregue para cada equipe (figura 13).

Previamente foram adquiridos cinco celulares com chip e pacote de *internet* para uso durante a captação. O aplicativo foi baixado, instalado e um usuário criado para uso. Essa etapa foi planejada para garantir que existissem aparelhos já em funcionamento no início do trabalho de captação. Entretanto, foi solicitado a todos que caso desejassem poderiam instalar o aplicativo e criar um usuário para uso. O líder do grupo foi aconselhado a rotacionar o aparelho do projeto para garantir que todos os alunos tivessem a oportunidade de fazer um cadastro de ocorrência.

Figura 14 - Rota de captura de dado



Fonte: Autor

Dentre as ocorrências disponíveis no aplicativo para cadastro, três são referentes a moradias: Água potável, coleta e tratamento de esgoto, coleta de lixo, e três referentes às vias: Calçadas e asfalto, iluminação pública e limpeza urbana. Desta forma, recomendou-se ao grupo a separação em dois subgrupos onde um cobriria as moradias e o registro das vias.

Para garantir a segurança dos alunos durante a execução da atividade, foi requisitado aos professores que identificasse no mapa os pontos de risco onde não deveríamos nos deslocar.

No que tange os recursos técnicos, durante a atividade foi planejado para que um computador com acesso à *internet* estivesse disponível para que fosse possível acessar o banco de dados validando assim os dados inseridos e o funcionamento do aplicativo.

4.6 Enquetes

Para analisar a experiência do usuário durante o uso do aplicativo e do *site* na captação de dados, elaborou-se duas enquetes estruturadas, onde a primeira com cinco perguntas focada no aplicativo móvel, e a segunda com seis perguntas focada no *site*. O público-alvo na enquete do aplicativo são os alunos participantes da atividade de coleta de dados. Já o público-alvo da enquete do *site* são os professores(as) participantes das oficinas da disciplina eletiva. O objetivo das enquetes não é realizar uma análise do impacto da pesquisa no aluno ou demais participantes, mas sim entender de forma quantitativa se o aplicativo e o *site*, da forma que foram desenvolvidos, atendem seus requisitos especificados. Também espera-se identificar pontos de melhoria que apenas poderiam ser identificados durante o uso em massa, pois um mesmo ponto, mesmo que superficial, quando apontado por muitos, pode indicar que existe um potencial de melhoria.

Como metodologia, utilizou-se a *Scale* de *Likert* para classificar o resultado de cada pergunta de 1 a 5, sendo 1 o valor para representar que não atendeu o requisito e 5 para representar que o requisito foi bem aceito. Este já é um modelo de avaliação com uma lógica difundida e de fácil entendimento, minimizando chances de má interpretação pelo participante.

Figura 15 – Enquete do aplicativo móvel

Considerando as perguntas abaixo, selecione um valor na escala de 1 a 5							
Quão fácil é operar o aplicativo	Difícil	1	2	3	4	5	Fácil
Sobre a clareza das perguntas	Confuso	1	2	3	4	5	Claro
Sobre o tempo utilizado para efetuar um cadastro	Lento	1	2	3	4	5	Rápido
Sobre interface/telas	Pouco agradável	1	2	3	4	5	Muito agradável
Quão relevante você considera as informações contidas no aplicativo?	Pouco relevante	1	2	3	4	5	Muito relevante

Fonte: Autor

Figura 16 – Enquete do *site*

Considerando as perguntas abaixo, selecione um valor na escala de 1 a 5						
Você tinha conhecimento da dimensão quantitativa dos problemas de infraestrutura apresentados site site?	Pouco	1	2	3	4	5 Muito
Você acredita que o site pode colaborar com informações estruturadas sobre infraestrutura para criação de uma base de conhecimento pública?	Não acredito	1	2	3	4	5 Acredito
Você acredita que a funcionalidade de recomendação de notícia pode colaborar com a criação de conteúdo pelos alunos informativo como notícias e postes nas redes sociais?	Não acredito	1	2	3	4	5 Acredito
Os dados apresentados geolocalizados em um mapa dão uma melhor visibilidade aos problemas da comunidade?	Não	1	2	3	4	5 Sim
Você acredita que o site pode colaborar na formação cidadã dos alunos e residentes do bairro?	Não acredito	1	2	3	4	5 Acredito
Quão relevante para sociedade você considera as informações contidas no site?	Pouco	1	2	3	4	5 Muito

Fonte: Autor

No caso do aplicativo móvel, a enquete foi desenvolvida em arquivo *Excel* e impressa para uso durante a captação. Ela foi anônima para incentivar respostas sinceras e livre de julgamento. Optou-se pela aplicação presencial e em papel para garantir maior número de respostas e evitar complexidades sistêmica que pode ser uma barreira para alguns alunos. No formato *online*, para garantir que um aluno responda apenas uma vez, teríamos que exigir um *e-mail*, impactando a característica anônima da enquete e dificultando para aqueles que não possuem.

Para a enquete do *site*, escolheu-se pela versão *online* utilizando *Google Forms*, também anônima pois não foi exigido a identificação por *e-mail*. Foi optado por este formato pois como o público-alvo foram os professores(as), entendesse que existe uma maior facilidade para uso dessas ferramentas. Outro fator foi a necessidade de esperar as oficinas subsequentes a atividade de coleta de dados, pois seriam quando os professores(as) teriam acesso ao *site* com conteúdo. Nesse período o time da pesquisa não estaria presente mais no local, sendo necessário o formato *online*.

Capítulo 5

RESULTADOS E DISCUSSÃO

5.1 Teste de desempenho

O objetivo de um teste de desempenho é mapear como uma solução se comporta ao receber um grande número de requisições em um curto espaço de tempo. Com este tipo de teste é possível identificar pontos onde a solução apresenta lentidão e quais os recursos mínimos de servidor para a sua execução.

Para a execução desta atividade utilizou-se a ferramenta *Postman*. Foi necessário definir uma configuração de carga, onde os parâmetros exigidos foram o alvo, tempo de execução, número de usuário e *rampup* (quadro 11). O alvo foi definido como as APIs que enviam e recebem os dados do aplicativo e *site*. Tempo de execução foi definido como 10 minutos, e representa o tempo que os testes ficarão enviando requisições as APIs. Número de usuários foi definido como 20, que representa o número de execuções paralelas. *Rampup* foi definido como 5 minutos, que é o tempo para atingir o número máximo de usuários definidos.

Quadro 11 - Configuração de carga

Alvo	APIs
Tempo de execução	10 minutos
Número de usuários	20
<i>Rampup</i>	5 minutos (Inicia o teste com 1 usuário e aumenta gradativamente até 20 usuários aos 5 minutos).

Fonte: Autor

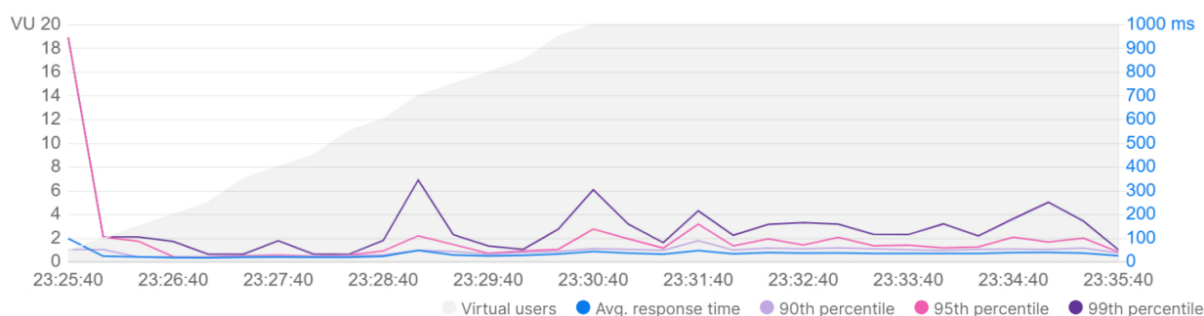
As execuções e seus resultados estão representados nas tabelas 1, 2, 3, 4, 5, 6 e 7, e figura 15, 16, 17, 18, 19, 20 e 21:

Asfalto e calçadas

Tabela 1 - Resultado do desempenho da API de asfalto e calçadas

Total requisições	Requisições por segundo	Mínimo (MS)	Média (MS)	90% (MS)	Max (MS)	Erros %
8.181	13.34	15	35	50	1.723	0

Fonte: Autor

Figura 17 - Gráfico do desempenho da API de asfalto e calçadas

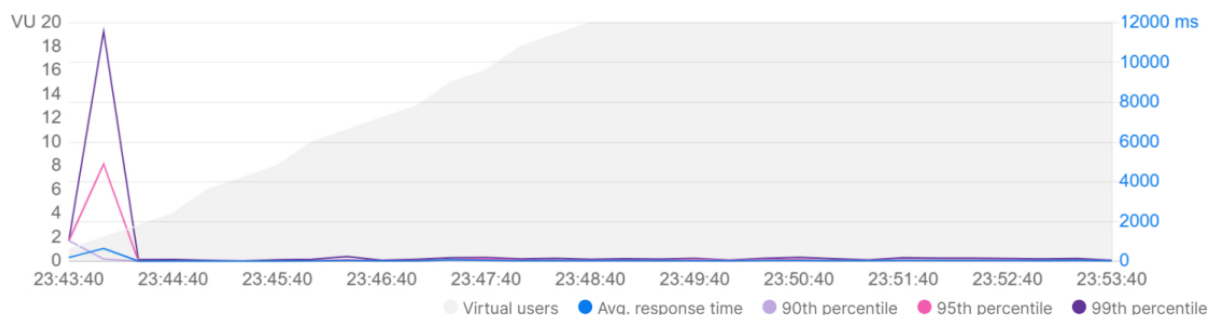
Fonte: Autor

Coleta de lixo

Tabela 2 - Resultado do desempenho da API de coleta de lixo

Total requisições	Requisições por segundo	Mínimo (MS)	Média (MS)	90% (MS)	Max (MS)	Erros %
8.031	13.04	15	40	53	11.561	0

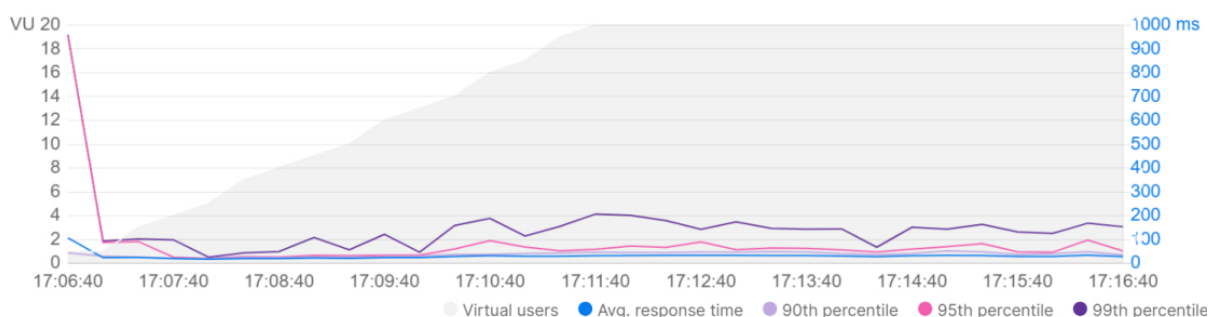
Fonte: Autor

Figura 18 - Gráfico do desempenho da API de coleta de lixo**Fonte:** Autor

Iluminação pública

Tabela 3 - Resultado do desempenho da API de iluminação pública

Total requisições	Requisições por segundo	Mínimo (MS)	Média (MS)	90% (MS)	Max (MS)	Erros %
8.168	13.36	14	30	42	960	0

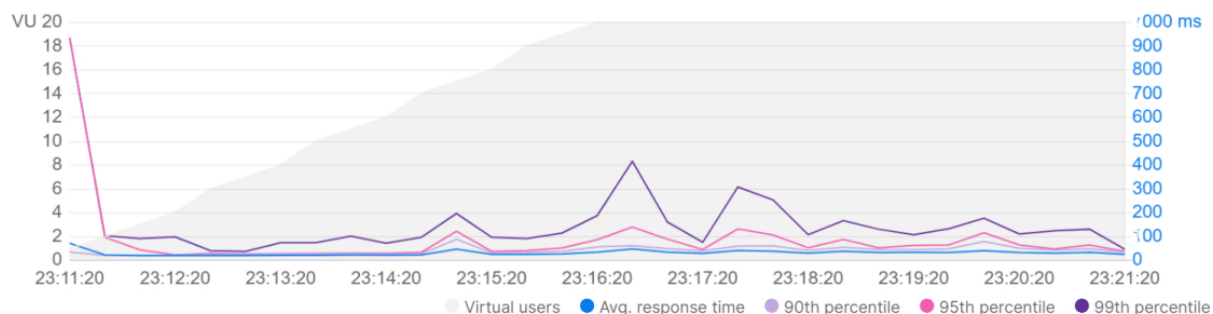
Fonte: Autor**Figura 19 - Gráfico do desempenho da API de iluminação pública****Fonte:** Autor

Coleta de esgoto

Tabela 4 - Resultado do desempenho da API de coleta de esgoto

Total requisições	Requisições por segundo	Mínimo (MS)	Média (MS)	90% (MS)	Max (MS)	Erros %
8.145	13.31	14	32	45	1.739	0

Fonte: Autor

Figura 20 - Gráfico do desempenho da API de coleta de esgoto

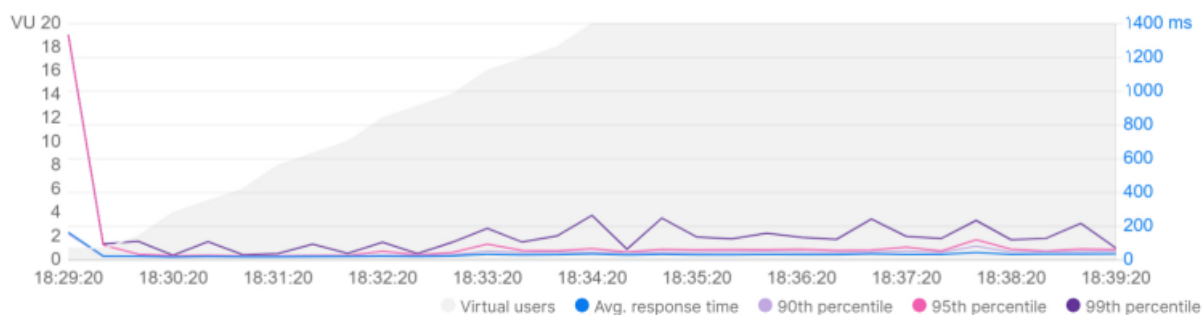
Fonte: Autor

Limpeza Urbana

Tabela 5 - Resultado do desempenho da API de limpeza urbana

Total requisições	Requisições por segundo	Mínimo (MS)	Média (MS)	90% (MS)	Max (MS)	Erros %
8.178	13.39	15	32	46	1.334	0

Fonte: Autor

Figura 21 - Gráfico do desempenho da API de limpeza urbana

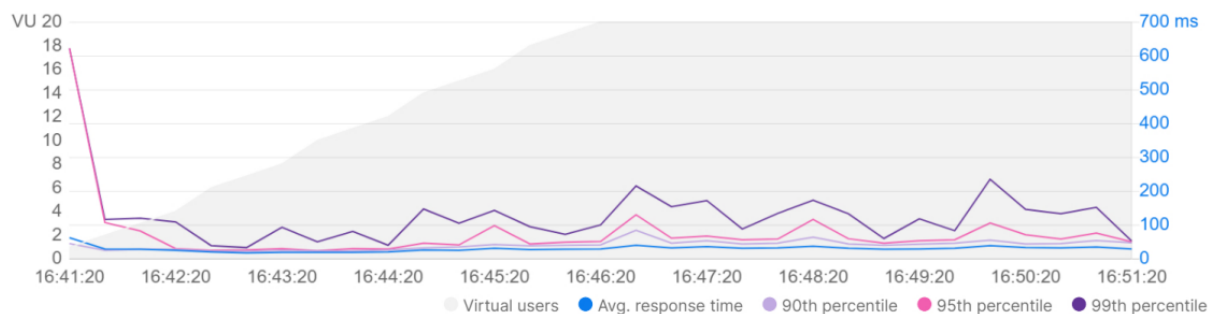
Fonte: Autor

Água potável

Tabela 6 - Resultado do desempenho da API de água potável

Total requisições	Requisições por segundo	Mínimo (MS)	Média (MS)	90% (MS)	Max (MS)	Erros %
8.181	13.35	14	31	44	623	0

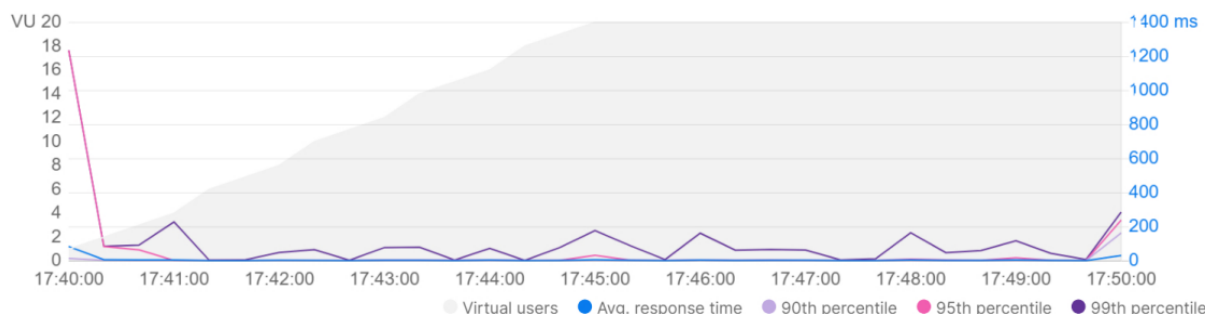
Fonte: Autor

Figura 22 - Gráfico do desempenho da API de água potável**Fonte:** Autor

Algoritmo de recomendação de criação de notícia

Tabela 7 - Resultado do desempenho da API de recomendação de criação notícia

	Total requisições	Requisições por segundo	Mínimo (MS)	Média (MS)	90% (MS)	Max (MS)	Erros %
API	8.421	13.80	2	5	4	1.236	0

Fonte: Autor**Figura 23 - Gráfico do desempenho da API de recomendação de criação notícia****Fonte:** Autor

Os resultados apresentados mostram que não houve interrupção no sistema durante toda a execução do teste (Erros 0%). O desempenho de cada uma das API se manteve constante durante todo o período do teste (10 minutos) mesmo com o aumento do número de usuários (de 1 até 20). A média mais alta foi 53 milissegundos para Coleta de Lixo. Considerando 90% das requisições, o valor mais alto foi 53 milissegundos também para coleta de Lixo. Com esse desempenho o número de requisições por segundo foi aproximadamente 13 para todas as API, exceto de recomendação de criação de notícia que foi aproximadamente 8.

Para os resultados da API de recomendação de criação de notícia é necessário fazer algumas considerações. O algoritmo realiza integrações externas com *Google Maps API* e *ChatGPT*. Cada uma dessas integrações possui seu tempo de resposta que pode variar de acordo com a região de acesso, licença utilizada e versão disponível pelo fornecedor. Realizar testes de desempenho nessas plataformas não é viável pois ao identificar múltiplas requisições, por segurança, ela pode bloquear o acesso. Outro ponto a considerar é que este projeto utiliza licença teste para ambas as plataformas, onde o número de requisições simultâneas e cota gratuitas disponíveis são limitados. Tendo isso, a tabela com resultados representa apenas o código executado pela API, sem os acréscimos das integrações externas.

5.2 Atividade de captação de dados

Como planejado, realizou-se a captação de dados nos dias 16, 17 e 18 de novembro de 2023. Os encontros com os alunos aconteceram em dois períodos do dia. Um pela manhã, das 08:00 às 10:00, e outro pela tarde, das 16:00 às 18:00. O local de encontro foi uma sala de aula cedida pela escola onde dispunha-se de cadeiras, mesas, rede de *internet* sem fio e ar-condicionado.

No primeiro dia de atividade os alunos foram convidados a sala para auto apresentação e primeiras formalidades como objetivos e funcionamento da dinâmica dos encontros. Foram distribuídos os crachás e realçado a importância da presença de cada um e o valor do dado capturado por eles no processo de mapeamento da situação real do bairro em que convivem.

O público-alvo foram os alunos de período integral, portanto esperava-se um número constante de presentes nas atividades. Entretanto, como mostrado na tabela 8, a presença no período da tarde do dia 17 e manhã do dia 18 foi consideravelmente menor. Isso se deve ao fato de excepcionalmente nesses dias os alunos foram dispensados das aulas devido a atividades administrativas da escola. Ainda assim, conseguiu-se engajar alguns alunos para irem à escola para participar da captação de dados. Na manhã do dia 17 houve um número maior de presentes,

pois, por iniciativa do diretor da escola, foram convidados alguns alunos de outras classes para participar da atividade.

Tabela 8 - Distribuição de alunos presentes por dia e período

Dia	Manhã	Tarde
16	26	24
17	39	9
18	14	

Fonte: Autor

Foram criados 4 grupos para alocação de todos os alunos. Estabeleceu-se este número pois foi o número de aparelhos móveis disponibilizado pelo projeto com *internet*. Embora os alunos pudessem baixar e instalar em seus aparelhos móveis, optou-se por não depender dessa ação afim de evitar constrangimento a aqueles que não possuíam *internet* ou provocar a prática do aluno pedir dinheiro ao(s) responsável(s) para colocar crédito e participar da tarefa.

Por iniciativa dos alunos cada grupo recebeu um nome definido por eles mesmos. Estes foram: Furacão, IBGE, Bacanas e PJJ (Projeto Jovens Lupa). Um grupo adicional foi criado no período da manhã do dia 17 para alocar os alunos adicionais das outras turmas.

Toda a atividade de organização, avisos, divisão dos grupos e treinamento prático em sala foi realizado no dia 16 no período da manhã. Apenas no período da tarde do mesmo dia foi executado a primeira saída às ruas.

Entre os pontos positivos foi notado que muitos alunos conheciam as ruas e algumas das residências, facilitando assim o processo de contato com o morador. Um fator motivador foram os grupos terem sido formados pelos próprios alunos, o que ocasionou em uma seleção por afinidade. Muitos alunos estavam entre amigos e isso ajudou a minimizar a timidez e aumentar a iniciativa.

Entre às dificuldades identificadas durante às saídas, a de maior recorrência foi a ausência de morador nas residências. Isso ocorreu principalmente na manhã do dia 17. Isso se deve ao fato de a captação estar sendo realizada em dia útil. Outra dificuldade notada foi o calor. Embora tenha-se optado por horários alternativos, o calor ainda causou um impacto no ânimo dos alunos. Alguns alunos reportaram a

demora dos moradores em atender a porta. Outros reportaram que alguns moradores queriam saber mais sobre a pesquisa e prolongavam a conversa. Essas situações estenderam a execução da captação de alguns grupos, gerando uma menor quantidade de registro para uma determinada rota.

O aplicativo móvel funcionou sem apresentar nenhum erro. Foram reportados por alguns alunos que em determinados momentos ao clicar no botão para salvar, a mensagem de confirmação levava alguns segundos para ser apresentada. Isso causou dúvidas pois não estava claro se o botão não tinha funcionado ou se era uma lentidão ao salvar. Todos os aparelhos que houve a instalação funcionaram dentro do esperado. Os alunos reportaram não terem dificuldades técnicas na utilização devido principalmente ao *layout* simplificado apenas com os botões para a execução do registro. O *login* automático também colaborou para a agilidade e menor tempo de interrupção. Alguns alunos fechavam o aplicativo ou deixavam em segundo plano após um registro. A tela do mapa onde é possível visualizar os ícones com as ocorrências funcionou como um motivador, pois causou um impacto visual e gerou interesse em acompanhar a alteração no mapa após cada registro.

Em relação as opções de ocorrências disponíveis para cadastro, todas foram utilizadas e não houve questionamentos. Muitas perguntas foram feitas em relação as opções internas de cada ocorrência. Por exemplo, para água potável, a pergunta “Há água encanada?”. O intuito dessa pergunta era identificar se havia disponível na residência o serviço público oferecido pelo município. Entretanto, na percepção dos alunos a água do poço era entendida como água encanada, pois era tão comum para a região, que isso foi incorporado como o padrão. O mesmo aconteceu para tratamento de esgoto e a pergunta “Há coleta ou tratamento de esgoto”. Nesse caso, a fossa era entendida como um tratamento existente. Alguns termos também causaram dúvidas como “fossa”. Alguns mencionaram que os moradores utilizam “sumidouro”. Outro termo foi o “amazônico” da pergunta “Possui poço amazônico?”, onde alguns questionavam se era um tipo específico de poço. Por último a pergunta “Alguma obra de saneamento está sendo executada” contida na ocorrência de água potável e tratamento de esgoto. Alguns moradores respondiam que havia obra na região ou em ruas próximas. Porém o objetivo da pergunta era na residência especificamente.

Figura 24 - Fotos da captação de dados



Fonte: Autor

Dado esses relatos, uma lista de melhoria foi elaborada para evolução futura dos *softwares* e garantir uma melhor acuracidade do processo de registro de ocorrência, como apresentado no quadro 12.

Quadro 12 - Lista de melhorias

	Descrição	Melhoria proposta
1	Lentidão ao clicar no botão salvar. Dúvida se o aluno tinha realmente clicado no botão pois a tela permanece estática até a apresentação da mensagem de confirmação.	Tudo indica que o problema ocorreu devido a rede de <i>internet</i> , pois foi em determinados momentos e não foi relatado por todos os grupos. Para minimizar o impacto foi proposto colocar uma imagem animada após clicar no botão salvar para indicar que o dado

		está sendo processado e deve-se aguardar.
2	Textos das perguntas ambíguas.	Os textos atualmente estão dentro do código fonte do aplicativo, sendo necessário gerar uma nova versão caso seja desejado trocar alguma palavra. Foi proposto mover os textos para um banco de dados para serem carregados de forma dinâmica. Desta forma, caso alguma atualização de termo seja necessária, pode ser realizada de forma rápida e refletida em todos os aplicativos de forma automática.
3	Depoimentos do impacto dos problemas no dia a dia.	Devido a muitos relatos dos moradores, foi proposto a adição de um campo de texto opcional em cada tela de ocorrência para adição de comentários e observações gerais.
4	Adição de imagem para melhor identificação e impacto do dado captado.	Foi proposto um campo opcional para captura de imagem da ocorrência na tela de registro. Após preencher os campos, o usuário clica em um botão para abrir a câmera e registrar uma foto. A imagem também deve ser apresentada no <i>site</i> ao visualizar os detalhes de uma ocorrência.
5	Permitir cadastro <i>offline</i> e sincronismo automático ao identificar <i>internet</i> .	Para remover a dependência de <i>internet</i> foi proposto permitir cadastro onde o dado é salvo no aparelho do usuário. Ao identificar <i>internet</i> , o aplicativo envia os dados para o servidor automaticamente.

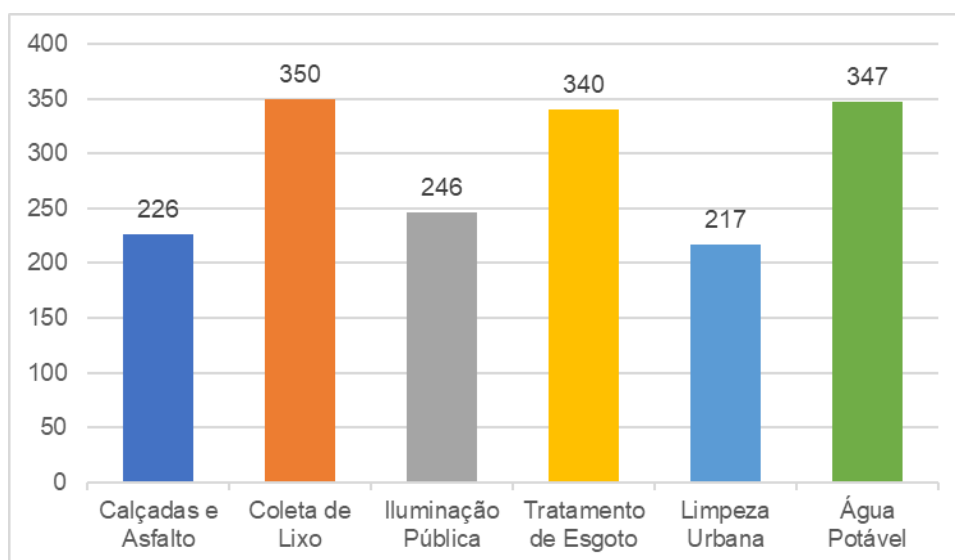
Fonte: Autor

Outro fator motivador dessa etapa foi a identificação de dados capturados fora do período da atividade em grupo. Isso significa que os alunos exploraram o aplicativo por iniciativa própria e realizaram cadastros de acordo com seu entendimento. Esse comportamento leva a reflexão de que tecnologias móveis combinada com projetos que focam na coparticipação pode despertar o interesse do cidadão em interagir com sua comunidade. Isso também pode ser percebido pelo lado do cidadão que participou respondendo às perguntas, pois, alguns descreviam com detalhes diversos problemas da região e seus impactos no dia a dia. Pode-se perceber uma necessidade de expressar e ser ouvido, gerando oportunidade para jovens jornalistas na coleta de dados.

5.3 Métricas da atividade de captação de dados

No total foram captadas 1.726 ocorrências nos três dias de atividade (Figura 22).

Figura 25 - Total de ocorrências captadas

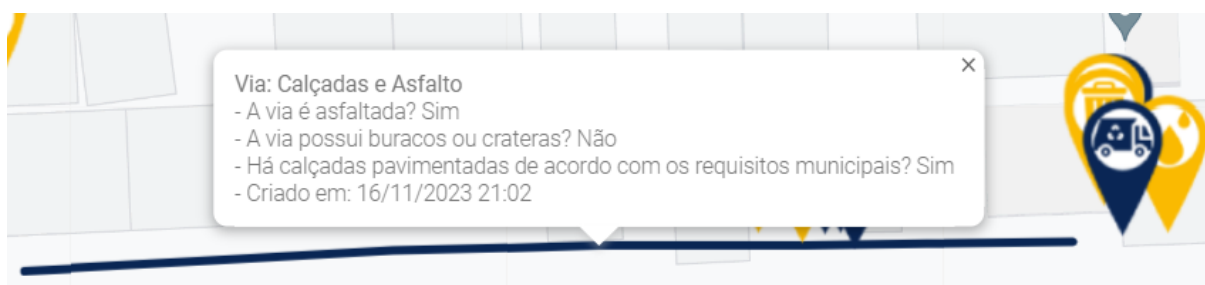


Fonte: Autor

Coleta de lixo, tratamento de esgoto e água potável foram as ocorrências mais captadas. Calçadas e asfalto, iluminação pública e limpeza urbana obtiveram

menos ocorrências. Isso se deve ao fato dos três primeiros serem itens cadastrados por residência. Ou seja, cada residência gera três cadastro, sendo um para cada. Já os três últimos são referentes a ruas e algumas vezes representa um trecho completo. Na figura 23 temos um exemplo. A linha em azul representa uma ocorrência onde o aluno optou por cadastrar um trecho, e não um ponto específico para calçadas e asfalto. Neste caso temos um registro para a rua toda. Já no caso de residências temos um aglomerado de ocorrências.

Figura 26 - Cadastro de trecho



Fonte: Autor

Ao considerarmos a classificação da ocorrência temos a distribuição apresentada na tabela 9. Coleta de lixo foi a ocorrência com menor classificação de problemas. Muitos moradores de fato reportaram que há coleta três vezes por semana em muitas das ruas. Tratamento de esgoto foi a ocorrência com maior classificação de problema. O bairro atualmente possui rede de esgoto apenas nas vias principais. Os moradores em sua maioria reportaram utilizar fossa como meio alternativo. Água potável também apresentou um número alto de classificação de problema. Similar a rede de esgoto, a disponibilização de água está presente somente nas ruas principais. Muitos moradores reportaram utilizar poço amazônico para retirada de água potável. Alguns moradores chegaram a reportar ter optado pelo poço para não pagar o valor da água e terem um serviço com maior disponibilidade, pois o serviço de distribuição pela fornecedora apresentava alguns problemas como falha na distribuição e ar no encanamento.

Um fator que contribuiu para menor classificação de problemas para as ocorrências de ruas foi o fato dos alunos se dedicarem a cadastrar apenas quando esta apresentava algum problema. Ou seja, quando uma rua era calçada, com

iluminação e limpa, alguns alunos não registravam a ocorrência pois estava em conformidade.

Tabela 9 - Total de ocorrência por classificação

	Em conformidade		Problemas		Total
	Qtde	%	Qtde	%	
Calçadas e Asfalto	43	19%	183	81%	226
Coleta de Lixo	348	99.4%	2	0.6%	350
Iluminação Pública	76	30.9%	170	69.1%	246
Tratamento de Esgoto	52	15.3%	288	84.7%	340
Limpeza Urbana	17	7.8%	200	92.2%	217
Água Potável	92	26.5%	255	73.5%	347
Total	628	36.4%	1.098	63.6%	1.726

Fonte: Autor

Deve-se analisar também o alcance do aplicativo a partir da sua disponibilização na loja da *Google*. Estes números são relevantes para medir a qualidade do *software* desenvolvido pois ele identifica como foi o comportamento em diferentes modelos e fabricantes.

O aplicativo foi instalado em 25 aparelhos móveis no total. Todos *Android* nas versões 8, 12 e 13. Utilizando o relatório do *Google Play* pode-se identificar os modelos específicos e como se comportou em diferentes configurações. Na tabela 10 nota-se que *Samsung Galaxy A14* foi o modelo mais utilizado com 5 instalações. Entre os fabricantes encontra-se *Samsung*, *LG*, *Asus* e *Motorola*. A taxa de erro de instalação foi 0%. Este número identifica os aparelhos que iniciaram a instalação após o *download*, porém o processo foi interrompido por algum erro. Em relação a resolução dos aparelhos móveis, aponta-se que houve instalações desde modelos pequenos com *Galaxy M02 720px* até os modelos maiores disponíveis no mercado como *Galaxy S22 Ultra 1.440px*. Ao que se refere a memória, também se identificou uma variação entre 2GB e 8GB. Para o uso de *internet* a rede móvel mais utilizada foi Vivo, porém alguns aparelhos também utilizaram TIM e Claro.

Tabela 10 - Dados de instalação do aplicativo

Fabricante	Modelo	Qtde	Taxa de erro	Resolução da tela	Memória
<i>Samsung</i>	<i>Galaxy A14</i>	5	0.00%	1080x2408	4GB
<i>Asus</i>	<i>Zenfone 9</i>	2	0.00%	1080x2400	8GB
<i>Samsung</i>	<i>Galaxy M02</i>	2	0.00%	720x1600	2GB
<i>Samsung</i>	<i>Galaxy Note10</i>	2	0.00%	1080x2280	8GB
<i>LGE</i>	<i>QStylus</i>	2	0.00%	1080x2160	2GB
<i>Motorola</i>	<i>moto g71 5G</i>	1	0.00%	1080x2400	6GB
<i>Samsung</i>	<i>Galaxy A14 5G</i>	1	0.00%	1080x2408	8GB
<i>Samsung</i>	<i>Galaxy S22+</i>	1	0.00%	1080x2340	8GB
<i>Motorola</i>	<i>motorola edge 30</i>	1	0.00%	1080x2400	6GB
<i>LGE</i>	<i>LG K40S</i>	1	0.00%	720x1560	2GB
<i>Samsung</i>	<i>Galaxy S22 Ultra</i>	1	0.00%	1440x3088	4GB
<i>Samsung</i>	<i>Galaxy A33 5G</i>	1	0.00%	1080x2400	6GB
<i>Samsung</i>	<i>Galaxy Quantum3</i>	1	0.00%	1080x2400	8GB
<i>Asus</i>	<i>ZenFone Max Pro M2</i>	1	0.00%	1080x2280	6GB
<i>Motorola</i>	<i>moto g22</i>	1	0.00%	720x1600	4GB
<i>Motorola</i>	<i>moto g52</i>	1	0.00%	1080x2400	4GB
<i>Samsung</i>	<i>Galaxy A31</i>	1	0.00%	1080x2400	8GB

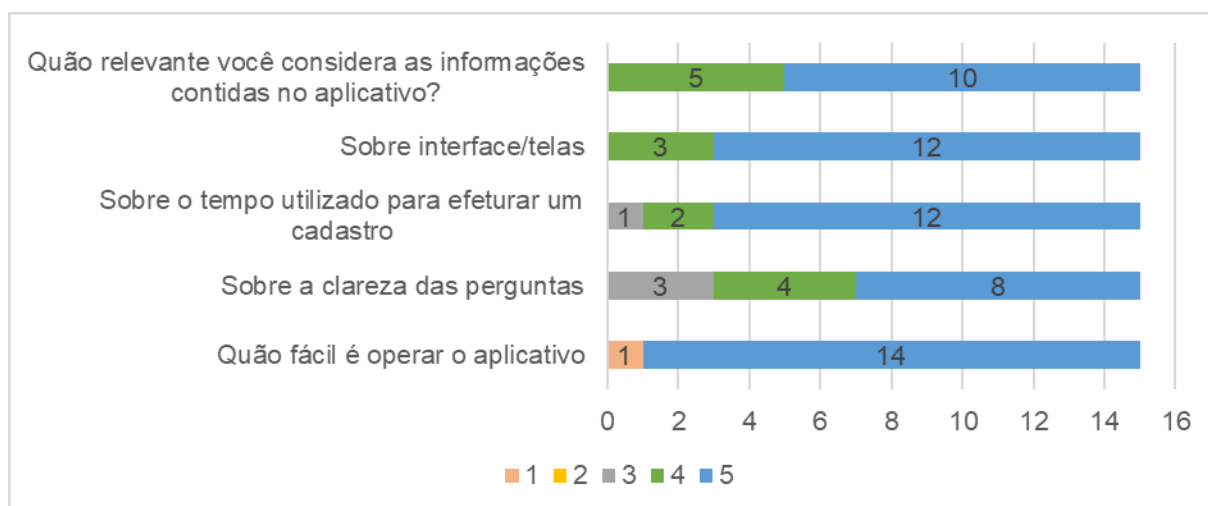
Fonte: Autor

5.4 Enquete do aplicativo

Aplicou-se a enquete durante a atividade de captação de dados para medir a percepção dos alunos em relação ao aplicativo utilizado. Como a enquete era opcional, alguns alunos não se sentiram à vontade para preencher, mesmo realçando que ela era anônima. Obteve-se 15 respostas. Como mostrado na figura 24, a pergunta melhor avaliada foi sobre a facilidade de utilizar o aplicativo. Essa era uma grande preocupação do projeto. Garantir que a usabilidade não fosse um fator que impedisse o uso, e sim, instigasse-o. Em relação ao *layout* e tempo de resposta, ambas receberam 12 notas máximas. Tempo de resposta teve uma resposta com

valor 1 (mínimo). Como apresentado no item 5.2, alguns alunos reportaram lentidão em alguns cadastros. Esse indicador reflete esse percentual de alunos que experienciaram essa dificuldade. Em relação a relevância das informações obtidas, todas as notas foram acima da média. Esse é um quesito importante pois fez-se um esforço de justificar e reafirmar em diversos momentos a razão da atividade em execução. O indicador que obteve menor desempenho foi em relação a clareza das perguntas, que recebeu 3 notas abaixo da média. Isso também está alinhado com as dificuldades reportadas pelos alunos durante a atividade e apresentados no item 5.2. Espera-se que a melhoria proposta para facilitar a atualização do texto das perguntas de forma dinâmica possa minimizar esse impacto, pois assim que identificado, pode ser sanado rapidamente.

Figura 27 - Respostas da enquete do aplicativo



Fonte: Autor

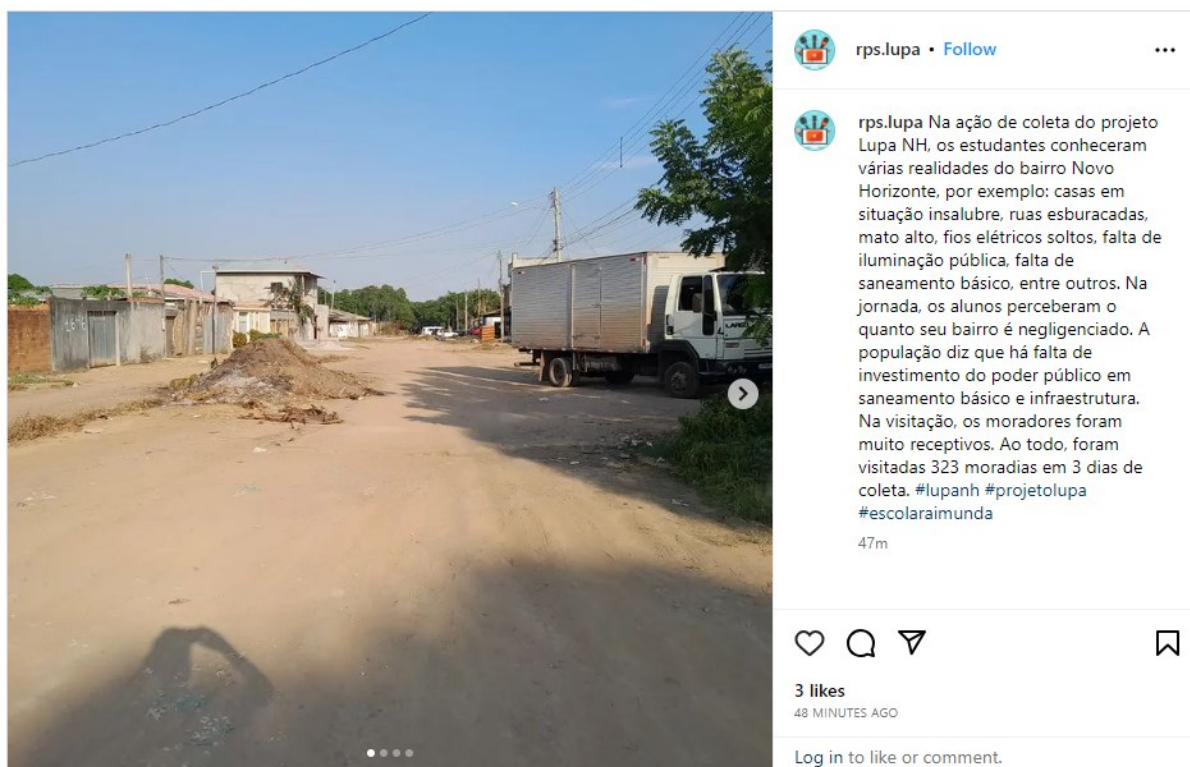
5.5 Site e algoritmo de recomendação de geração de notícia

A partir do dado captado na atividade em campo montou-se uma base de dados que representa as características do bairro Novo Horizonte. O *site* foi utilizado durante os 3 dias de evento como um meio de acompanhar o volume de dados e possíveis inconsistências.

Através do contador de acesso disponível foi possível identificar 28 visitas únicas entre os dias 16 de novembro e 15 de dezembro de 2023. Visitas únicas se caracteriza quando o servidor de destino, no caso o *site*, identifica que quem acessa é o mesmo IP já acessado anteriormente. Esse recurso minimiza a contagem de múltiplas visitas para uma mesma pessoa. Entretanto, não é garantido que sejam realmente únicas, pois uma pessoa pode acessar via computador do trabalho, pessoal e celular. Abas anônimas dos navegadores também contam como única.

O intervalo de data foi definido baseado na agenda de aulas programada para a turma da disciplina eletiva de jornalismo. Eram esperadas aulas nos dias 24 de novembro e 1 de dezembro (primeira e segunda semana pós atividade de captação). Entretanto, a aula do dia 24 foi cancelada pela escola. Esperou-se que fosse repostado para o dia 8 ou 15 de dezembro de 2023, porém a direção da escola reportou problemas de falta de água onde as aulas tiveram que ser canceladas. Logo, o número de acesso representa o uso entre os dias da captação de dados e atividade em sala do dia 1 de dezembro. O resultado do trabalho em sala foi compartilhado via *post* em rede social (figura 25).

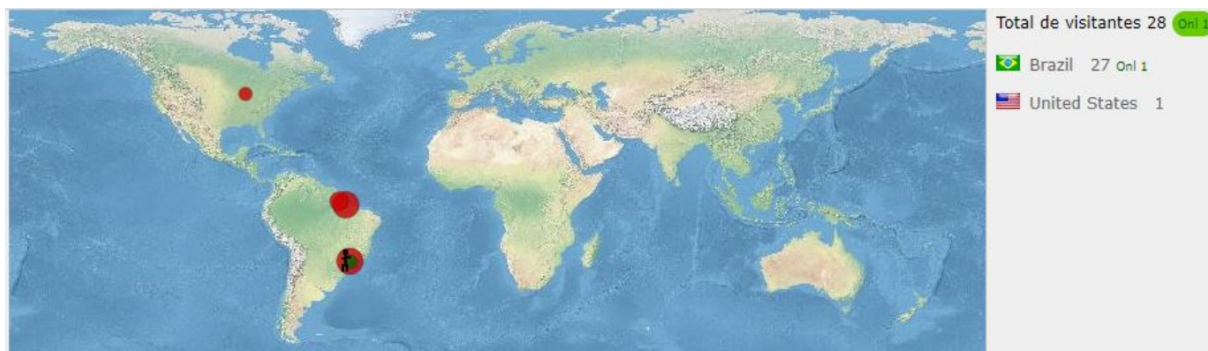
Figura 28 - Post Instagram sobre o projeto



Fonte: Instagram @rps.lupa

Na figura 26, vê-se concentração de acesso principalmente na região de Macapá, Belém e São Paulo. São Paulo e Belém devido aos pesquisadores voluntários do projeto LupaNH que foram convidados a participarem da atividade.

Figura 29 - Acesso ao *site*

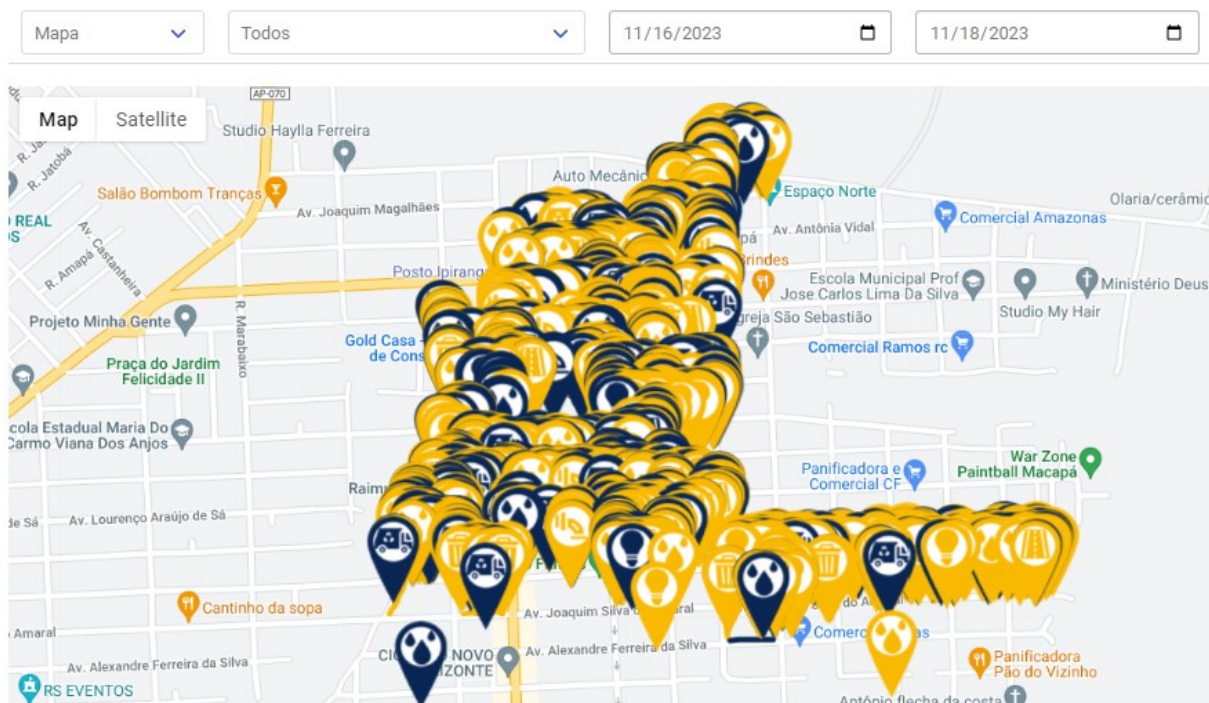


Fonte: Ferramenta webcontadores.com

As figuras 22 e 23 apresentadas no capítulo 5.3 foram imagem retiradas do *site* onde os dados são visualizados em forma gráfica. Abaixo temos a figura 27 com a visualização em mapa.

Figura 30 - Visualização de dados em mapa

Mapa

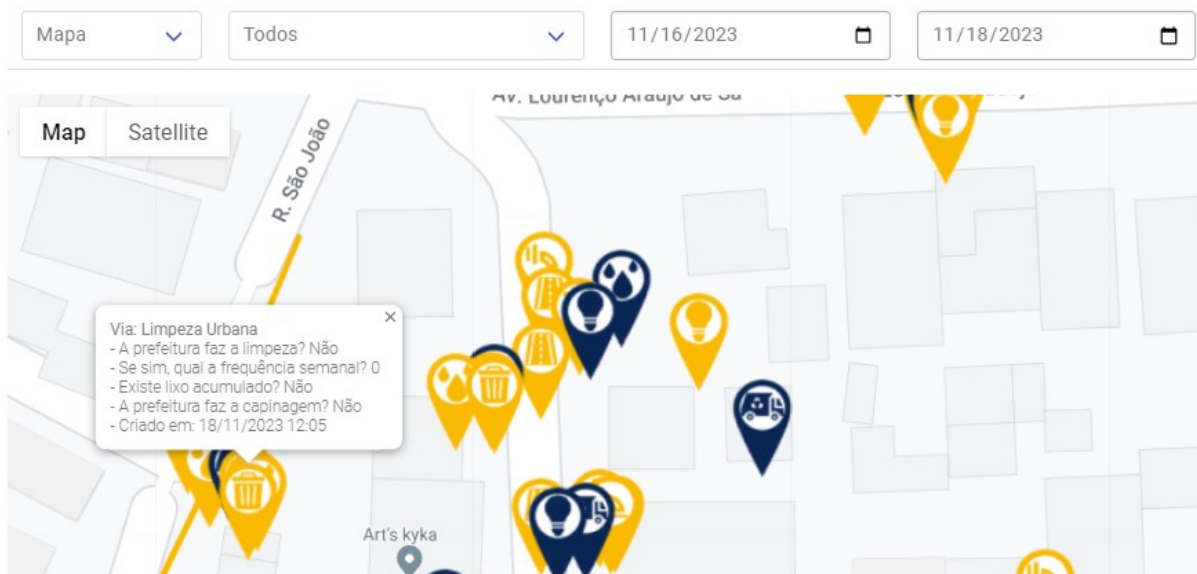


Fonte: Autor

Ao aproximar das ocorrências, pode-se clicar nos ícones para verificar seu detalhamento, como mostrado na figura 28. A cor azul e amarelo representam respectivamente ocorrências classificadas como em conformidade ou problema.

Figura 31 - Visualização do detalhe de uma ocorrência no mapa

Mapa



Fonte: Autor

A listagem em modo de tabela, como apresentado na figura 29, não se mostrou viável pois gera-se uma barra de rolagem muito longa quando o volume de dado para o período filtrado é grande. Para verificação em lista a opção de exportar em arquivo CSV mostrou-se mais conveniente pois pode-se utilizar o *Excel* para manipulação dos dados. Todas as visualizações e recursos do *site* foram desenvolvidas e testados no modo navegador e aparelho móvel, buscando garantir uma maior disponibilidade do dado.

Figura 32 – Visualização de dados em tabela

Mapa

Taba

Todos

11/16/2023

11/18/2023

Todos

Filtrar

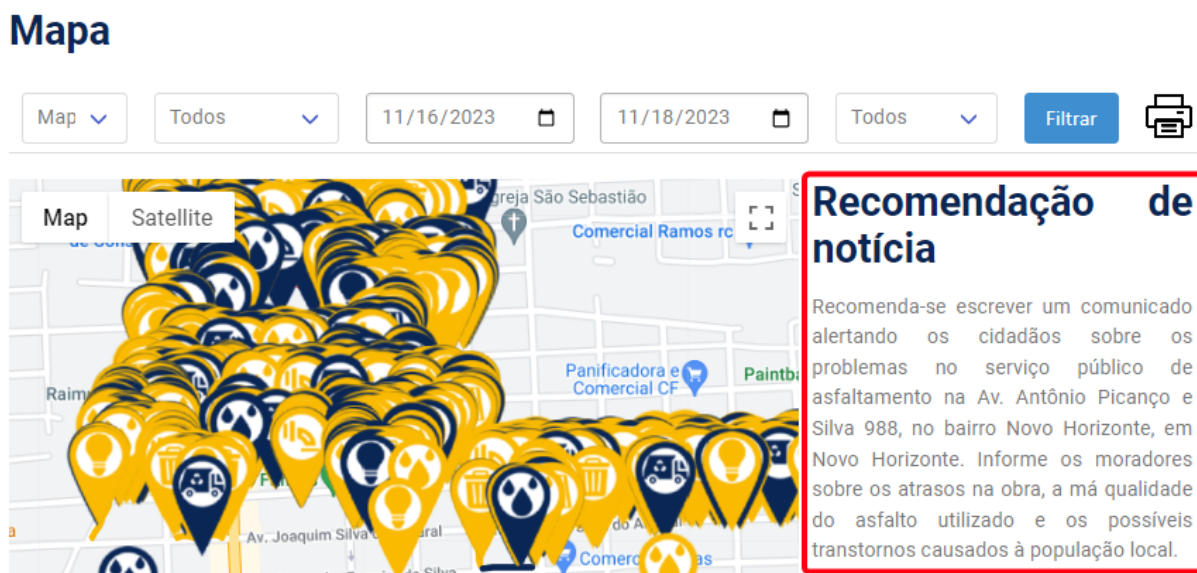
Iluminação Pública

Id	Título	Problema	Latitude	Longitude	Possui poste?	As luzes estão funcionando?	Há fios elétricos soltos?	Criado em
48	Iluminação Pública	Sim	-0.0088305	-51.0798717	Sim	Sim	Sim	11/16/2023
49	Iluminação Pública	Não	0.0955936	-51.0499227	Sim	Sim	Não	11/16/2023

Fonte: Autor

Ao carregar a página pode-se notar ao lado esquerdo da tela a funcionalidade de recomendação de criação de notícia e o texto gerado a partir do uso da base de dados, *Google Maps API* e *Chat GPT API* (figura 30). O dado é gerado a partir do algoritmo de clusterização e geração de prioridades. Ao atualizar a página repetidamente o texto se mantém o mesmo devido ao *cache* implementado. O novo texto é gerado apenas após a expiração que acontece a cada 1 minuto. Durante o uso do *cache* o carregamento do *site* mostrou-se alinhado com o resultado do teste de desempenho executado no capítulo 5.1. A cada renovação, devido ao acesso as integrações, o carregamento se mostrou mais lento.

Figura 33 - Texto de recomendação de criação de notícia



Fonte: Autor

Abaixo tem-se mais alguns exemplos de recomendações geradas pelo algoritmo utilizando a base de dados captada.

“A falta de manutenção e a obstrução constante do sistema de esgoto público na Av. Hermes Monteiro da Silva 1265 - Novo Horizonte têm gerado transtornos e insatisfação para os moradores do bairro. É urgente que o poder público tome providências para solucionar os problemas e garantir um serviço de qualidade na região.”

“Uma grave situação envolvendo o serviço público de esgoto no bairro de Novo Horizonte requer uma reportagem para

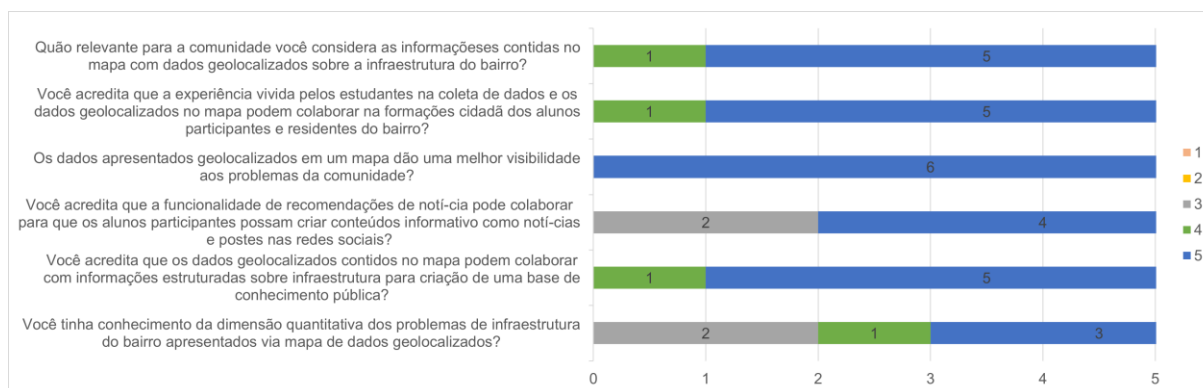
expor os problemas e encontrar soluções urgentes, a fim de garantir a saúde e qualidade de vida dos moradores.”

“Há uma urgente necessidade de se reportar os problemas relacionados ao serviço público de asfaltamento na Av. Joaquim Silva do Amaral 2037-1973, localizada no bairro Novo Horizonte. Moradores têm enfrentado dificuldade de acesso, devido à má conservação e falta de manutenção desta importante via, prejudicando a mobilidade e segurança da comunidade.”

As recomendações geradas utilizando o *ChatGPT* se mostraram interessantes e com grande potencial de uso, pois o texto possui um formato de sugestão, além de ressaltar o impacto e o benefício do dado. Em alguns exemplos também é possível identificar recomendações de hashtags. O risco do uso dessas ferramentas é na utilização de alguns termos que podem não refletir a real necessidade, como no primeiro exemplo acima onde usa-se “a obstrução constante do sistema de esgoto...”. Não é certo que este dado é sobre obstrução, por mais que isso aconteça no bairro em alguns casos, também existem casos em que há apenas a falta do esgoto. Mas ainda assim, o resultado é bastante satisfatório.

5.6 Enquete do *site*

De todos os professores(as) que participaram em algum momento da disciplina eletiva em jornalismo, durante o ano letivo de 2023, seis participaram da atividade de captação de dados. Portanto, estes foram os elegíveis para responder a enquete do *site*, pois possuíam o conhecimento do projeto e também dos dados em questão. O contato foi realizado via *whatsapp* onde foi enviado o *link* para o *site* e outro para a enquete. No formulário *online* também foi reforçado ao participante a acessar o *site* e navegar nas funcionalidades antes de responder as perguntas. Todos os seis participantes responderam prontamente a enquete onde o resultado se encontra na figura 32.

Figura 34 - Respostas da enquete do *site*

Fonte: Autor

Na primeira questão, 5 dos 6 participantes consideraram muito relevantes as informações presentes no site. Na segunda questão, também 5 dos 6 participantes acreditam que os dados apresentados podem colaborar com a formação cidadã dos alunos participantes da pesquisa. Em relação ao formato do dado disponibilizado em mapa, todos acreditam que isso pode melhorar a visibilidade dos problemas identificados. Em relação a funcionalidade de geração de recomendação de notícia, 4 participantes concordam totalmente que ela colaboraria no processo de criação de notícia pelos alunos, contra 2 que não concordam e nem discordam. 5 dos 6 participantes também concordam que os dados contidos no site podem colaborar para a criação de uma base de dados pública. Sobre o conhecimento dos participantes sobre a situação do bairro apresentada no site, todos os participantes responderam que tinham algum nível de conhecimento dos problemas apresentados.

O resultado da enquete apresentou que em geral as funcionalidade e formato de conteúdo apresentado são relevantes e podem colaborar com fornecimento de conteúdo informativo.

5.7 Discussão

Através de tecnologias móveis hiperlocal é possível mobilizar comunidades e criar capacidade de resolução de problemas sociais (Metzgar; Kurpius; Rowley,

2011). O arcabouço tecnológico e processos metodológicos executados nesse projeto jogou luz sobre os resultados que este tipo de ação social pode alavancar. Apesar de não possibilitar uma generalização, o estudo permite uma análise de uma situação da vida real de uma área e grupo.

O processo de coleta de dados através de participação pública utilizando uma escola local possibilitou o avanço da pesquisa de forma organizada e com um cronograma factível. Portanto o contato direto e o envolvimento da comunidade no processo se mostraram positivo nesse quesito. Outro fator importante foi a vinculação da atividade de captação e uso do dado a uma atividade escolar. Essa ação trouxe um viés educativo e formativo. Isso realça a questão de que a tecnologia combinada com participação pública pode ser um instrumento de promoção de cidadania e consciência coletiva (Lima Junior; Santos Coelho, 2015).

O número de dados captados e estudantes envolvidos mostrou que existe um público que está disposto a serem agentes participativos. O envolvimento da escola estadual do bairro também realça essa perspectiva. Como visto na RSL desenvolvida no capítulo 2.2.5, os estudos de Liu (2018), Robaeyst (2021), ambos categorizados como impacto da tecnologia hiperlocal na comunidade, mostram que essas ações colaboram com o engajamento civil e aumentam a coesão social, pois combinam a troca de experiência entre as pessoas de uma mesma região e os problemas sociais que as permeiam. A solução integrada desenvolvida também está alinhada com os resultados do estudo de Mukhtar (2021) sobre o perfil das tecnologias sociais, pois segue os requisitos e necessidades da região, maximizando as chances de impacto positivo. Han (2014) comenta que as possibilidades são vastas para o uso dessas tecnologias, e vimos isso através da diversidade de projetos encontrados na RSL e das inúmeras possibilidades de projetos futuros identificadas a partir desta pesquisa.

O que diferencia esse projeto dos demais é sua amplitude em relação a possibilidade de uso das ferramentas desenvolvidas. Através do aplicativo para coleta de dados, *site* para visualização estruturada e o uso de AI generativa para sugerir textos, possibilita-se o ciclo completo do processo de geração de conteúdo informativo. A ferramenta mostrou-se viável para o processo de educação e formação crítica de jovens em áreas periféricas assim como para geração de base de conhecimento sobre a situação atual do bairro. Nesse estudo específico o

público-alvo foram estudantes com interesse na área jornalística, entretanto, vê-se potencial para uso por outras disciplinas como português, geografia e estudos sociais. Adicionalmente, as ferramentas foram desenvolvidas, testadas e executadas em um cenário real, minimizando a possibilidade de não adequação a outros cenários. Seu código fonte foi documentado e disponibilizado em um repositório público e aberto, para que outros projetos possam reutilizar.

Na enquete de percepção focado no aplicativo foi possível identificar como os alunos reagiram em relação ao seu uso. O resultado mostrou-se alinhado com o que foi registrado durante a atividade de captação de dados. O aplicativo se comportou bem, de forma segura e intuitiva. As dificuldades apresentadas relacionam-se ao desempenho e texto das perguntas. Esses itens não foram impeditivos para execução da atividade, porém mostram a importância de considerar o máximo de aspectos possíveis para execução de pesquisa aplicada. O reporte sobre desempenho está fortemente ligado a disponibilidade de baixa qualidade de rede de *internet* na região. Entende-se agora que soluções tecnológicas móveis devem conter a possibilidade de uso *offline* para remover essa dependência. No que tange o texto das perguntas, entende-se também que cada região, mesmo que próxima pode possuir gírias, jargões e expressões regionais. A forma mais efetiva identificada é possibilitar rápidas atualizações dos *softwares* para evitar dificuldades e aumentar adaptação.

Em relação a enquete do *site*, notou-se que os participantes concordam com a relevância das informações e formato apresentado. Entretanto, em uma análise geral da pesquisa, o uso da funcionalidade de geração de notícia poderia ter sido mais bem explorado nas atividades em classe. Isso se deve ao fato de ter havido apenas uma aula com os alunos para interação com os dados, antes da conclusão do ano letivo. Esse é um ponto de lição aprendida em relação ao cronograma executado. Vale ressaltar que inicialmente essas atividades deveriam ter acontecido no meio do ano letivo. Porém diversos impeditivos não relacionados a esta pesquisa como greves, agenda escola, feriados, férias e entre outros, culminaram na sua execução apenas no final do ano. Essas dificuldades devem ser sempre consideradas em pesquisas desta magnitude e formato.

Reportar e denunciar são um nível de interação que caracteriza mídia hiperlocal, e a instrumentação, seja ela qual for, potencializa essa ação (Metzgar;

Kurpius; Rowley, 2011). A opção de utilizar conceitos de GIS neste projeto foi devido a sua capacidade de adicionar valor a um dado hiperlocal e beneficiar do uso do GPS dos aparelhos móveis. Considerando que o Brasil em 2023 possui cerca de 440 milhões de dispositivos²⁶, gerando uma média de 2 por habitantes, isso possibilita uma vantagem de instrumentação em massa. O relatório da *GooglePlay* apresenta o alcance que essa pesquisa teve em relação a diversidade de aparelhos móveis utilizados e suas diferentes configurações. Essa pesquisa se mostrou assertiva nas escolhas tecnológicas feitas tanto para a aplicação móvel quanto via navegador. Vale ressaltar que a escola e os alunos envolvidos devem ser replicadores da iniciativa caso desejam que a captação de dado, uso do *site* e recurso de recomendação de criação de notícia sejam utilizados por outras atividades da escola.

²⁶ FGV: Fundação Getúlio Vargas. Disponível em <https://eaesp.fgv.br/producao-intelectual/pesquisa-anual-uso-ti>. Acessado em 29 de janeiro de 2024.

Capítulo 6

CONCLUSÃO

Esta pesquisa teve como objetivo principal demonstrar que por intermédio de tecnologias móveis e participação pública, pode-se construir ações de levantamento e captação de dados hiperlocal e utilizá-los como base para geração de notícia. O aplicativo móvel, *site* e algoritmo de recomendação de criação de notícia foram construídos com o propósito de validar essa questão utilizando como público-alvo o bairro Novo Horizonte. O estudo de caso apresentado descreveu o potencial dessas tecnologias e ações locais, quando executadas de forma estruturada e colaborativa.

Dentre os principais resultados destaca-se o envolvimento dos alunos na ação de captação, onde através de uma ação coordenada foi possível uma imersão no ambiente real da pesquisa. Essa experiência trouxe uma relação mais íntima sem perder a abordagem científica. Confrontar as teorias e conceitos com a prática em campo possibilita um olhar privilegiado para descrever de forma mais realista os resultados.

Possibilitou-se também fazer a validação dos instrumentos técnicos da pesquisa como o aplicativo móvel e o *site*. Esses são os meios que possibilitaram uma execução harmoniosa e segura da pesquisa. O aplicativo móvel funcionou corretamente em registrar de forma precisa os locais onde as ocorrências foram identificadas. A API do *GoogleMaps* e *ChatGPT* utilizada se mostrou integrada de forma eficiente com o GPS dos aparelhos mesmo estando em uma região periférica.

Acerca das limitações presentes nesta pesquisa, ressaltamos que este estudo se limitou a proporcionar tecnologias móveis para realizar a captação, visualização dos dados e recomendação de criação de notícia, sendo assim, não foi executado

uma análise qualitativa do impacto dos dados resultantes. Outra limitação identificada foi a não realização de uma pesquisa de satisfação das tecnologias criadas. Como alternativa foi optado por uma enquete para captura de percepção do usuário devido a seu escopo menor. Essas limitações são justificadas devido ao tempo necessário para sua realização e ao desafio de englobar uma etapa de análise de satisfação em um estudo de caso, onde o pesquisador tem contato direto com o público-alvo e região, podendo gerar conflitos ou vieses nas execuções das avaliações. Deve-se considerar também o público-alvo serem estudantes menores de 18 anos, onde para uma abordagem assertiva, requeria-se uma etapa de preparação extensa para identificação dos melhores métodos de trabalho.

6.1 Contribuições

Estes resultados levam às contribuições práticas como os dois *softwares* e o algoritmo de recomendação de geração de notícia, pronto para uso, de código aberto, e escalável para outras regiões. Estas tecnologias permitem que grupos criem meios para aumentar engajamento social, através de atividades de identificação, monitoramento e divulgação de conteúdos informativo hiperlocal.

6.2 Inserção Social

Como contribuição social, considera-se a ação comunitária como um meio de difusão do conhecimento científico, empoderamento social e autogerenciamento, além da geração de uma base de dados sobre infraestrutura deficitária que possibilita os moradores a como prestação de conta.

6.3 Perspectivas Futuras

Recomenda-se que as análises dos dados captados sejam executadas por pesquisas futuras, onde pesquisadores poderão focar na análise do impacto dos dados e sua relevância, gerando planos de ações para mudanças e melhoria, tanto na comunidade, como nas tecnologias. Outra possibilidade está baseada na funcionalidade de geração de recomendação de notícia. O uso de inteligência artificial generativa combinada com base de dados hiperlocal, pode ser explorado no processo de coparticipação ou apoio a alunos na criação de textos diversos. Entende-se também que *softwares* devem ser constantemente atualizados para não caírem em desuso. É recomendável uma reavaliação das funcionalidades para estar alinhado com as necessidades de cada região e público-alvo.

É fundamental destacar o cunho interdisciplinar desta pesquisa, onde sem este não seria possível sua conclusão. Tecnologia sem propósito tende a ser abandonada pelos usuários que não veem ali uma razão para o uso. Outras áreas do conhecimento, sem o apoio da tecnologia, não sobrevivem a concorrência e dinamicidade das interações. O aplicativo móvel, *site* e o algoritmo de recomendação de criação de notícia só foram desenvolvidos, pois, existe um arcabouço conceitual que envolve diferentes disciplinas como comunicação, tecnologia e ciência sociais, para sustentar as decisões e caminhos percorridos.

Por último, ressalta-se que ferramentas não funcionam sem o apoio de pessoas. Logo o envolvimento da comunidade e uma constante estimulação para incentivar o uso são fundamentais para a execução e manutenção deste tipo de projeto.

REFERÊNCIAS

BASS, L.; CLEMENTS, P.; KAZMAN, R. **Software Architecture in Practice**. AddisonWesley, 2003. (SEI series in software engineering). ISBN 9780321154958. Disponível em: <http://books.google.fi/books?id=mdilu8Kk1WMC>).

BRANNOCK COX, J.; POEPSEL, M. A. Deep Participation in Underserved Communities: A Quantitative Analysis of Hearken's Model for Engagement Journalism. **Journalism Practice**, v. 14, n. 5, p. 537–555, 21 fev. 2020.

BRASIL. Senado federal. **Projeto de Lei nº 678, de 2021, que regula o artigo 193 da Constituição Federal Brasileira de 1988**. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=8935892&disposition=inline>. Acessado em 9 de julho de 2023.

BUCAY, Y.; ELLIOTT, V.; KAMIN, J.; PARK, A.; CJR. America's Growing News Deserts. **Columbia Journalism Review**. 2017 Disponível em https://www.cjr.org/local_news/american-news-deserts-donuts-local.php. Acessado em 9 de julho de 2023.

CARROLL, J. M. et al. Reviving community networks: hyperlocality and suprathresholding in Web 2.0 designs. **Personal and Ubiquitous Computing**, v. 19, n. 2, p. 477–491, 11 out. 2014.

CLAES WOHLIN. **Experimentation in software engineering**. Heidelberg: Springer, 2012.

DAGNINO, R. **Tecnologia Social**: Contribuições conceituais e metodológicas, v.2, p.19-33, 2014.

DELMELLE, E. C. **Giscience and neighborhood change**: Toward an understanding of processes of change. Transactions in GIS, John Wiley and Sons Inc, v. 26, p. 567–584, 4 2022. ISSN 14679671.

DOWLING, M.; LUCEY, B. ChatGPT for (Finance) research: The Bananarama Conjecture. **Finance Research Letters**, v. 53, p. 103662, jan. 2023.

DWIVEDI, Y. K. et al. “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. **International Journal of Information Management**, v. 71, n. 0268-4012, p. 102642, 1 ago. 2023.

ESRI. **What is GIS?** 2023. Disponível em: <https://www.ESRI.com/en-us/what-is-gis/overview>.

FRAGA, L. **Autogestão e tecnologia social**: Utopia e engajamento, v.1, 2011.

GLASER, M. Citizen Journalism: Widening World Views, Extending Democracy. In: **The Routledge Companion to News and Journalism**. London: Routledge, 2019.

GOODCHILD, M. F. **Geographic information systems**, p. 194-200, 1991.

HAN, K. **Studying the application of mobile technology to local communities**. 15 fev. 2014.

HAN, K.; SHIH, P. C.; CARROLL, J. M. Local News Chatter: Augmenting Community News by Aggregating Hyperlocal Microblog Content in a Tag Cloud. **International Journal of Human-Computer Interaction**, v. 30, n. 12, p. 1003–1014, 13 out. 2014.

HARRIS, T. M.; WEINER, D. Community Participation and Geographic Information Systems. [S.l.]: **Taylor and Francis**, 2002. 0-140 p. ISSN 0-203-46948-8.

HEPP, A.; LOOSEN, W. Molo.news: Experimentally Developing a Relational Platform for Local Journalism. **Media and Communication**, v. 7, n. 4, p. 56–67, 17 dez. 2019.

HIGGINBOTHAM, J.; SAFARI, a. O. M. C. **Designing Great Web APIs**. O'Reilly Media, Incorporated, 2015. ISBN 9781491924594. Disponível em: <https://books.google.com.br/books?id=Xuu6vQEACAAJ>.

IBGE 2010– Instituto Brasileiro de Geografia e Estatística. **Censo 2010**. Disponível em <https://censo2010.ibge.gov.br/resultados.html>. Acessado em 2 de julho de 2023.

INSTITUTO DE TECNOLOGIA SOCIAL (ITS). **Caderno de debate: tecnologia social: direito à ciência e ciência para a cidadania**. p. 26, 2004.

JANGDAL, L. Hyperlocals Matter: Prioritising Politics When Others Don't. **Journalism Practice**, p. 1–18, 11 mar. 2020.

JOHNSON, P. A. et al. **The cost(s) of geospatial open data**. Transactions in GIS, Blackwell Publishing Ltd, v. 21, p. 434–445, 6 2017. ISSN 14679671.

KLEINHANS, R.; HAM, M. V.; EVANS-COWLEY, J. **Using Social Media and Mobile Technologies to Foster Engagement and Self-Organization in Participatory Urban Planning and Neighbourhood Governance**. [S.l.]: Routledge, 2015. 237-247 p.

LEYSHON, M.; ROGERS, M. Designing for Inclusivity: Platforms of Protest and Participation. **Urban Planning**, v. 5, n. 4, p. 33–44, 14 out. 2020.

LIMA JUNIOR, W. T. Projeto experimental lupa nh. comunicação & inovação. **Revista Comunicação Inovação**, v. 19, p. 67–81, 2018.

LIMA JUNIOR, W. T. **JORNALISMO HIPERLOCAL E DISPOSITIVOS MÓVEIS**. [S.l.: s.n.], 2017.

LIMA JUNIOR, W. T.; COELHO, A. A. dos S. Captação de dados pela comunidade para a formação de inteligência social hiperlocal. **Intercom – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação**. In: . [S.l.: s.n.], 2015.

LIU, W. et al. Connecting With Hyperlocal News Website: Cause or Effect of Civic Participation? **American Behavioral Scientist**, v. 62, n. 8, p. 1022–1041, 19 mar. 2018.

LOELIGER, J.; MCCULLOUGH, M. **Version Control with Git**. Second edition, O'Reilly, 2012.

LOKMAN, A. S.; AMEEDDEEN, M. A. Modern Chatbot Systems: A Technical Review. **Proceedings of the Future Technologies Conference (FTC) 2018**, p. 1012–1023, 20 out. 2018.

LONGLEY, P. A. **The academic success of GIS in geography**: Problems and prospects. [S.l.]: Springer-Verlag, 2000.

MEIRELLES, F. S. FVG: Fundação Getúlio Vargas. **Pesquisa anual uso TI. 2023**. Disponível em <https://eaesp.fgv.br/producao-intelectual/pesquisa-anual-uso-ti>. Acessado em 21 de julho de 2023.

METZGAR, E. T.; KURPIUS, D. D.; ROWLEY, K. M. Defining hyperlocal media: Proposing a framework for discussion. **New Media and Society**, v. 13, p. 772–787, 8 2011. ISSN 14614448.

MIEL P.; FARIS R. **News and information as digital media** come of age. Berkman Center for Internet and Society at Harvard University. 2008. Disponível em: http://www.cyber.law.harvard.edu/sites/cyber.law/harvard.edu/files/Overview_MR.pdf. Acessado em 9 de julho de 2023.

MUKHTAR, M. et al. Conceptualizing Hyperlocal Information Systems for Developing Countries. **Proceedings of the ACM on Human-Computer Interaction**, v. 5, n. CSCW2, p. 1–26, 13 out. 2021.

Natural Language Processing (NLP) Solutions | IBM. Disponível em: <https://www.ibm.com/natural-language-processing> Acessado em 5 de janeiro de 2024.

NOCERA, L. et al. **Crosstown Foundry**: A Scalable Data-driven Journalism Platform for Hyper-local News. 9 jun. 2021.

OLIVEIRA, R. 30 milhões de pessoas vivem em desertos de notícias. **ABRAJI - Associação brasileira de Jornalismo Investigativo**. 2018. Disponível em <https://www.abraji.org.br/noticias/30-milhoes-de-pessoas-vivem-em-desertos-de-noticia-mostra-atlas-da-noticia-2-0>. Acessado em 9 de julho de 2023.

PAVLIK, J. V. A tecnologia digital e o jornalismo: as implicações para a Democracia. **Brazilian Journalism Research**, v. 7, n. 2, p. 94–118, 9 dez. 2011.

PINA, I.; PAINHO, M. **Participação pública e sistemas de informação geográfica**. 2012.

RADCLIFFE, D. **Here and Now UK hyperlocal media today**. 2013.

RADFORD, A. et al. **Improving Language Understanding by Generative Pre-Training. Preprint**. 1-12. 2018.

RICHARDSON, C. **Microservices Patterns**: With examples in Java. Manning, 2018. ISBN 9781617294549. Disponível em:
<https://books.google.com.br/books?id=UeK1swEACAAJ>.

ROBAEYST, B. et al. Online Neighborhood Networks: The Relationship Between Online Communication Practices and Neighborhood Dynamics. **Media and Communication**, v. 10, n. 2, p. 108–118, 26 maio 2022.

SAAR, C. M. A. de A.; LIMA JUNIOR, W. T.; FERREIRA, W. P. Lupa 2.0: technological evolution for strengthening of the hyperlocal social communication. **Revista Observatório Universidade Federal do Tocantins**, v. 7, p. a12en, 7 2021.

SALDIVAR, J. et al. Civic Technology for Social Innovation. **Computer Supported Cooperative Work (CSCW)**, v. 28, n. 1-2, p. 169–207, 23 maio 2018.

SAMPIERI, R. H.; COLLADO, C. F. **Metodologia de pesquisa** /. 5. ed.. ed. Porto Alegre:: Penso, 2013. Título original: Metodología de la investigacion.

SATU JUMISKO-PYYKKÖ; HELI VÄÄTÄJÄ; JAAKOLA, M. S. Quality management of user-generated content in participatory journalism. **Tampere University Institutional Repository (Tampere University)**, 2 dez. 2014.

SIEBER, R. Public participation geographic information systems: A literature review and framework. **Annals of the Association of American Geographers**, v. 96, p. 491–507, 2006.

STANFORD GRADUATE SCHOOL OF BUSINESS. **Stanford social innovation review (Online)**. Center for Social Innovation, 2003.

TRASEL, M. et al. **Pensar em Rede: pesquisa aplicada em jornalismo e tecnologias digitais**. [S.l.]: Unifap, 2021. -316 p. ISBN 978-65-89517-16-0.
TSHABANGU, T. Making News with the Citizens! Audience Participation and News-making Practices at the AMH Group. **African Journalism Studies**, p. 1–16, 12 jan. 2022.

WALKTER, A. J. **New fast method for generating discrete random numbers with arbitrary frequency distributions**. 1974.

WAZLAWICK, R. S. **Metodologia de pesquisa para ciência da computação**. Rio De Janeiro: Elsevier, 2009.

WRIGHT, D. J.; DUNCAN, S. L.; LACH, D. Social power and gis technology: A review and assessment of approaches for natural resource management. **Annals of the Association of American Geographers**, v. 99, p. 254–272, 4 2009. ISSN 00045608.

Apêndice A

TABELAS DO BANCO DE DADOS

<i>collect</i>	
Nome da coluna	Tipo de dado
id	inteiro
longitude	texto
latitude	texto
<i>hasCollect</i>	bool
<i>howManyTimes</i>	inteiro
<i>user</i>	inteiro
<i>createdAt</i>	data e hora

<i>asphalt</i>	
Nome da coluna	Tipo de dado
id	inteiro
longitude	texto
latitude	texto
<i>isPaved</i>	bool
<i>hasPavedSideWalks</i>	bool
<i>user</i>	inteiro
<i>createdAt</i>	data e hora

<i>light</i>	
Nome da coluna	Tipo de dado
Id	inteiro
Longitude	texto
Latitude	texto
<i>hasLight</i>	bool
<i>isItWorking</i>	bool

<i>hasLoosesCable</i>	bool
<i>user</i>	inteiro
<i>createdAt</i>	data e hora

<i>trash</i>	
Nome da coluna	Tipo de dado
id	inteiro
longitude	texto
latitude	texto
<i>hasRoadCleanUp</i>	bool
<i>howManyTimes</i>	inteiro
<i>hasAccumulatedTrash</i>	bool
<i>user</i>	inteiro
<i>createdAt</i>	data e hora

<i>sewer</i>	
Nome da coluna	Tipo de dado
id	inteiro
longitude	texto
latitude	texto
<i>hasHomeSewer</i>	bool
<i>hasHomeCesspool</i>	bool
<i>doesCityHallCleanTheSewer</i>	bool
<i>user</i>	inteiro
<i>createdAt</i>	data e hora

<i>water</i>	
Nome da coluna	Tipo de dado
id	inteiro
longitude	texto
latitude	texto
<i>hasCollect</i>	bool
<i>waterMissedInAWeek</i>	inteiro
<i>hasWell</i>	bool
<i>user</i>	inteiro
<i>createdAt</i>	data e hora