



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE UBERLÂNDIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL



MOISÉS DE SOUZA TEIXEIRA

**POTENCIAL DE APLICAÇÃO DE TÉCNICAS DE
MAPEAMENTO COLABORATIVO PARA O REGISTRO DE
SINISTROS DE TRÂNSITO**

Uberlândia, 2024

MOISÉS DE SOUZA TEIXEIRA

**POTENCIAL DE APLICAÇÃO DE TÉCNICAS DE
MAPEAMENTO COLABORATIVO PARA O REGISTRO DE
SINISTROS DE TRÂNSITO**

Dissertação apresentada à Faculdade de Engenharia Civil da Universidade Federal de Uberlândia, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Civil.

Área de Concentração: Engenharia Urbana,
Recursos Hídricos e Saneamento Ambiental

Orientador: Prof. Dr. Márcio Augusto Reolon
Schmidt

Uberlândia, 2024

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

T266p
2024 Teixeira, Moisés de Souza, 1987-
 Potencial de aplicação de técnicas de mapeamento colaborativo para
 o registro de sinistros de trânsito [recurso eletrônico] / Moisés de Souza
 Teixeira. - 2024.

 Orientador: Márcio Augusto Reolon Schmidt.
 Dissertação (Mestrado) - Universidade Federal de Uberlândia,
 Programa de Pós-graduação em Engenharia Civil.

 Modo de acesso: Internet.

 Disponível em: <http://doi.org/10.14393/ufu.di.2024.5057>

 Inclui bibliografia.

 Inclui ilustrações.

 1. Engenharia Civil. I. Schmidt, Márcio Augusto Reolon, 1978-,
 (Orient.). II. Universidade Federal de Uberlândia. Programa de Pós-
 graduação em Engenharia Civil. III. Título.

CDU: 624

 André Carlos Francisco
 Bibliotecário Documentalista - CRB-6/3408



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Civil				
Defesa de:	Dissertação de Mestrado Acadêmico, 301, PPGEC				
Data:	10 de abril de 2024	Hora de início:	14: 18h	Hora de encerramento:	17: 15 h
Matrícula do Discente:	12212ECV011				
Nome do Discente:	Moisés de Souza Teixeira				
Título do Trabalho:	Potencial de aplicação de técnicas de mapeamento colaborativo para o registro de sinistros de trânsito				
Área de concentração:	Análise Espacial, Engenharia Urbana, Recursos Hídricos e Saneamento Ambiental				
Linha de pesquisa:	Análise Espacial ambiental				
Projeto de Pesquisa de vinculação:	Visualização Cartográfica				

Reuniu-se, em sessão pública por webconferência e na Sala de Reuniões, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Civil, assim composta pelos Professores Doutores: Silvana Philippi Camboim - UFPR, Rogério Lemos Ribeiro - UFU e Marcio Augusto Reolon Schmidt, orientador do candidato.

Iniciando os trabalhos o presidente da mesa, Marcio Augusto Reolon Schmidt, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Marcio Augusto Reolon Schmidt, Professor(a) do Magistério Superior**, em 11/04/2024, às 14:55, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Rogério Lemos Ribeiro, Professor(a) do Magistério Superior**, em 11/04/2024, às 15:12, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Silvana Philippi Camboim, Usuário Externo**, em 15/04/2024, às 15:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5286229** e o código CRC **9436B0A9**.

Referência: Processo nº 23117.021312/2024-21

SEI nº 5286229

À Deus.

À Ketheleen Souza, minha esposa.

À Geraldo Teixeira e Eneida Estevam, meus pais.

AGRADECIMENTOS

À Deus, pela proteção e por guiar meus passos.

À minha esposa Ketheleen Souza dos Santos, pelo companheirismo, pelo apoio e por sempre acreditar em meu potencial.

À toda minha família, em especial aos meus pais e irmãos, pela assistência durante toda minha vida e por auxiliarem na construção de meu caráter.

Aos técnicos e docentes do Programa de Pós-graduação em Engenharia Civil da Universidade Federal de Uberlândia, em especial ao meu orientador Prof. Dr. Marcio Augusto Reolon Schmidt, por realmente me instruir durante o Mestrado e colaborar de forma essencial na produção do projeto de pesquisa.

Aos Professores Dr. Rogério Lemos Ribeiro e Dr. Nassau de Nogueira Nardez, pelas contribuições na etapa de qualificação do Mestrado.

Aos amigos e colegas de trabalho, pela boa convivência e experiências profissionais, em especial ao Gustavo Eugênio de Freitas Faria, por me estimular a ingressar em estudos na área da Geomática.

À todos os profissionais e acadêmicos que, de alguma forma, cooperaram com o desenvolvimento do trabalho, em especial à Dra. Denise Labrea Ferreira e à Dra. Maria Cecília de Sousa.

Ao 17º Batalhão de Polícia Militar de Minas Gerais, em especial ao Cabo Carlos Eduardo de Souza, pela prestatividade.

.

RESUMO

A atualização de dados espaciais, em especial aqueles destinados a mapeamento do território, é uma tarefa reservada às agências de mapeamentos nacionais. Sendo uma tarefa dispendiosa, é comum que os intervalos entre as campanhas sejam grandes. Nesse contexto, o mapeamento colaborativo se apresenta como uma alternativa interessante na coleta de dados, podendo contribuir de forma substancial no processo de análise espacial de dados geográficos. Diversas aplicações tem se aproveitado do voluntariado na contribuição de informações, como respostas a emergências, mapeamento de grande escala, entre outros. Desta forma, essa pesquisa buscou avaliar o potencial da aplicação de técnicas de mapeamento colaborativo para o registro de sinistros de trânsito no município de Uberlândia/MG. Foram avaliados os critérios de acurácia posicional e completude de atributos, representadas pela indicação do evento na base cartográfica da cidade e pela possível sub-notificação do registro de dados junto aos órgãos competentes, e proposto uma metodologia de avaliação extrínseca de qualidade para o parâmetro de qualidade temporal, sendo representada pela data e hora comunicadas na ocorrência. Compreendendo a potencialidade do uso do mapeamento colaborativo na captação de dados espaciais abertos, foi realizado um estudo quanto à disponibilização dessas informações em plataformas de acesso gratuito para o cadastro e armazenamento de dados referentes aos sinistros de trânsito. Para tanto, foi desenvolvido um formulário para o registro de dados colaborativos de sinistros, sendo posteriormente submetido a um processo de validação de conteúdo de instrumento de pesquisa. Os resultados obtidos foram satisfatórios, confirmando que a proposta apresentada é viável. Através do uso dessa técnica, espera-se refinar o mapeamento dos sinistros no município de Uberlândia/MG. Esse ganho de informações pode favorecer o acesso democrático a esse tipo de conteúdo por meio dos mapas digitais e fomentar a realização de análises técnicas mais assertivas sobre o comportamento do trânsito no município, apoiar políticas públicas e prevenir novos sinistros.

Palavras-chave: mapeamento colaborativo ; VGI; sinistros de trânsito ; acidentes de trânsito; qualidade dos dados; *OpenStreetMap*; sistema de informação geográfica.

ABSTRACT

The updating of spatial data, especially those intended for mapping the territory, is a task reserved for national mapping agencies. Being an expensive task, it is common that the intervals between campaigns are large. In this context, the collaborative mapping presents itself as an interesting alternative in data collection, and can substantially contribute to the process of spatial analysis of geographic data. Several applications have taken advantage of volunteering in the contribution of information, such as emergency response, large-scale mapping, among others. Thus, this research sought to evaluate the potential application of collaborative mapping techniques for the registration of traffic accidents in the municipality of Uberlândia/MG. The criteria of positional accuracy and attribute completeness were evaluated, represented by the indication of the event in the cartographic base of the city and the possible under-reporting of data records with the official agencies, and an extrinsic quality assessment methodology was proposed for the temporal quality parameter, being represented by the date and time reported in the occurrence. Understanding the potentiality of the use of collaborative mapping in the capture of open spatial data, a study was conducted regarding the availability of this information on free access platforms for the registration and storage of data related to traffic accidents. To this end, a form was developed to record collaborative accident data, which was subsequently subjected to a research instrument content validation process. The results obtained were satisfactory, confirming that the proposal presented is viable. Through the use of this technique, it is expected to refine the mapping of accidents in the city of Uberlândia/MG. This information acquisition can favour the democratic access to this type of content through digital maps and foster the realization of more assertive technical analyses about the traffic behaviour in the municipality, support public policies and prevent new accidents.

Keywords: collaborative mapping; VGI; traffic accidents; data quality; OpenStreetMap; geographic information system.

LISTA DE FIGURAS

Figura 1 - Tela de edição de objetos da plataforma <i>OpenStreetMap</i>	26
Figura 2 - Painel RENAEST	42
Figura 3 - Fluxo de dados e estatísticas de sinistros de trânsito - RENAEST	43
Figura 4 - Mapa de localização do município de Uberlândia.....	48
Figura 5 - Fluxograma com as etapas da metodologia proposta	50
Figura 6 - Estrutura dos formulários para a coleta de dados colaborativos de sinistros de trânsito.....	58
Figura 7 - Interface gráfica do serviço de geolocalização da versão móvel do aplicativo Epicollect5.....	60
Figura 8 - Interface gráfica do serviço de geolocalização da versão <i>web</i> do aplicativo Epicollect5.....	60
Figura 9 - Resultados de um sinistro de trânsito fictício registrado na versão móvel do aplicativo Epicollect5.....	62
Figura 10 - Estrutura revisada dos formulários para a coleta de dados colaborativos de sinistros de trânsito	72
Figura 11 - Resultados de um sinistro de trânsito fictício registrado na versão móvel do aplicativo Epicollect5 após ajustes de estrutura e de linguagem.....	73
Figura 12 - Mapa de localização da UFU Campus Santa Mônica.....	79
Figura 13 - Registro fotográfico do Local 1	80
Figura 14 - Registro fotográfico do Local 2.....	80
Figura 15 - Registro fotográfico do Local 3	80
Figura 16 - Registro fotográfico do Local 4.....	81
Figura 17 - Registro fotográfico do Local 5	81
Figura 18 - Representação gráfica das amostras coletadas	82
Figura 19 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas X do conjunto de dados do Local 1	83
Figura 20 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas Y do conjunto de dados do Local 1	84
Figura 21 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas X do conjunto de dados do Local 2	84
Figura 22 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas Y do conjunto de dados do Local 2	84
Figura 23 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas X do conjunto de dados do Local 3	85
Figura 24 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas Y do conjunto de dados do Local 3	85
Figura 25 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas X do conjunto de dados do Local 4	85

Figura 26 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas Y do conjunto de dados do Local 4	86
Figura 27 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas X do conjunto de dados do Local 5	86
Figura 28 - <i>Scatterplot</i> e <i>Box Plot</i> das coordenadas Y do conjunto de dados do Local 5	86
Figura 29 - <i>Scatterplot</i> e <i>Box Plot</i> da amostra coletada, após remoção de <i>outliers</i>	88
Figura 30 - Representação gráfica dos pontos amostrais e dos pontos homólogos de referência.....	90
Figura 31 - Representação gráfica dos registros fictícios	100
Figura 32 - Representação gráfica da mediana dos registros fictícios	102

LISTA DE QUADROS

Quadro 1 - Comparação entre as plataformas do <i>Google Maps</i> e do <i>OpenStreetMap</i>	26
Quadro 2 - Dados coletados em diferentes sistemas oficiais de registro de sinistros de trânsito.....	46
Quadro 3 - Descrição dos especialistas participantes do processo de validação de conteúdo	63
Quadro 4 - Configuração do questionário elaborado para a avaliação dos itens do instrumento de pesquisa	64
Quadro 5 - Configuração do questionário elaborado para a avaliação geral do instrumento de pesquisa	65
Quadro 6 - Relação de causas presumidas para a ocorrência de um sinistro de trânsito	70
Quadro 7 - Dados coletados em diferentes sistemas oficiais e colaborativos de registro de sinistros de trânsito	75
Quadro 8 - Medida de qualidade temporal - horário do sinistro de trânsito	93
Quadro 9 - Registro fictício de dados referentes a data e hora do sinistro.....	94
Quadro 10 - Medida de tendência central para a especificação do valor válido de um grupo de dados	98
Quadro 11 - Relação de peso dos dados registrados para cada nível de usuário.....	99
Quadro 12 - Valores dos atributos de cada registro fictício	101
Quadro 13 - Cálculo dos valores válidos de cada atributo	101

LISTA DE TABELAS

Tabela 1 - Quantidade de classes de objetos por categorias de informações	28
Tabela 2 - Quantidade de possíveis valores associados às principais chaves sugeridas para uso na plataforma OSM	30
Tabela 3 - Padrão de exatidão cartográfica da planimetria dos produtos cartográficos digitais em grandes escalas.....	54
Tabela 4 - Valores do CVC referentes aos itens do instrumento de pesquisa.....	66
Tabela 5 - Valores do CVC referentes ao instrumento de pesquisa como um todo	67
Tabela 6 - Grau de completude para cada sistema de registro de sinistros de trânsito.....	76
Tabela 7 - Quantidade de campos habilitados para coleta no formulário colaborativo para cada item que compõe um RAT	76
Tabela 8 - Grau de completude no formulário colaborativo para cada item que compõe um RAT	77
Tabela 9 - Resultados estatísticos básicos das amostras coletadas.....	83
Tabela 10 - Resultados estatísticos básicos da amostra coletada, após remoção de <i>outliers</i> ..	88
Tabela 11 - Medidas de dispersão das amostras coletadas, após remoção de <i>outliers</i>	89
Tabela 12 - Mediana das coordenadas e medidas para avaliação da acurácia posicional dos pontos amostrais	91
Tabela 13 - Média das dimensões das quadras de Bairros do município de Uberlândia	96

LISTA DE SÍMBOLOS

%	- por cento
e_{Hi}	- distância euclidiana (componente horizontal do erro)
e_{xi}	- erro da componente planimétrica na direção x
e_{yi}	- erro da componente planimétrica na direção y
EMQ_H	- erro médio quadrático
h	- hora
m	- metro
σ	- desvio padrão populacional
x_i	- dado de uma população
μ	- média populacional
N	- número total de dados ou registros

LISTA DE ABREVIATURAS E SIGLAS

ABNT	- Associação Brasileira de Normas Técnicas
ANTT	- Agência Nacional de Transportes Terrestres
BO	- Boletim de Ocorrência
CBMMG	- Corpo de Bombeiros Militar de Minas Gerais
CONCAR	- Comissão Nacional de Cartografia
CONTRAN	- Conselho Nacional de Trânsito
CTB	- Código de Trânsito Brasileiro
CVC	- Coeficiente de Validade de Conteúdo
DATASUS	- Departamento de Informática do Sistema Único de Saúde
DC	- Diagrama de Classes
Denatran	- Departamento Nacional de Trânsito
DNIT	- Departamento Nacional de Infraestrutura de Transportes
DPVAT	- Danos Pessoais Causados por Veículos Automotores de Via Terrestre
DSG	- Diretoria de Serviço Geográfico
EADI	- Estação Aduaneira do Interior
EP	- Erro-Padrão
ET-ADGV	- Especificações Técnicas para Aquisição de Dados Geoespaciais Vetoriais
ET-CQDG	- Especificação Técnica para Controle de Qualidade de Dados Geoespaciais
ET-EDGV	- Especificações Técnicas para Estruturação de Dados Geoespaciais Vetoriais
FCA	- Ferrovia Centro-Atlântica
GNSS	- Global Navigation Satellite System
GPS	- Global Positioning System
IBGE	- Instituto Brasileiro de Geografia e Estatística
IEC	- Inteligência Espacial Coletiva
ISO	- International Organization for Standardization
LGPD	- Lei Geral de Proteção de Dados Pessoais
MG	- Minas Gerais
NBR	- Norma Brasileira
OMS	- Organização Mundial de Saúde
ONSV	- Observatório Nacional de Segurança Viária
ONU	- Organização das Nações Unidas

OSM	- OpenStreetMap
OPAS	- Organização Pan-Americana da Saúde
PCMG	- Polícia Civil do Estado de Minas Gerais
PEC	- Padrão de Exatidão Cartográfica
PEC-PCD	- Padrão de Exatidão Cartográfica dos Produtos Cartográficos Digitais
Perfil MGB	- Perfil de Metadados Geospaciais do Brasil
PMMG	- Polícia Militar de Minas Gerais
PMU	- Prefeitura Municipal de Uberlândia
PNATRANS	- Plano Nacional de Redução de Mortes e Lesões no Trânsito
PRF	- Polícia Rodoviária Federal
RAT	- Relatório de Acidente de Trânsito
REDS	- Registro de Eventos de Defesa Social
RENACH	- Registro Nacional de Carteira de Habilitação
RENAEST	- Registro Nacional de Acidentes e Estatísticas de Trânsito
RENAINF	- Registro Nacional de Infrações de Trânsito
RENAVAM	- Registro Nacional de Veículos Automotores
RCO	- Relação de Classes de Objetos
SAMU	- Serviço de Atendimento Móvel de Urgência
SCN	- Sistema Cartográfico Nacional
SIDS	- Sistema Integrado de Defesa Social
SIG	- Sistema de Informação Geográfica
SIH/SUS	- Sistema de Informações Hospitalares do Sistema Único de Saúde
SIM	- Sistema de Informações sobre Mortalidade
Sisp	- Sistema Integrado de Segurança Pública
SNT	- Sistema Nacional de Trânsito
SP	- São Paulo
UFU	- Universidade Federal de Uberlândia
UTM	- Universal Transversa de Mercator
VGI	- Volunteered Geographic Information

SUMÁRIO

1	INTRODUÇÃO	16
1.1	OBJETIVOS	20
1.1.1	<i>Objetivo Geral.....</i>	20
1.1.2	<i>Objetivos Específicos</i>	20
1.2	JUSTIFICATIVA.....	20
1.3	ESTRUTURA DO TRABALHO.....	21
2	REVISÃO BIBLIOGRÁFICA.....	23
2.1	CONCEITO DE DADOS COLABORATIVOS E ORIGEM DO TERMO	23
2.1.1	<i>A Plataforma OpenStreetMap.....</i>	24
2.2	MAPEAMENTO COLABORATIVO DE EVENTOS EFÊMEROS.....	31
2.2.1	<i>Mapeamento Colaborativo Aplicado a Sinistros de Trânsito</i>	33
2.3	AVALIAÇÃO DA QUALIDADE DE DADOS COLABORATIVOS	35
2.3.1	<i>Parâmetros da Qualidade dos Dados Geográficos</i>	36
2.4	COLETA DE DADOS DE SINISTROS DE TRÂNSITO.....	39
2.4.1	<i>Sinistros de Trânsito no Brasil e no Mundo</i>	39
2.4.2	<i>Métodos de Coleta de Dados de Sinistros de Trânsito</i>	41
2.4.3	<i>Dados Estatísticos de Sinistros de Trânsito</i>	44
3	MATERIAIS E MÉTODOS	48
3.1	ÁREA DE ESTUDO	48
3.2	MATERIAIS UTILIZADOS	49
3.3	MÉTODO	50
3.3.1	<i>Construção do Questionário no Epicollect5 para o Registro Voluntário.....</i>	51
3.3.2	<i>Proposta de Garantia da Qualidade dos Dados Colaborativos.....</i>	52
3.3.3	<i>Critérios de Análise e Filtragem de Registros.....</i>	56
4	RESULTADOS E DISCUSSÕES	57
4.1	FORMULÁRIO PARA O REGISTRO COLABORATIVO DE SINISTROS.....	57
4.1.1	<i>Elaboração do Formulário.....</i>	57
4.1.2	<i>Validação de Conteúdo</i>	63
4.2	AVALIAÇÃO DA COMPLETUDE DE ATRIBUTOS.....	74
4.3	AVALIAÇÃO DA ACURÁCIA POSICIONAL	78
4.3.1	<i>Precisão da Coleta In Loco - Análise do Sensor GPS</i>	78

4.3.2	<i>Avaliação da Acurácia Posicional</i>	89
4.4	AVALIAÇÃO DA ACURÁCIA TEMPORAL	92
4.5	CRITÉRIOS DE ANÁLISE E FILTRAGEM DE REGISTROS	94
4.5.1	<i>Critérios de Análise e Filtragem de Registros Múltiplos</i>	94
4.5.2	<i>Critério para Definição dos Atributos Válidos</i>	98
4.5.3	<i>Exemplo Prático do Método Proposto</i>	100
5	CONCLUSÕES E RECOMENDAÇÕES	103
	REFERÊNCIAS BIBLIOGRÁFICAS	107
	APÊNDICE A - RESULTADOS RETORNADOS PELOS JUÍZES NA AVALIAÇÃO DOS ITENS DO INSTRUMENTO DE PESQUISA	124
	APÊNDICE B - RESULTADOS RETORNADOS PELOS JUÍZES NA AVALIAÇÃO GERAL DO INSTRUMENTO DE PESQUISA	125

1 INTRODUÇÃO

Em uma pesquisa sobre a cobertura do mapeamento brasileiro, Silva e Camboim (2020) identificaram que apenas 5,5% do território brasileiro apresenta mapeamento vinculado ao Sistema Cartográfico Nacional (SCN) na escala 1:25.000, sendo que esse valor decresce muito à medida que a escala aumenta. Por ser uma tarefa cara e dispendiosa, o tempo para a atualização do mapeamento em escala regional no Brasil pode variar de 10 a 15 anos e, por essa razão, agências oficiais de mapeamento devem estabelecer medidas alternativas para a identificação de alterações em feições territoriais (CAROTI; PIEMONTE; NESPOLI, 2017). Visto que a estrutura tradicional de mapeamento é caracterizada por imagens aéreas, fotogrametria e trabalho de campo, apoiados em metodologias com especificações técnicas detalhadas, o processo de produção cartográfica requer muito tempo para a sua finalização, o que causa impacto direto na regularidade de mapas atualizados (SANTOS *et al.*, 2011).

Neste contexto, Goodchild (2007) apresentou o conceito de Informações Geográficas Voluntárias, do inglês *Volunteered Geographic Information* (VGI), em que os dados geoespaciais passam a ser produzidos e distribuídos por pessoas com dispositivos móveis, desde que tenham acesso a internet, a receptores GNSS (*Global Navigation Satellite System*) e a imagens de alta resolução via *internet*. Segundo Hecht, Kunze e Hahmann (2013), muitas das informações, que tradicionalmente necessitavam ser coletadas e tratadas por técnicos especialistas, passaram a ser disponibilizadas por diferentes usuários nas mídias digitais. Isso criou uma rede de voluntários que coletam e publicam informações geográficas sobre determinado local e também validam as informações postadas por outros usuários em plataformas *online*, como o *OpenStreetMaps*, *Wikimapia*, *Wikiloc* e muitos outros.

Yan *et al.* (2020) ressaltam que a importância no que tange o VGI alcança toda a sociedade, visto que estas fontes de dados são indispensáveis para compreender a dinâmica espacial de uso e ocupação do solo. Uma vez que existem áreas no Brasil sem nenhum mapeamento ou com dados oficiais desatualizados, os dados colaborativos podem ser uma alternativa viável ao governo para aumentar a cobertura de seus dados de referência (SILVA; CAMBOIM, 2020).

Contudo, a utilização de dados geoespaciais colaborativos requer o conhecimento da sua qualidade. A qualidade é um aspecto importante a ser considerado em dados geoespaciais, pois permite determinar sua adequação para uso em propósitos, sendo que para dados colaborativos, essa questão se torna ainda mais mitigadora, pois os dados possuem aspectos heterogêneos e podem variar de acordo com a localização, características das colaborações e perfil dos colaboradores (ELIAS *et al.*, 2021b).

A ISO 19157 (2013) e a norma nacional baseada nesta, a ET-CQDG (BRASIL, 2016), apresentam os parâmetros que podem identificar a qualidade do dado espacial: a acurácia posicional, a acurácia temática, a qualidade temporal, a completude, a consistência lógica e a usabilidade. Cada uma dessas classes tem critérios e análises que ainda estão na fase de desenvolvimento. Pesquisas como Elias *et al.* (2020, 2021a), Girres e Touya (2010), Haklay *et al.* (2010), Hecht, Kunze e Hahmann (2013), Minghini e Frassinelli (2019) e Paiva (2021) tem sido realizadas no sentido de estimar parâmetros intrínsecos e extrínsecos na tentativa de estabelecer parâmetros qualitativos e semânticos para as contribuições.

Uma das aplicações que podem se beneficiar da participação da sociedade civil é a geografia urbana, responsável pelo estudo do espaço urbano e de todas as suas características. A mobilidade urbana, isto é, o trânsito de veículos e de pedestres, configura um de seus principais aspectos. De acordo com Pereira *et al.* (2021), o Brasil tem experimentado mudanças consideráveis dos padrões de mobilidade urbana nas últimas décadas, causadas principalmente pela queda da demanda por transporte público e o aumento significativo do uso do transporte individual. Segundo os autores, a frota de veículos individuais motorizados (automóveis e motocicletas) das cidades brasileiras aumentou em média 331% no período de 2001 a 2020, sendo esse crescimento mais expressivo nos municípios com até 20 mil habitantes. Essas mudanças trazem diversas consequências negativas à sociedade, como o aumento dos níveis de poluição na cidade, congestionamentos, aumento do número de sinistros de trânsito, dentre outros.

Um dos parâmetros relacionado ao estudo desse tema que merece destaque diz respeito aos sinistros de trânsito. De acordo com a ABNT NBR 10697 (2020), sinistro de trânsito é todo evento que resulte em dano ao veículo ou à sua carga e/ou em lesões a pessoas e/ou animais, e que possa trazer dano material ou prejuízos ao trânsito, à via

ou ao meio ambiente, em que pelo menos uma das partes está em movimento nas vias terrestres ou em áreas abertas ao público.

Neste contexto, a ABNT NBR 10697 (2020) apresenta alguns termos e definições importantes. São eles: fatores geradores de sinistro de trânsito e fatores contribuintes do sinistro de trânsito. Enquanto o primeiro se refere a cada fator que contribui diretamente para a ocorrência de um sinistro de trânsito, o segundo diz respeito a cada fator que contribui para o aumento da probabilidade de sinistros de Trânsito. Os principais tipos de sinistros de trânsito são: atropelamento de animal, atropelamento de pessoa, capotamento, choque, colisão (frontal, lateral, transversal ou traseira), engavetamento, queda e tombamento (ABNT NBR 10697, 2020).

Há várias formas de coletar dados referentes a acidentes de trânsito, assim como também há várias maneiras de se analisar esses dados. Segundo Torres, Pereira e Machado (2020), algumas das possíveis fontes de dados para o estudo dos acidentes de trânsito são: Boletins de Ocorrência (BO), registros de ocorrências atendidas pelo Corpo de Bombeiros, registros de atendimentos pré-hospitalares e registros hospitalares das vítimas de acidentes.

Contudo, Souza (2011) destaca que o preenchimento dos formulários, a depender da ocasião, pode ocorrer ou não no local do acidente, podendo comprometer a qualidade dos dados coletados. Para o autor, as principais circunstâncias relacionadas ao preenchimento que podem prejudicar a qualidade dos dados são: a) as condições emocionais dos envolvidos, o que pode reduzir a vontade destas em preencher questionários e responder perguntas; b) atropelamentos em que a vítima requer cuidados e os danos ao veículo são mínimos, sendo comum nesses casos a anotação tardia do registro na entrada do pronto-socorro ou em delegacia de polícia; c) cadastro da ocorrência por parte de profissionais da área de pronto-socorro e de saúde que, muitas vezes, não se atentam ao preenchimento dos dados referentes ao local do acidente; d) indicação equivocada de nomes das ruas e da numeração dos imóveis, devido à ausência de placas indicativas no local.

Dentre os parâmetros que identificam a qualidade dos dados espaciais, valem ser destacadas a acurácia posicional, a qualidade temporal e a completude de atributos. Neste sentido, para cada um desses parâmetros, é possível identificar algumas limitações que afetam na qualidade dos dados coletados por meio dos métodos oficiais de registro, tais como os citados no parágrafo anterior. No Brasil, é

comum que a posição dos sinistros seja indicada pelo endereço e/ou ponto de referência. Entretanto, muitas vezes os boletins são preenchidos com endereços incompletos ou com pontos de referência que já não existem mais (ROCHA, 2015). Além disso, o processo de transformação de endereço em coordenadas (geocodificação) de eventos dinâmicos já é por si só um desafio (GUILLEN, 2015). Erros de posicionamento também podem ocorrer, pois a posição em muitos casos é estimada, principalmente em vias muito movimentadas, onde os veículos envolvidos no sinistro precisam se mover para um local seguro antes de tomar as devidas providências. A temporalidade também é um aspecto suscetível a inconsistências em registros de sinistros de trânsito, visto que, quando as informações são tomadas a *posteriori*, ficam sujeitas a julgamento e impressões do declarante que, muitas vezes, pode fornecer dados incorretos. Quanto à completude da informação, vale destacar que, em geral, pequenos sinistros não são reportados. Costa (2016) divide esse mesmo pensamento ao afirmar que os sinistros leves têm uma elevada taxa de sub-notificação, uma vez que é comum as partes envolvidas optarem por não consumir o registro, o que prejudica no processo de quantificação e qualificação dos eventos.

Em seus estudos sobre análise espacial de sinistros de trânsito, uma das conclusões de Rocha (2015) enfatiza o problema da sub-amostragem do número de registros de sinistros de trânsito, especialmente em regiões de baixa renda, pois há uma maior possibilidade de as pessoas envolvidas nos acidentes apresentarem irregularidades com os documentos de habilitação e do veículo, além da falta de seguro, o que reduz o interesse do indivíduo em contatar os órgãos responsáveis para o registro de sinistros. Esta realidade não é exclusiva do Brasil. Segundo Abegaz *et al.* (2014), o grau de completude de registro de sinistros de trânsito nos países em desenvolvimento, em geral, é sub-reportado.

Por isso, esta pesquisa adota a hipótese de que o uso do mapeamento colaborativo na captação de dados espaciais referentes aos sinistros de trânsito por meio de contribuição voluntária tem potencial para complementar e atualizar informações inerentes ao ocorrido de forma mais próxima da realidade, minimizando as sub-notificações e equívocos de dados coletados.

Para isso, foi apresentado uma proposta de avaliação de qualidade de dados colaborativos a partir da análise da forma atual de coleta e registro de sinistros em Uberlândia/MG e, posteriormente, a disponibilização dos resultados das contribuições

validadas por meio de testes com uso de plataformas de acesso gratuito. Os resultados podem apoiar mudanças no trânsito por meio da regionalização de ocorrências, melhorar a completude das informações por meio da análise de quantidade de contribuições e posicionamento e, finalmente, tem potencial de apoiar análises técnicas dos setores responsáveis.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

O objetivo desta pesquisa é avaliar o potencial de aplicação de técnicas de mapeamento colaborativo para a captação de dados espaciais referentes a ocorrência de sinistros de trânsito do município de Uberlândia/MG.

1.1.2 Objetivos Específicos

- Definir os atributos mínimos comuns nos diferentes métodos de registro de sinistro de trânsito realizados por órgãos oficiais do município de Uberlândia/MG;
- Definir o conjunto de dados colaborativos com potencial de complementar os registros;
- Propor uma metodologia de avaliação extrínseca para assegurar a qualidade de dados colaborativos referentes a sinistros de trânsito;
- Disponibilizar um protótipo por meio de testes com uso de plataformas de acesso gratuito;
- Definir regras para a integração de dados abundantes quando relacionados a um mesmo evento.

1.2 JUSTIFICATIVA

Compreender o comportamento do trânsito de um município do porte de Uberlândia/MG é de extrema importância para o planejamento e gestão do trânsito e do transporte da cidade. Dentre os diversos problemas existentes, a falta de estatísticas

e dados que apoiem esses sistemas é um desafio para a manutenção da qualidade e segurança do trânsito. Destacam-se, nesse contexto, a sub-notificação de ocorrências e as discrepâncias das informações em função da dinâmica dos sinistros de trânsito. De acordo com Souza (2011), mapear a distribuição dos sinistros é fundamental para estabelecer políticas públicas destinadas a atenuar esse grave problema contemporâneo de saúde pública.

Costa e Moraes (2009) afirmam que a análise espacial é uma importante ferramenta para subsidiar pesquisadores com dados mais exatos referentes à localização dos pontos críticos de sinistros de trânsito. Isto possibilita estudos de engenharia de tráfego e o planejamento do trânsito, além de conduzir ações de promoção à saúde e assistência às vítimas.

Desta forma, o uso do mapeamento colaborativo para o registro de sinistros de trânsito poderá viabilizar, a qualquer pessoa interessada no assunto, um acesso prático às informações referentes a sinistros de trânsito via *webmapas*, além de proporcionar aos setores responsáveis o desenvolvimento mais preciso de estudos técnicos voltados à análise espacial dos dados geográficos coletados. Salienta-se a importância da qualidade dos dados geoespaciais. Ou seja, é necessário definir a posição correta dos dados coletados, registrar com maior precisão a data e o horário das ocorrências, e aumentar a capacidade de captação dos acidentes, independente das circunstâncias em que os eventos ocorreram. Elevar o grau de atendimento desses parâmetros de qualidade traduz em um mapeamento mais preciso, melhorando a eficiência das análises técnicas realizadas a partir das informações cadastradas.

Entretanto, a limitação está na qualidade dos dados espaciais informados. Dentre as avaliações intrínsecas e extrínsecas, essa pesquisa restringe a avaliação para o segundo caso, comparando avaliações posicional, temporal e completude. Essas características dos dados, classificadas como fatores extrínsecos, são possíveis de traduzir a adequação dos dados VGI após uma avaliação que expresse sua qualidade baseada nos fatores selecionados para esta pesquisa.

1.3 ESTRUTURA DO TRABALHO

O conteúdo da dissertação é estruturado em cinco capítulos.

- Capítulo 1: dedicado à introdução, aos objetivos, à justificativa da pesquisa e à delimitação da estrutura do trabalho;
- Capítulo 2: apresenta uma revisão bibliográfica de conceitos importantes relacionados ao trabalho, sendo abordado assuntos como mapeamento colaborativo, avaliação da qualidade de dados colaborativos, relação de dados estatísticos de sinistros de trânsito e métodos de coleta dos dados oficiais;
- Capítulo 3: descreve a área de estudo, especifica os materiais utilizados para o desenvolvimento do trabalho e apresenta os procedimentos metodológicos adotados na pesquisa;
- Capítulo 4: exhibe os resultados e discussões da pesquisa, sendo eles a elaboração e validação de conteúdo do formulário para o registro colaborativo de sinistros de trânsito, a descrição dos procedimentos para a avaliação da qualidade dos dados colaborativos e o detalhamento dos critérios para a análise e filtragem de registros;
- Capítulo 5: expõe as conclusões da pesquisa e apresenta sugestões para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

2.1 CONCEITO DE DADOS COLABORATIVOS E ORIGEM DO TERMO

Dados são qualquer conjunto de elementos associados e ordenados que possibilitam a percepção de um fato. Sua estrutura é formada por três elementos básicos: posição, descrição e forma. Os dados são parte essencial do conceito de informação, uma vez que a informação é um produto resultante da observação dos dados. Entende-se por dados geoespaciais aqueles que retratam os fenômenos geográficos espacialmente distribuídos na superfície terrestre com base em um sistema de coordenadas, podendo compreender demais atributos (não espaciais, temporais, dentre outros). O processamento e análise dos dados o transformam em informação geoespacial, adicionando significado e contexto ao dado. E o uso da informação para algum propósito produz conhecimento (JUNIOR; RIBEIRO, 2012).

Nas últimas décadas, muitos países passaram a desenvolver e padronizar plataformas na web para arquivamento da infraestrutura dos dados espaciais, facilitando o acesso a tais informações (UREÑA-CÁMARA, 2019). Diante deste cenário, nota-se o surgimento e evolução de um fenômeno no mundo do geoprocessamento: com a democratização das tecnologias digitais e do acesso à internet, um grande número de pessoas, independente de sua experiência com processamento de dados, passaram a produzir e distribuir, de forma voluntária, dados geoespaciais, uma função que durante muito tempo foi responsabilidade exclusiva de órgãos oficiais (GOODCHILD, 2007). Esse é o conceito de informação geográfica voluntária, do inglês *Volunteered Geographic Information* (VGI), também denominado em diversos estudos por mapeamento colaborativo (BRAVO; SLUTER, 2018; DORN; TORNROS; ZIPF, 2015; JASIM; AL-HAMADANI, 2020; MACHADO; CAMBOIM, 2019; SILVA; CAMBOIM, 2021; TEIXEIRA; SCHMIDT, 2023).

Em 2017, um artigo publicado no jornal britânico *The Economist* apontou que os dados se tornaram o recurso mais valioso do mundo, passando a ser conhecido como o petróleo da era digital. Em diversos estudos recentes revela-se o valor econômico da informação geográfica, além dos benefícios à sociedade pelo uso dos

diversos serviços geoespaciais disponíveis. Também vale destacar que diversas organizações internacionais consideram informação geográfica uma ferramenta essencial para garantir a boa governança (ARIZA-LÓPEZ *et al.*, 2020).

Muitas são as políticas públicas que demandam produção e utilização de informação geográfica, como é o caso da educação, planejamento urbano e rural, gestão ambiental, agricultura e mineração. Entende-se que a qualidade de dados geoespaciais tem relação direta com o grau de eficiência da informação geográfica, isto é, fatores como disponibilidade do dado geográfico, organização, metodologia de interpretação e técnicas de processamento adequadas geram geoinformação mais produtiva, eficaz e confiável (JUNIOR; RIBEIRO, 2012).

O VGI é o resultado da crescente variedade de interações possíveis devido à evolução da *Web* (GOODCHILD, 2007). O mapeamento colaborativo se desenvolveu a partir da chegada da plataforma *Web 2.0*, pois seu sistema passou a permitir que geoinformação fosse gerada por cidadãos comuns sem grandes conhecimentos em produção e processamento de dados geoespaciais. No sistema *Web 1.0*, o usuário não tinha permissão para editar o conteúdo visualizado na internet, ao passo que na plataforma *Web 2.0* passa a ser possível que o usuário não só modifique informações oferecidas como também produza e disponibilize novos conteúdos. São exemplos as plataformas de redes sociais (como o *Facebook* e o *Instagram*) e as ferramentas de adição, edição e distribuição de conteúdos variados (como o projeto *Wikipedia*) (BRAVO; SLUTER, 2018). Na era da *Web 2.0*, o ramo da internet usado para a produção e propagação de informações geoespaciais é também chamado de *Geoweb*, concentrando um grande número de ferramentas e dados georreferenciados, usualmente organizados em plataformas *online*, e que podem ser utilizados para abordar temas diversos e apoiar vários tipos de iniciativas (JOHNSON; SIEBER, 2013).

Exemplos de sistemas da *Geoweb* são as plataformas de mapeamento colaborativo, como o *OpenStreetMap*, o *Ushahidi* e o *Wikimapia*, sistemas de *software* livres destinados à criação e distribuição voluntária de conteúdo geográfico na internet (BRAVO; SLUTER, 2018).

2.1.1 A Plataforma *OpenStreetMap*

Embora a qualidade de dados colaborativos não seja homogênea, diversos estudos apontam que ela é satisfatória, a depender do objetivo. Nesse sentido, o VGI se apresenta como uma alternativa interessante no apoio ao desenvolvimento de trabalhos voltados ao planejamento urbano dos municípios e também na complementação de dados referentes ao mapeamento oficial (MACHADO; CAMBOIM, 2019).

Em seus estudos, Johnson e Sieber (2013) apontam dois motivos que incentivam as instituições públicas a utilizarem dados provenientes do mapeamento colaborativo. A primeira razão diz respeito ao potencial dos cidadãos atuarem como sensores de seu ambiente (GOODCHILD, 2007): cada vez mais os municípios desempenham atividades de gestão e planejamento urbano, entretanto sofrem com a falta de recursos e mão de obra para a produção de dados. Essa demanda em aberto gera oportunidade às instituições usufruírem do VGI, alimentando sua infraestrutura de dados geoespaciais com informações produzidas e distribuídas em plataformas na internet de forma voluntária. O segundo motivo refere-se ao fato de que o uso do VGI por parte das instituições públicas fortalece a relação entre o cidadão e o Governo, estimulando a população a contribuir, de forma mais ativa, com os projetos econômicos, sociais e ambientais da sua região. Com frequência, os usuários que praticam o mapeamento colaborativo são moradores locais que, tendo um maior conhecimento da dinâmica do espaço em que vivem, podem identificar e relatar mudanças de forma mais ágil do que empresas privadas ou servidores públicos que, muitas vezes, têm disponíveis somente dados desatualizados (JOHNSON; SIEBER, 2013).

Dentre as plataformas de mapeamento colaborativo, o *OpenStreetMap* (OSM) é o mais conhecido, uma vez que dispõe de um grande número de usuários, e é composto por elementos bem documentados, cujas tags e descrições semânticas são organizadas e disponibilizadas na internet através de wikis. Por apresentar uma concepção de armazenamento de informações com lógica semelhante ao conceito de mapeamento de referência oficial, a plataforma OSM vem sendo avaliada por diversos trabalhos científicos (MACHADO; CAMBOIM, 2019).

O projeto OSM foi criado pelo britânico Steve Coast no ano de 2004. O desenvolvimento do geoportal foi motivado, dentre outras razões, pelo aumento do número de usuários detentores de dispositivos móveis equipados com receptores

GNSS e pelo acesso factível a imagens aéreas de alta resolução via internet (HECHT; KUNZE; HANMANN, 2013). A plataforma dispõe de mosaicos de imagens de satélite de média resolução e ortofotos, que funcionam como camadas base que propiciam a vetorização dos dados geográficos usando as primitivas geométricas ponto, linha ou polígono (Figura 1).

Figura 1 - Tela de edição de objetos da plataforma *OpenStreetMap*



Fonte: *OpenStreetMap* (2023)

Diferentemente do *Google Maps* (*software* proprietário), o OSM é uma plataforma de uso livre sob uma licença aberta, e todos os dados geográficos produzidos e distribuídos na plataforma são livres e editáveis, podendo ser usados para qualquer fim, desde que o usuário atribua os devidos créditos ao *OpenStreetMap* e aos seus contribuidores (OSM, 2023). O Quadro 1 demonstra um comparativo entre as plataformas citadas.

Quadro 1 - Comparação entre as plataformas do *Google Maps* e do *OpenStreetMap*

Google Maps	OpenStreetMap
<i>Software</i> proprietário (os dados são de propriedade da empresa)	<i>Software</i> livre (os dados são abertos)
Licença comercial	Licença aberta
Os usuários contribuem com informações, podendo apenas visualizá-las	Os usuários contribuem com informações, podendo visualizá-las e acessar seus dados brutos

A elevada disponibilidade de informações geográficas, frequentemente atualizadas e revisadas pelos colaboradores, torna o *OpenStreetMap* escopo de

diversos estudos. De acordo com Yan *et al.* (2020), *OpenStreetMap* é o termo mais mencionado em artigos científicos internacionais sobre VGI. Em seu trabalho, Pizzolotto (2022) realizou um levantamento do número de publicações científicas no mundo sobre a plataforma OSM. Os resultados apontam um total de 2337 documentos no período de 2004 a 2021, com um volume de publicações crescente com o tempo. Em ambos os estudos, o Brasil está entre os 13 países com o maior número de artigos publicados sobre os temas (VGI e a plataforma *OpenStreetMap*).

Em 2020, o *OpenStreetMap* já dispunha de mais 7 milhões de colaboradores cadastrados, sendo 1,1 milhão de usuários efetivos, contribuindo com uma média de 4,5 milhões de edições por dia (SILVA; CAMBOIM, 2021). Atualmente, o OSM conta com mais de 11 milhões de usuários registrados em sua plataforma (OSMstats, 2023). Representando cerca de 2% do total de contribuição, o Brasil vem se destacando nos últimos anos no número de colaborações à plataforma, equiparando-se em participação a países como Canadá e Espanha (OLIVEIRA et al., 2022)

As normas brasileiras são apresentadas por meio das especificações técnicas homologadas pela Comissão Nacional de Cartografia (CONCAR). Estas especificações baseiam-se nas normas ISO e expandem os critérios para aquisição, estruturação e controle de qualidade de dados geoespaciais, para o contexto brasileiro. Para este tópico do trabalho, vale citar a ET-EDGV (BRASIL, 2017), versão 3.0.

A ET-EDGV (Especificações Técnicas para Estruturação de Dados Geoespaciais Vetoriais) tem como principal objetivo padronizar as estruturas de dados geoespaciais vetoriais que compõem produtos cartográficos digitais do mapeamento de referência produzidos em pequenas e grandes escalas (BRASIL, 2017). A norma propõe um modelo conceitual para a estruturação de dados vetoriais na qual as classes de objetos são agrupadas por categorias de informação, observando o aspecto funcional comum. As classes de objetos reúnem instâncias de dados geoespaciais com características e semântica semelhantes. Podem ser do tipo convencional (sem propriedades geométricas) ou do tipo georreferenciada (com representação espacial e associadas às primitivas geométricas ponto, linha e/ou polígono) e são compostas por atributos e comportamentos (BRASIL, 2017). Os atributos detalham as características estáticas dos objetos, e podem ser do tipo alfanumérico ou com geometria. Já os comportamentos especificam a dinâmica dos objetos. Classes de uma mesma categoria podem ser replicadas nos diagramas de outras categorias, para instanciar a

feição de interesse, sem implicar na duplicação desta classe. As categorias de informação são divididas em dois grupos (ver Tabela 1), os produzidos em pequenas escalas (MapTopoPE) e os produzidos em grandes escalas (MapTopoGE) (BRASIL, 2017).

Tabela 1 - Quantidade de classes de objetos por categorias de informações

Grupo	Categorias de Informações		Classes de Objetos (quantidade)
	Abreviação	Descrição	
MapTopoPE	ENC	Energia e Comunicações	16
	ECO	Estrutura Econômica	3
	HID	Hidrografia	20
	LML	Limites e Localidades	25
	PTO	Pontos de Referência	5
	REL	Relevo	18
	SAB	Saneamento Básico	4
	TRA	Sistema de Transporte	16
	AER	Sistema de Transporte / Subsistema Aeroportuário	2
	DUT	Sistema de Transporte / Subsistema Dutos	2
	FER	Sistema de Transporte / Subsistema Ferroviário	5
	HDV	Sistema de Transporte / Subsistema Hidroviário	7
	ROD	Sistema de Transporte / Subsistema Rodoviário	3
	VEG	Vegetação	13
MapTopoGE	VER	Área Verde	4
	CBGE	Classes Base do Mapeamento Topográfico em Grandes Escalas	29
	LAZ	Cultura e Lazer	9
	EDF	Edificações	35
	EMU	Estrutura de Mobilidade Urbana	9

Fonte: Adaptada de ET-EDGV (BRASIL, 2017).

Segundo a ET-EDGV 3.0 (BRASIL, 2017), as categorias de informação são visualizadas como pacotes, constituindo os chamados diagramas de classes (DC), que descrevem a estrutura e os relacionamentos espaciais e não espaciais das classes de objetos correspondentes e seus atributos associados. Em seus anexos, a ET-EDGV 3.0 apresenta a estruturação e descrição semântica das classes de objetos propostas pela norma. O Anexo A exhibe os diagramas de classe simplificados para cada categoria de informação relacionada na norma, e as correspondentes relações de classes de objetos (RCO) com a especificação dos atributos que compõem a estrutura de cada classe. O Anexo B exhibe as listas de domínio referentes aos valores que os atributos podem

assumir na estrutura das RCO's. Desta forma, qualquer elemento geográfico representado no mapa pode ser identificado, classificado e descrito (BRASIL, 2017).

Já a plataforma *OpenStreetMap* apresenta um modelo conceitual constituído por 3 tipos de elementos (nós, caminhos e relações), considerados os componentes básicos para a estruturação de seus dados. Os nós simbolizam pontos no espaço. Os caminhos descrevem elementos lineares ou o limite de áreas (caminhos fechados). E as relações são usadas para representar associações lógicas ou geográficas entre os elementos, podendo ser nós, caminhos e/ou outras relações (OSM WIKI, 2023).

No OSM, a aplicação da propriedade semântica dos seus elementos ocorre através das etiquetas (*tags*). As etiquetas são compostas por pares de chaves e valores, informações que expressam o significado de cada feição adicionada ao mapa. A chave representa um tópico ou categoria. O valor detalha de forma mais específica a informação apresentada pela chave (OSM WIKI, 2023).

De forma comparativa, a chave tem função semântica análoga à categoria de informações e o valor se assemelha à classe de objetos, classificações propostas pela ET-EDGV 3.0 (MACHADO, 2020). As chaves e os valores são sempre representados na língua inglesa, idioma nativo da plataforma OSM. E de forma a uniformizar os dados e informações produzidos e distribuídos na plataforma, recomenda-se seguir o padrão de rotulagem proposto pelo *OpenStreetMap*, documentado e disponibilizado na página *Map Features*, que apresenta as combinações de chave e valor para as *tags* mais usadas para a descrição das feições que compõem o mapa (OSM WIKI, 2023). A Tabela 2 aponta as principais categorias (chaves) relacionadas na página *Map Features* e a quantidade de objetos (valores) disponíveis para a composição das *tags* de classificação.

O OSM permite incluir um número ilimitado de atributos a cada feição, e cada etiqueta descreve um atributo geográfico ou uma propriedade. Além das *tags* de classificação, é possível compor etiquetas que descrevam propriedades adicionais das feições, como nome, endereço, anotações, referências e restrições. Neste caso os valores são textos livres, respostas qualitativas (variável categorizada) e respostas quantitativas (variável numérica) (OSM WIKI, 2023).

Todas as informações referentes à plataforma OSM são documentadas de forma detalhada pelos seus usuários efetivos e disponibilizados na internet através de *Wikis*, páginas *web* editáveis. Seu repositório de documentação é chamado de

OpenStreetMap Wiki, com o objetivo de orientar os usuários da plataforma e suprir dúvidas técnicas dos contribuidores (OSM WIKI, 2023). Como bem exposto por Machado (2020), cada chave e cada etiqueta está associada a uma *Wiki*, através da qual é possível acessar outros links com conteúdos relacionados ao tema. Nas páginas das chaves são exibidas descrições específicas e a relação de todas as etiquetas que podem ser compostas. E as *Wikis* das etiquetas apresentam orientações de como mapear, *tags* opcionais (semelhantes) e sugestões de *tags* complementares (combinações úteis) para enriquecer a descrição da feição mapeada.

Tabela 2 - Quantidade de possíveis valores associados às principais chaves sugeridas para uso na plataforma OSM

Principais Feições	Chaves	Valores (quantidade)	Principais Feições	Chaves	Valores (quantidade)
Aerialway	aerialway	14	Leisure	leisure	33
Aeroway	aeroway	13	Man made	man_made	57
Amenity	amenity	126	Military	military	14
Barrier	barrier	39	Natural	natural	48
Boundary	boundary	16	Office	office	53
Building	building	87	Place	place	28
Craft	craft	85	Power	power	19
Emergency	emergency	17	Power	line	2
Emergency	lifeguard	1	Public transport	public_transport	5
Geological	geological	7	Railway	railway	31
Healthcare	healthcare	20	Railway	public_transport	3
Highway	highway	53	Railway	landuse	1
Highway	footway	2	Route	route	22
Highway	cycleway	8	Shop	shop	169
Highway	busway	1	Sport	sport	125
Highway	parking	9	Telecom	telecom	5
Highway	emergency	1	Tourism	tourism	22
Historic	historic	47	Water	water	15
Landuse	landuse	37	Waterway	waterway	19

Fonte: Adaptada de OSM WIKI (2023).

Cada elemento (nó, caminho ou relação) adicionado ao mapa possui um ID, numeração que o identifica exclusivamente. Além disso, cada conjunto de edições realizado por um mesmo usuário em um curto período de tempo também é salvo no banco de dados do OSM e associado a um ID único. Isto permite que todo o histórico

de edições de cada elemento inserido na plataforma fique registrado e disponível para consulta (MINGHINI; FRASSINELLI, 2019).

Ao se comparar a ET-EDGV 3.0 e a plataforma OSM, é possível identificar semelhanças nas estruturas conceituais e semânticas propostas por seus modelos, sendo este assunto amplamente explorado por Machado (2020). De fato, o conceito de armazenamento e documentação de informações do *OpenStreetMap* tem lógica similar ao praticado pelo mapeamento de referência, porém com nível de detalhamento superior ao proposto pela ET-EDGV 3.0 (MACHADO; CAMBOIM, 2019). Entretanto, questões como a liberdade para a edição e inserção de um número ilimitado de *tags* para a descrição de um mesmo elemento gera heterogeneidade semântica (ELIAS; FERNANDES, 2021; SILVA; CAMBOIM, 2021). Neste sentido, a própria comunidade de usuários do OSM desenvolve sistemas que possam auxiliar na análise dos dados produzidos pela plataforma como, por exemplo, a Taginfo, *software* que disponibiliza estatísticas sobre o uso e a combinação de *tags* (MINGHINI; FRASSINELLI, 2019).

2.2 MAPEAMENTO COLABORATIVO DE EVENTOS EFÊMEROS

Além da mobilidade urbana e da segurança pública, diversas outras áreas podem se beneficiar com a prática do mapeamento colaborativo, tais como saúde, economia, saneamento, cultura e educação (COSTA; FERNANDES; JUNIOR, 2015). E nas últimas décadas, o uso de VGI tem se apresentado bastante pertinente em diversas aplicações, como planejamento espacial, estudos ambientais e gerenciamento de crises e emergências (POSER; DRANSCH, 2010). Algumas plataformas colaborativas buscam mapear lugares ou elementos mais perenes, isto é, que possuem um longo ciclo de vida, como vias, edificações e pontos turísticos. Já outros sistemas VGI visam registrar eventos efêmeros, isto é, episódios cujo ciclo de vida é curto. São exemplos: sinistros de trânsito, crimes, crises, engarrafamentos, deslizamentos, epidemias e qualquer outro evento cujo mapeamento é de interesse da população ou de instituições (BORBA *et al.*, 2017; ROCHA *et al.*, 2013).

Para que um sistema VGI tenha êxito em seu propósito, é necessário a contribuição dos voluntários com a produção de dados espaciais que irão alimentar a plataforma. Logo, é fundamental compreender a motivação dos usuários em participar

do projeto (GÓMEZ-BARRÓN *et al.*, 2016). Nesse sentido, algumas iniciativas têm se dedicado ao apoio a causas sociais. Em seus estudos, Liu e Palen (2010) propuseram a produção de dados espaciais voluntários em episódios de desastres naturais. Recentemente Birkin, Vasileiou e Stagg (2021) investiram seus esforços no uso de VGI para o estudo epidemiológico e combate à COVID-19.

A popularização do uso de *smartphones* estimulou o desenvolvimento de aplicações e serviços que utilizam técnicas de mapeamento colaborativo para a navegação baseada em GPS, isto é, sistemas que utilizam o sensoriamento dos *smartphones* para mapear, em tempo real, alguma atividade em específico, bastando apenas o usuário trafegar conectado ao aplicativo (OVELGONNE; SONNENBICHLER; GEYER-SCHULZ, 2010). São exemplos o *Strava* e o *Waze*. O *Strava* é um aplicativo muito utilizado por ciclistas e corredores, no qual é possível monitorar os dados dos treinos dos usuários, como tempo, rota e velocidade, permitindo uma posterior avaliação de desempenho. Os dados dos usuários ficam hospedados na plataforma, o que permite gerar resultados estatísticos, como as rotas mais populares para cada tipo de atividade. Já o *Waze* utiliza o sensoriamento dos dispositivos móveis para mapear o trânsito de veículos, sendo bastante usado como ferramenta de navegação pelos condutores, com sugestões otimizadas de rotas. Além disso, o *Waze* permite que o usuário mapeie eventos circunstanciais que podem, de alguma maneira, interferir no fluxo de veículos, como buracos na via, engarrafamentos, blitz policial e sinistros de trânsito. Usualmente, os dados desses eventos ficam disponíveis durante curtos períodos, o que demanda atualizações constantes (COSTA; FERNANDES; JUNIOR, 2015).

Aplicações em VGI destinados a segurança pública e a apoio a desastres e emergências estão sendo desenvolvidas. Rocha *et al.* (2013) produziram o protótipo de um sistema que alerta qualquer cidadão portando um *smartphone* quanto a presença de problemas sociais urbanos na região em que transitam, como a violência e a carência de infraestrutura viária, relatados por outros cidadãos. Tal iniciativa teria um potencial para evitar acidentes ou ocorrências indesejadas. Furtado *et al.* (2010) e Shah *et al.* (2011) dedicaram seus estudos especificamente no desenvolvimento de ferramentas para o registro e pesquisa de ocorrências de crimes em centros urbanos.

Há também projetos desenvolvidos para o mapeamento colaborativo de fenômenos naturais de recorrência cotidiana ou esporádica. Hirata (2013) produziu

um protótipo para o mapeamento voluntário de pontos de alagamento na cidade de São Paulo, implementado através da plataforma *Crowdmap* do *Ushahidi*. Além de registrar as ocorrências, o usuário pode consultar os demais pontos de alagamento mapeados e receber alertas. Goodchild e Glennon (2010) destacaram em seu estudo as iniciativas de VGI utilizadas para o mapeamento de focos de incêndio na cidade de Santa Bárbara, Califórnia. Neste trabalho foi identificado que repositórios locais e *websites* específicos administrados por jornais locais ou grupos comunitários eram formas mais eficazes de disseminar as informações. Em ambos os casos citados (pontos de alagamento e focos de incêndio), o tempo de representação dos dados no sistema é não acumulativo.

2.2.1 Mapeamento Colaborativo Aplicado a Sinistros de Trânsito

Alguns sistemas de navegação por GPS como o *Google Maps*, o *Inrix* e o *Waze* disponibilizam a seus usuários uma interface para relatar de forma voluntária e em tempo real, eventos circunstanciais às condições de trânsito como, por exemplo, os sinistros de trânsito (YUNDUAN; RUIMIN, 2020). Nestas plataformas, o sinistro de trânsito é apresentado como uma das variáveis que afeta no fluxo de veículos naquele instante. Desta forma, o registro das ocorrências se faz valer como uma informação complementar ao usuário, uma vez que a proposta principal dessas aplicações é servir como ferramenta de navegação, sugerindo ao condutor rotas otimizadas em tempo real.

Apesar disso, em aplicativos como o *Waze* é possível minerar os dados históricos de sinistros de trânsito, podendo ser usados para estudos independentes voltados à segurança viária (COSTA; FERNANDES; JUNIOR, 2015). Recentemente, o próprio *Waze* disponibilizou em seu aplicativo de navegação uma nova funcionalidade, que cria alertas de histórico de acidentes para informar ao condutor trechos de vias propensas a ocorrência de sinistros de trânsito. Para tanto, o sistema faz uso de inteligência artificial para processar os dados de acidentes fornecidos pelos usuários do *Waze* durante o uso do aplicativo.

Pesquisas buscam mensurar a aplicabilidade de dados de sinistros de trânsito oriundos de mapeamento colaborativo. Flynn, Gilmore e Sudderth (2018) empenharam esforços no desenvolvimento um modelo para estimar o volume de

sinistros de trânsito relatados pela polícia, tendo como base os alertas de sinistros de trânsito do *Waze* registrados nas rodovias do estado de Maryland, Estados Unidos. Santos, Davis Jr. e Smarzaró (2017) investigaram o grau de equivalência entre os dados de sinistros de trânsito reportados por usuários do *Waze* e os dados registrados por agências oficiais no município de Belo Horizonte/MG. Turner *et al.* (2020) avaliaram o potencial do uso de dados de sinistros de trânsito do *Waze* para identificar locais de alto risco em uma rodovia do estado do Texas, Estados Unidos.

São escassas as aplicações voltadas ao registro de problemas sociais urbanos, tal como segurança pública, infraestrutura e trânsito (ROCHA *et al.*, 2013). E poucos estudos foram realizados para produzir uma plataforma exclusiva para o mapeamento colaborativo de sinistros de trânsito. Cinnamon e Schuurman (2010) desenvolveram um projeto piloto de mapeamento voluntário em um hospital na Cidade do Cabo, capital da África do Sul. Tal projeto de VGI buscou avaliar a viabilidade do uso de ferramentas gratuitas da *Geoweb* para o registro de lesões em países de baixa renda, em grande parte resultantes de violência interpessoal ou de sinistros de trânsito. Entretanto, conta somente com a participação dos profissionais de saúde do hospital, não estando aberto à contribuição da população em geral.

Já Medeiros (2022) desenvolveu uma plataforma digital colaborativa para o registro exclusivo de eventos relacionados à violência no trânsito em Manaus. Disponível nas versões *web* e *mobile*, o aplicativo é baseado no conceito de infovigilância, isto é, busca promover práticas de comunicação participativa. Através da plataforma, qualquer cidadão pode tanto realizar o registro georreferenciado do evento quanto gerar e exportar relatórios das ocorrências já registradas e validadas pelos administradores. Como estímulo ao uso da plataforma, os usuários são pontuados por cada registro válido de sinistro de trânsito. O projeto ainda se encontra em fase de testes.

De forma geral, a criação de sistemas VGI para o registro de eventos efêmeros se mostra um processo desafiador e muitas vezes complexo, dado as características diferenciadas de ordem, variação e granularidade. Além disso, por se tratar muitas vezes de um projeto dinâmico para o registro de eventos em tempo quase real, questões como as particularidades temporais do fenômeno mapeado e o tempo de representação do dado no sistema devem ser bem discutidas para a otimização da plataforma colaborativa (HIRATA, 2013).

2.3 AVALIAÇÃO DA QUALIDADE DE DADOS COLABORATIVOS

As informações referentes à qualidade dos dados geográficos são essenciais para a definição do critério de seleção de um grupo de dados, uma vez que o valor do dado relaciona-se diretamente com a sua qualidade (ISO 19157, 2013). O mapeamento colaborativo produz e disponibiliza uma grande quantidade de dados, atualizados de forma constante em suas plataformas. Entretanto sua qualidade é heterogênea e, por esse motivo, se torna necessário elaborar métodos para avaliar a qualidade de seus dados (MACHADO; CAMBOIM, 2019).

A qualidade da informação geográfica pode ser definida como o total de características de um produto que influenciam em sua capacidade de cumprir a sua finalidade (BROVELLI, ZAMBONI, 2018). Dentre os possíveis indicadores utilizados para avaliar a qualidade de dados na cartografia, dois parâmetros se destacam: a qualidade posicional e a semântica. A maioria dos estudos avaliam a acurácia posicional dos dados, e poucos são os estudos que tratam sobre análise da qualidade semântica e dos fatores humanos a ela relacionados (BRAVO; SLUTER, 2015).

Devido à grande disponibilidade de informações geográficas produzidas e distribuídas de forma voluntária, o *OpenStreetMap* se tornou objeto de inúmeras pesquisas científicas voltadas à análise de seus dados. Na maioria desses estudos, os dados colaborativos são comparados com a base de dados oficial. Nesse método, em que dois conjuntos de dados diferentes são comparados, ocorre a chamada avaliação extrínseca. Já na avaliação intrínseca, a análise é realizada com base nos próprios dados, isto é, apenas um conjunto de dados é o suficiente para aferir a qualidade dos dados geográficos, não sendo necessário a aplicação de métodos comparativos para conceder crédito aos dados colaborativos (BARRON; NEIS; ZIPF, 2013; DORN; TORNROS; ZIPF, 2015).

Pelo fato da cobertura do mapeamento de referência não ser homogêneo, métodos de avaliação intrínseca passaram a ser desenvolvidos, gerando novas alternativas para a análise de qualidade dos dados voluntários. Esse procedimento se tornou necessário, pois diversos estudos apontaram que, em avaliações extrínsecas, o grau de qualidade dos dados do OSM está diretamente ligado com a disponibilidade

dos dados oficiais. E na realidade há diversos locais na superfície terrestre que sequer dispõem de dados geográficos, e quando possuem são de baixa qualidade, não sendo adequados para comparação (MINGHINI; FRASSINELLI, 2019; PAIVA, 2021).

A percepção de qualidade dos usuários que produzem e consomem informações espaciais, o conhecimento dos princípios para descrever a qualidade de dados geográficos e a compreensão das normas nacionais e internacionais de qualidade são indispensáveis para a análise de um conjunto de dados geográficos (SANTOS, 2013). O fato é que as normas de avaliação de qualidade são, via de regra, elaboradas para a análise dos dados oriundos de mapeamentos oficiais, enquanto que os dados produzidos em plataformas de mapeamento colaborativo não se adequam completamente às métricas dessas normas, uma vez que são produzidos voluntariamente por usuários diversos com experiências variadas, nutridos por diferentes motivações (PAIVA, 2021). Entretanto, dentre as pesquisas que utilizam métodos extrínsecos, usualmente a qualidade dos dados colaborativos é mensurada com base em parâmetros de qualidade descritos nas normas técnicas convencionais, como a ISO 19157 (2013) e suas derivações (ELIAS *et al.*, 2020).

2.3.1 Parâmetros da Qualidade dos Dados Geográficos

Para aferir a qualidade dos elementos integrantes de um produto cartográfico, são aplicados indicadores determinados por normas da ISO (*International Organization for Standardization*) (ELIAS; FERNANDES, 2021).

A partir do ano de 2002 foram publicadas as primeiras versões da ISO tratando sobre o tema, a ISO 19113 (2002), a ISO 19114 (2003) e a ISO/TS 19138 (2006), que estabelecem, respectivamente, os princípios, os procedimentos e as medidas para a avaliação da qualidade de dados geográficos (ARIZA-LÓPEZ *et al.*, 2020). Posteriormente, as três normas sofreram revisão, e em 2013 foram mescladas em um único documento técnico, dando origem à ISO 19157 (2013) (ISO 19157, 2013).

A ISO 19157 (2013) fornece, de forma consistente e padronizada, os princípios para a descrição dos dados geográficos, as diretrizes para o tratamento das informações e os procedimentos para a avaliação quantitativa da qualidade desses dados, definindo as estruturas de medida de qualidade (ISO 19157, 2013).

Os parâmetros de qualidade estabelecidos pela ISO 19157 (2013) são a acurácia posicional, acurácia temática, completude, consistência lógica, qualidade temporal e usabilidade. A acurácia posicional está associada à qualidade posicional, enquanto que os demais são elementos indicadores da qualidade semântica (BRAVO; SLUTER, 2015).

Segue a descrição dos parâmetros de qualidade (ISO 19157, 2013):

- Acurácia posicional: precisão da posição das feições em um dado sistema de referência espacial. Classificações: acurácia absoluta (ou externa), acurácia relativa (ou interna) e acurácia em *grid* (malha regular);
- Acurácia temática: exatidão dos atributos das feições, das classes e suas relações. Classificações: acurácia de classificação, acurácia de atributo não quantitativo e acurácia de atributo quantitativo;
- Completude: presença ou ausência de feições, seus atributos e relações. Classificações: comissão e omissão;
- Consistência lógica: nível de concordância com as regras lógicas da estrutura dos dados (conceitual, lógica ou física), seus atributos e relações. Classificações: consistência conceitual, consistência de domínio, consistência de formato e consistência topológica;
- Qualidade temporal: qualidade dos atributos temporais e da relação temporal das feições. Sua medida pode ser pontual ou um intervalo de tempo. Classificações: precisão de uma medida de tempo, consistência temporal e validade temporal;
- Usabilidade: requisitos do usuário e nível de concordância entre as informações e as necessidades do usuário. Todos os elementos de qualidade podem ser usados para avaliar a usabilidade.

No Brasil, as normas sobre qualidade de dados geoespaciais têm como base a ISO 19157 (2013). Merecem destaque a Especificação Técnica para Controle de Qualidade de Dados Geoespaciais (ET-CQDG) (BRASIL, 2016), homologada pela Comissão Nacional de Cartografia (CONCAR), e o Manual Técnico de Avaliação da Qualidade de Dados Geoespaciais do IBGE (IBGE, 2019) (BRAVO; SLUTER, 2018; ELIAS; FERNANDES, 2021; MACHADO, 2020). E o padrão nacional de documentação de metadados geoespaciais é determinado pelo Perfil de Metadados Geoespaciais do Brasil (Perfil MGB 2.0) (IBGE, 2021), documento elaborado pelo

IBGE em parceria com a Diretoria de Serviço Geográfico (DSG) do Exército Brasileiro.

Parâmetros de qualidade especificados em normas ISO usualmente utilizam de métodos extrínsecos para a avaliação da qualidade de dados geográficos (BROVELLI; ZAMBONI, 2018). Pelo fato de não haver normas técnicas específicas para a análise de dados colaborativos, pesquisas sobre o VGI adotam as premissas estabelecidas em normas como a ISO 19157 (2013) para a avaliação extrínseca da qualidade de seus dados, comparando-os com a base de dados do mapeamento de referência (BRAVO; SLUTER, 2015; BROVELLI; ZAMBONI, 2018; MINGHINI; FRASSINELLI, 2019).

De acordo com Minghini e Frassinelli (2019), em pesquisas sobre os dados VGI produzidos e distribuídos na plataforma OSM, os parâmetros de qualidade mais investigados por métodos extrínsecos são a acurácia posicional, a completude e a acurácia temática. E as feições mais analisadas nestes estudos são as estradas, as edificações, o uso e ocupação do solo e os pontos de interesse. Silva e Camboim (2021) também destacam o aumento de estudos que avaliam a qualidade de dados colaborativos com base na acurácia posicional e completude.

Em sua pesquisa, Haklay (2010) avaliou a acurácia posicional, a completude e a consistência lógica dos dados VGI no Reino Unido. Girres e Touya (2010) estenderam o trabalho de Haklay para a França, porém com uma avaliação mais completa dos dados do *OpenStreetMap*, uma vez que foram considerados em seus estudos oito parâmetros de qualidade. Na Alemanha, Hecht, Kunze e Hahmann (2013) analisaram a completude das feições de edificações na plataforma OSM. Brovelli e Zamboni (2018), na Itália, também estudaram os elementos de edificações, mensurando a acurácia posicional e a completude de suas feições. No Iraque, Jasim e Al-Hamadani (2020) avaliaram a acurácia posicional de feições de estradas. E no Brasil, Elias e Fernandes (2021) aferiram a acurácia posicional, a completude e a acurácia temática de feições diversas no OSM.

Já os métodos intrínsecos são aplicados para inferir a qualidade de dados colaborativos nos locais em que há carência ou inexistência de conjunto de dados de referência para proceder com métricas comparativas (MINGHINI; FRASSINELLI, 2019; PAIVA, 2021). Algumas propriedades intrínsecas apontadas por pesquisadores para a avaliação da qualidade dos dados colaborativos são: quantidade de

atualizações, densidade de feições, quantidade de contribuidores, análise da origem dos dados, diferenças demográficas ou socioeconômicas locais e IEC (Inteligência Espacial Coletiva) (MACHADO; CAMBOIM, 2019; PAIVA, 2021).

Haklay *et al.* (2010) investigaram o número de colaboradores como indicador de qualidade para avaliação da acurácia posicional e da completude de dados do *OpenStreetMap*. Spielman (2014) pesquisou a influência da IEC na qualidade de dados colaborativos, isto é, o grau de conhecimento e experiência do grupo de usuários que contribuem com o mapeamento de uma área. Minghini e Frassinelli (2019) apresentam um aplicativo de código aberto para avaliar a qualidade de dados do OSM baseando-se na análise do histórico de edições de seus dados. Elias *et al.* (2020, 2021a) concentraram seus estudos em identificar possíveis padrões espaço-temporais relacionados à qualidade de dados colaborativos. E Paiva (2021) propõe uma metodologia para estimar a acurácia posicional e a completude de dados VGI com base em um conjunto de características intrínsecas e indicadores sociodemográficos.

2.4 COLETA DE DADOS DE SINISTROS DE TRÂNSITO

2.4.1 Sinistros de Trânsito no Brasil e no Mundo

Segundo a OPAS (Organização Pan-Americana da Saúde) da OMS (Organização Mundial de Saúde), mais de 3,5 mil pessoas no mundo perdem a vida diariamente no trânsito, isto é, cerca de 1,3 milhão de mortes por ano. Acidentes de trânsito se tornaram a principal causa de morte de crianças e adolescentes no mundo. Nestas condições, é previsto mais de 13 milhões de mortes e mais de 500 milhões de lesões por acidentes de trânsito nos próximos 10 anos, principalmente em países em desenvolvimento (OPAS/OMS, 2021).

Atualmente, o Brasil possui uma frota ativa de mais de 78 milhões de veículos (RENAEST, 2023), considerando frota ativa os veículos com o último licenciamento ou com infrações nos últimos 10 anos, constantes no sistema do RENAINF. Porém, sua frota total estimada é de mais de 115 milhões de veículos (RENAVAM, 2023). Isto é, a cada dois brasileiros, há um veículo compondo a frota do país.

De acordo com os dados apresentados pelo Mapa da Violência (levantamento baseado no SIM - Sistema de Informações de Mortalidade, gerido pelo Ministério da Saúde), no Brasil, houve um aumento de 38,3% no número de mortos por sinistro de trânsito no período de 2002 a 2012. Em 2018, o Relatório de Status Global sobre Segurança Rodoviária, publicado pela Organização Mundial da Saúde, apontou o Brasil como o terceiro país com o maior número de mortes registradas no trânsito, sendo superado somente pela China e a Índia (WHO, 2018). No ano de 2022, o país registrou 1.020.926 sinistros de trânsito. Mais 20.852 pessoas perderam a vida (RENAEST, 2023). Isto é, o Brasil aferiu uma média de, ao menos, 57 óbitos por dia decorrentes do trânsito.

Em 2010, a Assembleia Geral da ONU (Organização das Nações Unidas) publicou uma resolução buscando conscientizar os países a traçarem ações e adotarem medidas de segurança no trânsito. Com este documento, a ONU estabeleceu uma meta global de redução em 50% da mortalidade no trânsito até o fim da década, ficando definido o período de 2011 a 2020 como a "Década de Ação para a Segurança no Trânsito" (ONU, 2010).

O Brasil aderiu à campanha, com a participação de grande parte dos órgãos e entidades de trânsito federais, estaduais e municipais. Em uma nota técnica, Carvalho e Guedes (2023) realizaram um balanço desse período (2010 a 2019), comparando-o com a década anterior (2000 a 2009). Segundo os dados (extraídos do sistema SIM/DATASUS) houve um aumento de 13,5% no número de óbitos em termos absolutos, ultrapassando a marca de 390 mil mortes no período 2010 a 2019. Já a taxa de mortalidade, calculada por cada 100 mil habitantes (índice utilizado pela ONU), apontou uma redução significativa a partir de 2014. Conforme análise do Observatório Nacional de Segurança Viária (ONSV), o Brasil apresentou uma redução de quase 25% no número de mortes em sinistros de trânsito no período de 2011 a 2018. Apesar disso, a taxa de mortalidade experimentou uma leve variação positiva de 2,3% ao se comparar as duas décadas, frustrando a meta global de redução em 50% (CARVALHO; GUEDES, 2023).

Em 2018, o país aprovou a Lei nº 13614, que criou o PNATRANS (Plano Nacional de Redução de Mortes e Lesões no Trânsito) e acrescentou dispositivo ao CTB (Código de Trânsito Brasileiro) para incluir metas objetivas a fim de reduzir o índice de óbitos no trânsito (BRASIL, 2018a). Em 2021, a ONU proclamou o período

de 2021 a 2030 como a "Nova Década de Ação para a Segurança no Trânsito". Neste mesmo ano, o PNATRANS foi alinhado ao plano global, passando por uma revisão que contou com a contribuição de mais de 100 especialistas. A meta deste programa é a mesma que a estipulada para a década anterior, isto é, reduzir no mínimo à metade o índice de mortes por grupo de habitantes (BRASIL, 2021). O PNATRANS encontra-se em andamento.

2.4.2 Métodos de Coleta de Dados de Sinistros de Trânsito

De acordo com o Código de Trânsito Brasileiro, compete ao Denatran (Departamento Nacional de Trânsito), órgão máximo executivo de trânsito da União, estabelecer modelo padrão de coleta de informações sobre as ocorrências de sinistros de trânsito e as estatísticas de trânsito. E cabe aos órgãos ou entidades executivos rodoviários e de trânsito dos Estados e Municípios, no âmbito de sua circunscrição, coletar dados estatísticos e elaborar estudos sobre os sinistros de trânsito e as suas causas (BRASIL, 1997).

No Brasil, os sinistros de trânsito urbano ou rodoviário são registrados pelas unidades policiais em Boletins de Ocorrência (BO). E em alguns Municípios, a depender de sua estrutura, o registro de acidentes em áreas urbanas é realizado por agentes municipais de trânsito (ROMÃO; CAMPOS, 2011).

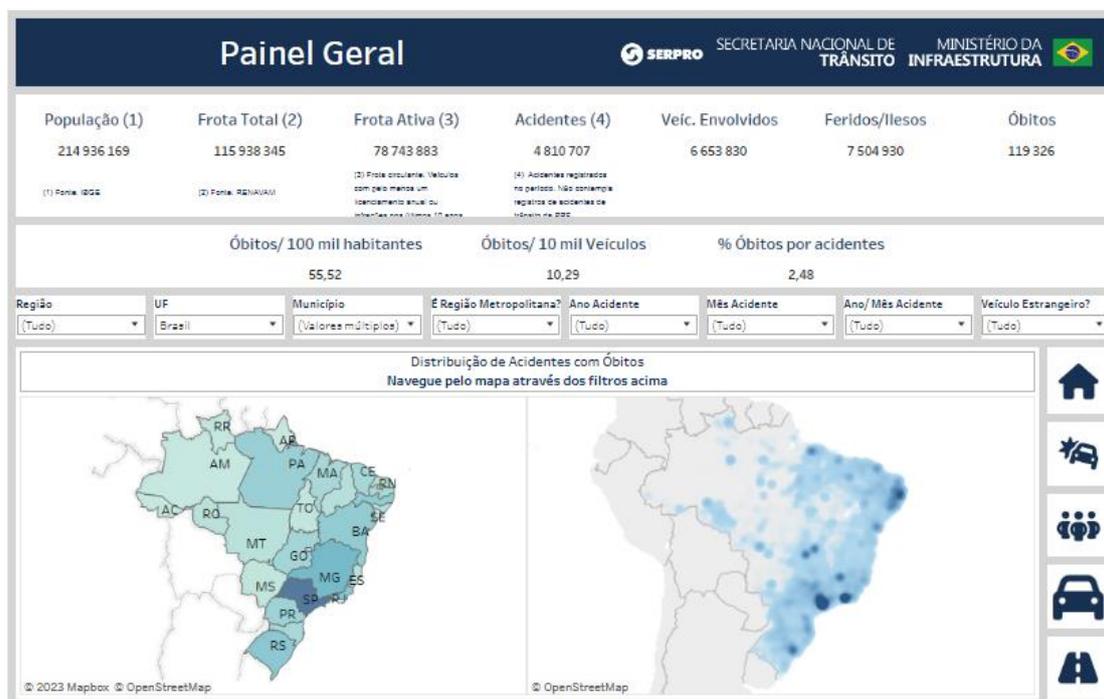
De forma geral, para o registro de ocorrência de sinistros de trânsito sem vítima, o interessado tem o direito de solicitar a prestação de atendimento em unidades policiais de forma presencial. Quando necessário, os envolvidos no acidente podem solicitar o comparecimento de viaturas no local da ocorrência para proceder com o registro do evento. Além desses métodos, de forma a facilitar e agilizar o processo, alguns Estados permitem que os próprios envolvidos procedam com o registro do sinistro de trânsito de forma virtual por meio do sítio eletrônico oficial do órgão do Estado, sendo necessária a posterior validação por parte da equipe da Delegacia Virtual do Estado em que o evento ocorreu (ESTADO DE MINAS GERAIS, 2023).

Além dos Boletins de Ocorrência, os dados referentes aos sinistros de trânsito podem ser anotados em outras fontes, como registros de atendimento pré-hospitalares realizados pelo Serviço de Atendimento Móvel de Urgência (SAMU), registros de

ocorrências atendidas pelo Corpo de Bombeiros e registros hospitalares das vítimas de acidente (TORRES; PEREIRA; MACHADO, 2020).

A Resolução CONTRAN n° 808 de 15 de dezembro de 2020 dispõe sobre o Registro Nacional de Acidentes e Estatísticas de Trânsito (RENAEST). Organizado e mantido pelo Denatran, o RENAEST é a primeira base nacional de registro, gestão e controle de dados de sinistros de trânsito. Seu sistema consolida dados coletados nas vias urbanas e nas rodovias estaduais e federais do país, registrados pelos órgãos que compõem o Sistema Nacional de Trânsito (SNT) (BRASIL, 2020). As estatísticas estão disponíveis a todos, podendo ser consultadas em um painel *online* (Figura 2), que exibe as informações de forma gráfica e interativa, subsidiando estudos e pesquisas com vista à melhoria da segurança viária do país.

Figura 2 - Painel RENAEST



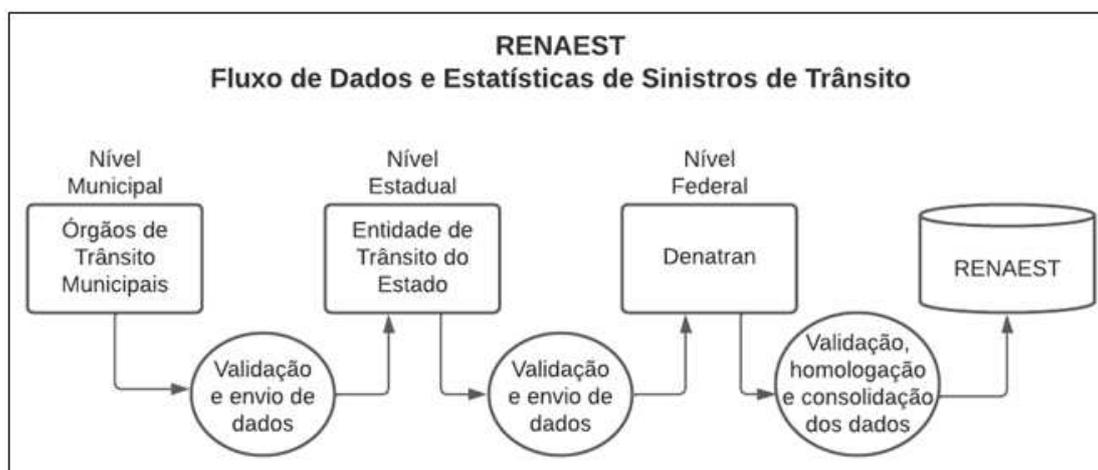
Fonte: Brasil (2023b)

Podem compor o SNT, qualquer órgão e entidade que efetue o registro de ocorrências de sinistros de trânsito, investigue suas circunstâncias ou preste atendimento às vítimas, entre os quais: Ministério da Saúde, secretarias de saúde dos Estados e Municípios, SAMU, Polícias Cíveis, Corpo de Bombeiros Militares, administradora do Seguro DPVAT (Danos Pessoais Causados por Veículos

Automotores de Via Terrestre), PRF (Polícia Rodoviária Federal), DNIT (Departamento Nacional de Infraestrutura de Transportes) e ANTT (Agência Nacional de Transportes Terrestres) (BRASIL, 2020).

Cabe aos órgãos de trânsito dos Municípios validar e enviar à entidade de trânsito do respectivo Estado os dados e estatísticas referentes a sinistros de trânsito coletados em nível municipal. Os Estados, por sua vez, validam as informações coletadas em nível estadual e as enviam ao Denatran, além dos dados de seus Municípios. Ao Denatran, cabe validar os dados coletados em nível federal e, de posse de todos os demais dados, o órgão deve homologar e consolidar todo o conteúdo na base nacional do RENAEST (BRASIL, 2020). A Figura 3 apresenta um diagrama que descreve o fluxo de dados e estatísticas de sinistros de trânsito, com destino ao RENAEST. O CONTRAN (Conselho Nacional de Trânsito) orienta que estas informações sejam atualizadas mensalmente no sistema e divulgadas no sítio eletrônico do órgão (BRASIL, 2020).

Figura 3 - Fluxo de dados e estatísticas de sinistros de trânsito - RENAEST



Fonte: Adaptada de Ministério da Infraestrutura (BRASIL, 2020).

O RENAEST incorpora outros três sistemas: o RENAVAM (Registro Nacional de Veículos Automotores), o RENACH (Registro Nacional de Carteira de Habilitação) e o RENAINF (Registro Nacional de Infrações de Trânsito) (BRASIL, 2020).

Além da consulta a dados referentes aos sinistros de trânsito no Brasil, o painel RENAEST (Figura 2) também disponibiliza outras informações, como a população e a frota de veículos. Além disso, seu sistema gera, de forma interativa,

gráficos com análises comparativas dos dados de acordo com a filtragem definida pelo usuário. O painel permite consultar a quantidade de acidentes agrupado por município, e os resultados são acumulados por mês. Os dados brutos de sinistros de trânsito do RENAEST, como coordenadas de cada sinistro de trânsito e a data (dia e hora) em que cada evento ocorreu, estão disponíveis para *download* em formato CSV, viabilizando ao usuário a realização de análises geoespaciais da informação geográfica (BRASIL, 2023a).

Não integram o RENAEST os dados de sinistros de trânsito registrados em rodovias federais, cuja coleta é atribuição da PRF (BRASIL, 2023a). Apesar disso, seus dados brutos também estão disponíveis para *download* em formato CSV (BRASIL, 2023c).

Dados pessoais sensíveis, protegidos pela LGPD (Lei Geral de Proteção de Dados Pessoais) não constam na relação de atributos que compõem os dados abertos de acidentes de trânsito do RENAEST e da PRF.

O painel RENAEST vem sendo desenvolvido desde o início dos anos 2000, sendo parcialmente concretizado em 2021. E, por ser um projeto recente, o programa ainda necessita de ajustes, o que torna esta ferramenta ainda limitada ao desenvolvimento de ações e políticas públicas (CARVALHO; GUEDES, 2023). Além disso, outro desafio para esse projeto consiste em especificar um padrão para o registro de sinistros de trânsito no Brasil, uma vez que os Boletins de Ocorrência não seguem um modelo estadual ou nacional para a coleta de dados de acidentes (TORRES; PEREIRA; MACHADO, 2020).

2.4.3 *Dados Estatísticos de Sinistros de Trânsito*

Como elucidado no tópico anterior, ainda não há especificação de um padrão nacional para o registro de sinistros de trânsito no Brasil. Isto é, até o momento cada Estado define o modelo de Boletim de Ocorrência que julgar mais adequado para a coleta de dados de acidentes. Portanto, ficou definido que as análises realizadas nesta pesquisa contemplam os documentos utilizados pelos órgãos no estado de Minas Gerais.

A seguir é apresentada uma análise por meio do cruzamento de dados estatísticos de sinistros de trânsito realizado a partir da listagem comparativa de todos

os atributos habilitados para preenchimento, relacionados a um sinistro de trânsito oficial.

Em sua primeira coluna, o Quadro 2 lista os principais dados estatísticos de sinistros de trânsito que compõem um RAT (Relatório de Acidente de Trânsito) fixado pela ABNT NBR 12898 (1993). E em suas colunas adjacentes são apresentados sistemas oficiais de registro de sinistros de trânsito (em cada um dos quais são apontados a relação de atributos disponíveis para coleta). São eles:

- i. Boletim de Ocorrência (BO) da Polícia Militar de Minas Gerais (PMMG), vinculado ao Sistema Integrado de Defesa Social (SIDS). Plataforma utilizada pelo policial para o registro de sinistros de trânsito com ou sem vítimas;
- ii. Delegacia Virtual da Polícia Civil do Estado de Minas Gerais (PCMG), vinculado ao SIDS. Plataforma utilizada pelo envolvido no sinistro para o registro *online* de sinistros de trânsito simples, sem vítimas e não relacionado a ilícito penal, dispensando a necessidade de se apresentar a uma unidade policial.

Vale ressaltar que o que está sendo avaliado no Quadro 2 é a existência, nos boletins/formulários, de campo específico para o registro, quando pertinente, de cada um dos dados estatísticos de sinistros de trânsito listados. Entretanto, nada impede que algum dado que não tenha local para seu cadastro seja registrado em campos que exigem uma resposta aberta (frase ou texto elaborado com as próprias palavras do respondente) como, por exemplo, campo de observações ou o campo para registro do histórico da ocorrência.

Toda informação que não seja possível ser colocada nos campos apropriados, mas que tenha importância para esclarecer as circunstâncias ou causas presumidas do acidente, deve ser detalhada no Relatório do Policial. No Boletim de Ocorrência da PMMG há campo específico para o registro da dinâmica do acidente com ou sem vítima, sendo realizado pelo agente de polícia. Em sinistros de trânsito sem vítima registrados via Delegacia Virtual também há campo específico para a anotação do histórico da ocorrência, porém as informações são apontadas pelo usuário (civil envolvido no acidente).

Quadro 2 - Dados coletados em diferentes sistemas oficiais de registro de sinistros de trânsito

Dados Estatísticos de Sinistros de Trânsito (RAT - NBR 12898:1993) ¹		BO PMMG ²	Delegacia Virtual PCMG ³
Itens	Atributos		
Dados gerais	Data e hora do acidente	X	X
	Tipo de acidente	X	X
	n° de veículos envolvidos	X	X
	n° de passageiros mortos, feridos ou ileso	X	X
	n° de Pedestres mortos ou feridos	X	
	Outras propriedades atingidas	X	
Condições do local	Características da via	X	
	Condições da pista e calçada	X	
	Condições da sinalização	X	
	Condições do tempo	X	
Detalhes do veículo	Dados do proprietário	X	X
	Dados do veículo	X	X
	Espécie de veículo	X	
	Categoria de veículo	X	
	Tipo de veículo	X	X
	Condições do veículo (geral, pneus, luzes) ⁴		
	Danos ao veículo decorrentes do acidente	X	X
	Manobra do veículo no momento do acidente	X	X
	Situação do veículo ⁵	X	
Condutor do veículo	Dados do condutor	X	X
	Gravidade das lesões	X	
	Condições adversas alegadas	X	X
	Infração constatada		
	Atendimento médico ⁶	X	
Passageiro vitimado	Dados do passageiro morto ou ferido	X	
	Gravidade das lesões	X	
	Atendimento médico ⁶	X	
Pedestre vitimado	Dados do pedestre morto ou ferido	X	
	Gravidade das lesões	X	
	Atendimento médico ⁶	X	
Localização do acidente		X	X
Croqui e fotografias do acidente		X	X
Testemunhas	Dados das testemunhas	X	X
Providências	Perícia técnica no local	X	
	Presença de outros órgãos no local	X	
Relatório policial		X	
Identificação do policial		X	
Notas:			
¹ Principais dados estatísticos de sinistros de trânsito que compõem um RAT (Relatório de Acidente de Trânsito) fixado pela NBR 12898:1993.			
² Boletim de Ocorrência da PMMG vinculado ao SIDS para o registro, pelo policial, de sinistros de trânsito com ou sem vítimas.			
³ Delegacia Virtual, sistema <i>online</i> da PCMG vinculado ao SIDS para o registro, pelo envolvido, de sinistros de trânsito sem vítimas.			
⁴ Condições aparentes do veículo antes do acidente.			
⁵ Destino (status final) do veículo após o registro da ocorrência.			
⁶ Local para onde e por quem a vítima foi removida para receber atendimento médico.			

Para sinistros de trânsito sem vítima, o formulário de registro da Delegacia Virtual não contabiliza pedestres, uma vez que o pedestre só é qualificado como envolvido em um acidente na condição de vítima, isto é, quando morto ou ferido.

Apesar de o registro de sinistro de trânsito na Delegacia Virtual ser realizado por um dos envolvidos no acidente, a ocorrência necessitará passar por uma triagem para sua validação, realizada por uma equipe de policiais e analistas após o envio dos dados ao sistema (ESTADO DE MINAS GERAIS, 2023).

Em Minas Gerais, os dados registrados tanto no Boletim de Ocorrência da PMMG quanto na Delegacia Virtual da PCMG estão associados ao SIDS, sistema que permite a gestão integrada das informações de defesa social produzidas pelos órgãos de segurança pública do Estado, como a PMMG, a PCMG e o CBMMG (MINAS GERAIS, 2004). Em 2022 o SIDS passou ser denominado Sistema Integrado de Segurança Pública (Sisp) (MINAS GERAIS, 2022).

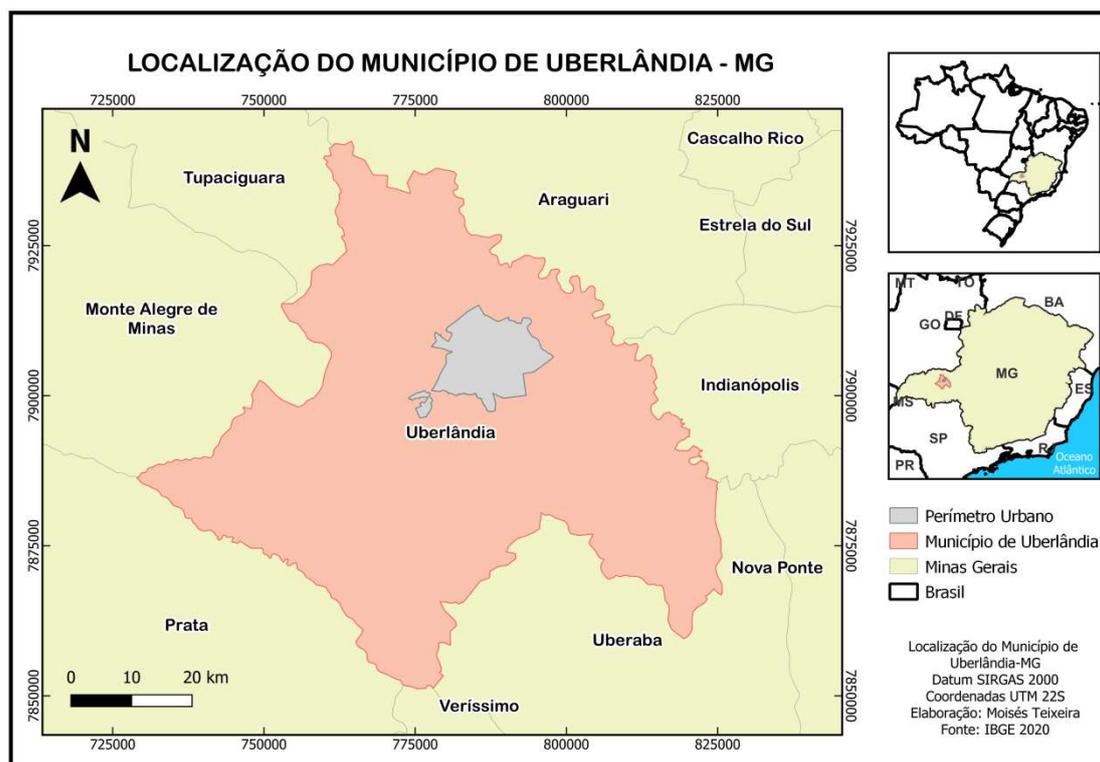
Por fim, vale destacar que, mesmo entre órgãos de um mesmo Estado, o registro de ocorrências relacionadas a sinistros de trânsito é realizado de formas diferentes. Cada entidade possui competências e responsabilidades distintas, sem contar as diferentes finalidades a que os registros se destinam, o que torna o processo de padronização da coleta de informações algo ainda mais desafiador (MEDEIROS, 2022).

3 MATERIAIS E MÉTODOS

3.1 ÁREA DE ESTUDO

Localizado na mesorregião do Triângulo Mineiro e Alto Paranaíba, o município de Uberlândia (Figura 4) é considerado um pólo logístico e um grande centro de distribuição e comércio atacadista. Sua infraestrutura viária merece destaque, uma vez que algumas das mais importantes rodovias federais (BR-50 e BR-365) e estaduais (MGC-452, MG-455 e MGC-497) da região central do Brasil cruzam o perímetro urbano do Município. Sua malha urbana é composta, de forma geral, por traçados ortogonais, mesclado de praças (SILVA; GARREFA, 2020).

Figura 4 - Mapa de localização do município de Uberlândia



Fonte: Elaboração própria

Uberlândia faz parte do corredor centro-sudeste da Ferrovia Centro-Atlântica (FCA), a maior ferrovia do país e uma das principais rotas de escoamento das exportações do agronegócio brasileiro, com destino ao Porto de Santos (SP). Além

disso, o Município mineiro dispõe de uma Estação Aduaneira do Interior (EADI), o Porto Seco do Cerrado, favorecendo a produtividade da região.

O município de Uberlândia é o mais populoso da região do Triângulo Mineiro, sendo a segunda cidade mais populosa do estado de Minas Gerais e um dos maiores do interior do Brasil. Em um período de 30 anos (1991 a 2022), a população de Uberlândia dobrou (IBGE, 2022), ultrapassando a marca de 713 mil habitantes. Este fato é reflexo do potencial econômico da cidade que, com sua infraestrutura urbana sólida e outros fomentos, atrai cada vez mais investimentos, propiciando ainda mais seu crescimento e ampliando o fluxo de mercadoria e informações por suas vias.

Conforme dados do Ministério da Infraestrutura (BRASIL, 2022), Minas Gerais é o estado que detém a segunda maior frota de veículos do país. Nesse contexto, Uberlândia, a segunda maior cidade do estado, merece destaque. A frota estimada de veículos em Uberlândia no ano de 2022 era de mais de 505.000 veículos (IBGE, 2022), isto é, uma média de 2 veículos para cada 3 habitantes. Essa proporção é significativa, maior do que a média brasileira de 1 veículo para cada 2 habitantes. Em 2022, o Município de Uberlândia registrou 13.705 sinistros de trânsito, resultando em 40 mortes (RENAEST, 2023). Ferreira e Ribeiro (2006) informavam quanto aos elevados índices de acidentes que o Município experimentava no início dos anos 2000: cerca de 41 acidentes por dia em 2004 e pouco mais de 48 em 2005. Segundo os autores, o aumento da frota de veículos e o crescimento da população têm influência direta no aumento de acidentes de trânsito.

3.2 MATERIAIS UTILIZADOS

Para a produção, tratamento e análise de dados geoespaciais, foi utilizado o programa QGIS (versão 3.24.1). *Software* de SIG gratuito e de código aberto, o QGIS dispõe de diversas ferramentas de geoprocessamento que permitem a visualização, edição e análise de dados georreferenciados.

A coleta e hospedagem de dados foi realizada através do aplicativo *web* Epicollect5, versão 5.1.52. Esta plataforma gratuita, disponível para os sistemas Android e iOS em sua versão móvel, permite a criação de formulários (questionários) para pesquisa de campo. Os dados são processados e armazenados na internet,

podendo ser visualizados em um servidor central (mapas, gráficos e tabelas). Além disso, os dados podem ser exportados nos formatos CSV e JSON.

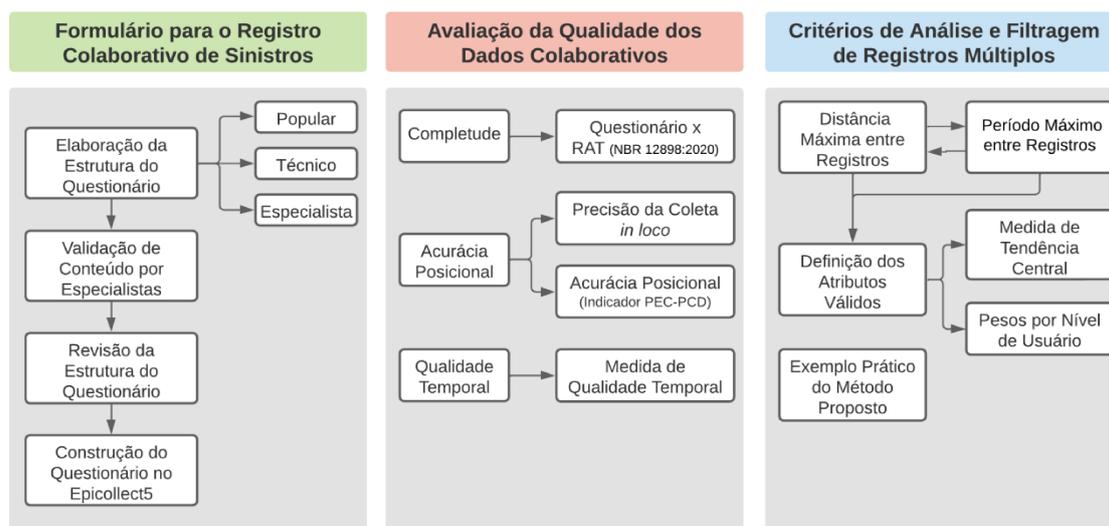
O dispositivo móvel utilizado para coletar os dados de campo via Epicollect5 foi o *smartphone* da marca Motorola, modelo Moto G60, sistema operacional Android (versão 12), dotado de receptor GPS (*Global Positioning System*), sistema de navegação via satélite.

Para a aplicação das análises estatísticas, foi utilizado o programa STATISTICA (*StatSoft*), versão 7.1. Este *software* contém uma ampla variedade de ferramentas para análise, gestão e visualização de bases de dados. Abriga milhares de algoritmos, funções, testes e métodos estatísticos, permitindo a mineração de dados de forma mais eficiente.

3.3 MÉTODO

A metodologia proposta na pesquisa contempla três grandes etapas: construção do questionário para o registro colaborativo de sinistros de trânsito, avaliação da qualidade dos dados colaborativos e especificação dos critérios de análise e filtragem de registros múltiplos. O fluxograma da Figura 5 sistematiza tais etapas.

Figura 5 - Fluxograma com as etapas da metodologia proposta



3.3.1 Construção do Questionário no Epicollect5 para o Registro Voluntário

Fundamentado nas normas ABNT NBR 10697 (2020) e ABNT NBR 12898 (1993), e com base na configuração dos Boletins de Ocorrência utilizado pelas unidades policiais de Minas Gerais (Polícia Militar e Polícia Civil) para o registro de sinistros de trânsito, foi elaborada a estrutura do questionário sugerido para uso no registro colaborativo de sinistros no município de Uberlândia. Para a seleção dos atributos válidos do questionário, foi avaliado campo a campo da estrutura de um RAT (Relatório de Acidente de Trânsito), fixado pela ABNT NBR 12898 (1993). O critério de seleção dos atributos válidos seguiu a premissa de que o formulário proposto visa acumular a maior quantidade de informações possível, limitada aos aspectos objetivos básicos, que viabilize uma experiência prática e intuitiva de coleta colaborativa de dados de sinistro de trânsito. Foram considerados níveis distintos de usuários, sendo desenvolvido um tipo de formulário para cada natureza de colaborador. A base de todos os formulários é a mesma, diferindo-se quanto ao grau de descrição dos atributos. Após a estruturação dos dados, um modelo de formulário foi construído no aplicativo Epicollect5 (disponível nas versões *web* e *móvel*). Concluído a construção do formulário no aplicativo, foram realizados testes visando avaliar a usabilidade da plataforma para a coleta e hospedagem de dados colaborativos.

Para atribuir confiabilidade e credibilidade ao estudo, foi realizado o processo de validação do formulário elaborado para o registro voluntário de sinistros de trânsito. A validação de um instrumento de pesquisa consiste em submeter seu roteiro a uma avaliação criteriosa da qualidade, seguindo etapas estratégicas de verificação e autocorreção, buscando assim reduzir a subjetividade da qualidade da pesquisa. Portanto, nesta etapa do trabalho, cada item que compõe a estrutura do formulário foi submetido a uma análise de juízes (análise de constructo), isto é, a avaliação por um grupo de especialistas na área. Para a análise dos resultados retornados por oito especialistas, foi aplicada a metodologia proposta por Hernández-Nieto (2002), com o cálculo do Coeficiente de Validade de Conteúdo (CVC), que mede o nível de concordância entre os juízes.

Ao final dessa etapa, com base nos resultados, comentários e sugestões dos especialistas, a estrutura do formulário foi alterada em alguns pontos e inserida no

aplicativo Epicollect5 para a realização de testes de campo. O produto desse processo foi o formulário validado para o pré-teste com uma amostra do público alvo e posterior aplicação em pesquisa de campo (registros colaborativos de sinistros de trânsito).

3.3.2 *Proposta de Garantia da Qualidade dos Dados Colaborativos*

Alinhado aos conceitos apresentados na ISO 19157 (2013), a qualidade dos dados deve ser especificada por meio de indicadores (parâmetros) de qualidade. Para esta pesquisa, os três parâmetros selecionados para a avaliação da qualidade dos dados colaborativos foram: a completude de atributos, a acurácia posicional e a qualidade temporal.

a) Completude de Atributos

A completude indica o excesso (comissão) ou a ausência (omissão) de feições, seus atributos e relações. Geralmente, a medição da completude de dados colaborativos ocorre pela sua comparação com um conjunto homólogo de dados de referência, calculando-se o total de elementos existentes em uma área pré-definida.

O presente trabalho concentra esforços em um estudo de viabilidade para o mapeamento colaborativo de sinistros de trânsito, com um apontamento das etapas preliminares necessárias à aplicação da proposta e simulações de registro de acidentes para testes. Isto é, não integra o escopo deste projeto a produção, armazenamento, e tratamento de dados colaborativos de sinistros de trânsito reais. Este fato impede a avaliação extrínseca do registro em si para o indicador de completude, o que permitiria, por exemplo, inferir uma possível sub-notificação dos registros junto aos órgãos competentes (suposição sustentada até o momento por evidências empíricas).

Entretanto, nesta etapa da pesquisa, foi avaliada a completude de atributos da estrutura dos formulários elaborada para a coleta de dados colaborativos de sinistros de trânsito, considerando níveis distintos de usuários. Para tanto, o formulário proposto foi comparado com a estrutura de um RAT (Relatório de Acidente de Trânsito) fixado pela ABNT NBR 12898 (1993), e com alguns sistemas oficiais de registro de sinistros de trânsito do estado de Minas Gerais, a saber, Boletim de

Ocorrência (BO) da PMMG e Delegacia Virtual da PCMG. Ao RAT foi atribuído 100% de completude, tornando-se a referência para as análises comparativas.

b) Acurácia Posicional - Análise do Sensor GPS

A primeira etapa necessária à avaliação da qualidade posicional consistiu na aquisição dos dados. Para gerar os pontos amostrais de dados colaborativos de sinistro de trânsito, foram simulados registros de acidentes nas dependências do Campus Santa Mônica da Universidade Federal de Uberlândia (UFU), localizado no Bairro Santa Mônica, Zona Leste do município de Uberlândia. A coleta de dados das ocorrências fictícias foi realizada por receptor GPS de um dispositivo móvel, sendo o registro das coordenadas efetuado com uso do aplicativo Epicollect5. Para inferir a acurácia posicional dos dados coletados, suas coordenadas foram comparadas com as coordenadas dos pontos homólogos extraídos da base de referência (Mapa georreferenciado do Campus da Universidade).

No Brasil, a acurácia posicional de produtos cartográficos é regulamentada pelo Decreto-lei nº 89817 (BRASIL, 1984). Esta norma é complementada por especificações técnicas homologadas pela Comissão Nacional de Cartografia (CONCAR) e publicadas pela Diretoria de Serviço Geográfico (DSG) do Exército Brasileiro, merecendo destaque a ET-ADGV (BRASIL, 2018b) e a ET-CQDG (BRASIL, 2016).

O Decreto-lei nº 89817 determina as diretrizes a serem respeitadas para a elaboração de normas técnicas da cartografia nacional. Em seu texto, o decreto classifica o produto cartográfico em três classes (A, B e C) para a avaliação de sua acurácia posicional. Esta classificação tem como base as tolerâncias PEC (Padrão de Exatidão Cartográfica) e EP (Erro-Padrão) para uma escala predefinida (BRASIL, 1984). De forma a também abranger produtos cartográficos digitais, a ET-ADGV (Especificação Técnica para Aquisição de Dados Geoespaciais Vetoriais) introduziu um novo indicador estatístico, o PEC-PCD (Padrão de Exatidão Cartográfica dos Produtos Cartográficos Digitais). Estruturado de forma mais rigorosa, o PEC-PCD integra uma nova classe (classe D) para a avaliação da acurácia posicional dos dados (BRASIL, 2018b). A Tabela 3 apresenta, para cada classe dos indicadores estatísticos

PEC e PEC-PCD, os valores das tolerâncias PEC e EP para produtos cartográficos digitais em escalas que variam de 1:1000 a 1:10000.

Tabela 3 - Padrão de exatidão cartográfica da planimetria dos produtos cartográficos digitais em grandes escalas

Classes PEC	Classes PEC- PCD	Escala 1:1000		Escala 1:2000		Escala 1:5000		Escala 1:10000	
		PEC (m)	EP (m)	PEC (m)	EP (m)	PEC (m)	EP (m)	PEC (m)	EP (m)
-	A	0,28	0,17	0,56	0,34	1,40	0,85	2,80	1,70
A	B	0,50	0,30	1,00	0,60	2,50	1,50	5,00	3,00
B	C	0,80	0,50	1,60	1,00	4,00	2,50	8,00	5,00
C	D	1,00	0,60	2,00	1,20	5,00	3,00	10,00	6,00

Fonte: ET-CQDG (BRASIL, 2016).

A ET-CQDG (Especificação Técnica para Controle de Qualidade de Dados Geoespaciais) foi desenvolvida em conformidade com a ISO 19157 (2013), e apresenta os passos para a avaliação da qualidade de produtos de conjuntos de dados geoespaciais usando procedimentos padronizados, o que permite comparar os conjuntos de dados analisados usando uma estrutura comum (BRASIL, 2016).

A ET-CQDG apresenta a descrição das medidas relativas ao cálculo da acurácia posicional absoluta para produtos cartográficos digitais na planimetria. Seguem os passos (BRASIL, 2016):

Calcular a componente horizontal dos erros para cada par de ponto "i" homólogo da amostra, isto é, a distância euclidiana (e_{Hi}) (equação 1) entre os elementos de cada par de feições pontuais.

$$e_{Hi} = \sqrt{e_{xi}^2 + e_{yi}^2} \quad (1)$$

Neste caso, para o presente estudo, e_x e e_y são, respectivamente, as diferenças entre as coordenadas X e Y dos pontos amostrais do dado colaborativo e das suas coordenadas homólogas na base de referência.

Em seguida, é necessário calcular o Erro Médio Quadrático (EMQ) dos erros da amostra (equação 2).

$$EMQ_H = \sqrt{\frac{\sum_{i=1}^n (e_{Hi}^2)}{n}} \quad (2)$$

O Decreto-lei 89817 (BRASIL, 1984) e a ET-ADGV (BRASIL, 2018b) apontam duas condições que devem ser atendidas para a determinação da acurácia posicional (escala e classe) das feições pontuais. São elas:

- i. No mínimo, 90% dos pontos coletados devem apresentar erros menores ou iguais ao valor do indicador PEC-PDC;
- ii. O EMQ dos erros da amostra deve ser menor ou igual à tolerância EP.

Dessa forma, o próximo passo consiste em comparar cada valor de e_{Hi} com o erro máximo admissível (tolerância PEC) da Tabela 3 para a escala e classe consideradas. O produto se enquadra quando ao menos 90% dos pontos apresentar erro inferior à tolerância PEC. Se nenhum valor for correspondente, o resultado será considerado "não conforme", e encerra-se a medida.

Por fim, o último passo resulta em comparar o valor de EMQ_H com o Erro-Padrão (tolerância EP) da Tabela 3 para a escala e classe apontadas no passo anterior. Caso o valor seja menor, o resultado será a classe e escala em análise. Se for maior, caminha-se na tabela para o próximo par de valores PEC/EP até que as duas condições sejam atendidas. Perdurando a inconsistência, o resultado será considerado "não conforme".

c) Qualidade Temporal

A qualidade temporal corresponde à exatidão encontrada no componente temporal do dado geoespacial. Isto é, indica a qualidade dos atributos temporais e da relação temporal das feições.

Para o mapeamento colaborativo de sinistros de trânsito, a qualidade temporal dos dados está diretamente ligada à maneira que os registros são realizados: a depender do *software* a ser utilizado para a coleta e hospedagem de dados, os registros podem ser realizados *in loco* (via aplicativo móvel e com uso de receptor GNSS) ou *posteriori* (via aplicativo *web*). Usualmente, a versão móvel permite somente o registro da localização atual do GPS, ao passo que na versão *web* a geolocalização é editável. Normalmente, em ambos os casos, o campo para preenchimento da data e

horário do acidente é editável, isto é, os atributos temporais poderiam ser produzidos em tempo diferente do envio. Entretanto, a restrição relacionada ao serviço de geolocalização via aplicativo móvel incentiva o registro do evento apenas no local e em tempo real, o que pode favorecer tanto a acurácia posicional quanto a qualidade temporal dos dados.

Em uma proposta de mapeamento colaborativo de sinistros de trânsito, espera-se que um mesmo sinistro seja registrado mais de uma vez por diferentes usuários. Sendo assim, nesta etapa do trabalho foi especificado o método para a análise extrínseca do componente temporal (horário da comunicação da ocorrência) da informação geográfica voluntária, sendo a mediana a medida de tendência central definida para representar um conjunto de dados na hipótese de registros múltiplos de um mesmo acidente de trânsito.

3.3.3 Critérios de Análise e Filtragem de Registros

Nesta etapa da pesquisa, foram apresentados os critérios de análise e filtragem de registros múltiplos de sinistro de trânsito para aferir se pertencem a um mesmo evento, e os parâmetros para a definição dos atributos válidos.

Localização, data e horário do sinistro de trânsito: A posição e o tempo são atributos intimamente ligados. Em uma dinâmica de mapeamento colaborativo, para identificar quais registros de acidentes se referem a um mesmo evento, é necessário avaliar, simultaneamente, a localização geográfica e o horário. Neste sentido, foi determinada uma distância em metros (raio) entre a posição (latitude e longitude) dos registros. De forma análoga, foi especificado um intervalo de tempo máximos, dentro dos quais diferentes registros são associados a um mesmo evento. Não cumprindo ambas as premissas, os registros serão considerados oriundos de eventos distintos.

Níveis de usuários: a cada usuário cadastrado na plataforma, foi atribuído um nível associado ao seu grau de conhecimento. Vale ressaltar que, quanto maior o nível de experiência do usuário, mais assertivo serão os dados por eles coletados e, conseqüentemente, maior a qualidade da informação geográfica. Neste sentido, identificado quais registros se referem a um mesmo evento, o nível do usuário terá um peso maior para o critério de definição dos atributos válidos.

4 RESULTADOS E DISCUSSÕES

4.1 FORMULÁRIO PARA O REGISTRO COLABORATIVO DE SINISTROS

4.1.1 *Elaboração do Formulário*

Com base nas normas ABNT NBR 10697 (2020) e ABNT NBR 12898 (1993), e levando em consideração a configuração do formulário utilizado atualmente pelas unidades policiais de Minas Gerais (PMMG e PCMG) para o registro de ocorrências, foi elaborada a estrutura do formulário sugerido para uso por cidadãos comuns no registro colaborativo de sinistros de trânsito no município de Uberlândia.

A Figura 6 apresenta o resultado final referente à proposta de estruturação dos formulários para a coleta de dados colaborativos de sinistros de trânsito, a ser utilizado por cidadãos comuns, considerando níveis distintos de usuários.

O Formulário Especialista compreende os profissionais que compõem o quadro de funcionários dos órgãos públicos que atuam nos registros oficiais de sinistros de trânsito em Uberlândia. São exemplos, a Polícia Militar de Minas Gerais (PMMG) e o Corpo de Bombeiros Militar de Minas Gerais (CBMMG).

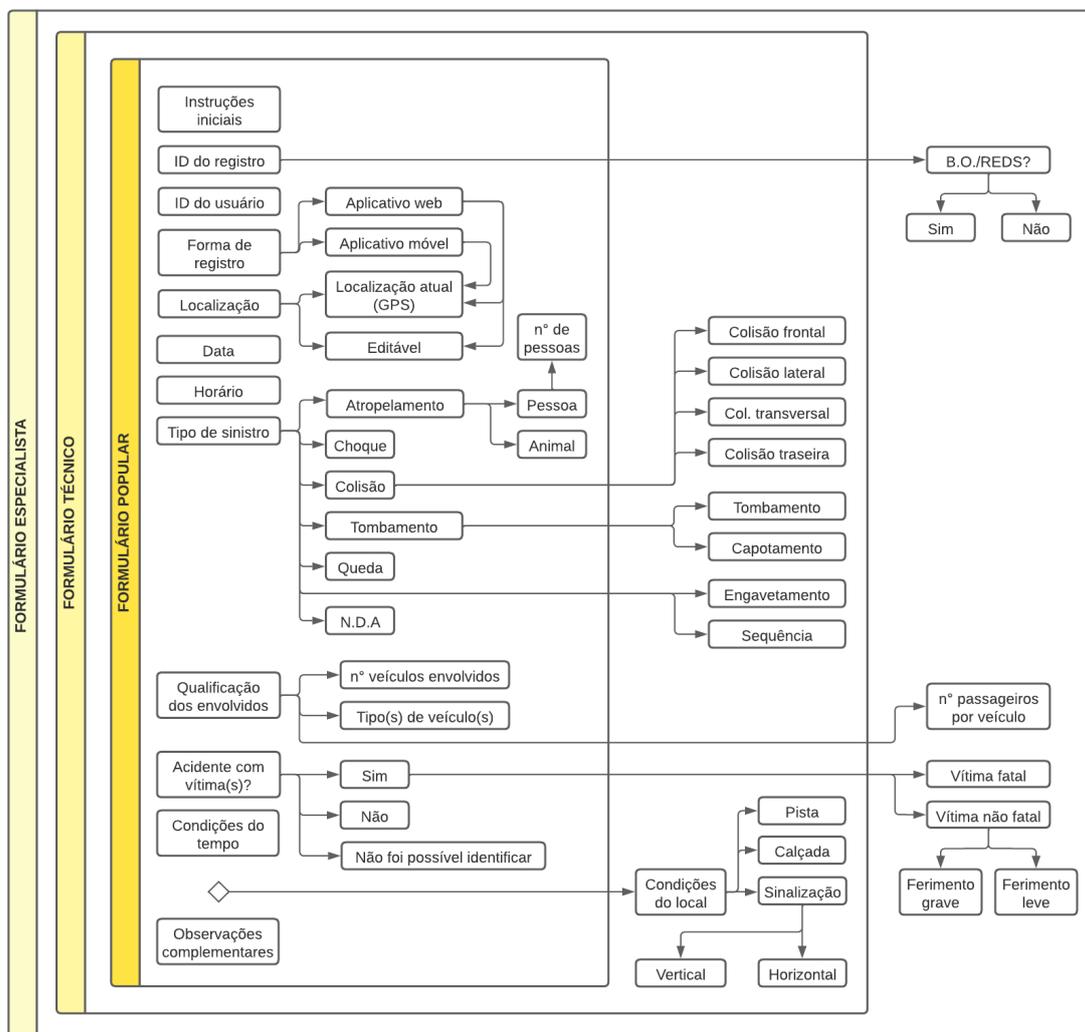
Já o Formulário Técnico foi elaborado para uso dos colaboradores que possuem certo conhecimento técnico sobre sinistros de trânsito. No caso de Uberlândia, integram essa categoria os profissionais que compõem os quadros técnicos (ex.: arquitetos, engenheiros civis e geógrafos) e operacionais (ex.: agentes de trânsito) da Secretaria Municipal de Trânsito e Transportes.

O Formulário Popular é composto pelos atributos básicos mínimos julgados necessários a compor a base de dados estatísticos de um sinistro de trânsito. Neste sentido, o questionário foi confeccionado para a utilização por parte de qualquer usuário que se interesse em colaborar, de forma voluntária, com o registro de dados de acidentes no município de Uberlândia, independente de seu grau de conhecimento técnico sobre o assunto.

Com a consolidação dos formulários, um modelo de questionário foi construído na versão *web* do aplicativo Epicollect5. Segue, nos próximos parágrafos, a explicação de cada um dos elementos que compõem a estrutura dos formulários

apresentados na Figura 6, levando em consideração as funcionalidades do sistema do Epicollect5.

Figura 6 - Estrutura dos formulários para a coleta de dados colaborativos de sinistros de trânsito



Fonte: Elaboração própria

Instruções iniciais: Nesta etapa, são apresentadas duas orientações preliminares importantes ao usuário. A primeira solicita que o usuário ative o serviço de geolocalização (GPS) de seu dispositivo móvel antes de iniciar o registro. A segunda o instrui a se aproximar o máximo possível do local do acidente antes de coletar a localização do sinistro de trânsito, caso o registro esteja sendo realizado via aplicativo móvel. Este procedimento visa garantir a precisão posicional do dado. Se esta ação não for possível, o usuário é orientado a não realizar o registro via aplicativo

móvel, e sim via aplicativo web, no qual é possível editar a geolocalização da informação coletada.

ID do registro: Após realizar o *upload*, os dados coletados ficam armazenados em um servidor central no sistema *web* do aplicativo Epicollect5, podendo ser visualizados, editados e retirados pelo criador do formulário. E ao serem exportados (formatos CSV e JSON) o aplicativo gera automaticamente um ID único para cada registro, formado por uma sequência alfanumérica. Para o caso do Formulário Especialista, cuja utilização limita-se a profissionais com acesso ao sistema de registros oficiais de sinistros de trânsito, estimula-se que o mesmo agregue uma informação adicional a este tópico: se foi gerado REDS (ou BO) para o acidente em questão. Esse dado extra poderá auxiliar em análises estatísticas quanto ao grau de completude dos dados colaborativos.

ID do usuário: Para acessar o sistema do Epicollect5, tanto para criar um novo projeto quanto para contribuir com a coleta de novos dados, é necessário que o usuário se registre por meio de seu e-mail. Sem esse cadastro, o site só permite acessar projetos definidos como públicos e listados, e o usuário poderá somente visualizar os dados. No *login*, o Epicollect5 requer o e-mail cadastrado do usuário e o código numérico (válido por tempo determinado) gerado pelo programa e enviado ao e-mail no momento da solicitação de *login*. Para cada projeto, o criador do formulário define se o acesso será público ou privado. Quando privado, o acesso ao formulário e habilitação para coleta de dados se torna restrito aos usuários (e-mails) predefinidos pelo criador. E quando os dados são exportados, o aplicativo relaciona cada usuário (e-mail) responsável pelos dados coletados.

Forma de registro: O Epicollect5 permite duas formas de registro, via aplicativo *web* e via aplicativo móvel. Ao efetuar o registro via aplicativo *web*, os dados são automaticamente armazenados no sistema *web* do programa (servidor central). Já quando o aplicativo móvel é utilizado para a coleta dos dados, as informações ficam, provisoriamente, armazenadas na memória do app. Neste caso, em algum momento, deverá ser realizado o *upload* do conjunto de dados para o servidor central (sistema *web*) para que os dados fiquem disponíveis, de forma definitiva, na interface do programa para visualização, geoprocessamento e exportação. Entende-se que, ao informar em que via o registro foi realizado (*web* ou móvel), os usuários contribuem com informação que pode favorecer estudos estatísticos referentes a

acurácia posicional e qualidade temporal do conjunto de dados colaborativos, uma vez que o registro da geolocalização do evento é limitado em um dos casos: o aplicativo móvel possibilita somente o registro da localização atual via GPS (Figura 7), enquanto que o aplicativo *web* permite a edição da geolocalização (Figura 8). Isto é, o aplicativo móvel incentiva o registro direto e em tempo real ao evento, enquanto que o aplicativo *web* permite a conclusão do registro posteriormente.

Figura 7 - Interface gráfica do serviço de geolocalização da versão móvel do aplicativo Epicollect5

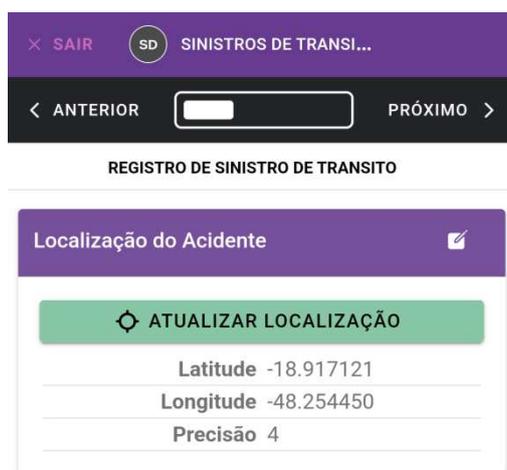


Figura 8 - Interface gráfica do serviço de geolocalização da versão *web* do aplicativo Epicollect5



Localização: Neste momento (Figuras 7 e 8), deve ser realizado a coleta da geolocalização do sinistro de trânsito. *A priori*, o Epicollect realiza o registro em coordenadas geodésicas (latitude e longitude). E ao exportar os dados, o programa

também gera e disponibiliza, automaticamente, as coordenadas UTM (N e E) de cada registro.

Data: Informar a data em que o sinistro de trânsito ocorreu.

Horário: Informar o horário em que o sinistro de trânsito ocorreu.

Tipo de sinistro: Selecionar qual o tipo de sinistro de trânsito de acordo com a norma ABNT NBR 10697 (2020). No Formulário Popular, as opções são apresentadas de forma mais simples e genérica, favorecendo uma melhor interpretação por parte do usuário que irá realizar a coleta dos dados. A partir do Formulário Técnico, eleva-se a descrição do tipo de acidente e são relacionadas as demais opções apontadas na Norma que dependem de uma maior expertise técnica do usuário para a especificação correta do dado.

Qualificação dos envolvidos: Informar o número de veículos envolvidos no sinistro de trânsito e especificar o(s) tipo(s) de veículo(s) de acordo com a norma ABNT NBR 10697 (2020). Além disso, no Formulário Especialista, espera-se que o usuário também informe o número de passageiros por veículo, uma vez que, usualmente, são profissionais que procedem, no exercício de sua profissão, com entrevistas aos envolvidos no acidente.

Acidente com vítimas: Caso seja possível identificar, deve-se informar se o sinistro de trânsito resultou em vítimas. Se sim, no Formulário Especialista também é solicitado especificar a gravidade da lesão segundo os parâmetros apresentados pela ABNT NBR 12898 (1993), uma vez que os usuários com acesso a esse formulário são profissionais com conhecimento técnico para esse tipo de análise e com habilitação para o contato direto com as vítimas do acidente.

Condições do tempo: Indicar o tempo atmosférico do local no momento da ocorrência do sinistro do trânsito, de acordo com a norma ABNT NBR 12898 (1993).

Condições do local: Informar as condições da pista, da calçada e das sinalizações (vertical e horizontal) nas proximidades do local onde o sinistro de trânsito foi registrado, de acordo com a norma ABNT NBR 12898 (1993). A especificação desse dado deverá ser realizada a partir do Formulário Técnico, pois entende-se que é necessário que o usuário tenha conhecimento técnico para avaliar dados dessa natureza.

Observações complementares: O objetivo do mapeamento voluntário é proporcionar ao usuário uma experiência de registro de sinistro de trânsito prática,

ágil e intuitiva. Desta forma, somente se julgar necessário, o usuário deverá descrever, de forma simples e objetiva, informações complementares aos dados do acidente.

É importante destacar que dados pessoais sensíveis da população são protegidos pela Lei Geral de Proteção de Dados Pessoais (LGPD). Por esse motivo, os dados dos envolvidos no acidente (condutores, passageiros, pedestres e testemunhas) não devem ser coletados na proposta de registro colaborativo de sinistros de trânsito, assim como a gravação de mídia (fotos e vídeos) do evento.

A Figura 9 apresenta a interface da versão móvel do aplicativo Epicollect5 com os resultados de um registro de sinistro de trânsito fictício. Neste exemplo, a estruturação considerada refere-se ao Formulário Popular.

Figura 9 - Resultados de um sinistro de trânsito fictício registrado na versão móvel do aplicativo Epicollect5

The figure displays two side-by-side screenshots of the Epicollect5 mobile application interface for recording a traffic accident. Both screens feature a purple header with a hamburger menu icon, a circular 'SD' icon, and the text 'SINISTROS DE TRANSITO_UBERLA...'. Below the header is a dark navigation bar with a left arrow and 'VOLTAR' (Return) and a trash icon with 'EXCLUIR' (Exclude).

Left Screenshot (Warning and Information):

- ATENÇÃO:** Antes de iniciar o registro, ative o serviço de localização (GPS) de seu dispositivo!
- IMPORTANTE:** Caso o registro esteja sendo realizado pelo aplicativo móvel: Antes de coletar a localização do sinistro de trânsito, se aproxime o máximo possível do local do acidente para garantir a precisão posicional do dado. Se isso não for possível, o registro não deverá ser realizado via aplicativo móvel, e sim via aplicativo web (no qual é possível editar a geolocalização da informação coletada).
- Forma de Registro:** via aplicativo móvel
- ID Ponto:** 1
- Localização do Acidente:**
 - LATITUDE: -18.917095
 - LONGITUDE: -48.254577
 - PRECISÃO: 4
- Data do Acidente:** 20/08/2023

Right Screenshot (Form Fields):

- Horário do Acidente:** 20:28:57
- Tipo de Sinistro de Trânsito:** colisão
- n° de Veículos Envolvidos:** 2
- Envolvido 1:** automóvel
- Envolvido 2:** motocicleta
- Acidente com Vítima(s):** sim
- Condições do Tempo:** nublado
- Alguma Observação Complementar?:** (Empty field)

O questionário é composto por perguntas fechadas. Isto é, para cada atributo que compõe o formulário são oferecidas opções de resposta previamente definidas. O usuário pode selecionar somente uma opção de resposta, e o formulário só permite ir à próxima questão após a definição do valor do atributo da questão em aberto. Dessa forma, cada sinistro de trânsito, registrado de forma voluntária na plataforma, fica associado a um conjunto completo de valores dos atributos estabelecidos pelo usuário responsável pela coleta.

4.1.2 Validação de Conteúdo

Nesta etapa do trabalho, o formulário elaborado para o registro voluntário de sinistros de trânsito (Figura 6) foi submetido a um processo de validação de conteúdo de instrumento de pesquisa. Para isso, cada item que compõe a estrutura do formulário passou por uma análise de especialistas (teste de juízes) para o julgamento de sua representatividade.

Para a composição do comitê de juízes foram convidados profissionais e acadêmicos (com titulação de especialistas, mestres e/ou doutores) com experiência em estudos relacionados a sinistros de trânsito. Oito juízes aceitaram o convite e participaram do processo de validação de conteúdo, atendendo à meta de contribuição de ao menos cinco especialistas para imprimir maior rigor ao método de análise (COLUCI; ALEXANDRE; MILANI, 2015). O Quadro 3 apresenta a descrição dos especialistas.

Quadro 3 - Descrição dos especialistas participantes do processo de validação de conteúdo

Juiz	Formação Acadêmica e Especialização	Titulação	Instituição
1	Geógrafa - Especialista em Planejamento Urbano	Doutora	USP
2	Geógrafa - Especialista em Mobilidade Urbana	Doutora	UFU
3	Engenheiro Cartógrafo - Especialista em Ciências Geodésicas	Doutor	UFPR
4	Engenheiro Civil - Especialista em Transportes	Doutor	UnB
5	Engenheiro Civil - Especialista em Transportes	Doutor	UnB
6	Geógrafo - Especialista em Mobilidade Urbana	Especialização	UFU
7	Economista - Especialista em Trânsito e Transporte	Especialização	CEFET-MG
8	Tecnólogo em Trânsito e Transportes Terrestres - Especialista em Mobilidade Urbana	Especialização	Uniube

Aos oito juízes foi enviado e-mail com o arquivo do texto base contendo as informações essenciais à compreensão da proposta de pesquisa (formulário para

registro colaborativo de dados de sinistros de trânsito), sendo orientado que, após a leitura do material, fosse preenchido um questionário *online*, elaborado no Google Forms, para a avaliação do instrumento de pesquisa.

O método de medição utilizado no questionário foi a escala numérica tipo *Likert* graduada em cinco pontos (LIKERT, 1932), sendo que o número um indica concordância mínima e o número cinco aponta concordância máxima. E, para mensurar a representatividade de cada item que compõe o formulário em análise, foram definidos três critérios para avaliação, a saber, a clareza da linguagem (grau de compreensão do conteúdo apresentado), a pertinência prática (grau de relevância ao objetivo da pesquisa) e a aderência à estrutura de formulários proposta (grau de adequação dos atributos distribuídos pela estrutura dos formulários Popular, Técnico e Especialista). Além disso, ao final da avaliação de cada item, foi disponibilizado campo para a livre escrita de comentários, críticas e sugestões, de forma a contribuir para a identificação de lacunas na pesquisa. A configuração do questionário é apresentada no Quadro 4.

Quadro 4 - Configuração do questionário elaborado para a avaliação dos itens do instrumento de pesquisa

Item do Formulário	Critérios para Avaliação															Comentários, Críticas e Sugestões
	Clareza da Linguagem					Pertinência Prática					Aderência à Estrutura					
	O item está claro?					O item é relevante ao objetivo do formulário?					A distribuição do item pelos formulários (Popular, Técnico e Especialista) está adequada?					
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	
Instruções Iniciais																
ID do Registro																
ID do Usuário																
Forma de Registro																
Localização																
Data																
Horário																
Tipo de Sinistro																
Qualificação dos Envolvidos																
Acidente com Vítima(s)?																
Condições do Tempo																
Condições do Local																
Observações Complementares																

Fonte: Elaboração própria

O Apêndice A apresenta as respostas retornadas pelos juízes quanto à avaliação dos itens que compõem o instrumento de pesquisa. Buscando também avaliar a estrutura do formulário como um todo, foi elaborado um questionário complementar (Quadro 5), apresentado aos especialistas para preenchimento após a conclusão do questionário de avaliação dos itens que o compõem.

Quadro 5 - Configuração do questionário elaborado para a avaliação geral do instrumento de pesquisa

Formulário para o Registro Colaborativo de Dados de Sinistros de Trânsito	O Formulário cumpre o objetivo proposto?					A quantidade de itens do Formulário (13 itens) está adequada?					Comentários, Críticas e Sugestões
	1	2	3	4	5	1	2	3	4	5	

Fonte: Elaboração própria

O Apêndice B revela os resultados retornados pelos especialistas referentes à avaliação geral do instrumento de pesquisa.

Para a análise dos resultados retornados pelos especialistas, foi calculado o Coeficiente de Validade de Conteúdo (CVC) proposto por Hernández-Nieto (2002). O CVC é um coeficiente que mede o nível de concordância entre os juízes, permitindo o cálculo de validade do conteúdo para cada item que compõe o instrumento (CVC_i) e a validade do instrumento de pesquisa como um todo (CVC_t). Seguem os passos para o cálculo (HERNÁNDEZ-NIETO, 2002):

Calcular a média das notas dos juízes para cada item (M_x), dado pela divisão entre a soma das notas dos juízes e o número total de juízes (equação 3).

$$M_x = \frac{\sum_{i=1}^J x_i}{J} \quad (3)$$

Em seguida é calculado o CVC para cada item (CVC_i) dividindo-se o resultado obtido na equação 3 pela nota máxima permitida na escala numérica de medição (equação 4).

$$CVC_i = \frac{M_x}{v_{máx}} \quad (4)$$

Para o cálculo do CVC final de cada item do instrumento (CVC_c) (equação 5), é necessário descontar o erro atribuído ao item (Pe_i) (equação 6), dado pelo inverso do número de total de juízes elevado pelo mesmo número de juízes. De acordo com Hernández-Nieto (2002), o desconto do erro é necessário para reduzir o efeito de possíveis tendências na avaliação dos especialistas.

$$CVC_c = CVC_i - Pe_i \quad (5)$$

$$Pe_i = \left(\frac{1}{J}\right)^J \quad (6)$$

Por fim, a validade do conteúdo do instrumento como um todo (CVC_t), para cada um dos critérios de avaliação definidos na pesquisa, é dado pela média do CVC de cada item menos a média dos respectivos erros atribuídos aos itens (equação 7).

$$CVC_t = M_{CVC_i} - M_{Pe_i} \quad (7)$$

A Tabela 4 apresenta, para cada um dos critérios de avaliação definidos na pesquisa, os valores do CVC final de cada item do formulário (CVC_c) e do CVC do instrumento como um todo (CVC_t), calculados com base nos resultados retornados pelos especialistas.

Tabela 4 - Valores do CVC referentes aos itens do instrumento de pesquisa

Item do Formulário	Critérios para Avaliação			Houve Comentários, Críticas ou Sugestões? Quantos(as)?	
	Clareza da Linguagem	Pertinência Prática	Aderência à Estrutura		
Instruções Iniciais	CVCc =	0,90	0,95	0,75	Sim (5)
ID do Registro	CVCc =	0,92	0,97	0,77	Sim (3)
ID do Usuário	CVCc =	0,95	1,00	0,87	Sim (2)
Forma de Registro	CVCc =	0,92	0,92	0,92	Sim (1)
Localização	CVCc =	0,95	0,92	0,90	Sim (1)
Data	CVCc =	0,95	1,00	0,90	Sim (2)
Horário	CVCc =	1,00	0,97	0,92	Sim (1)
Tipo de Sinistro	CVCc =	0,80	1,00	0,75	Sim (7)
Qualificação dos Envolvidos	CVCc =	0,90	0,92	0,77	Sim (4)
Acidente com Vítima(s)?	CVCc =	1,00	1,00	0,80	Sim (4)
Condições do Tempo	CVCc =	0,90	0,95	0,90	Sim (6)
Condições do Local	CVCc =	0,95	1,00	0,80	Sim (5)
Observações Complementares	CVCc =	0,90	0,87	0,90	Sim (2)
	CVCt =	0,93	0,96	0,84	

De forma semelhante, a Tabela 5 exhibe os valores do CVC oriundos das notas dos juizes apontadas no questionário elaborado para a avaliação geral do instrumento de pesquisa.

Tabela 5 - Valores do CVC referentes ao instrumento de pesquisa como um todo

Formulário para o Registro Colaborativo de Dados de Sinistros de Trânsito	O Formulário cumpre o objetivo proposto?	A quantidade de itens do Formulário (13 itens) está adequada?	Houve Comentários, Críticas ou Sugestões? Quantos(as)?
	CVCt =	0,86	0,86

Em consonância com outras pesquisas (DAVIS, 1992; GRANT; DAVIS, 1997; PASQUALI, 1998; POLIT; BECK, 2006; RUBIO *et al.*, 2003), Hernández-Nieto (2002) sugere uma concordância mínima de 0,80 (80%) entre os juizes para validar o conteúdo dos itens e do instrumento de pesquisa. Desta forma, itens com valores de CVC menores que o especificado sofrerão reformulação ou serão eliminados do instrumento, a depender do critério de avaliação. Para esta pesquisa, seguindo a metodologia proposta, itens com clareza da linguagem ou aderência à estrutura inferiores a 0,80 deverão ser reformulados, enquanto que itens com pertinência prática menor que 0,80 serão considerados não representativos ao instrumento de pesquisa, devendo ser excluídos.

Ao examinar os resultados expostos na Tabela 4, todos os itens atingiram CVC maior ou iguais a 80% quanto à clareza da linguagem e à pertinência prática. Para o critério de aderência à estrutura, apesar de um CVC_t de 84%, quatro itens obtiveram CVC_c abaixo da concordância mínima estipulada. Para cada um desses, seguem as considerações e decisão do presente autor (ante as observações apontadas pelos especialistas).

Item "Instruções Iniciais": CVC_c = 75%. As cinco observações apontadas pelos especialistas expõem pareceres gerais sobre o formulário, e não uma análise do item indicado. De qualquer forma, dado o valor do CVC_c, todas as cinco observações dos juizes foram apreciadas. Os impactos dessas considerações na estrutura do formulário são apresentados ainda neste tópico, alguns parágrafos abaixo, logo após a análise dos itens com CVC_c abaixo de 80%.

Item "ID do registro": CVC_c = 77%. As três observações registradas neste item expõem considerações do instrumento como um todo, e não sobre o item em questão.

Para atender ao método de análise de resultados selecionado para esta pesquisa, as três observações foram consideradas. E, semelhante ao item anterior, a implicação dessa análise é exposta no decorrer do tópico, logo após a avaliação dos itens com baixo valor do CVC_c .

Item "Tipo de sinistro": $CVC_c = 75\%$. Das sete observações apontadas pelos especialistas, somente uma expressa consideração quanto a aderência à estrutura. Nesta, o juiz sugeriu que o registro do tipo de sinistro não seja disponibilizado no Formulário Popular, mas somente nos formulários Técnico e Especialista. Por julgar que este item é um dado estatístico essencial à natureza da informação e que é esperado que o usuário popular seja maioria dentre o voluntariado, decidiu-se pela manutenção da distribuição previamente apresentada. Cinco observações fazem alusão à clareza da linguagem ($CVC_c = 80\%$) e merecem destaque, uma vez que expõem a visão quase que unânime dos juízes de que é necessário elevar a descrição de cada tipo de sinistro (seja por definições, pictogramas ou imagens) de forma a orientar principalmente o usuário comum na classificação do sinistro de trânsito por ele testemunhado. Entretanto, o aplicativo *web* selecionado nesta pesquisa para os testes de campo (Epicollect5) não possui essa funcionalidade. Neste sentido, compreendendo a importância da informação apontada pelos juízes, fica recomendada a inserção desses dados complementares em uma possível pesquisa futura que, estimulada pelo presente trabalho, que avalia a viabilidade da proposta, buscará efetivamente desenvolver um sistema para a coleta colaborativa, hospedagem e geoprocessamento de dados de sinistros de trânsito. Para os testes em campo da presente pesquisa, apesar das limitações do aplicativo Epicollect5, foi adicionado ao formulário (em tópico à parte) as definições de alguns tipos de sinistros de trânsito para consulta, em caso de dúvidas do usuário.

Item "Qualificação dos envolvidos": $CVC_c = 77\%$. Do total de quatro observações registradas pelos juízes, três exprimem considerações quanto a aderência à estrutura, sendo uma referente à distribuição e duas relativas ao nível de discriminação dos atributos. Quanto à distribuição, o comentário apontado pelo juiz sugere que não seja coletado dado algum (referente à qualificação dos envolvidos) no Formulário Popular. Por julgar que os atributos propostos para registro no Formulário Popular (nº de veículos envolvidos e tipos de veículos) são de simples identificação e que é esperado que o usuário popular seja maioria dentre o voluntariado, decidiu-se

pela manutenção da distribuição previamente apresentada. Já as duas observações referentes à discriminação dos atributos sugerem a inserção de mais dados dos envolvidos para elevar o detalhamento do evento e subsidiar estudos voltados à análise comportamental. Como o formulário proposto visa transmitir ao voluntariado uma experiência sucinta, prática e intuitiva, os novos atributos inseridos em sua versão atualizada seguem a premissa de serem dados básicos (protegidos pela LGPD) e de captação presumida, a saber, sexo e faixa etária aproximada do envolvido (condutor ou pedestre), dispostos para coleta nos formulários Popular, Técnico e Especialista. Por fim, um dos juízes apresentou uma observação apontando uma possível inadequação dos atributos ao item. Tal composição se justifica ao pressuposto de que sempre um veículo envolvido em um acidente está associado a um condutor. Dessa forma, buscando otimizar o fluxograma dos dados estatísticos de sinistros de trânsito sugeridos para coleta, os envolvidos no acidente (quando condutores) são identificados no formulário pelo veículo que dirigiam no momento da ocorrência. Por esse motivo, os atributos "nº de veículos envolvidos" e "tipo(s) de veículo(s)" estão incorporados ao item "qualificação dos envolvidos".

Buscando perfazer uma versão consistente do instrumento de pesquisa, foram apreciados todos os comentários, críticas e sugestões dos especialistas, independente dos valores de CVC calculados, reservado o entendimento de que o pesquisador deve ter autonomia nas decisões, desde que apoiadas pela teoria (TORLIG *et al.*, 2022). Seguem algumas das observações dos juízes consideradas pertinentes ao instrumento em estudo, seguidas dos comentários do presente autor.

Alguns comentários sugerem a exibição de todos os valores selecionáveis dos atributos do formulário. Segue resposta: há situações em que o número valores selecionáveis é elevado (exemplo, "tipo de veículo"). Nestes casos, para fins práticos, foi apresentada na estrutura revisada do formulário a quantidade de valores disponíveis para seleção, antecedido do termo "Diversos". Além disso, as fontes de referência de cada grupo de valores selecionáveis dos atributos também passou a ser exibida na própria estrutura do formulário.

Três comentários dos especialistas apontaram a necessidade de ser criado um item específico para o registro da causa presumida do acidente, perante o consenso de este ser um dado estatístico de grande valia a estudos voltados ao planejamento e gestão de segurança viária. A proposta inicial do formulário seria acumular no item

"observações complementares" qualquer informação adicional que o usuário julgasse relevante à ocorrência, inclusive indicações quanto às prováveis causas do sinistro de trânsito. Entretanto, alinhado à opinião dos especialistas, foi incluído na versão atualizada do formulário um item específico para a coleta dessa informação, nomeado "Causa presumida", disponibilizado em todos os formulários (Popular, Técnico e Especialista). Os valores selecionáveis serão os dispostos no sistema da Delegacia Virtual da PCMG (Quadro 6). Por permitir ao cidadão comum a opção de registro (devido a sua esperada representatividade no montante de usuários), foi incluída a opção de seleção "não foi possível identificar" como resposta ao item. Essa decisão se deve ao fato de que são esperados casos em que o usuário não terá a oportunidade de criar diálogo com os envolvidos ou as testemunhas do acidente para presumir a causa do sinistro.

Quadro 6 - Relação de causas presumidas para a ocorrência de um sinistro de trânsito

Motivos Aparentes para o Ocorrência de um Sinistro de Trânsito			
1	Agressões/ brigas em decorrência de manobras	15	Dirigir sob efeito de medicamentos
2	Animal na pista	16	Dormir ao volante
3	Aquaplanagem	17	Falta de atenção
4	Avanço de sinal	18	Linha com cerol (papagaio/pipa)
5	Carga mal acondicionada	19	Má visibilidade
6	Conduta inadequada do pedestre	20	Não dar Preferência
7	Contramão de direção	21	Não manter distância de segurança
8	Defeito na sinalização	22	Outros - causas relacionadas a trânsito
9	Defeito na via	23	Retorno/ conversão em local proibido
10	Defeito no veículo	24	Ultrapassagem forçada
11	Derrapagem	25	Ultrapassagem proibida
12	Desobedecer parada obrigatória	26	Uso de celular
13	Dirigir embriagado/alcoolizado	27	Velocidade incompatível
14	Dirigir sob efeito de droga ilícita		

Fonte: Delegacia Virtual da PCMG (ESTADO DE MINAS GERAIS, 2023)

Dois especialistas apresentaram comentários orientando a possibilidade de inserção de registros fotográficos durante o processo de coleta de dados. Apesar de o Epicollect5 disponibilizar essa funcionalidade, na proposta apresentada e nos testes de campo essa opção não foi aberta ao usuário devido a questões de ética e em atendimento a LGPD.

Em dois comentários, os especialistas sugeriram a padronização do uso de um único termo (sinistro ou acidente) no formulário. Tendo em vista que a revisão da norma ABNT NBR 10697 (2020) determinou a substituição da expressão "acidentes de trânsito" para "sinistros de trânsito" em pesquisas acadêmicas e relatórios técnicos, optou-se em padronizar o uso do termo "sinistros" na versão atualizada da estrutura do formulário (instrumento de pesquisa) apresentada neste trabalho. Já para o formulário replicado no aplicativo Epicollect5 para os testes de campo, optou-se pelo uso do termo "acidentes" uma vez que o processo de substituição da expressão ainda é recente, e o uso no termo "sinistro" poderia gerar dúvidas ao usuário comum.

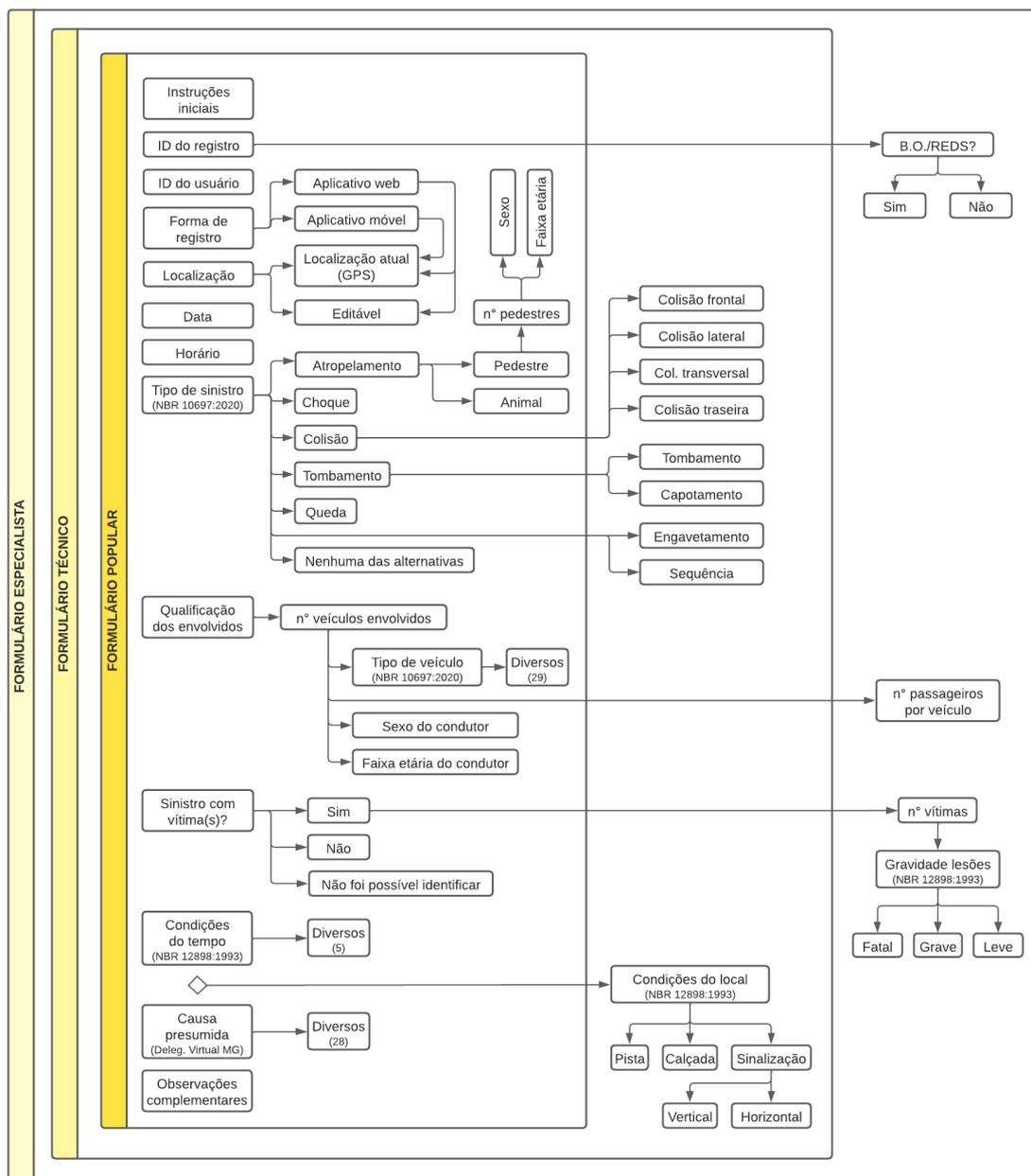
Como resultado, segue a Figura 10, que apresenta a versão atualizada da proposta de estruturação dos formulários para a coleta de dados colaborativos de sinistros de trânsito, a ser utilizado por cidadãos comuns, estratificado por tipo de usuário.

Por fim, a estrutura do questionário construída no aplicativo Epicollect5 sofreu ajustes (conforme a versão atualizada do formulário) para a realização de testes de campo. Foi definido o uso do Formulário Popular para os testes, com uso da versão móvel do aplicativo Epicollect5. Logo, como o público alvo deste formulário representa o nível de usuário comum, foi identificado a necessidade de adequar a linguagem do questionário a ser disponibilizado ao voluntariado para elevar o grau de compreensão do que cada item solicita.

A Figura 11 apresenta a interface da versão móvel do aplicativo Epicollect5 com os resultados de um registro de sinistro de trânsito fictício realizado pelo autor após os ajustes de linguagem do questionário, referente ao Formulário Popular.

Com a conclusão da etapa de análise de especialistas e os ajustes de estrutura, a versão atualizada do formulário e os atributos que o compõem atingiram um grau satisfatório, validado para o pré-teste com uma amostra do público alvo e posterior aplicação em pesquisa de campo (registros colaborativos de sinistros de trânsito).

Figura 10 - Estrutura revisada dos formulários para a coleta de dados colaborativos de sinistros de trânsito



Fonte: Elaboração própria

Figura 11 - Resultados de um sinistro de trânsito fictício registrado na versão móvel do aplicativo Epicollect5 após ajustes de estrutura e de linguagem

☰
SD SINISTROS DE TRANSITO_UBERLA...

< VOLTAR
EXCLUIR

ATENÇÃO:
Antes de iniciar o registro, ative o serviço de localização (GPS) de seu dispositivo!

IMPORTANTE:
Caso o registro esteja sendo realizado pelo aplicativo móvel: Antes de coletar a localização do acidente de trânsito, se aproxime o máximo possível do local do acidente para garantir a precisão posicional do dado.
Se isso não for possível, o registro não deverá ser realizado via aplicativo móvel, e sim via aplicativo web (no qual é possível editar a geolocalização da informação coletada).

Forma de Registro

via aplicativo móvel (app)

Localização do Acidente

LATITUDE	-18.918418
LONGITUDE	-48.254209
PRECISÃO	4

Data do Acidente

19/02/2024

Horário do Acidente

22:51:26

ATENÇÃO:
Na próxima etapa será solicitado a identificação do tipo de acidente de trânsito. Seguem algumas definições que podem ser úteis. Em caso de dúvida, retorne a este tópico para consulta.

Veículo: todo e qualquer meio de transporte existente, motorizado ou não, que transite por quaisquer vias

Atropelamento: quando pedestre(s) ou animal(is) sofre(m) o impacto de um veículo em movimento

Choque: quando há impacto de um veículo contra qualquer objeto fixo ou objeto móvel sem movimento

Colisão: quando um veículo em movimento sofre o impacto de outro veículo também em movimento

☰
SD SINISTROS DE TRANSITO_UBERLA...

< VOLTAR
EXCLUIR

Tipo de Acidente de Trânsito

choque

n° de Veículos Envolvidos

1

Envolvido n°1 (Veículo)

automóvel

Sexo do Condutor (Envolvido n°1)

masculino

Faixa Etária do Condutor (Envolvido n°1)

adulto jovem (25 a 44 anos)

Acidente com Mortos ou Feridos?

não

Condições do Tempo

chuva

Provável Motivo do Acidente

velocidade incompatível

Alguma Observação Complementar?

O veículo se chocou contra um poste. O motorista não se feriu.

4.2 AVALIAÇÃO DA COMPLETUDE DE ATRIBUTOS

A completude é um parâmetro de qualidade que aponta o excesso (comissão) ou a ausência (omissão) de feições, seus atributos e relações. Nesta etapa do trabalho foi avaliada a completude de atributos da estrutura revisada dos formulários para a coleta de dados colaborativos de sinistros de trânsito, considerando níveis distintos de usuários.

Vale rememorar que, para aplicações em campo, a pesquisa propõe o desenvolvimento do questionário composto exclusivamente por perguntas fechadas, salve exceção ao atributo "Observações complementares", que permite uma resposta aberta. Logo, para cada atributo que compõe o formulário são oferecidas opções de resposta previamente definidas. O usuário pode selecionar somente uma opção de resposta, e o questionário só permite ir à próxima questão após a definição do valor do atributo da questão em aberto. Isso se torna viável uma vez que é disponibilizado no questionário opções de escolha como "nenhuma das alternativas" e "não foi possível identificar", valores úteis para os casos de dúvidas ou dificuldades encontradas pelo voluntariado durante o processo de registro. Dessa forma, cada sinistro de trânsito registrado de forma voluntária fica associado a um conjunto completo de valores dos atributos estabelecidos pelo usuário responsável pela coleta. Em outras palavras, a qualquer registro colaborativo estará atribuído 100% de completude, sem casos de comissão ou omissão de dados.

Esta pesquisa buscou proceder com uma avaliação extrínseca de completude. Para tal, o formulário proposto foi comparado com a estrutura de um RAT (Relatório de Acidente de Trânsito) fixado pela ABNT NBR 12898 (1993), e com alguns sistemas oficiais de registro de sinistros de trânsito do estado de Minas Gerais, a saber, Boletim de Ocorrência (BO) da PMMG e Delegacia Virtual da PCMG. Ao RAT foi atribuído 100% de completude, tornando-se a referência para as análises comparativas.

O Quadro 7 lista os principais dados estatísticos de sinistros de trânsito que compõem um RAT. E em suas colunas adjacentes são apresentados os sistemas (oficiais ou colaborativos) de registro de sinistros de trânsito (em cada um dos quais são apontados a relação de atributos habilitados para coleta).

Quadro 7 - Dados coletados em diferentes sistemas oficiais e colaborativos de registro de sinistros de trânsito

Dados Estatísticos de Sinistros de Trânsito (RAT - NBR 12898:1993) ¹		BO PMMG ²	Delegacia Virtual PCMG ³	Formulário para Registro Colaborativo ⁴		
Itens	Atributos			Especialista	Técnico	Popular
Dados gerais	Data e hora do acidente	X	X	X	X	X
	Tipo de acidente	X	X	X	X	X
	nº de veículos envolvidos	X	X	X	X	X
	nº de passageiros mortos, feridos ou ileso	X	X	X		
	nº de Pedestres mortos ou feridos	X		X	X	X
	Outras propriedades atingidas	X				
Condições do local	Características da via	X				
	Condições da pista e calçada	X		X	X	
	Condições da sinalização	X		X	X	
	Condições do tempo	X		X	X	X
Detalhes do veículo	Dados do proprietário	X	X			
	Dados do veículo	X	X			
	Espécie de veículo	X				
	Categoria de veículo	X				
	Tipo de veículo	X	X	X	X	X
	Condições do veículo (geral, pneus, luzes) ⁵					
	Danos ao veículo decorrentes do acidente	X	X			
	Manobra do veículo no momento do acidente	X	X			
	Situação do veículo ⁶	X				
Condutor do veículo	Dados do condutor	X	X			
	Gravidade das lesões	X		X		
	Condições adversas alegadas	X	X			
	Infração constatada					
	Atendimento médico ⁷	X				
Passageiro vitimado	Dados do passageiro morto ou ferido	X				
	Gravidade das lesões	X		X		
	Atendimento médico ⁷	X				
Pedestre vitimado	Dados do pedestre morto ou ferido	X				
	Gravidade das lesões	X		X		
	Atendimento médico ⁷	X				
Localização do acidente		X	X	X	X	X
Croqui e fotografias do acidente		X	X			
Testemunhas	Dados das testemunhas	X	X			
Providências	Perícia técnica no local	X				
	Presença de outros órgãos no local	X				
Relatório policial		X				
Identificação do policial		X				
<p>Notas:</p> <p>¹ Principais dados estatísticos de sinistros de trânsito que compõem um RAT fixado pela NBR 12898:1993.</p> <p>² Boletim de Ocorrência da PMMG vinculado ao SIDS para o registro, pelo policial, de sinistros de trânsito com ou sem vítimas.</p> <p>³ Delegacia Virtual, sistema <i>online</i> da PCMG vinculado ao SIDS para o registro, pelo envolvido, de sinistros de trânsito sem vítimas.</p> <p>⁴ Formulários para o registro colaborativo, pelo usuário, de sinistros de trânsito com ou sem vítimas (vide item 4.1 deste trabalho).</p> <p>⁵ Condições aparentes do veículo antes do acidente.</p> <p>⁶ Destino (status final) do veículo após o registro da ocorrência.</p> <p>⁷ Local para onde e por quem a vítima foi removida para receber atendimento médico.</p>						

Como exposto no Quadro 7, o RAT é composto por trinta e sete atributos, agrupados em doze itens, sendo eles: dados gerais, condições do local, detalhes do veículo, condutor do veículo, passageiro vitimado, pedestre vitimado, localização do acidente, *croqui* e fotografias do acidente, testemunhas, providências, relatório policial e identificação do policial. Cada atributo equivale a 1/37 do formulário como um todo, isto é, cerca de 2,7%. Desta forma, segue abaixo (Tabela 6) o grau de completude de atributos para cada um dos sistemas de registro elencados nesta análise, quando comparados com o RAT fixado pela ABNT NBR 12898 (1993), tendo como base de cálculo o número de atributos habilitados para coleta.

Tabela 6 - Grau de completude para cada sistema de registro de sinistros de trânsito

Sistema de Registro de Sinistros de Trânsito		Grau de Completude (%)
Sistema	BO da PMMG	94,6
Oficial	Delegacia Virtual da PCMG	37,8
Sistema	Formulário Especialista	35,1
Colaborativo	Formulário Técnico	24,3
	Formulário Popular	18,9

Para cada item que compõe um RAT, a Tabela 7 apresenta a quantidade de campos habilitados para coleta no sistema colaborativo de registro de sinistros.

Tabela 7 - Quantidade de campos habilitados para coleta no formulário colaborativo para cada item que compõe um RAT

Itens (RAT - NBR 12898:1993)	Formulário para Registro Colaborativo					
	Especialista		Técnico		Popular	
	quant.	%	quant.	%	quant.	%
Dados gerais	5	38,4	4	44,5	4	57,1
Condições do local	3	23,1	3	33,3	1	14,3
Detalhes do veículo	1	7,7	1	11,1	1	14,3
Condutor do veículo	1	7,7	0	0,0	0	0,0
Passageiro vitimado	1	7,7	0	0,0	0	0,0
Pedestre vitimado	1	7,7	0	0,0	0	0,0
Localização do acidente	1	7,7	1	11,1	1	14,3
Croqui e fotografias do acidente	0	0,0	0	0,0	0	0,0
Testemunhas	0	0,0	0	0,0	0	0,0
Providências	0	0,0	0	0,0	0	0,0
Relatório policial	0	0,0	0	0,0	0	0,0
Identificação do policial	0	0,0	0	0,0	0	0,0
Total	13	100,0	9	100,0	7	100,0

Avaliando a Tabela 7, é possível perceber que os campos habilitados para coleta no sistema colaborativo são oriundos principalmente dos itens "Dados gerais" e "Condições do local", representando mais que 50,0% do total de campos para os três formulários (Especialista, Técnico e Popular).

A Tabela 8 apresenta o grau de completude dos formulários para cada um dos doze itens listados, quando comparados com a estrutura de um RAT.

Tabela 8 - Grau de completude no formulário colaborativo para cada item que compõe um RAT

Itens (RAT - NBR 12898:1993)	Grau de completude (%)		
	Formulário para Registro Colaborativo		
	Especialista	Técnico	Popular
Dados gerais	83,3	66,7	66,7
Condições do local	75,0	75,0	25,0
Detalhes do veículo	11,1	11,1	11,1
Condutor do veículo	20,0	0,0	0,0
Passageiro vitimado	33,3	0,0	0,0
Pedestre vitimado	33,3	0,0	0,0
Localização do acidente	100,0	100,0	100,0
Croqui e fotografias do acidente	0,0	0,0	0,0
Testemunhas	0,0	0,0	0,0
Providências	0,0	0,0	0,0
Relatório policial	0,0	0,0	0,0
Identificação do policial	0,0	0,0	0,0

Embora a estrutura proposta para o mapeamento colaborativo de sinistros de trânsito permita o registro do sexo e da faixa etária do pedestre vitimado e do condutor, os atributos "dados do condutor" e "dados do pedestre vitimado" não foram considerados como habilitados para registro na tabela comparativa com o RAT, uma vez que a relação de dados dos envolvidos solicitada em um RAT é extensa, tornando o registro somente de sexo e faixa etária não significativo.

Segundo a ABNT NBR 12898 (1993), toda informação que não seja possível ser colocada nos campos apropriados, mas que tenha importância para esclarecer as circunstâncias ou causas presumidas do acidente, deve ser detalhada no campo "Relatório policial". Porém, ainda que não componha a relação de atributos em um RAT, a causa presumida do sinistro possui campo específico para registro em todos os sistemas de registro (oficiais e colaborativos).

Apesar do baixo grau de completude de atributos do sistema colaborativo, a relação de atributos habilitados para coleta em seus formulários contém os principais valores comumente utilizados em pesquisas de análise espacial de sinistros de trânsito. Dentre eles, se destacam a data e hora do sinistro, localização do sinistro, tipo de sinistro, sexo e faixa etária dos envolvidos, quantificação e qualificação de vítimas e tipo de veículo.

A maior parte dos atributos não disponíveis para coleta nos formulários do sistema colaborativo usualmente não são utilizados em estudos de engenharia de tráfego e planejamento de trânsito. São exemplos os dados protegidos pela Lei Geral de Proteção de Dados Pessoais (LGPD), outras propriedades atingidas, dados das testemunhas, espécie e categoria de veículo, condições aparentes do veículo antes do acidente, destino do veículo após o registro da ocorrência, infração constatada e identificação do policial.

Enfim, o formulário proposto para o registro colaborativo de dados de sinistros de trânsito acumula uma quantidade de informações considerada suficiente para contribuir com os estudos estatísticos, ao mesmo tempo que permite um experiência sucinta, prática e intuitiva ao voluntariado.

4.3 AVALIAÇÃO DA ACURÁCIA POSICIONAL

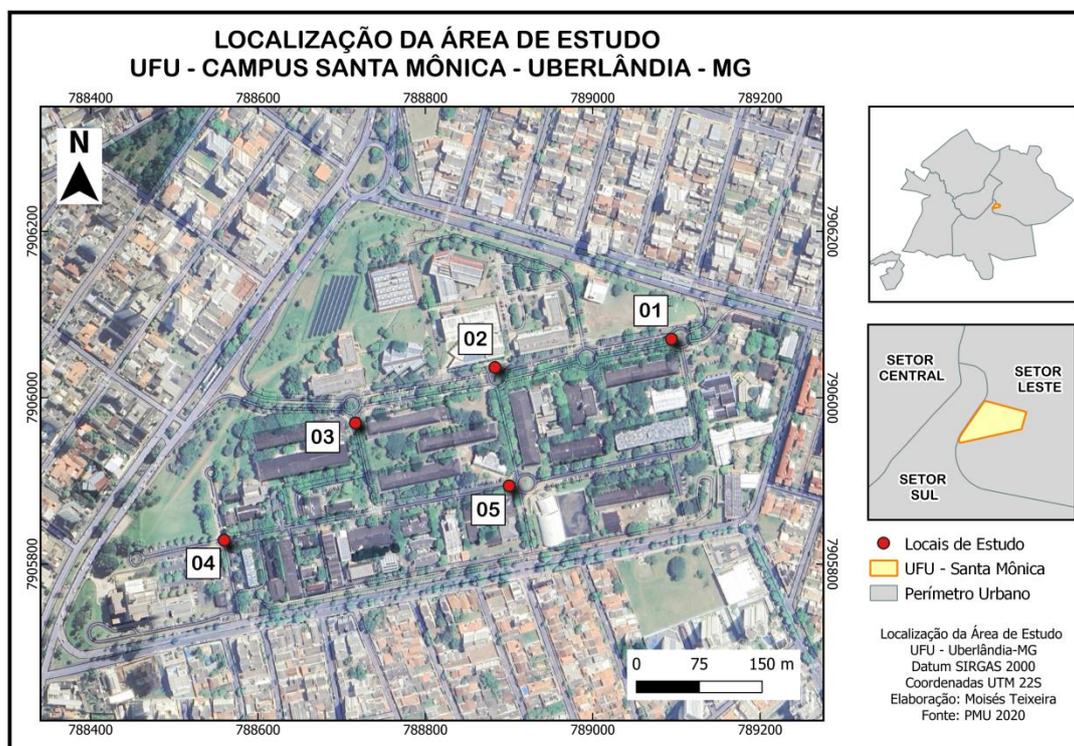
4.3.1 *Precisão da Coleta In Loco - Análise do Sensor GPS*

Uma etapa preliminar à avaliação da acurácia posicional dos dados colaborativos consiste em aferir a precisão do processo de coleta *in loco* de dados via versão móvel do aplicativo Epicollect5. O método utilizado se traduz em efetuar uma série de registros de geolocalização, por GPS, de um mesmo ponto em intervalos de tempo uniforme (20 minutos), o que permite verificar a dispersão dos dados da amostra com o tempo, causada por variações incontrolláveis e aleatórias dos instrumentos de medida e de condições externas.

Foram definidos um total de 5 locais contidos na área de estudo, neste caso, as dependências do Campus Santa Mônica da Universidade Federal de Uberlândia (UFU), situado no Bairro Santa Mônica, Zona Leste do município de Uberlândia. A

Figura 12 revela a disposição dos 5 pontos no Campus da Universidade. As Figuras 13 a 17 apresentam o registro fotográfico dessas regiões.

Figura 12 - Mapa de localização da UFU Campus Santa Mônica



Fonte: Elaboração própria

A escolha das regiões de estudo buscou selecionar locais expostos a diferentes coberturas, de forma a avaliar se a presença de elementos diversos (como edificações e copas de árvores) próximo ao receptor interfere na orientação do GPS e, por consequência, na precisão dos dados. Os Locais 1 (Figura 13) e 2 (Figura 14) situam-se no desenvolvimento de rotatórias. Os Locais 3 (Figura 15) e 5 (Figura 17) estão posicionados sobre faixas para travessia de pedestres. E o Local 4 (Figura 16) encontra-se no cruzamento entre duas vias.

Figura 13 - Registro fotográfico do Local 1



Figura 14 - Registro fotográfico do Local 2



Figura 15 - Registro fotográfico do Local 3



Figura 16 - Registro fotográfico do Local 4



Figura 17 - Registro fotográfico do Local 5



O dispositivo móvel utilizado para coletar os dados de campo foi o *smartphone* Motorola Moto G60, dotado de sistema operacional Android. Para cada um dos 5 pontos, foram coletados 12 amostras a cada 20 minutos, perfazendo um total de 60 registros. As coletas foram realizadas na manhã do dia 08 de agosto de 2023, das 07h00 às 11h00. Com a conclusão do trabalho de campo, procedeu-se com *upload* para o armazenamento dos dados no servidor central do aplicativo Epicollect5. Logo após, os dados foram exportados no formato CSV.

No *software* QGIS (versão 3.24.1), os dados foram importados e, após definição de geometria e especificação do SRC (Sistema de Referência de Coordenadas) da camada, a amostra foi convertida para uma camada vetorial tipo

ponto composta por dados georreferenciados. Na Figura 18 observa-se a representação gráfica dos dados, sendo possível visualizar sua dispersão.

Figura 18 - Representação gráfica das amostras coletadas



Fonte: Elaboração própria

Utilizando a ferramenta de análise vetorial do QGIS, aplicou-se o algoritmo para extrair informações estatísticas da coordenada X (*East*) e da coordenada Y (*North*) do sistema UTM, para cada conjunto de dados dos 5 locais selecionados para análise. Os principais resultados da estatística básica são apresentados na Tabela 9.

Para a análise da dispersão dos dados, foram utilizadas as medidas de dispersão e as representações gráficas de dados estatísticos. Ao consultar a Tabela 9, as medidas de dispersão apresentadas são a amplitude e o desvio padrão. A amplitude representa a diferença entre o maior e o menor valor do conjunto de observações. Entretanto, a amplitude não informa como os valores variam entre as extremidades. Já o desvio padrão é uma medida de dispersão amplamente utilizada para acompanhar uma medida de tendência central e sumarizar um conjunto de dados, uma vez que se trata de uma grandeza que leva em consideração toda a informação contida na amostra.

Tabela 9 - Resultados estatísticos básicos das amostras coletadas

	Coordenada	Mínimo (m)	Máximo (m)	Amplitude (m)	Média (m)	Mediana (m)	Desvio Padrão (m)
Local 1	X	789091,577	789096,009	4,432	789093,608	789093,260	1,254
	Y	7906062,474	7906072,997	10,523	7906069,396	7906069,691	2,509
Local 2	X	788880,561	788883,476	2,915	788882,287	788882,247	0,814
	Y	7906034,849	7906039,041	4,192	7906037,102	7906037,447	1,329
Local 3	X	788714,688	788717,639	2,951	788716,364	788716,570	0,817
	Y	7905966,847	7905969,951	3,104	7905968,294	7905968,125	0,964
Local 4	X	788559,401	788563,933	4,532	788561,406	788560,955	1,391
	Y	7905823,148	7905826,698	3,550	7905825,317	7905825,537	0,949
Local 5	X	788897,645	788902,386	4,741	788899,937	788899,979	1,305
	Y	7905893,978	7905899,085	5,107	7905896,293	7905895,846	1,631

De forma a complementar a análise, foram geradas representações gráficas de dados estatísticos para cada conjunto de observações (Figuras 19 a 28), sendo elas o gráfico de dispersão (*scatterplot*) e o diagrama de caixa (*Box Plot*). O *software* utilizado para a produção dos gráficos foi o programa STATISTICA (*StatSoft*), versão 7.1.

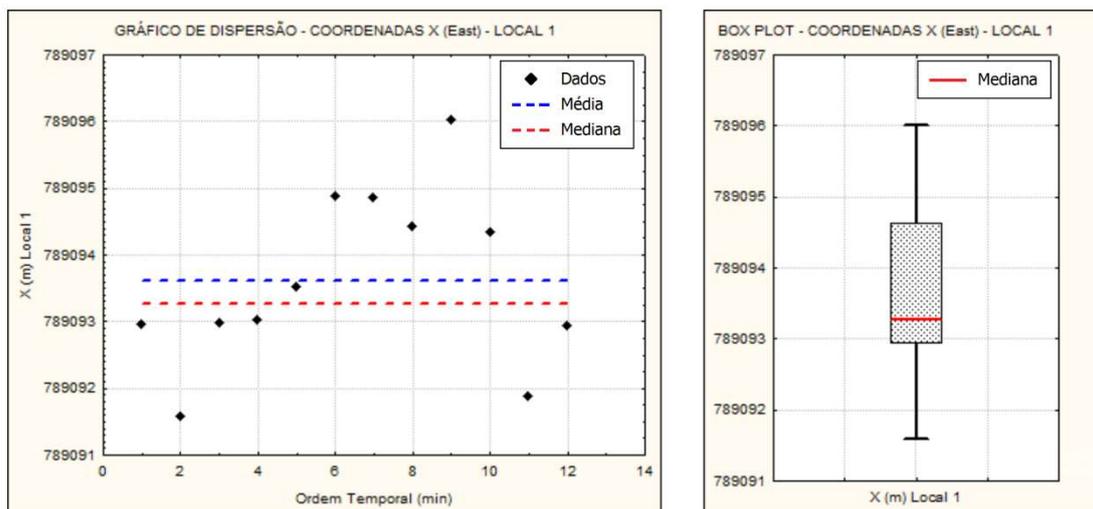
Figura 19 - *Scatterplot* e *Box Plot* das coordenadas X do conjunto de dados do Local 1

Figura 20 - Scatterplot e Box Plot das coordenadas Y do conjunto de dados do Local 1

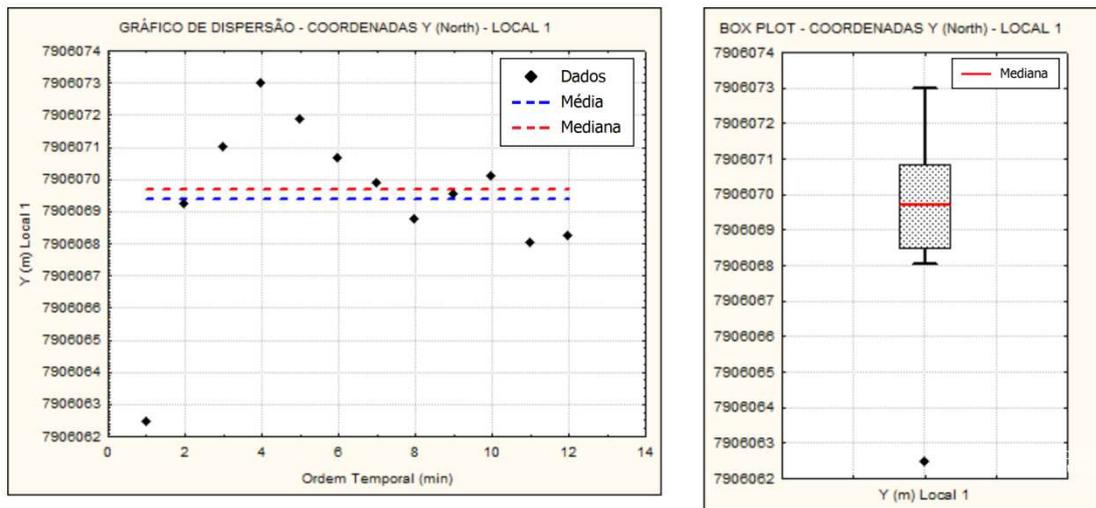


Figura 21 - Scatterplot e Box Plot das coordenadas X do conjunto de dados do Local 2

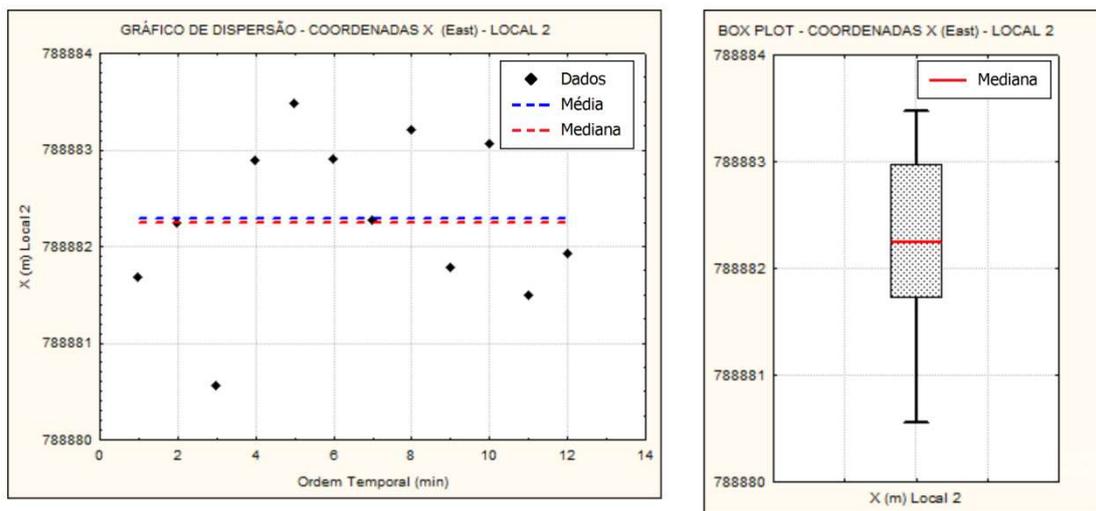


Figura 22 - Scatterplot e Box Plot das coordenadas Y do conjunto de dados do Local 2

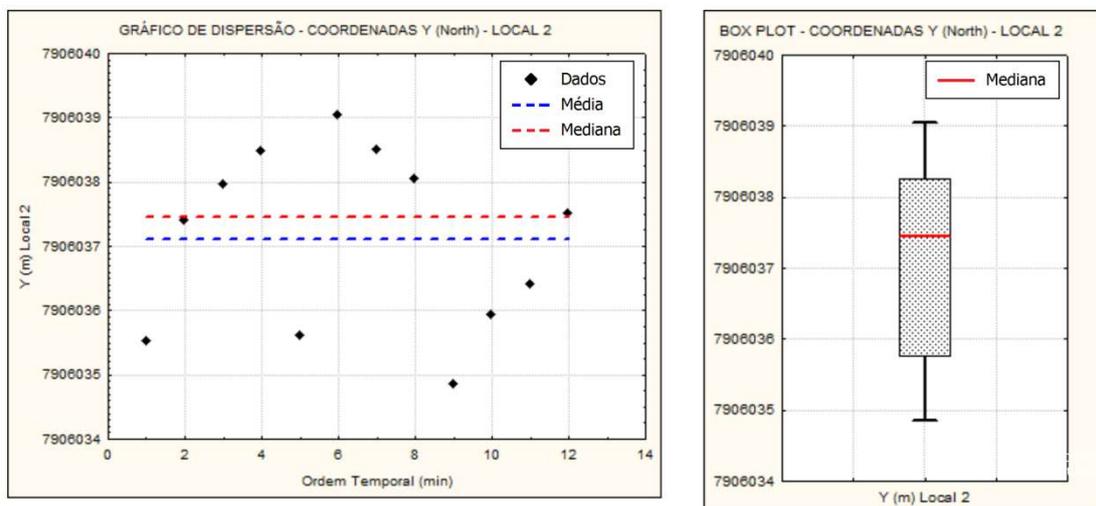


Figura 23 - Scatterplot e Box Plot das coordenadas X do conjunto de dados do Local 3

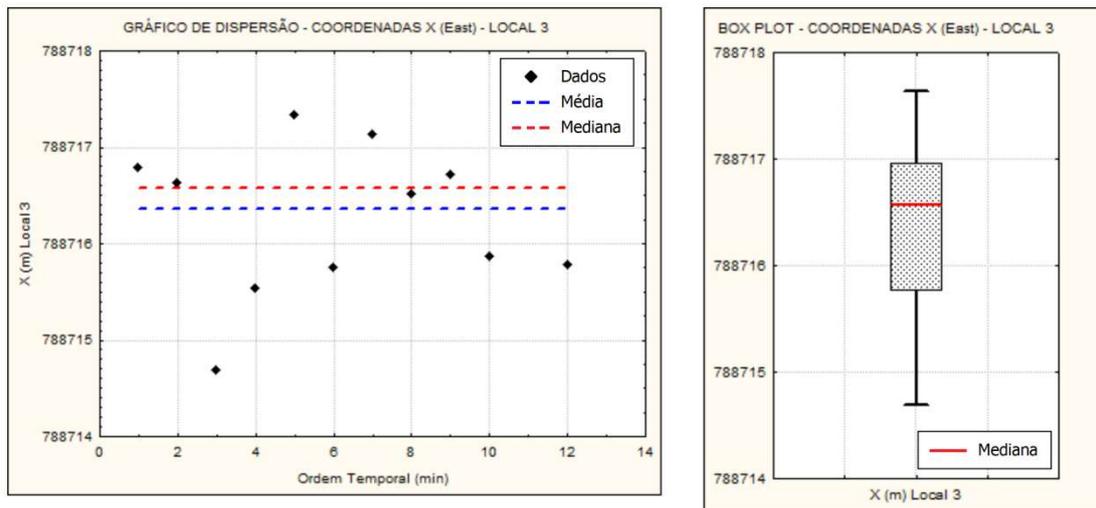


Figura 24 - Scatterplot e Box Plot das coordenadas Y do conjunto de dados do Local 3

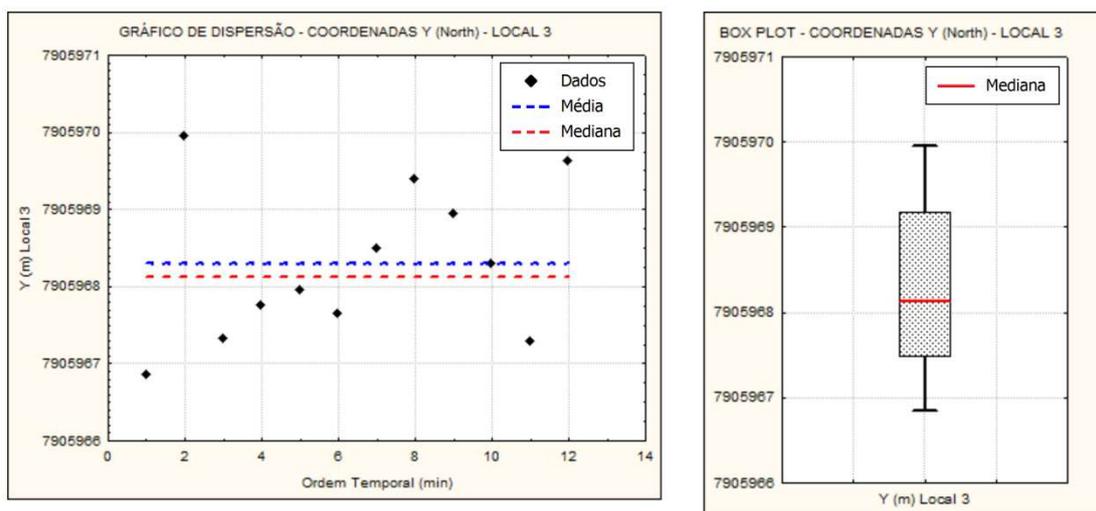


Figura 25 - Scatterplot e Box Plot das coordenadas X do conjunto de dados do Local 4

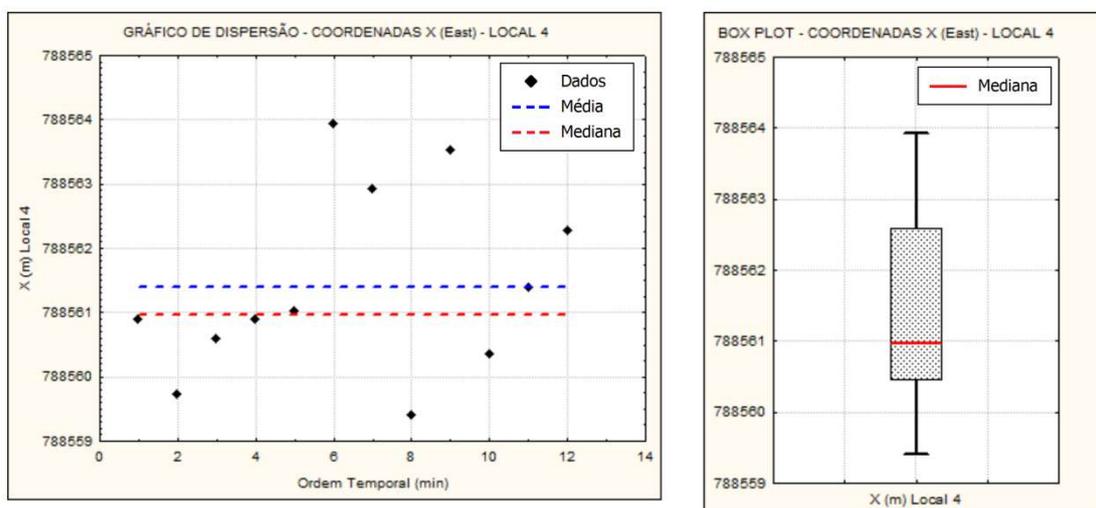


Figura 26 - Scatterplot e Box Plot das coordenadas Y do conjunto de dados do Local 4

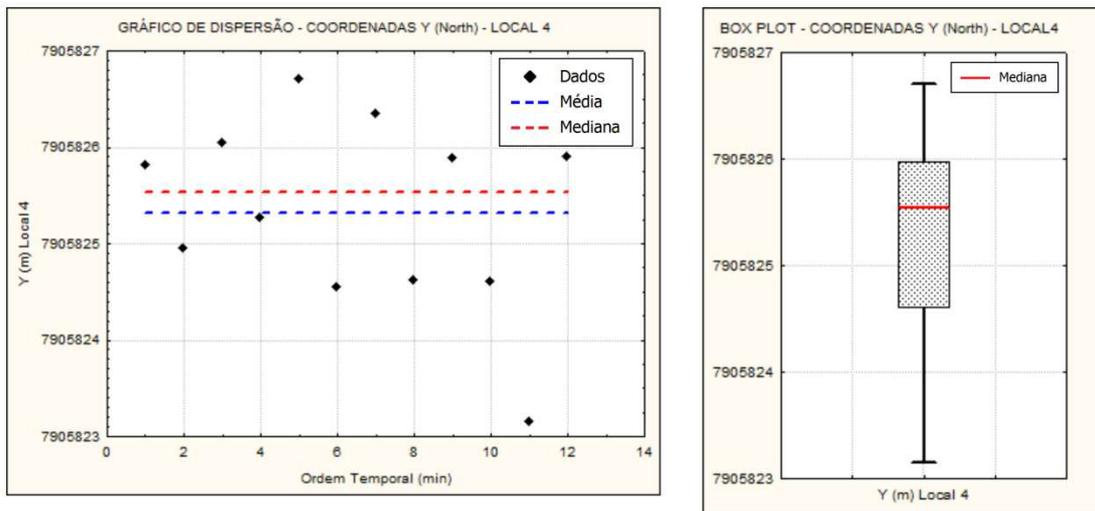


Figura 27 - Scatterplot e Box Plot das coordenadas X do conjunto de dados do Local 5

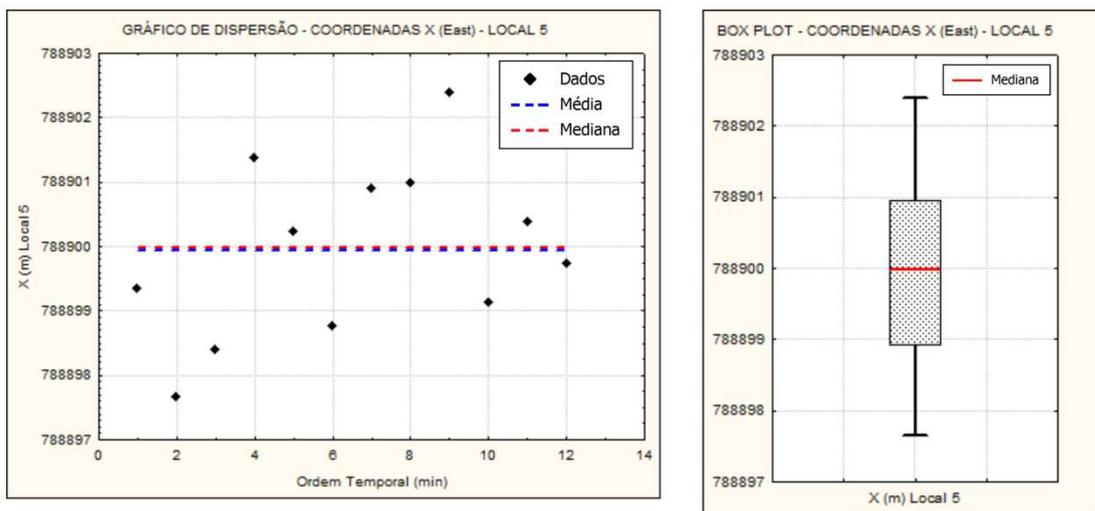
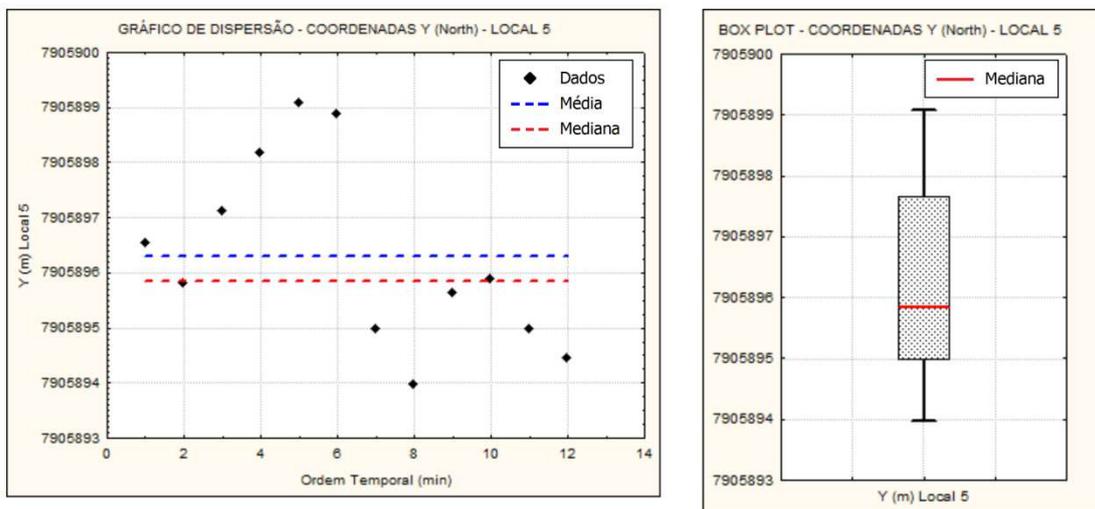


Figura 28 - Scatterplot e Box Plot das coordenadas Y do conjunto de dados do Local 5



O conceito estatístico de dispersão diz respeito à variabilidade dos dados em torno da média central. Neste sentido, o gráfico de dispersão (*scatterplot*) é uma ferramenta gráfica utilizada para representar a relação entre duas variáveis quantitativas e a consistência dos dados, o que permite detectar possíveis padrões e tendências nos dados ou correlações entre as variáveis.

Já o diagrama de caixa (*Box Plot*) é um gráfico que permite avaliar a simetria e distribuição dos dados por meio de extremos e quartis, e revela a presença de possíveis dados discrepantes (*outliers*) no conjunto de dados. As cinco medidas estatísticas descritivas representadas em um Box Plot são, de baixo para cima, o limite inferior, o primeiro quartil, o segundo quartil (ou mediana), o terceiro quartil e o limite superior. Se algum valor ultrapassar esses limites (inferior ou superior), serão pontos atípicos do conjunto de dados da amostra, isto é, *outliers*.

Ao analisar a distribuição de cada conjunto de dados por meio dos Box Plots, percebe-se casos de simetria, assimetria positiva e assimetria negativa. Esta aleatoriedade é um indício de que os dados, quando coletados em locais distintos, não apresentam tendências. Foi identificado somente um *outlier* em todos os conjuntos de dados. O episódio ocorreu durante a coleta de uma das coordenadas Y (norte) do Local 1.

Quanto menor o desvio padrão, maior a homogeneidade dos dados da amostra. Logo, ao analisar os resultados estatísticos apresentados na Tabela 9, o conjunto de dados mais homogêneo são as coordenadas X (*East*) referentes à amostra do Local 2, apresentando um desvio padrão de 0,814 m e amplitude de 2,915 m. E o conjunto mais heterogêneo é representado pelas coordenadas Y (*North*) dos dados da amostra do Local 1, com um desvio padrão de 2,509 m e amplitude de 10,523 m. Entretanto, como informado, foi detectado a presença de *outlier* no conjunto de dados, o que influencia nos resultados aqui apresentados.

O próximo passo consiste em remover da série de dados originais os *outliers* identificados, e executar os cálculos estatísticos para uma nova análise da dispersão dos dados. Porém, é importante compreender que um *outlier* pode ser um erro da base, sendo necessário seu descarte, mas também há situações em que os *outliers* são considerados componentes importantes na representação de um conjunto de dados, merecendo permanecer na série, principalmente nos casos em que a quantidade de

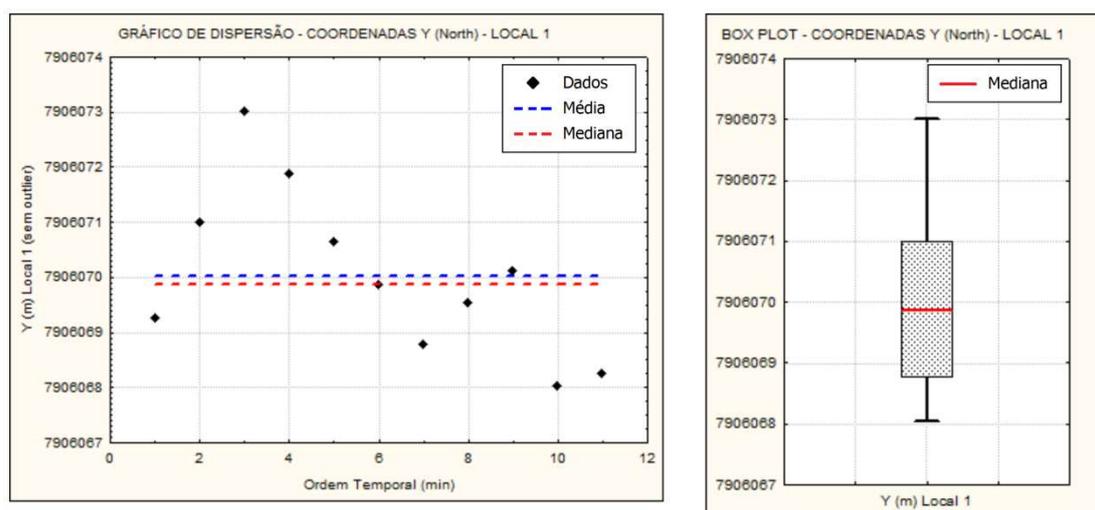
dados coletados por amostra é pequena. Enfim, cabe ao analista de dados avaliar cada caso e definir como tratar os *outliers*.

Para fins de comparação e compreensão do impacto nos resultados, seguem as estatísticas básicas das coordenadas Y dos dados remanescentes da amostra coletados no Local 1, após remoção do *outlier* (Tabela 10). A Figura 29 apresenta suas representações gráficas.

Tabela 10 - Resultados estatísticos básicos da amostra coletada, após remoção de *outliers*

	Coordenada	Mínimo (m)	Máximo (m)	Amplitude (m)	Média (m)	Mediana (m)	Desvio Padrão (m)
Local 1	Y	7906068,029	7906072,997	4,968	7906070,026	7906069,866	1,455

Figura 29 - Scatterplot e Box Plot da amostra coletada, após remoção de *outliers*



Com a supressão dos *outliers*, as amostras da pesquisa passaram a conter diferentes números de observações. Por esse motivo, o erro padrão será o parâmetro estatístico empregado na análise dos resultados.

O erro padrão da média é uma medida de dispersão utilizada para comparar conjunto de dados com tamanhos amostrais diferentes. O valor do erro padrão é dado pela divisão do desvio padrão pela raiz quadrada do tamanho amostral. Isto é, o desvio padrão é ponderado pelo número de observações da amostra. Após os cálculos, a Tabela 11 apresenta a relação de todas as medidas de dispersão dos dados das amostras da pesquisa após remoção dos *outliers*.

Tabela 11 - Medidas de dispersão das amostras coletadas, após remoção de *outliers*

	Coordenada	Tamanho Amostral	Amplitude (m)	Desvio Padrão (m)	Erro Padrão (m)
Local 1	X	12	4,432	1,254	0,362
	Y	11	4,968	1,455	0,439
Local 2	X	12	2,915	0,814	0,235
	Y	12	4,192	1,329	0,384
Local 3	X	12	2,951	0,817	0,236
	Y	12	3,104	0,964	0,278
Local 4	X	12	4,532	1,391	0,402
	Y	12	3,550	0,949	0,274
Local 5	X	12	4,741	1,305	0,377
	Y	12	5,107	1,631	0,471

Ao analisar os novos resultados estatísticos apresentados na Tabela 11, o conjunto de dados mais homogêneo permanece sendo as coordenadas X (*East*) referentes à amostra do Local 2, apresentando um erro padrão de 0,235 m. E o conjunto mais heterogêneo passa a ser representado pelas coordenadas Y (*North*) dos dados da amostra do Local 5, com um erro padrão de 0,471 m.

A dispersão de dados está relacionada com a qualidade do experimento. Por isso é tão importante a realização desta etapa do trabalho, que consiste em compreender a precisão da coleta *in loco*. E por ser uma etapa preliminar à avaliação da acurácia posicional, que é um parâmetro de qualidade (ISO 19157, 2013), os resultados e discussões deste tópico da pesquisa serão essenciais para o andamento dos estudos.

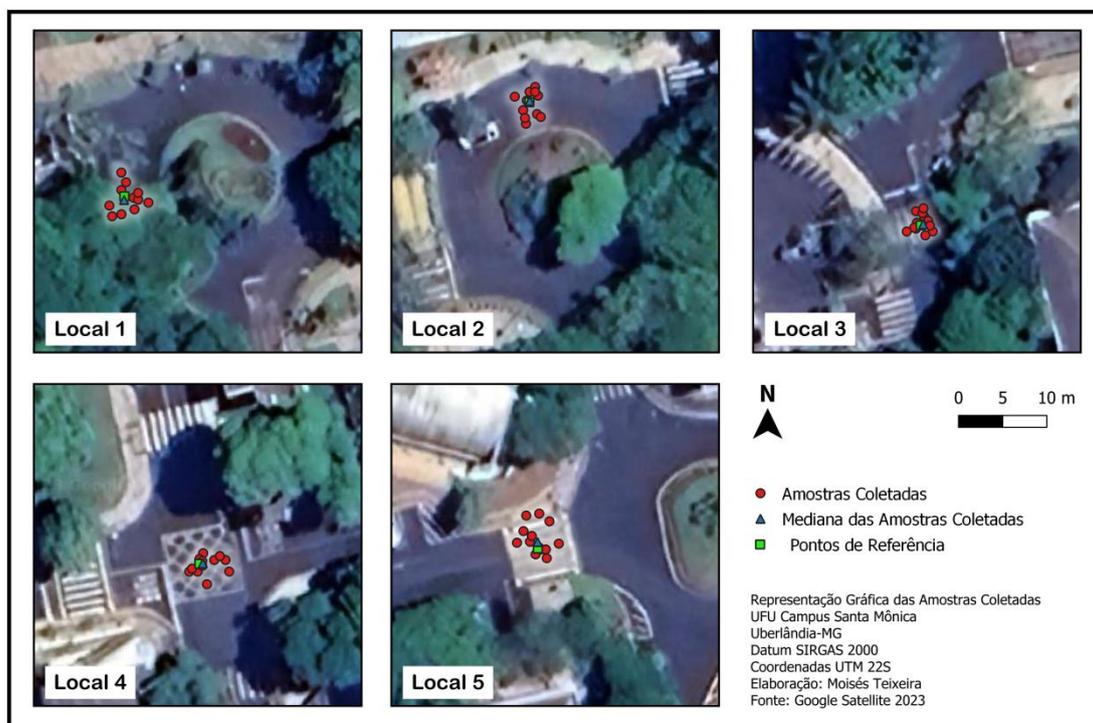
Dentre os agregadores comumente utilizados para compreender a tendência central de um conjunto de dados, há a média e a mediana. A média é influenciada por todos os valores da amostra. Isto é, conjunto de dados que contêm valores discrepantes (*outliers*) interfere diretamente no resultado da média. Ao analisar a Figura 20, é possível perceber esse efeito. Nestes casos, cujas distribuições numéricas são distorcidas, indica-se o uso da mediana para retornar a tendência central. Por esse motivo, a mediana das coordenadas UTM será o valor utilizado na etapa de avaliação da acurácia posicional de dados colaborativos.

4.3.2 Avaliação da Acurácia Posicional

Seguindo a metodologia proposta nesta pesquisa, para mensurar a acurácia posicional dos dados coletados, as coordenadas dos pontos amostrais de dados colaborativos de sinistro de trânsito foram comparadas com as coordenadas dos pontos homólogos extraídos da base de referência, aplicando-se os passos e procedimentos estabelecidos pelo Decreto-lei nº 89817 (BRASIL, 1984) e as especificações técnicas da DSG do Exército Brasileiro que a complementam.

Como informado ao final do item anterior, o valor utilizado para os pontos amostrais do dado colaborativo (pontos avaliados) foi a mediana das coordenadas UTM dos registros de 5 ocorrências fictícias de sinistros de trânsito, coletados por um receptor GPS de um dispositivo móvel nas dependências do Campus Santa Mônica da UFU para a análise de precisão da coleta *in loco* (etapa preliminar à avaliação da acurácia posicional). Já os pontos homólogos de referência são as coordenadas UTM exatas dos 5 locais das ocorrências fictícias, extraídas da base oficial (Mapa georreferenciado do Campus da Universidade). A Figura 30 apresenta a representação gráfica desses dados.

Figura 30 - Representação gráfica dos pontos amostrais e dos pontos homólogos de referência



Fonte: Elaboração própria

Após a realização dos cálculos preliminares, seguem os resultados (Tabela 12):

Tabela 12 - Mediana das coordenadas e medidas para avaliação da acurácia posicional dos pontos amostrais

	Y Ref (m)	X Ref (m)	Y Aval (m)	X Aval (m)	e_H (m)	EMQ_H (m)
Local 1	7906070,275	789093,794	7906069,866	789093,260	0,673	0,559
Local 2	7906037,300	788882,690	7906037,447	788882,247	0,467	0,559
Local 3	7905967,990	788716,780	7905968,125	788716,570	0,250	0,559
Local 4	7905825,490	788561,060	7905825,537	788560,955	0,115	0,559
Local 5	7905895,100	788900,490	7905895,846	788899,979	0,904	0,559

O resultado da distância euclidiana (e_{Hi}) entre os pontos de cada par de pontos homólogos variou de 0,12 m a 0,90 m. E o Erro Médio Quadrático (EMQ) dos erros da amostra calculado foi de 0,559 m.

Com base na Tabela 3, para a escala 1:2000, foi possível constatar que a acurácia posicional dos pontos amostrais de dados colaborativos de sinistros de trânsito atingiu a classe B do indicador estatístico PEC-PCD, uma vez que:

- i. Pelo menos 90% dos pares de pontos homólogos apresentam distâncias euclidianas menores que 1 metro. Isto é, em 100% das amostras os erros são menores ou iguais ao valor da tolerância PEC, atendendo a primeira condição;
- ii. O Erro Médio Quadrático dos erros da amostra (0,559 metro) é menor ou igual à tolerância EP (0,60 metro), atendendo a segunda condição.

A escala (1:2000) e classe (B) reveladas são compatíveis com pesquisas brasileiras, como Oliveira e Brito (2019) e Silva *et al.* (2018), que orientam empregar escalas iguais ou maiores a 1:10000 para a representação de feições com alto nível de detalhamento, a exemplo de cartas cadastrais, frequentemente utilizadas em área urbana para a definição dos limites legais das propriedades. Para escalas iguais ou maiores a 1:2000 este grau de detalhamento é ainda maior, expondo as menores entidades contidas no conjunto de dados (BALLATORE; BERTOLOTTI, 2011). Desta forma, a qualidade posicional dos dados colaborativos em estudo é julgada satisfatória.

4.4 AVALIAÇÃO DA ACURÁCIA TEMPORAL

A qualidade temporal dos dados colaborativos está relacionada à exatidão encontrada no componente temporal do dado geoespacial. Em outras palavras, mensura a qualidade dos atributos temporais e da relação temporal das feições.

Para a avaliação da qualidade temporal, foi levado em consideração as funcionalidades do sistema do Epicollect5, aplicativo *web* utilizado nesta pesquisa para a construção do modelo de questionário e aplicação de testes em campo. Neste, o campo para preenchimento da data e horário do sinistro é editável, independente do *software* a ser utilizado para a coleta e hospedagem de dados (a saber, via aplicativo móvel, que permite registro *in loco*, em tempo real, com uso de receptor GNSS, ou via aplicativo *web*, que permite registros *a posteriori*).

Esta circunstância, associada ao fato de que é esperado que um mesmo sinistro seja registrado mais de uma vez por diferentes usuários, tende a gerar resultados heterogêneos, o que pode comprometer a precisão dos dados coletados.

Tendo em vista que a dispersão de dados está relacionada com a qualidade do experimento, segue a especificação do método para a análise exploratória do componente temporal e o tratamento matemático-estatístico de um conjunto de dados.

Medida de tendência central para especificação do valor do atributo temporal validável: mediana. Motivo: a mediana reduz a influência de valores discrepantes (*outliers*) na medida de tendência central de um conjunto de dados.

Já o procedimento proposto para a medida da qualidade temporal é expresso no Quadro 8, estruturado segundo os descritores definidos na ET-CQDG (BRASIL, 2016), em conformidade com a ISO 19157 (2013), para a composição de uma medida de qualidade.

O desvio padrão foi a medida básica escolhida uma vez que se trata de uma grandeza que leva em consideração toda a informação contida na amostra, sendo eficaz em sumarizar um conjunto de dados.

Referente à análise do resultado: quanto maior o valor do desvio padrão, menor será a qualidade dos dados, uma vez que o desvio-padrão elevado expressa uma maior variabilidade do conjunto de dados, o que caracteriza uma informação mais heterogênea.

Quadro 8 - Medida de qualidade temporal - horário do sinistro de trânsito

Linha	Componente	Descrição
1	Nome	Horário do sinistro de trânsito
2	Nome alternativo	-
3	Elemento de qualidade	Exatidão de uma medida de tempo
4	Medida básica	Desvio padrão populacional
5	Definição	Medida de dispersão dos dados em torno da média populacional, dado pela raiz quadrada da variância da população
6	Descrição	<p>Seguir os seguintes passos:</p> <p>1) Converter o formato dos dados de hh:mm:ss para número decimal.</p> <p>2) Aplicar a seguinte fórmula ao conjunto de dados.</p> $\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$ <p>sendo: σ: desvio padrão populacional x_i: cada dado da população μ: média populacional N: número de dados</p> <p>3) Retornar o formato dos dados de número decimal para hh:mm:ss.</p>
7	Parâmetro	-
8	Tipo de valor	Hora
9	Estrutura do valor	hh:mm:ss (horas:minutos:segundos)
10	Referência da medida	-
11	Exemplo	-
12	Identificador	1

Fonte: Adaptada de ET-CQDG (BRASIL, 2016)

É importante destacar que, antes do cálculo tanto da mediana quanto do desvio padrão, é necessário converter o formato dos dados de "hh:mm:ss" para "número decimal". Após a conclusão dos cálculos, o dado deve ser retornado ao seu formato de origem (hh:mm:ss) para a exposição dos resultados.

Outro ponto que vale destaque diz respeito à análise de conjunto de dados registrados em dias subsequentes, isto é, na hipótese de ocorrência de sinistro de trânsito próximo à meia-noite (24:00:00 h), situação em que parte dos usuários podem acabar registrando o horário do sinistro de trânsito no dia seguinte. Este cenário será contemplado no exemplo a seguir.

Segue abaixo a aplicação da metodologia, apresentada nesta etapa do trabalho, a um exemplo hipotético de um conjunto de dados referentes a registros múltiplos associados a um mesmo sinistro de trânsito (Quadro 9). Para a compreensão do critério de análise e filtragem de registros múltiplos de sinistro de trânsito para aferir se pertencem a um mesmo evento, vide item 4.5.

Quadro 9 - Registro fictício de dados referentes a data e hora do sinistro

ID do Usuário	Data e Hora do Sinistro de Trânsito
Usuário A	30/12/2023 às 00:02:45
Usuário B	29/12/2023 às 23:57:00
Usuário C	29/12/2023 às 23:50:30
Usuário D	29/12/2023 às 23:55:12

Com base nos valores apresentados no Quadro 9, seguem os resultados:

Valor validável do Atributo:

Conversão do formato dos dados:

$$00:02:45 = 24:02:45 = 24,045833$$

$$23:57:00 = 23:57:00 = 23,950000$$

$$23:50:30 = 23:50:30 = 23,841667$$

$$23:55:12 = 23:55:12 = 23,920000$$

Valor validável = Mediana = 23,935 = 23:56:06 do dia 29/12/2023

Desvio padrão populacional:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}} = \sqrt{\frac{0,021368}{4}} = 0,073089 = 00:04:23$$

4.5 CRITÉRIOS DE ANÁLISE E FILTRAGEM DE REGISTROS

Como exposto anteriormente, esta pesquisa apresenta uma proposta para o mapeamento colaborativo em que qualquer voluntário possa contribuir com o registro de sinistros de trânsito, independente de sua experiência na área. Em uma livre proposta de mapeamento colaborativo de sinistros de trânsito, espera-se que um mesmo sinistro de trânsito seja registrado mais de uma vez por diferentes usuários.

Tendo em vista a dinâmica proposta para produção de dados, em um primeiro momento foram definidos os critérios de análise e filtragem de registros múltiplos para aferir se pertencem a um mesmo evento. Em um segundo momento, estando determinado do conjunto de registros a ser explorado, foram especificados os critérios para a definição dos atributos válidos.

4.5.1 Critérios de Análise e Filtragem de Registros Múltiplos

Em uma ocasião de registros múltiplos, para aferir quais pertencem ao mesmo evento deve-se avaliar os valores dos atributos referentes à localização, data e horário do sinistro de trânsito, uma vez que a posição e o tempo são atributos fortemente associados. Ou seja, passa a ser necessário avaliar, simultaneamente, a localização geográfica e o horário dos registros voluntários.

Para tanto, o critério definido para análise dos atributos envolveu a determinação de uma dimensão (raio) e especificação de um intervalo de tempo máximos dentro dos quais diferentes registros serão considerados pertencentes ao mesmo evento. Não cumprindo ambas as condições, os registros serão considerados oriundos de eventos distintos.

i. Determinação da distância máxima (raio) entre registros:

Em ocorrências de sinistros de trânsito com vítima, os envolvidos devem aguardar no local do acidente até que sejam concluídos os trabalhos da polícia e da perícia. Caso o sinistro comprometa a segurança e fluidez do trânsito, somente o agente policial tem o poder de autorizar a remoção dos veículos envolvidos, mesmo que a perícia no local ainda não tenha sido realizada (BRASIL, 1973).

Nas hipóteses de sinistro de trânsito sem vítima, é indicado que os envolvidos solicitem a prestação de atendimento em unidades policiais de forma presencial ou procedam com o registro do sinistro de trânsito de forma virtual, quando disponível, dispensando-se o comparecimento de viaturas no local da ocorrência. Desta forma, para não comprometer o tráfego, os condutores envolvidos no sinistro de trânsito devem, por lei, remover o veículo do local do acidente, sob pena de cometer infração de trânsito (BRASIL, 1997). Recomenda-se que, antes de remover o veículo para um local seguro, os condutores registrem mídia do evento de forma célere, para acúmulo de provas que podem ser úteis posteriormente.

Em ambos os casos, presume-se que os veículos envolvidos no sinistro sejam deslocados o mínimo possível do local do sinistro para um ponto seguro, seja para a polícia dar andamento em seus trabalhos para a apuração dos fatos (sinistros com vítima), seja para os condutores procederem com diálogos para solução do impasse (sinistros sem vítima). Ou seja, após o deslocamento, é provável que o veículo ainda permaneça na mesma quadra onde o sinistro ocorreu.

Por esse motivo, a geometria das quadras é o principal indicador a ser considerado na determinação da distância máxima (raio) dentro da qual diferentes registros serão considerados pertencentes ao mesmo evento em uma livre proposta de mapeamento colaborativo de sinistros de trânsito.

Para o cálculo desse raio, foi levado em consideração a dimensão média das quadras de alguns dos principais Bairros do município de Uberlândia com histórico de grande número de ocorrência de sinistros de trânsito registrados pelos órgãos oficiais (BRASIL, 2023a).

Segue a média das dimensões das quadras dos Bairros (Tabela 13), calculadas com o auxílio da ferramenta de medição de distância do *Google Maps*.

Tabela 13 - Média das dimensões das quadras de Bairros do município de Uberlândia

Bairro	Dimensão Média (m)	Bairro	Dimensão Média (m)
Centro	135	Saraiva	103
Nossa Sra. Aparecida	117	Oswaldo Rezende	113
Santa Mônica	138	Brasil	110
Martins	117	Presidente Roosevelt	125
Tibery	140	Tabajaras	100
Média Geral (m)			120

Como exposto, o resultado encontrado para a média geral das dimensões das quadras foi de 120 metros. Desta forma, esta pesquisa adotará o raio de 120 metros como a dimensão máxima dentro da qual diferentes registros serão considerados pertencentes ao mesmo evento em uma livre proposta de mapeamento colaborativo de sinistros de trânsito.

ii. Especificação do intervalo de tempo máximo entre registros:

Embora não haja legislação vigente que defina um limite de tempo para a liberação do trânsito após um acidente, o Código de Trânsito Brasileiro (BRASIL, 1997) estabelece penalidades e medidas administrativas aos condutores envolvidos em sinistros de trânsito que não adotarem as devidas providências para remover o veículo do local do acidente.

Em casos de sinistros de trânsito sem vítimas, a pronta remoção do veículo é imprescindível para assegurar a segurança e a fluidez do trânsito, devendo ocorrer tão logo seja possível.

Já para os casos de sinistros com vítima, o CTB imputa ao condutor a obrigação de preservar o local do acidente para facilitar os trabalhos da polícia e da perícia, devendo adotar as providências para remover o veículo do local somente quando determinadas por policial ou agente da autoridade de trânsito.

De forma geral, é comum que os condutores envolvidos em um sinistro de trânsito sem vítima demandem algum tempo para efetuarem os diálogos preliminares, registrarem provas e procederem com as medidas necessárias para retirar o veículo da via quando o veículo não possuir condições de transitar. Esse tempo passa a ser ainda maior quando o sinistro de trânsito culmina em vítimas, uma vez que a ocorrência demanda a sinalização do local do acidente, serviços de atendimento médico pré-hospitalar, relatório policial e perícia, além do posterior processo de desobstrução de pista. Enquanto todos esses procedimentos ocorrem, os elementos que caracterizam o sinistro se mantêm, permitindo que qualquer testemunha retrate o acidente. Logo, o evento balizador para a especificação do tempo máximo para a identificação de um acidente são os casos de sinistros de trânsito com vítima.

No presente momento, tramita na Câmara dos Deputados o Projeto de Lei nº 1827/2023 (GAMBALE, 2023), cuja ementa determina o tempo máximo para a liberação das pistas após acidentes com vítimas nos casos em que a fluidez do trânsito fique prejudicada, e fixa outras providências. De acordo com a proposta, o prazo máximo estabelecido para a liberação das pistas será de até 3 horas em vias urbanas, e de até 5 horas em vias rurais. Se o acidente envolver mais de seis veículos, o prazo poderá ser dilatado em 2 horas. O texto também prevê a não aplicação dos prazos em casos de acidentes envolvendo dez ou mais veículos.

O conteúdo disposto no Projeto de Lei possui coerência. Tendo ciência de que são raros os casos de sinistros de trânsito envolvendo mais de seis veículos em vias urbanas, esta pesquisa adotará o prazo de 3 horas como o intervalo de tempo máximo dentro do qual diferentes registros serão considerados pertencentes ao mesmo evento em uma livre proposta de mapeamento colaborativo de sinistros de trânsito.

Três horas é, de fato, um prazo razoável, acima do qual a maioria dos cenários mais graves de sinistros de trânsito com vítima em vias urbanas provavelmente já

seriam totalmente alterados, não permitindo a identificação de elementos que caracterizem um acidente.

Por fim, fica definido que, para concluir que diferentes registros se referem a uma mesma ocorrência de sinistro de trânsito, os mesmos devem atender simultaneamente os seguintes critérios: as posições estarem contidas em um raio máximo de 120 metros, e o período entre o primeiro e o último registro não ser superior a 3 horas.

4.5.2 Critério para Definição dos Atributos Válidos

i. Medida de tendência central:

Para grupos de registros colaborativos que se referirem a um mesmo evento, serão utilizadas medidas de tendência central para retornar os valores válidos de cada atributo (Quadro 10).

Quadro 10 - Medida de tendência central para a especificação do valor válido de um grupo de dados

Atributo	Medida de Tendência Central
Localização	Mediana
Data e hora	Mediana
Demais atributos	Moda

Referente à escolha da mediana como medida de tendência central para os atributos de posição e de tempo: a mediana reduz a influência de valores discrepantes (*outliers*) na medida de tendência central de um conjunto de dados.

Referente à escolha da moda como medida de tendência central para os atributos cujos valores são predefinidos: a moda retorna o valor mais frequente de um grupo de valores, isto é, o valor de maior ocorrência dentre um conjunto de dados observados.

ii. Atribuição de pesos por nível de usuário:

A estrutura de formulários proposta para a coleta de dados colaborativos de sinistros de trânsito considera níveis distintos de usuários. A cada usuário cadastrado

na plataforma é associado um grau de conhecimento, sendo progressivo na seguinte ordem: Popular, Técnico e Especialista.

É esperado que, quanto maior o nível de experiência do usuário, mais assertivo seja os dados por eles coletados e, conseqüentemente, maior será a qualidade da informação geográfica produzida. Desta forma, a depender do nível do usuário, serão atribuídos pesos diferentes aos dados durante o processo de definição dos atributos válidos para grupos de registros colaborativos que se referirem a um mesmo evento. Neste sentido, dados oriundos de usuário com nível técnico serão considerados em dobro, e dados oriundos de usuário com nível de especialista serão considerados três vezes no conjunto de dados para o cálculo da medida de tendência central (Quadro 11).

Quadro 11 - Relação de peso dos dados registrados para cada nível de usuário

Nível de Usuário	Peso
Popular	1
Técnico	2
Especialista	3

Para os atributos em que a moda é a medida de tendência central que retorna os valores válidos: se após a aplicação dos critérios supracitados perdurar algum empate na quantidade de registros dos valores mais frequentes, o valor válido do atributo será aquele associado ao registro realizado pelo usuário com o maior nível de experiência. Persistindo o empate, o valor válido será aquele associado ao registro gerado pelo usuário com o maior tempo de cadastro na plataforma colaborativa.

A exceção ocorre no item do formulário "tipo de sinistro", momento em que pode haver dificuldades do usuário em classificar o tipo de sinistro testemunhado. Por esse motivo, se faz necessário formular um critério específico para a definição do valor válido do atributo "tipo de sinistro" em caso de registros múltiplos. Portanto, fica convencionado que o valor válido será aquele associado ao registro realizado pelo usuário com o maior nível de experiência, independente da quantidade de registros efetuados por usuários de níveis menores. Havendo somente um nível de usuário no conjunto de registros, valerá a regra geral (cálculo da moda).

4.5.3 Exemplo Prático do Método Proposto

Para melhor compreensão dos critérios de análise e filtragem de registros, segue abaixo o método aplicado a um exemplo prático. A Figura 31 exibe a representação gráfica de um total de cinco registros de um sinistro de trânsito fictício nas dependências do Campus Santa Mônica da Universidade Federal de Uberlândia, nas proximidades do cruzamento entre a Alameda das Amoras e a Alameda das Pitangueiras.

Figura 31 - Representação gráfica dos registros fictícios



Fonte: Elaboração própria

Para fins práticos, neste exemplo foi considerado uma versão resumida do formulário, porém com a quantidade suficiente de atributos que permita a demonstração do método proposto de análise e filtragem de registros. O Quadro 12 apresenta os valores dos atributos de cada registro.

Quadro 12 - Valores dos atributos de cada registro fictício

ID do Registro	ID do Usuário	Nível do Usuário	Localização (Coord. UTM)		Data e hora do Sinistro	Tipo de Sinistro	Condições do Tempo
			X (m)	Y (m)			
01	Usuário A	Técnico	788977,234	7906056,508	13/02/2024 às 20:32:45	Choque	Chuva
02	Usuário B	Técnico	788979,016	7906063,636	13/02/2024 às 20:37:02	Choque	Chuva
03	Usuário C	Popular	789008,797	7906056,389	13/02/2024 às 20:30:17	Tombamento	Chuva
04	Usuário D	Popular	788971,888	7906048,083	13/02/2024 às 21:45:33	Choque	Nublado
05	Usuário E	Popular	788958,441	7906058,452	13/02/2024 às 20:35:49	Choque	Chuva

A primeira etapa da análise consiste em aferir se todos os registros pertencem a um mesmo evento. De fato, todos os registros se referem ao mesmo sinistro de trânsito, uma vez que a distância máxima entre registros foi de 50 metros e o intervalo de tempo entre o primeiro e último registro foi de 01:15:16 (uma hora, quinze minutos e dezesseis segundos), valores menores que os limites estabelecidos (120 metros de raio e de 3 horas de período).

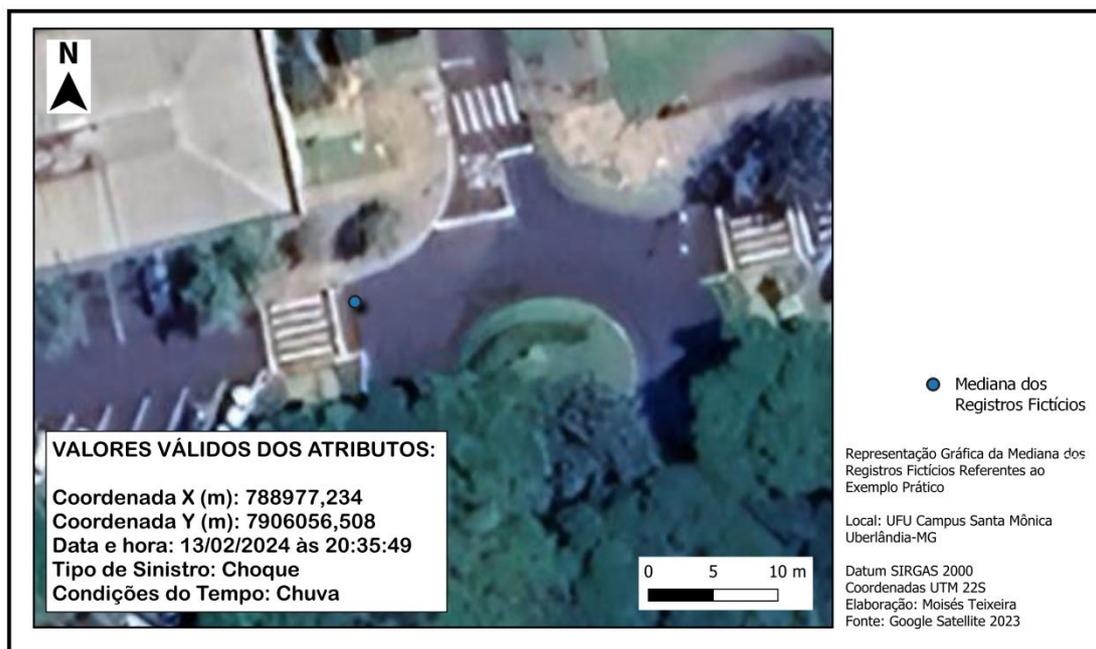
Na sequência, foi realizado o cálculo das medidas de tendência central para retornar os valores válidos de cada atributo. Uma vez que os usuários A e B pertencem ao nível técnico, seus dados foram duplicados na tabela. Para o atributo "tipo de sinistro", a moda foi calculada considerando somente os valores associados aos registros gerados pelos usuários com o maior nível de experiência, isto é, usuários A e B. O Quadro 13 exibe os resultados encontrados.

Quadro 13 - Cálculo dos valores válidos de cada atributo

ID do Registro	ID do Usuário	Nível do Usuário	Localização (Coord. UTM)		Data e hora do Sinistro	Tipo de Sinistro	Condições do Tempo
			X (m)	Y (m)			
01	Usuário A	Técnico	788977,234	7906056,508	13/02/2024 às 20:32:45	Choque	Chuva
			788977,234	7906056,508	13/02/2024 às 20:32:45		Chuva
02	Usuário B	Técnico	788979,016	7906063,636	13/02/2024 às 20:37:02	Choque	Chuva
			788979,016	7906063,636	13/02/2024 às 20:37:02		Chuva
03	Usuário C	Popular	789008,797	7906056,389	13/02/2024 às 20:30:17		Chuva
04	Usuário D	Popular	788971,888	7906048,083	13/02/2024 às 21:45:33		Nublado
05	Usuário E	Popular	788958,441	7906058,452	13/02/2024 às 20:35:49		Chuva
Valor Válido			788977,234	7906056,508	13/02/2024 às 20:35:49	Choque	Chuva

A Figura 32 mostra a representação gráfica da mediana dos registros fictícios do sinistro de trânsito, isto é, a posição válida para o conjunto de dados estudados.

Figura 32 - Representação gráfica da mediana dos registros fictícios



5 CONCLUSÕES E RECOMENDAÇÕES

Vários são os obstáculos encontrados pelas agências de mapeamento nacionais para manterem atualizadas suas bases de dados espaciais. Por esse motivo, o mapeamento colaborativo passou a ser visto como uma alternativa atrativa na coleta e compartilhamento de dados, estimulando estudos voltados a avaliar o potencial de integração dos dados voluntários ao mapeamento de referência. Porém o processo de compatibilização é desafiador, uma vez que a produção e distribuição de dados espaciais em plataformas de mapeamento colaborativo inevitavelmente gera resultados heterogêneos.

Há diversas áreas associadas à ciência de dados espaciais, elevando ainda mais a importância de estudos específicos. Buscando contribuir com essa lacuna, e por entender ser uma aplicação que pode se beneficiar da participação da sociedade civil, o presente trabalho concentrou seus estudos em sinistros de trânsito, um importante aspecto da mobilidade urbana.

Evidências empíricas sugerem haver uma sub-notificação dos registros de sinistros de trânsito junto aos órgãos competentes, principalmente em casos de ocorrências mais leves, uma vez que é comum as partes envolvidas optarem por não consumir o registro. Este fato prejudica o processo de quantificação e qualificação do fenômeno.

Diante disso, esta pesquisa avaliou o potencial do uso de técnicas de mapeamento colaborativo para a captação de dados espaciais referentes a ocorrência de sinistros de trânsito do município de Uberlândia/MG, de forma a contribuir com a base de dados oficial.

Além do esforço em manter a base de dados estatísticos atualizada, um dos maiores desafios dos órgãos competentes é estabelecer um modelo padrão de coleta de dados referentes a sinistros de trânsito, uma vez que os Boletins de Ocorrência não seguem um modelo estadual ou nacional para a coleta de dados de acidentes.

Tendo isso em vista, buscando elevar as condições que possam viabilizar a compatibilização de dados voluntário e oficial (alinhamento semântico), o formulário desenvolvido nesta pesquisa para a coleta colaborativa de dados de sinistros de trânsito foi concebido com base nas normas ABNT NBR 10697 (2020) e ABNT NBR

12898 (1993), e levando em consideração a configuração do formulário utilizado atualmente pelas agências oficiais de registro de ocorrências.

Para a definição dos atributos que iriam compor o formulário, seguiu-se a premissa de o registro colaborativo garantir ao usuário uma experiência sucinta, prática e intuitiva. Submetido a um processo de validação de conteúdo (análise de especialistas) ficou demonstrado que o produto final atingiu um grau adequado, sendo um roteiro validado para a realização de pesquisa de campo (pré-teste com uma amostra do público-alvo e posterior aplicação efetiva como ferramenta para o registro colaborativo de sinistros de trânsito).

Há de fato ganho de informação com a proposta de registro colaborativo. O formulário proposto acumula uma quantidade de informações considerada suficiente e com padrão adequado para contribuir com estudos de engenharia de tráfego e planejamento de trânsito, e estabelece um conjunto mínimo comum de informações que podem colaborar com a elaboração das análises estatísticas nas diferentes esferas de governo. A contribuição voluntária se mostra com potencial para complementar e atualizar dados com qualidade admissível, minimizando as sub-notificações de registro das ocorrências.

Apesar de algumas limitações, o aplicativo *web* selecionado nesta pesquisa para os testes de campo (Epicollect5) cumpriu seu objetivo, uma vez que a pesquisa avaliou a viabilidade da proposta, e não o desenvolvimento efetivo de um sistema para a coleta colaborativa, hospedagem e geoprocessamento de dados de sinistros de trânsito.

As plataformas de mapeamento colaborativo estão cada vez mais simples e intuitivas. Apesar disso, ainda é esperado haver equívocos no registro de dados por parte dos usuários, já que os sistemas permitem que qualquer pessoa contribua com a produção de informação, o que afeta na qualidade de seus dados.

Porém, vale ressaltar que a forma como o registro é feito atualmente pelas agências competentes também pode criar oportunidades que comprometam na qualidade do dado produzido. Como exemplo, o autor cita a Delegacia Virtual, funcionalidade disponibilizada por alguns Estados que traz grandes benefícios e praticidade ao cidadão. Entretanto, como a ferramenta permite que os próprios envolvidos procedam com o registro *online* do sinistro de trânsito sem vítima, é possível que o usuário cometa algum equívoco no preenchimento dos dados.

Tendo em vista que, em uma proposta de mapeamento colaborativo de sinistros de trânsito espera-se que um mesmo sinistro seja registrado mais de uma vez por diferentes usuários, um dos objetivos desta pesquisa buscou propor uma metodologia de avaliação extrínseca de qualidade de dados colaborativos referentes a sinistros de trânsito para o parâmetro de qualidade temporal.

Quanto à avaliação extrínseca da acurácia posicional, é evidente que os cálculos apontam um resultado positivo, levando em conta o baixo valor encontrado para o Erro Médio Quadrático (EMQ) da amostra. Na prática, os órgãos competentes registram a localização das ocorrências de sinistros de trânsito de forma aproximada, vinculando-a ao cruzamento mais próximo ou a um endereço de referência. Por esse motivo, o EMQ de 0,559 m se mostra irrelevante no impacto da precisão posicional da informação geográfica.

Na hipótese de registros múltiplos associados a um mesmo sinistro de trânsito, esta pesquisa buscou estabelecer os critérios de análise e filtragem dos registros e a definição dos atributos válidos. O método foi aplicado a um exemplo prático, demonstrando-se eficaz.

Em síntese, o autor conclui que os os objetivos desta pesquisa foram atingidos, uma vez que os resultados encontrados foram favoráveis, confirmando que a proposta apresentada é viável.

Como sugestão para trabalhos futuros, são citadas algumas lacunas identificadas ao longo da pesquisa e os próximos passos necessários à continuidade do projeto:

- Aprofundar estudos para inferir o potencial do usuário comum como avaliador de dados técnicos;
- Concluir o processo de validação de conteúdo (formulário para o registro colaborativo de sinistros de trânsito): pré-teste com uma amostra do público-alvo e estruturação do roteiro final validado, a ser aplicado em pesquisa de campo;
- Desenvolver um *software* para a coleta colaborativa, hospedagem e geoprocessamento de dados de sinistros de trânsito;
- Desenvolver algoritmos para a identificação de padrões nos registros, de forma a detectar automaticamente inconsistências nos mapas, além de auxiliar

no processo de análise e filtragem de registros abundantes e na definição dos valores válidos dos atributos;

- Aprofundar estudos sobre mecanismos de compensação e prática associada a gamificação para incentivar a contribuição das pessoas com a plataforma proposta.

REFERÊNCIAS BIBLIOGRÁFICAS

ABEGAZ, T.; BERHANE, Y.; WORKU, A.; ASSRAT, A.; ASSEFA, A. Road traffic deaths and injuries are under-reported in Ethiopia: a capture-recapture method. **PLOS ONE**, v. 9, n. 7, 2014. Disponível em: <https://doi.org/10.1371/journal.pone.0103001>. Acesso em: 13 jun. 2024.

ARIZA-LÓPEZ, F. J.; GONZÁLEZ, P. B.; PAU, J. M.; TORRES, A. Z.; PASCUAL, A. F. R.; VERGARA, G. M.; BALBOA, J. L. G. Calidad de datos geoespaciales (ISSO 19157-1): evoluciona o muere. **Revista Cartográfica**, n. 100, p. 129-154, 2020. Disponível em: <https://doi.org/10.35424/rcarto.i100.692>. Acesso em: 13 jun. 2024.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 6023**: informação e documentação: referências: elaboração: Versão Corrigida 2:2020. Rio de Janeiro: ABNT, 2018. 68 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 10520**: informação e documentação: citações em documentos: apresentação. Rio de Janeiro: ABNT, 2002. 7 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 10697**: pesquisa de sinistros de trânsito: terminologia. Rio de Janeiro: ABNT, 2020. 16 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 12898**: relatório de acidente de trânsito (RAT): procedimento. Rio de Janeiro: ABNT, 1993. 21 p.

BALLATORE, A. BERTOLOTTO, M. Semantically enriching VGI in support of implicit feedback analysis. *In*: TANAKA, K.; FRÖHLICH, P.; KIM, KS. Web and Wireless Geographical Information Systems: 10th International Symposium, W2GIS 2011. **Springer Berlin Heidelberg**, p. 78-93, 2011. Disponível em: https://doi.org/10.1007/978-3-642-19173-2_8. Acesso em: 04 jan. 2024.

BARRON, C.; NEIS, P.; ZIPF, A. Towards intrinsic quality analysis of OpenStreetMap datasets. Online proceedings of the international workshop on action and interaction in volunteered geographic information (ACTIVITY). **Agile**. 2013.

BIRKIN, L. J.; VASILEIOU, E.; STAGG, H. R. Citizen science in the time of COVID-19. **Thorax**, v. 76, n. 7, p. 636-637, 2021. Disponível em: <https://doi.org/10.1136/thoraxjnl-2020-216673>. Acesso em: 28 mai. 2024.

BORBA, R. L. R.; STRAUCH, J. C. M.; FILHO, A. S. L.; SOUZA, J. M. GeoStream-API: aplicação colaborativa para monitoramento e consumo de informação geográfica de mídia social. **Revista Brasileira de Cartografia**, v. 3, n. 69, p. 585-597, 2017. Disponível em: <https://doi.org/10.14393/rbcv69n3-44351>. Acesso em: 13 jun. 2024.

BRASIL. **Decreto nº 89.817, de 20 de junho de 1984**. Estabelece as Instruções Reguladoras das Normas Técnicas da Cartografia Nacional. Brasília, DF: Presidência da República, 1984. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/1980-1989/D89817.htm. Acesso em: 02 nov. 2022.

BRASIL. **Decreto-lei nº 243, de 28 de fevereiro de 1967**. Fixa as Diretrizes e Bases da Cartografia Brasileira e dá outras providências. Brasília, DF: Presidência da República, 1967. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto-lei/1965-1988/Del0243.htm. Acesso em: 27 abr. 2024.

BRASIL. **Lei nº 5.970, de 11 de dezembro de 1973**. Exclui da aplicação do disposto nos artigos 6º, inciso I, 64 e 169, do Código de Processo Penal, os casos de acidente de trânsito, e, dá outras providências. Brasília, DF: Presidência da República, 1973. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/1970-1979/L5970.htm. Acesso em: 02 mar. 2024.

BRASIL. **Lei nº 9.503, de 23 de setembro de 1997**. Institui o Código de Trânsito Brasileiro. Brasília, DF: Presidência da República, 1997. Disponível em:

http://www.planalto.gov.br/ccivil_03/leis/l9503compilado.htm. Acesso em: 15 set. 2023.

BRASIL. **Lei nº 13.614, de 11 de janeiro de 2018**. Cria o Plano Nacional de Redução de Mortes e Lesões no Trânsito (Pnatrans) e acrescenta dispositivo à Lei nº 9.503, de 23 de setembro de 1997 (Código de Trânsito Brasileiro), para dispor sobre regime de metas de redução de índice de mortos no trânsito por grupos de habitantes e de índice de mortos no trânsito por grupos de veículos. Brasília, DF: Presidência da República, 2018a. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13614.htm. Acesso em: 19 nov. 2023.

BRASIL. Ministério da Defesa. Exército Brasileiro. Departamento de Ciência e Tecnologia. Diretoria de Serviço Geográfico. **Norma da especificação técnica para aquisição de dados geoespaciais (ET-ADGV) versão 3.0**. Brasília, DF: MF, 2018b. Disponível em: https://bdgex.eb.mil.br/portal/media/adgv/ET-ADGV_3.0_211218.pdf. Acesso em: 27 out. 2022.

BRASIL. Ministério da Defesa. Exército Brasileiro. Departamento de Ciência e Tecnologia. Diretoria de Serviço Geográfico. **Norma da especificação técnica para controle de qualidade de dados geoespaciais (ET-CQDG)**. Brasília, DF: MF, 2016. Disponível em: https://bdgex.eb.mil.br/portal/media/cqdg/ET_CQDG_1a_edicao_2016.pdf. Acesso em: 06 mar. 2023.

BRASIL. Ministério da Defesa. Exército Brasileiro. Departamento de Ciência e Tecnologia. Diretoria de Serviço Geográfico. **Norma da especificação técnica para estruturação de dados geoespaciais vetoriais (ET-EDGV) versão 3.0**. Brasília, DF: MF, 2017. Disponível em: https://bdgex.eb.mil.br/portal/media/edgv/ET-EDGV-3_0_210518.pdf. Acesso em: 27 out. 2022.

BRASIL. Ministério da Infraestrutura. **Frota de Veículos - 2022**. Brasília, DF: Ministério da Infraestrutura, 02 mar. 2022. Disponível em:

<https://www.gov.br/infraestrutura/pt-br/assuntos/transito/conteudo-Senatran/frota-de-veiculos-2022>. Acesso em: 15 abr. 2023.

BRASIL. Ministério da Infraestrutura. **Resolução CONTRAN n° 808, de 15 de dezembro de 2020**. Dispõe sobre o Registro Nacional de Acidentes e Estatísticas de Trânsito (RENAEST). Brasília, DF: Diário Oficial da União, 2020.

BRASIL. Ministério dos Transportes. **Plano nacional de redução de mortes e lesões no trânsito (PNATRANS)**. Brasília, DF: Ministério dos Transportes, 2021. Disponível em: <https://www.gov.br/participamaisbrasil/revisao-pnatrans>. Acesso em: 20 nov. 2023.

BRASIL. Ministério dos Transportes. **Registro nacional de acidentes e estatísticas de trânsito: dados de acidentes de trânsito**. 2023a. Disponível em: <https://www.gov.br/transportes/pt-br/assuntos/transito/conteudo-Senatran/registro-nacional-de-acidentes-e-estatisticas-de-transito>. Acesso em 29 nov. 2023.

BRASIL. Ministério dos Transportes. **Registro nacional de acidentes e estatísticas de trânsito: painel geral**. 2023b. Disponível em: <https://www.gov.br/transportes/pt-br/assuntos/transito/arquivos-senatran/docs/renaest>. Acesso em: 16 jul. 2023.

BRASIL. Polícia Rodoviária Federal. **Dados abertos da PRF**. 2023c. Disponível em: <https://www.gov.br/prf/pt-br/aceso-a-informacao/dados-abertos/dados-abertos-da-prf>. Acesso em: 29 nov. 2023.

BRAVO, J. V. M.; SLUTER, C. R. O mapeamento colaborativo: seu surgimento, suas características e o funcionamento das plataformas. **Revista Brasileira de Geografia Física**, v. 11, n. 5, p. 1902-1916, 2018. Disponível em: <https://doi.org/10.26848/rbqf.v11.5.p1902-1916>. Acesso em: 23 out. 2022.

BRAVO, J. V. M.; SLUTER, C. R. O problema da qualidade de dados espaciais na era das informações geográficas voluntárias. **Boletim de Ciências Geodésicas**, v. 21,

n. 1, p. 56-73, 2015. Disponível em: <https://doi.org/10.1590/S1982-21702015000100005>. Acesso em: 17 out. 2022.

BROVELLI, M. A.; ZAMBONI, G. A new method for the assessment of spatial accuracy and completeness of OpenStreetMap building footprints. **ISPRS International Journal of Geo-Information**, v. 7, n. 8, 2018. Disponível em: <https://doi.org/10.3390/ijgi7080289>. Acesso em: 07 mai. 2023.

CAROTI, G.; PIEMONTE, A.; NESPOLI, R. **UAV-Borne photogrammetry: a low cost 3D surveying methodology for cartographic update**. In: MATEC Web of Conferences, vol. 120, pp. 09005. EDP Sciences, 2017. Disponível em: <https://doi.org/10.1051/mateconf/201712009005>. Acesso em: 13 jun. 2024.

CARVALHO, C. H. R.; GUEDES, E. P. **Balanco da primeira década de ação pela segurança no trânsito no Brasil e perspectivas para a segunda década**. Brasília, DF: Ipea, 2023. (Dirur: Nota Técnica, 42). Disponível em: https://repositorio.ipea.gov.br/bitstream/11058/12250/4/NT_42_Dirur_Balanco.pdf. Acesso em: 17 nov. 2023.

CINNAMON, J.; SCHUURMAN, N. Injury surveillance in low-resource settings using geospatial and social web technologies. **International Journal of Health Geographics**, v. 9, n. 25, 2010. Disponível em: <https://doi.org/10.1186/1476-072X-9-25>. Acesso em: 30 mai. 2024.

COLUCI, M. Z. O.; ALEXANDRE, N. M. C.; MILANI, D. Construção de instrumentos de medida na área da saúde. **Ciência & Saúde Coletiva**, v. 20, n. 3, p. 925-936, 2015. Disponível em: <https://doi.org/10.1590/1413-81232015203.04332013>. Acesso em: 10 fev. 2024.

COSTA, C. S.; FERNANDES, V. O.; JUNIOR, M. J. A. Aplicação de crowdsourcing na gestão e no planejamento de transportes: conceitos, potencialidades e parcerias do Waze. **Revista Brasileira de Geomática**, v. 3, n. 2, p. 68-80, 2015. Disponível em: <https://doi.org/10.3895/rbgeo.v3n2.5468>. Acesso em: 27 mai. 2024.

COSTA, Daniel Antônio Torno de Araújo. **A influência do policiamento e da fiscalização nos acidentes de trânsito**. 2016. Dissertação (Mestrado em Administração Pública) - EBAPE, Fundação Getúlio Vargas, Rio de Janeiro, 2016.

COSTA, D. C. S.; MORAES, R. M. Análise dos acidentes de trânsito utilizando técnicas de análise espacial. **Proceedings of Safety, Health and Environment World Congress**, v. 9, p. 81–84, 2009.

DAVIS, L. L. Instrument review: getting the most from a panel of experts. **Applied Nursing Research**, v. 5, n. 4, p. 194-197, 1992. Disponível em: [https://doi.org/10.1016/S0897-1897\(05\)80008-4](https://doi.org/10.1016/S0897-1897(05)80008-4). Acesso em: 13 fev. 2024.

DORN, H.; TÖRNROS, T.; ZIPF, A.. Quality evaluation of VGI using authoritative data - A comparison with land use data in Southern Germany. **ISPRS International Journal of Geo-Information**, v. 4, n. 3, p. 1657-1671, 2015. Disponível em: <https://doi.org/10.3390/ijgi4031657>. Acesso em: 15 nov. 2022.

ELIAS, E. N. N.; AMORIM, F. R.; SCHMIDT, M. A. R.; CAMBOIM, S. P. Data quality analysis tools and their use for the detection of spatio-temporal heterogeneities in VGI contributions. **GeoMundus**, 2020.

ELIAS, E. N. N.; AMORIM, F. R.; SCHMIDT, M. A. R.; CAMBOIM, S. P. Discovering spatiotemporal patterns on data quality assessment in collaborative mapping: a preliminary study in an area of Brazil. **Abstracts of the ICA**, v. 3, 2021a. Disponível em: <https://doi.org/10.5194/ica-abs-3-73-2021>. Acesso em: 13 jul. 2023.

ELIAS, E. N. N.; AMORIM, F. R.; SILVA, L. S.; SCHMIDT, M. A.; CAMBOIM, S. P.; FERNANDES, V. O. A proposal for a QGIS plugin for spatio-temporal analysis of OSM data quality: the case study for the city of Salvador, Brazil. *In*: MINGHINI, M.; LUDWING, C.; ANDERSON, J.; MOONEY, P.; GRINBERGER, A. Y. Proceedings of the Academic Track at the State of the Map 2021. **Zenodo**, p. 27-30, 2021b.

ELIAS, E. N. N.; FERNANDES, V. O. Qualidade dos dados geoespaciais do OpenStreetMap para os indicadores de acurácia posicional, acurácia temática e completude. **Geografia (Londrina)**, v. 30, n.2, p. 255-275, 2021.

ESTADO DE MINAS GERAIS. **Solicitar registro de ocorrência de acidente de trânsito sem vítima**. 2023. Disponível em: <https://www.mg.gov.br/servico/solicitar-registro-de-ocorrencia-de-acidente-de-transito-sem-vitima>. Acesso em: 29 nov. 2023.

FERREIRA, Denise Labrea; RIBEIRO, Lindioneza Adriano. **Acidentes de trânsito em Uberlândia: ensaios da epidemiologia e da geografia**. 1. ed. Uberlândia: Aline, 2006. 224 p. :il. Varios autores.

FLYNN, D. F. B.; GILMORE, M. M.; SUDDERTH, E. A. **Estimating traffic crash counts using crowdsourced data: pilot analysis of 2017 Waze data and police accident reports in Maryland**. Washington D.C.: United States Department of Transportation, 2018. Disponível em: <https://rosap.ntl.bts.gov/view/dot/37256>. Acesso em: 30 mai. 2024.

FURTADO, V.; AYRES, L.; OLIVEIRA, M.; VASCONCELOS, E.; CAMINHA, C.; D'ORLEANS, J.; BELCHIOR, M. Collective intelligence in law enforcement - The WikiCrimes system. **Information Sciences**, v. 180, n. 1, p. 4-17, 2010. Disponível em: <https://doi.org/10.1016/j.ins.2009.08.004>. Acesso em: 30 mai. 2024.

GAMBALE, R. **Projeto de Lei nº 1.827**. Acresce o artigo 301-A à Lei nº 9.503, de 23 de Setembro de 1997, que institui o Código de Trânsito Brasileiro, determinando o tempo máximo para a liberação das pistas após acidentes com vítimas e fixa outras providências. Brasília, DF: Câmara dos Deputados, 12 abr. 2023. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=235588> 1. Acesso em: 02 mar. 2024.

GIRRES, J-F.; TOUYA, G. Quality Assessment of the French OpenStreetMap dataset. **Transactions in GIS**, v. 14, n. 4, p. 435-459, 2010. Disponível em: <https://doi.org/10.1111/j.1467-9671.2010.01203.x>. Acesso em: 10 jul. 2023.

GÓMEZ-BARRÓN, J. P.; MANSO-CALLEJO, M. A.; ALCARRIA, R.; ITURRIOZ, T. Volunteered Geographic Information system design: project and participation guidelines. **ISPRS International Journal of Geo-Information**, v. 5, n. 7, 2016. Disponível em: <https://doi.org/10.3390/ijgi5070108>. Acesso em: 28 mai. 2024.

GOODCHILD, M. F. Citizens as sensors: the world of volunteered geography. **GeoJournal**, v. 69, n. 4, p. 211-221, 2007. Disponível em: <https://doi.org/10.1007/s10708-007-9111-y>. Acesso em: 16 mai. 2021.

GOODCHILD, M. F.; GLENNON, J. A. Crowdsourcing geographic information for disaster response: a research frontier. **International Journal of Digital Earth**, v. 3, n. 3, p. 231-241, 2010. Disponível em: <https://doi.org/10.1080/17538941003759255>. Acesso em: 30 mai. 2024.

GRANT, J. S.; DAVIS, L. L. Selection and use of content experts for instrument development. **Research in Nursing & Health**, v. 20, n. 3, p. 269-274, 1997. Disponível em: [https://doi.org/10.1002/\(SICI\)1098-240X\(199706\)20:3<269::AID-NUR9>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.0.CO;2-G). Acesso em: 13 fev. 2024.

GUILLEN, Luis Miguel Marquez. **Proposta metodológica de geocodificação para uso da estatística espacial na compreensão do fenômeno crime**. 2015. Dissertação (Mestrado em Ciências Geodésicas) - Programa de Pós-graduação em Ciências Geodésicas, Universidade Federal do Paraná, Curitiba, 2015.

HAKLAY, M.; BASIOUKA, S.; ANTONIOU, V.; ATHER, A. How many volunteers does it take to map an area well? The validity of Linus' law to volunteered geographic information. **The Cartographic Journal**, v. 47, n. 4, p. 315-322, 2010. Disponível em: <https://doi.org/10.1179/000870410X12911304958827>. Acesso em: 09 jul. 2023.

HAKLAY, M. How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. **Environment and Planning**

B: Planning and Design, v. 37, n. 4, p. 682–703, 2010. Disponível em: <https://doi.org/10.1068/b35097>. Acesso em: 10 mai. 2023.

HECHT, R.; KUNZE, C.; HAHMANN, S. Measuring completeness of building footprints in Openstreetmap over space and time. **ISPRS International Journal of Geo-Information**, v. 2, n. 4, p. 1066-1091, 2013. Disponível em: <https://doi.org/10.3390/ijgi2041066>. Acesso em: 09 nov. 2022.

HERNÁNDEZ-NIETO, R. A. **Contributions to statistical analysis**. Universidade de Los Andes: Mérida, 2002.

HIRATA, Eliane. **Proposta de um esquema conceitual para sistema dinâmico de mapeamento colaborativo de alagamentos e inundações na cidade de São Paulo**. 2013. Dissertação (Mestrado em Engenharia de Transportes) - Escola Politécnica, Universidade de São Paulo, São Paulo, 2013.

IBGE. INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. COORDENAÇÃO DE CARTOGRAFIA. **Acesso e uso de dados geoespaciais** (Manuais técnicos em geociências, ISSN 0103-9598; n. 14). Rio de Janeiro: IBGE, 2019, 143 p. ISBN 978-85-240-4508-0.

IBGE. INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, DIRETORIA DE GEOCIÊNCIAS, EXÉRCITO BRASILEIRO, DIRETORIA DE SERVIÇO GEOGRÁFICO. **Perfil de metadados geoespaciais do Brasil: perfil MGB 2.0**. Rio de Janeiro: IBGE, 2021, 106 p. ISBN 978-65-87201-63-4.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 19113**: geographic information - quality principles. Geneva, Switzerland, 2002.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 19114**: geographic information - quality evaluation procedures. Geneva, Switzerland, 2003.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 19157**: geographic information - data quality. Geneva, Switzerland, 2013.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/TS 19138**: geographic information - data quality measures. Geneva, Switzerland, 2006.

JASIM, M. S.; AL-HAMADANI, O. Positional accuracy assessment for updating authoritative geospatial datasets based on open source data and remotely sensed images. **Journal of Engineering**, v. 26, n. 2, p. 70-84, 2020. Disponível em: <https://doi.org/10.31026/j.eng.2020.02.06>. Acesso em: 07 mai. 2023.

JOHNSON, P. A.; SIEBER, R. E. Situating the adoption of VGI by government. In: **Crowdsourcing geographic knowledge: volunteered geographic information (VGI) in theory and practice**. Dordrecht: Springer Netherlands, p. 65-81, 2013. Disponível em: https://doi.org/10.1007/978-94-007-4587-2_5. Acesso em: 20 dez. 2023.

JUNIOR, W. M. S.; RIBEIRO, G. P. Qualidade dos dados geográficos disponibilizados em ambiente de sistema de informação geográfica. In: **IV SIMPÓSIO BRASILEIRO DE CIÊNCIAS GEODÉSICAS E TECNOLOGIAS DA GEOINFORMAÇÃO**, 2012, Recife. **Anais [...]**. Recife: UFPE, 2012.

LIKERT, Rensis. **A technique for the measurement of attitudes**. New York: Archives of Psychology, 1932.

LIU, S. B.; PALEN, L. The new cartographers: crisis map mashups and the emergence of neogeographic practice. **Cartography and Geographic Information Science**, v. 37, n. 1, p. 69-90, 2010. Disponível em: <https://doi.org/10.1559/152304010790588098>. Acesso em: 28 mai. 2024.

MACHADO, A. A.; CAMBOIM, S. P. Mapeamento colaborativo como fonte de dados para o planejamento urbano: desafios e potencialidades. **urbe. Revista**

Brasileira de Gestão Urbana, v. 11, 2019. Disponível em: <https://doi.org/10.1590/2175-3369.011.e20180142>. Acesso em: 24 out. 2022.

MACHADO, Adriana Alexandria. **Compatibilização semântica entre o modelo de dados do OpenStreetMap e a especificação técnica para estruturação de dados geoespaciais vetoriais (ET-EDGV)**. 2020. Tese (Doutorado em Ciências Geodésicas) - Programa de Pós-graduação em Ciências Geodésicas, Universidade Federal do Paraná, Curitiba, 2020.

MEDEIROS, Marcílio Sandro de. **Plataforma digital colaborativa para prevenção da violência no trânsito: perspectiva de uma infovigilância em saúde no Amazonas**. Manaus: Reggo/Edua, 2022. Disponível em: <https://doi.org/10.29327/5244358>. Acesso em: 21 mai. 2024.

MINAS GERAIS. **Decreto nº 43.778, de 12 de abril de 2004**. Institui o Sistema Integrado de Defesa Social de que trata o inciso I do art. 2º, da Lei Delegada nº 56, de 29 de janeiro de 2003, no âmbito da Secretaria de Estado de Defesa Social. Belo Horizonte, MG: Assembléia Legislativa de Minas Gerais, 2004. Disponível em: <https://www.almg.gov.br/legislacao-mineira/texto/DEC/43778/2004/>. Acesso em: 15 jan. 2024.

MINAS GERAIS. **Decreto nº 48.355, de 24 de janeiro de 2022**. Dispõe sobre o Sistema Integrado de Segurança Pública, institui a Base Integrada de Segurança Pública e dá outras providências. Belo Horizonte, MG: Assembléia Legislativa de Minas Gerais, 2022. Disponível em: <https://www.almg.gov.br/legislacao-mineira/texto/DEC/48355/2022/>. Acesso em: 15 jan. 2024.

MINGHINI, M.; FRASSINELLI, F. OpenStreetMap history for intrinsic quality assessment: is OSM up-to-date?. **Open Geospatial Data, Software and Standards**, v. 4, n. 9, 2019. Disponível em: <https://doi.org/10.1186/s40965-019-0067-x>. Acesso em: 25 out. 2022.

OLIVEIRA, D. F.; FERNANDES, V. O.; NAIM ELIAS, E. N.; MARTINS, V. E.; ALIXANDRINI JUNIOR, M. J. Avaliação intrínseca dos dados do OpenStreetMap e os aspectos espaço-temporais no estudo de caso do rompimento da barragem de Brumadinho. **Geosciences = Geociências**, v. 41, n. 3, p. 797-810, 2022. Disponível em: <https://doi.org/10.5016/geociencias.v41i3.16786>. Acesso em: 19 abr. 2023.

OLIVEIRA, D. V.; BRITO, J. L. S. Avaliação da acurácia posicional de dados gerados por aeronave remotamente pilotada. **Revista Brasileira de Cartografia**, v. 71, n. 4, p. 934-959, 2019. Disponível em: <https://doi.org/10.14393/rbcv71n4-50086>. Acesso em: 01 set. 2022.

ONU. Organização das Nações Unidas. **Resolução A/RES/64/255, de 2 de março de 2010**. ONU, 2010.

OPAS/OMS. **PAHO**, 2021. OMS lança década de ação pela segurança no trânsito 2021-2030. Disponível em: <https://www.paho.org/pt/noticias/28-10-2021-oms-lanca-decada-acao-pela-seguranca-no-transito-2021-2030>. Acesso em: 01 set. 2022.

OPENSTREETMAP. Disponível em: <https://www.openstreetmap.org>. Acesso em: 17 abr. 2023.

OSMSTATS. **Statistics of the free wiki world map**. Disponível em: <https://osmstats.neis-one.org>. Acesso em: 17 out. 2023.

OSM WIKI. OpenStreetMap Wiki. **Elements**. Disponível em: <https://wiki.openstreetmap.org/wiki/Elements>. Acesso em: 26 mai. 2023.

OSM WIKI. OpenStreetMap Wiki. **Map features**. Disponível em: https://wiki.openstreetmap.org/wiki/Map_features. Acesso em: 01 jun. 2023.

OSM WIKI. OpenStreetMap Wiki. **Wiki**. Disponível em: <https://wiki.openstreetmap.org/wiki/Wiki>. Acesso em: 09 jun. 2023.

OVELGONNE, M.; SONNENBICHLER, A. C.; GEYER-SCHULZ, A. Social emergency alert service - A location-based privacy - Aware personal safety service. *In: 2010 FOURTH INTERNATIONAL CONFERENCE ON NEXT GENERATION MOBILE APPLICATIONS, SERVICES AND TECHNOLOGIES*, p. 84-89, 2010, Amman, Jordan. Disponível em: <https://doi.org/10.1109/NGMAST.2010.27>. Acesso em: 30 mai. 2024.

PAIVA, Caio dos Anjos Paiva. **Inferência da qualidade de dados VGI a partir de métricas intrínsecas**. 2021. Tese (Doutorado em Ciências Geodésicas) - Programa de Pós-graduação em Ciências Geodésicas, Universidade Federal do Paraná, Curitiba, 2021.

PASQUALI, L. Princípios de elaboração de escalas psicológicas. **Revista de Psiquiatria Clínica**, v. 25, n. 5, p. 206-213, 1998.

PEREIRA, R. H. M. et al. Tendências e desigualdades da mobilidade urbana no Brasil I: O uso do transporte coletivo e individual, Texto para Discussão, **Instituto de Pesquisa Econômica Aplicada (IPEA)**, Brasília, 2021. Disponível em: <https://doi.org/10.38116/td2673>. Acesso em: 16 mai. 2021.

PIZZOLOTTO, R. Tracing the scientific trajectory of volunteered cartography: the case of OpenStreetMap. **ISPRS International Journal of Geo-Information**, v. 11, n. 7, 2022. Disponível em: <https://doi.org/10.3390/ijgi11070410>. Acesso em: 02 nov. 2022.

POLIT, D. F.; BECK, C. T. The content validity index: are you sure you know what's being reported? Critique and recommendations. **Research in Nursing & Health**, v. 29, n. 5, p. 489-497, 2006. Disponível em: <https://doi.org/10.1002/nur.20147>. Acesso em: 13 fev. 2024.

POSER, K.; DRANSCH, D. Volunteered Geographic Information for disaster management with application to rapid flood damage estimation. **Geomatica**, v. 64, n. 1, p. 89-98, 2010.

ROCHA, A. L. C.; VIEIRA, V.; PRAZERES, C. V. S.; SANTOS, C. A. S. YouOnAlert: um sistema para alertar cidadãos comuns acerca de problemas do cotidiano das cidades. *In: SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO UBÍQUA E PERVASIVA (SBCUP), 5., 2013, Maceió. Anais [...].* Porto Alegre: Sociedade Brasileira de Computação, 2013 . p. 2002-2011.

ROCHA, Marcos de Meneses. **Análise espacial de acidentes de trânsito no contexto de variáveis agregadas em áreas: proposta metodológica e aplicação na cidade do Rio de Janeiro.** 2015. Tese (Doutorado em Engenharia de Transportes) - COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015.

ROMÃO, M. N. P. V.; CAMPOS, C. I. Análise comparativa dos bancos de dados disponíveis no Brasil sobre vítimas fatais em acidentes de trânsito. *In: CONGRESSO BRASILEIRO DE TRANSPORTE E TRÂNSITO, 18., 2011, Rio de Janeiro.*

RUBIO, D. M.; BERG-WEGER, M.; TEBB, S. S.; LEE, E. S.; RAUCH, S. Objectifying content validity: conducting a content validity study in social work research. **Social Work Research**, v. 27, n. 2, p. 94-104, 2003. Disponível em: <https://doi.org/10.1093/swr/27.2.94>. Acesso em: 13 fev. 2024.

SANTOS, Alex da Silva. **Geoestatística aplicada na avaliação da qualidade de dados geoespaciais.** 2013. Dissertação (Mestrado) - Universidade Federal de Viçosa, Viçosa, 2013.

SANTOS, S. R.; DAVIS JR., C. A.; SMARZARO, R. Analyzing traffic accidents based on the integration of official and crowdsourced data. **Journal of Information and Data Management**, v. 8, n. 1, p. 67-82, 2017.

SANTOS, T.; FREIRE, S.; FONSECA, A.; TENEDÓRIO, J. A. Producing a building change map for urban management. **EARSeL EProceedings**, v. 10, n. 1, p. 56-65, 2011.

SHAH, S.; BAO, F.; LU, C.; CHEN, I. CROWDSAFE: crowd sourcing of crime incidents and safe routing on mobile devices. *In: GIS'11; PROCEEDINGS OF THE 19TH ACM SIGSPATIAL INTERNATIONAL CONFERENCE ON ADVANCES IN GEOGRAPHIC INFORMATION SYSTEMS*, p. 521-524, 2011, Chicago, Illinois. Disponível em: <https://doi.org/10.1145/2093973.2094064>. Acesso em: 30 mai. 2024.

SILVA, L. O.; GARREFA, F. As dinâmicas de fluxos do comércio atacadista de autosserviço e suas implicações na estruturação da periferia urbana de Uberlândia-MG. **Caminhos de Geografia**, v. 21, n. 75, p. 264-283, 2020. Disponível em: <https://doi.org/10.14393/RCG217552997>. Acesso em: 24 jul. 2023.

SILVA, L. S. L.; CAMBOIM, S. P. Authoritative cartography in Brazil and collaborative mapping platforms: challenges and proposals for data integration. **Boletim de Ciências Geodesicas**, 27(spe): e2021003, 2021. Disponível em: <https://doi.org/10.1590/s1982-21702021000100003>. Acesso em: 11 dez. 2023.

SILVA, L. S. L.; CAMBOIM, S. P. Brazilian nsdi ten years later: current overview, new challenges and propositions for national topographic mapping. **Boletim de Ciências Geodesicas**, 26(4): e2020018, 2020. Disponível em: <https://doi.org/10.1590/s1982-21702020000400018>. Acesso em: 16 mai. 2021.

SILVA, P. L.; SPERANDIO, V. G.; RODRIGUES, L. F.; DIAS, V.E.C.; LOTI, L.B.S.; LISBOA, M.H.M.; OLIVEIRA, G.D.; MIRANDA, G.H.B.; MARTINS, G.S.; MONTEIRO, C.R.; MENDES, V.F.; LISBOA-FILHO, J. Bases cartográficas para municípios de pequeno porte geradas por informação geográfica voluntária. **Revista Brasileira de Cartografia**, v. 70, n. 5, p. 1664-1685, 2018. Disponível em: <https://doi.org/10.14393/rbcv70n5-44572>. Acesso em: 20 nov. 2022.

SOUZA, G. A. Georreferenciamento de acidentes de trânsito: uma discussão metodológica. **ACTA Geográfica**, v. 5, n. 11, p. 31-40, 2011.

SPIELMAN, S. E. Spatial collective intelligence? Credibility, accuracy, and volunteered geographic information. **Cartography and Geographic Information Science**, v. 41, n. 2, p. 115-124, 2014. Disponível em: <https://doi.org/10.1080/15230406.2013.874200>. Acesso em: 13 jun. 2024.

TEIXEIRA, M. S.; SCHMIDT, M. A. R. Avaliação da acurácia posicional de dados colaborativos do OpenStreetMap: eixos viários de bairro no município de Uberlândia (MG). **Geosciences = Geociências**, v. 42, n. 2, p. 281-290, 2023. Disponível em: <https://doi.org/10.5016/geociencias.v42i2.17630>. Acesso em: 13 jun. 2024.

TORLIG, E. G. S.; RESENDE JUNIOR, P. C.; FUJIHARA, R. K.; MONTEZANO, L.; DEMO, G. Proposta de validação para instrumentos de pesquisa qualitativa (vali-quali). **Administração: Ensino e Pesquisa**, v. 23, n. 1, p. 5-31, 2022. Disponível em: <https://doi.org/10.13058/raep.2022.v23n1.2022>. Acesso em: 04 dez. 2023.

TORRES, T. B.; PEREIRA, B. M.; MACHADO, L. R. Panorama da prática de coleta, consolidação e uso de bancos de dados de acidentes de trânsito em cidades brasileiras. *In: CONGRESSO DE PESQUISA E ENSINO EM TRANSPORTE DA ANPET*, 34., 2020, 100% Digital.

TURNER, S. M.; MARTIN, M. W.; GRIFFIN, G. P.; LE, M.; DAS, S.; WANG, R.; DADASHOVA, B.; LI, X. **Exploring crowdsourced monitoring data for safety**. Blacksburg: Safety Through Disruption National University Transportation Center, 2020. Disponível em: <http://hdl.handle.net/10919/98795>. Acesso em: 30 mai. 2024.

UREÑA-CÁMARA, M. A.; NOGUERAS-ISO, J.; LACASTA, J.; ARIZA-LÓPEZ, F. J. A method for checking the quality of geographic metadata based on ISO 19157. **International Journal of Geographical Information Science**, v. 33, n. 1, p. 1-27, 2019. Disponível em: <https://doi.org/10.1080/13658816.2018.1515437>. Acesso em: 18 out. 2022.

WORLD HEALTH ORGANIZATION et al. **Global status report on road safety 2018: summary**. World Health Organization, 2018.

YAN, Y.; FENG, C. C.; HUANG, W.; FAN, H.; WANG, Y. C.; ZIPF, A. Volunteered geographic information research in the first decade: a narrative review of selected journal articles in GIScience. **International Journal of Geographical Information Science**, v. 34, n. 9, p. 1765-1791, 2020. Disponível em: <https://doi.org/10.1080/13658816.2020.1730848>. Acesso em: 20 abr. 2023.

YUNDUAN, L.; RUIMIN, L. Real-time traffic accidents post-impact prediction: based on crowdsourcing data. **Accident Analysis & Prevention**, v. 145, 2020. Disponível em: <https://doi.org/10.1016/j.aap.2020.105696>. Acesso em: 30 mai. 2024.

APÊNDICE A - Resultados Retornados Pelos Juizes na Avaliação dos Itens do Instrumento de Pesquisa

Item do Formulário	Critérios para Avaliação																							
	Clareza da Linguagem						Pertinência Prática						Aderência à Estrutura											
	O item está claro?						O item é relevante ao objetivo do formulário?						A distribuição do item pelos formulários (Popular, Técnico e Especialista) está adequada?											
	J1	J2	J3	J4	J5	J6	J7	J8	J1	J2	J3	J4	J5	J6	J7	J8	J1	J2	J3	J4	J5	J6	J7	J8
Instruções Iniciais	4	5	3	5	5	4	5	5	5	5	5	5	5	4	4	5	4	5	4	2	4	5	2	4
ID do Registro	5	5	3	5	5	4	5	5	5	5	4	5	5	5	5	5	3	5	3	2	5	4	4	5
ID do Usuário	3	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	4	2	5	5	5	5
Forma de Registro	5	5	5	2	5	5	5	5	5	5	3	4	5	5	5	5	5	5	5	2	5	5	5	5
Localização	5	4	5	5	4	5	5	5	5	5	5	2	5	5	5	5	5	4	5	2	5	5	5	5
Data	5	5	5	5	5	3	5	5	5	5	5	5	5	5	5	5	5	5	5	1	5	5	5	5
Horário	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	5	2	5	5	5	5
Tipo de Sinistro	4	3	4	3	5	5	3	5	5	5	5	5	5	5	5	5	2	5	3	1	5	5	5	4
Qualificação dos Envolvidos	5	5	2	5	4	5	5	5	5	5	4	3	5	5	5	5	3	5	3	2	4	5	5	4
Acidente com Vítima(s)?	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	4	2	3	5	5	4
Condições do Tempo	4	5	4	5	3	5	5	5	5	5	4	5	4	5	5	5	3	5	3	5	5	5	5	5
Condições do Local	4	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	3	5	5	2	3	5	5	4
Observações Complementares	5	5	5	1	5	5	5	5	5	5	5	1	4	5	5	5	5	5	5	1	5	5	5	5

APÊNDICE B - Resultados Retornados Pelos Juízes na Avaliação Geral do Instrumento de Pesquisa

Formulário para o Registro Colaborativo de Dados de Sinistros de Trânsito	O Formulário cumpre o objetivo proposto?								A quantidade de itens do Formulário (13 itens) está adequada?							
	J1	J2	J3	J4	J5	J6	J7	J8	J1	J2	J3	J4	J5	J6	J7	J8
	4	5	5	3	4	5	4	5	5	5	4	3	4	4	4	5