



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS GEODÉSICAS E
TECNOLOGIAS DA GEOINFORMAÇÃO

THIAGO ARRUDA SILVA

**IMPLEMENTAÇÃO DE PROCESSOS PARA VERIFICAÇÃO DA CONSISTÊNCIA
CONCEITUAL E DE DOMÍNIO CONFORME A ESPECIFICAÇÃO TÉCNICA PARA
CONTROLE DE QUALIDADE DE DADOS GEOESPACIAIS (ET-CQDG)**

Recife
2025

THIAGO ARRUDA SILVA

**IMPLEMENTAÇÃO DE PROCESSOS PARA VERIFICAÇÃO DA CONSISTÊNCIA
CONCEITUAL E DE DOMÍNIO CONFORME A ESPECIFICAÇÃO TÉCNICA PARA
CONTROLE DE QUALIDADE DE DADOS GEOESPACIAIS (ET-CQDG)**

Defesa de Mestrado apresentado ao Programa de Pós-Graduação em Ciências Geodésicas e Tecnologias da Geoinformação, do Departamento de Engenharia Cartográfica, da Escola de Engenharia de Pernambuco, do Centro de Tecnologia e Geociências, da Universidade Federal de Pernambuco.

Área de Concentração: Ciências Geodésicas e Tecnologias da Geoinformação

Orientadora: Prof^a. Dra. Simone Sayuri Sato

Coorientadora: Prof^a. Dra. Lígia Albuquerque de Alcântara Ferreira

Recife
2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Silva, Thiago Arruda.

Implementação de processos para verificação da consistência conceitual e de domínio conforme a especificação técnica para controle de qualidade de dados geoespaciais (ET-CQDG) / Thiago Arruda Silva. - Recife, 2025.

76f.: il.

Dissertação (Mestrado)- Universidade Federal de Pernambuco, Programa de Pós-Graduação em Ciências Geodésicas e Tecnologias da Geoinformação.

Orientação: Simone Sayuri Sato.

Coorientação: Lígia Albuquerque de Alcântara Ferreira.

1. Consistência lógica; 2. EDGV; 3. Automação; 4. QGIS; 5. PostGIS. I. Sato, Simone Sayuri. II. Ferreira, Lígia Albuquerque de Alcântara. III. Título.

UFPE-Biblioteca Central

THIAGO ARRUDA SILVA

IMPLEMENTAÇÃO DE PROCESSOS PARA VERIFICAÇÃO DA CONSISTÊNCIA CONCEITUAL E DE DOMÍNIO CONFORME A ESPECIFICAÇÃO TÉCNICA PARA CONTROLE DE QUALIDADE DE DADOS GEOESPACIAIS (ET-CQDG)

Defesa de Mestrado apresentado ao Programa de Pós-Graduação em Ciências Geodésicas e Tecnologias da Geoinformação, do Departamento de Engenharia Cartográfica, da Escola de Engenharia de Pernambuco, do Centro de Tecnologia e Geociências, da Universidade Federal de Pernambuco.

Projeto aprovado em: 29 de agosto de 2025.

BANCA EXAMINADORA

Prof^a. Dra. Simone Sayuri Sato (Orientadora)
Universidade Federal de Pernambuco

Prof^a. Dra. Andrea Lopes lescheck (Examinador Externo ao Programa)
Universidade Federal do Rio Grande do Sul

Prof. Dr. Fernando Luiz de Paula Santil (Examinador Externo ao Programa)
Universidade Federal de Uberlândia.

Dedico este trabalho à minha esposa, Rosy, à minha filha do coração, Rayanna, aos meus pais, Evandro e Angela, e aos meus filhos amados, Pedro e Cecília, que, juntos, sempre me apoiaram e são minha maior inspiração.

AGRADECIMENTOS

Primeiramente, agradeço a Deus, Criador dos céus e da Terra, pela vida, força e sabedoria concedidas ao longo desta caminhada. Em todos os momentos, especialmente nos mais desafiadores, senti Sua presença guiando meus passos e colocando pessoas especiais no meu caminho.

Aos meus pais e familiares, pelo carinho e por compreenderem os momentos de ausência durante este período, bem como pela força e apoio recebidos por meio de suas orações.

À minha orientadora, Profª. Dra. Simone Sayuri Sato, pela orientação, paciência e confiança ao longo desta pesquisa.

À minha coorientadora, Profª. Dra. Lígia Albuquerque de Alcantara Ferreira, pela dedicação e pelas valiosas contribuições, atuando também como minha orientadora durante grande parte desta jornada.

À Profª. Dra. Andrea Flávia Tenório Carneiro, pela confiança depositada em meu trabalho, pelas valiosas oportunidades concedidas e pelo incentivo constante. Sua orientação em diversas atividades de pesquisa e o apoio dedicado contribuíram de forma significativa para o meu crescimento acadêmico e profissional.

À Profª. Dra. Andrea Lopes lescheck, ao Prof. Dr. Fernando Luiz Paula Santil e ao Prof. Dr. José Luiz Portugal, pelas observações criteriosas e contribuições enriquecedoras apresentadas como membros da banca examinadora, que foram fundamentais para o aprimoramento deste trabalho.

Ao colega Joel, pelo auxílio fundamental no processo de ingresso no mestrado.

Aos meus colegas de curso do Programa de Pós-Graduação em Ciências Geodésicas e Tecnologias da Geoinformação, pela convivência, troca de conhecimentos e apoio mútuo ao longo dessa jornada. A amizade construída, o incentivo nos momentos de dificuldade e as experiências compartilhadas tornaram o percurso acadêmico mais leve e enriquecedor.

Aos colegas do 3º Centro de Geoinformação, pelo companheirismo e colaboração durante o desenvolvimento deste trabalho.

A todos que, de alguma forma, contribuíram para a realização deste trabalho, o meu sincero agradecimento.

RESUMO

A confiabilidade, interoperabilidade e aplicabilidade de dados geoespaciais estão condicionadas à qualidade desses dados e à estruturação do banco de dados que os armazena. Os dados geoespaciais são amplamente utilizados em diferentes contextos, como planejamento urbano, monitoramento ambiental, gestão territorial e análises espaciais baseadas em sistemas de informações geográficas. Manter a acurácia e eliminar erros são pré-requisitos essenciais para assegurar a qualidade das informações espaciais. Para isso, é necessário controlar os erros em cada etapa da manipulação dos dados, desde a modelagem até a preparação e análise. A Especificação Técnica para Controle de Qualidade de Dados Geoespaciais (ET-CQDG) estabelece diretrizes para a avaliação da qualidade de produtos cartográficos e de bancos de dados geoespaciais, por meio de cinco categorias: Completude, Consistência Lógica, Acurácia Posicional, Acurácia Temporal e Acurácia Temática. Esta pesquisa concentra-se na avaliação da consistência lógica, com ênfase nos subelementos consistência conceitual e de domínio, a fim de garantir a padronização e integridade da informação. A verificação da conformidade entre o modelo conceitual de referência e sua implementação física é essencial para a interoperabilidade dos dados. Neste contexto, o estudo propôs a automatização da verificação da consistência conceitual e de domínio de dados geoespaciais vetoriais, com base na ET-CQDG e na Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais (EDGV). Utilizaram-se ferramentas de software livre, como o QGIS e o banco de dados PostgreSQL/PostGIS. A metodologia envolveu a coleta e organização de bases públicas, o desenvolvimento de scripts automatizados para análise da conformidade de atributos e domínios, e a geração de relatórios sobre as inconsistências identificadas. A verificação automatizada permitiu identificar discrepâncias entre a estrutura conceitual prevista na EDGV e a implementação dos bancos analisados, considerando nomenclatura, tipos de atributos e coerência com os domínios predefinidos. Os resultados demonstram que a automação permitiu detectar inconsistências conceituais e estruturais, o que permitiu reduzir o tempo das análises manuais e trouxe confiabilidade na identificação de desvios em relação aos padrões normativos.

Palavras-chave: consistência lógica. EDGV. automação. QGIS. PostGIS.

ABSTRACT

The reliability, interoperability, and applicability of geospatial data are directly dependent on their quality and on the structure of the database in which they are stored. Geospatial data are widely used in diverse contexts, such as urban planning, environmental monitoring, territorial management, and spatial analysis within geographic information systems. Maintaining accuracy and eliminating errors are essential prerequisites to ensure the quality of spatial information. This requires controlling errors at every stage of data handling, from modeling to preparation and analysis. The *Especificação Técnica para Controle de Qualidade de Dados Geoespaciais* (ET-CQDG – Technical Specification for Geospatial Data Quality Control) establishes guidelines for assessing the quality of cartographic products and geospatial databases through five categories: Completeness, Logical Consistency, Positional Accuracy, Temporal Accuracy, and Thematic Accuracy. This research focuses on the evaluation of logical consistency, with an emphasis on the subelements of conceptual and domain consistency, to ensure standardization and information integrity. Verifying compliance between the reference conceptual model and its physical implementation is essential for data interoperability. In this context, the study proposes the automation of conceptual and domain consistency verification for vector geospatial data, based on the ET-CQDG and the *Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais* (ET-EDGV – Technical Specification for Structuring of Vector Geospatial Data). Open-source tools were used, such as QGIS and the PostgreSQL/PostGIS database. The methodology involved collecting and organizing public datasets, developing automated scripts to analyze attribute and domain conformity, and generating reports on the inconsistencies identified. The automated verification allowed detecting of discrepancies between the conceptual structure defined in the EDGV and the implementation in the analyzed databases, considering nomenclature, attribute types, and compliance with predefined domains. The results demonstrate that automation made it possible to identify conceptual and structural inconsistencies, reduce the time required for manual analyses, and improve reliability in detecting deviations from normative standards.

Keywords: logical consistency. EDGV. automation. QGIS. PostGIS.

LISTA DE FIGURAS

Figura 1 – Categorias e Elementos de Qualidade segundo a ISO 19157:2013.....	16
Figura 2 – Fluxo de avaliação da consistência lógica para os dados de um CDGV.	22
Figura 3 – Exemplos de inconsistências topológicas: sobreposição (a), lacunas (b) e conectividade (c).....	23
Figura 4 – Repositório oficial da Diretoria de Serviço Geográfico (DSG) no GitHub contendo as especificações EDGV.	38
Figura 5 – Fluxograma do Procedimento Metodológico.	39
Figura 6 – Integração de Ferramentas Geoespaciais.....	40
Figura 7 – Representação alternativa de domínio implementada por tabela de códigos e descrições.....	51
Figura 8 – Definição de domínio do tipo enumerado diretamente no modelo de dados.	52
Figura 9 – Ambiente QGIS com os scripts desenvolvidos para análise automatizada.	53
Figura 10 – Fluxo institucional de conferência e controle de qualidade de dados geoespaciais.	54

LISTA DE TABELAS

Tabela 1 – Componentes que descrevem uma medida de qualidade.	17
Tabela 2 – Exemplo de medida de qualidade adaptada.	18
Tabela 3 – Medida porcentagem de objetos com sobreposição inválida (CQDG: 215).	24
Tabela 4 – Medida porcentagem de objetos com sobreposição inválida (CQDG: 217).	25
Tabela 5 – Medida conflito de estrutura física.	27
Tabela 6 – Resumo dos principais tipos de domínios.	29
Tabela 7 – Detalhes da medida CQDG:201 (Consistência Conceitual).	30
Tabela 8 – Exemplos de erros de consistência conceitual.	31
Tabela 9 - Medida porcentagem de não conformidade com o domínio.	33
Tabela 10 – Exemplos de inconsistências de domínio em atributos textuais e numéricos.	35
Tabela 11 - Comparação de Atributos entre BancoX e BancoY.	42
Tabela 12 – Comparação de Atributos: EDGV 3.0 x Aracaju (CQDG 201)....	46
Tabela 13 – Resumo das Classes Avaliadas.	47
Tabela 14 – Exemplo de Verificação de Domínio (CQDG 204).	49

LISTA DE ABREVIATURAS E SIGLAS

CDGV – Conjuntos de Dados Geoespaciais Vetoriais

DSG – Diretoria de Serviço Geográfico

ET-CQDG – Especificação Técnica para Controle de Qualidade de Dados Geoespaciais

ET-EDGV – Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais

IBGE – Instituto Brasileiro de Geografia e Estatística

INDE – Infraestrutura Nacional de Dados Espaciais

ISO – International Organization for Standardization

OSM – OpenStreetMap

PostGIS – Extensão espacial para PostgreSQL

PyQGIS – Python para GIS

SIG – Sistema de Informação Geográfica

SHP – Shapefile

SQL – Structured Query Language

SUMÁRIO

1 INTRODUÇÃO	10
1.1 OBJETIVO GERAL	12
1.2 OBJETIVOS ESPECÍFICOS	12
1.3 ESTRUTURAÇÃO DA DISSERTAÇÃO	13
2. QUALIDADE E ESPECIFICAÇÕES TÉCNICAS PARA DADOS GEOESPACIAIS	14
2.1 QUALIDADE DOS DADOS GEOESPACIAIS	14
2.2 NORMAS E ESPECIFICAÇÕES TÉCNICAS PARA CONTROLE DE QUALIDADE DE DADOS GEOESPACIAIS	15
3. CONSISTÊNCIA LÓGICA	20
3.1 CONSISTÊNCIA TOPOLÓGICA	22
3.2 CONSISTÊNCIA DE FORMATO	26
3.3 CONSISTÊNCIA CONCEITUAL	28
3.4 CONSISTÊNCIA DE DOMÍNIO	32
4. PROCEDIMENTOS METODOLÓGICOS	36
4.1 ANÁLISE DE DADOS GEOESPACIAIS	36
4.2 DESENVOLVIMENTO E IMPLANTAÇÃO DE PROCESSOS AUTOMATIZADOS E PROCEDIMENTOS DE VERIFICAÇÃO	38
4.2.1 Desenvolvimento do Processo de Automatização das Verificações	39
4.2.2 Coleta e Procedimentos de Análise	42
4.2.2.1 Validação de Estruturas Conceituais	43
4.2.2.2 Análise de Consistência de Domínio	43
4.2.2.3 Automatização das Verificações	44
5 RESULTADOS E DISCUSSÕES	45
5.1. QUANTO A MEDIDA DE QUALIDADE CQDG 201	46
5.1.1 Comparação entre Modelagens (EDGV 3.0 versus outras versões)	46
5.1.2 Diferenças Estruturais e Conceituais	47
5.2 QUANTO A MEDIDA DE QUALIDADE CQDG 204	49
5.2.2 Verificação da consistência de domínio	51
5.3 CRIAÇÃO E VALIDAÇÃO DA APLICAÇÃO	52
5.3.1 Vantagens da processos automatizados	53
6 CONCLUSÕES E RECOMENDAÇÕES	56
REFERÊNCIAS	58
APÊNDICE A – CÓDIGO PYTHON: VERIFICAÇÃO CQDG 201	62
APÊNDICE B – CÓDIGO PYTHON: VERIFICAÇÃO CQDG 204	66
APÊNDICE C – RELATÓRIO DE VERIFICAÇÃO CQDG 201	71

1 INTRODUÇÃO

Os Sistemas de Informações Geográficas (SIG), integrados à análise de dados, têm contribuído para a governança pública e para a tomada de decisões, auxiliando no diagnóstico e prognóstico do território municipal com base em dados e informações públicas e acessíveis. Dessa forma, tornam-se instrumentos essenciais para a definição dos diversos planos que compõem o planejamento urbano. Conforme destacam Attah et al. (2024), essas tecnologias possibilitam a coleta, visualização e análise de dados espaciais, além de promoverem alocação otimizada de recursos, respostas mais eficazes a crises e melhor gestão de impactos ambientais. Ao facilitar o planejamento do uso do solo e o desenvolvimento de infraestrutura, os SIG contribuem para a construção de cidades inteligentes e sustentáveis. Ainda assim, os desafios relacionados à qualidade dos dados, limitações financeiras e técnicas e resistências institucionais limitam seu pleno aproveitamento. Assim, torna-se essencial investir em modelos colaborativos de governança e no avanço tecnológico para ampliar seu impacto.

Ademais, a crescente demanda por informações geoespaciais confiáveis é reflexo de sua relevância em diversas áreas, incluindo planejamento urbano, gestão ambiental, monitoramento de recursos naturais, análise de mudanças climáticas, entre outras. O avanço de tecnologias acessíveis de medição e o desenvolvimento de SIGs democratizaram o uso de dados geográficos, permitindo que mesmo usuários não especializados participem da geração e análise desses dados (Gutiérrez et al., 2020; Batista et al., 2019). No entanto, essa popularização de disponibilidade de acesso trouxe consigo desafios consideráveis relacionados à confiabilidade e à qualidade dos dados produzidos.

O controle de qualidade é um aspecto essencial na produção de dados geoespaciais. Estudos clássicos, como os de Galo e Camargo (1994), Sato (2003) e Santos (2010), já alertavam para a negligência na avaliação da qualidade em diversos níveis, como acurácia posicional, integridade geométrica, consistência temporal e atributos. Mais recentemente, o IBGE (2023) enfatiza que inconsistências geométricas e topológicas nas bases cartográficas podem comprometer diretamente a qualidade das análises espaciais e de mapeamentos temáticos, afetando a acurácia e a confiabilidade das decisões baseadas em geoinformação, seja em monitoramento ambiental, planejamento territorial ou gestão de recursos naturais. Ferreira et al.

(2018) corroboram essa preocupação ao identificar inconsistências topológicas em mapeamentos, que comprometeram a integridade de análises específicas.

Segundo Devendran e Lakshmanan (2014), a complexidade inerente aos dados geoespaciais agrava o desafio do controle de qualidade, uma vez que envolve a integração de informações provenientes de múltiplas fontes, escalas e níveis de detalhe. Esses fatores dificultam a padronização, aumentam as chances de inconsistências semânticas e estruturais e exigem o emprego de ferramentas especializadas capazes de verificar a acurácia e a confiabilidade em todas as etapas do processo. Tong et al. (2019) reforçam que um sistema automatizado para controle de qualidade é essencial, não apenas para detectar inconsistências, mas também para corrigi-las de forma eficiente.

A qualidade de dados geoespaciais envolve não apenas a conformidade com especificações técnicas, mas também sua capacidade de atender às necessidades dos usuários (Servigne et al., 2006). De acordo com Ariza-López (2011), essa qualidade é um fator diferenciador entre produtores, destacando a importância de normas e padrões robustos para garantir a confiabilidade dos dados. Nesse sentido, normas como a ISO 19157 (2013) e a Especificação Técnica para Controle de Qualidade de Dados Geoespaciais (ET-CQDG), de 2016, têm desempenhado papel fundamental na padronização de processos de validação.

A consistência lógica é um dos elementos mais críticos no controle de qualidade de dados geoespaciais. Dividida em quatro categorias principais – consistência conceitual, de domínio, de formato e topológica –, ela assegura que os dados sejam compatíveis com regras lógicas, estruturais e relacionais (ISO, 2013; França, 2021). A consistência conceitual, por exemplo, refere-se à aderência dos dados ao esquema conceitual definido, essencial para a organização e uso eficiente das informações (INSPIRE, 2014; Davis Jr. & Laender, 2000). Já a consistência topológica garante que as relações espaciais entre feições, como conectividade e adjacência, sejam preservadas, evitando erros que possam comprometer análises e aplicações (Câmara, 2005).

Hahmann et al. (2011) estimam que mais de 80% dos dados gerados no mundo possuem esse tipo de referência, o que evidencia a importância da acurácia e da consistência dos dados espaciais. No entanto, inconsistências em mapeamentos urbanos são frequentemente associadas à ausência de uma cultura institucional sólida de controle de qualidade. Essa fragilidade é observada por Elias (2021), que, ao

avaliar feições urbanas no OpenStreetMap (OSM), demonstra que falhas na qualidade geométrica e temática decorrem da carência de procedimentos formais e da descontinuidade institucional na verificação e validação dos dados colaborativos.

Diante desse cenário, este trabalho propõe a implementação de processos para a verificação da consistência lógica em dados geoespaciais vetoriais, fundamentado nas diretrizes da ET-CQDG. A proposta inclui a criação de scripts em Python, integrados ao ambiente QGIS, com o objetivo de identificar e documentar inconsistências conceituais e de domínio em bases estruturadas segundo diferentes versões da Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais (ET-EDGV).

A implementação de processos automatizados de verificação da consistência conceitual e de domínio, por meio de scripts em PyQGIS conformes à ET-CQDG e aplicados a bases modeladas segundo a ET-EDGV, tornam confiáveis o uso dos bancos geoespaciais, ao reduzir divergências estruturais e de domínio e ao padronizar a validação.

1.1 OBJETIVO GERAL

Desenvolver um método automatizado para controle de qualidade de conjuntos de dados geoespaciais vetoriais (CDGV), com foco na verificação da consistência lógica conceitual e de domínio, utilizando ferramentas de software livre.

1.2 OBJETIVOS ESPECÍFICOS

- a) Desenvolver e implementar um conjunto de scripts automatizados para verificar a conformidade dos atributos e domínios dos dados geoespaciais com as especificações da EDGV, visando a interoperabilidade das bases de dados; e
- b) Validar a implementação dos processos de verificação das medidas de qualidade da ET-CQDG, consistência conceitual e de domínio, usando dados geoespaciais disponíveis e confiáveis, a fim de identificar e documentar inconsistências, gerar automaticamente relatórios detalhados sobre os desvios encontrados.

1.3 ESTRUTURAÇÃO DA DISSERTAÇÃO

A presente dissertação está organizada em cinco capítulos, descritos a seguir:

Capítulo 1 – Introdução:

Apresenta o contexto do trabalho, a relevância do controle de qualidade em dados geoespaciais vetoriais, os objetivos geral e específicos, e a estruturação da dissertação.

Capítulo 2 – Fundamentação Teórica:

Discute os conceitos relacionados à qualidade de dados geoespaciais, com ênfase na consistência lógica conceitual e de domínio, e na aplicação das medidas 201 e 204 da ET-CQDG. Também aborda as principais normas e metodologias utilizadas no controle de qualidade.

Capítulo 3 – Metodologia:

Descreve o mapeamento dos processos de controle de qualidade, a elaboração de fluxogramas e os scripts desenvolvidos para avaliação das medidas de qualidade. O capítulo detalha as ferramentas utilizadas e os passos realizados no processo de análise.

Capítulo 4 – Resultados e Discussões:

Apresenta os resultados obtidos com a aplicação das medidas 201 e 204 da ET-CQDG em dados geoespaciais, destacando os erros identificados e a eficiência dos scripts utilizados.

Capítulo 5 – Conclusão:

Conclui o trabalho, destacando as contribuições para o campo do controle de qualidade de dados geoespaciais, as limitações do estudo e as possibilidades de trabalhos futuros.

2. QUALIDADE E ESPECIFICAÇÕES TÉCNICAS PARA DADOS GEOESPACIAIS

2.1 QUALIDADE DOS DADOS GEOESPACIAIS

A qualidade de dados geoespaciais é um dos pilares fundamentais para garantir a segurança de análises espaciais e a tomada de decisões em diversas áreas, como planejamento urbano, gestão ambiental, monitoramento de recursos naturais e produção cartográfica. A crescente disponibilidade de dados geoespaciais, oriundos de fontes variadas e muitas vezes gerados de forma colaborativa, reforça a necessidade de processos rigorosos de controle de qualidade. Segundo Bravo e Sluter (2015), a disseminação das informações geoespaciais proporcionada pela evolução tecnológica trouxe novos desafios para a avaliação da qualidade dos dados espaciais, especialmente devido à variabilidade das fontes e à crescente participação de usuários na produção de geoinformação. Esses fatores demandam métodos de validação eficientes que garantam a padronização e fidedignidade dos dados utilizados em diferentes aplicações.

De acordo com Teixeira et al. (2017), a avaliação da qualidade de dados geoespaciais deve seguir abordagens modernas que considerem tanto os avanços tecnológicos quanto às novas demandas por precisão e confiabilidade em diferentes aplicações. Esses autores destacam que as métricas de qualidade precisam ser adaptadas para atender às especificidades de cada tipo de dado geoespacial, sejam eles provenientes de sensores remotos, SIGs ou bancos de dados geoespaciais.

Em um contexto mais recente, Nascimento et al. (2022) exploraram o uso de aprendizado de máquina e modelagem estatística para a curadoria de dados geoespaciais, destacando que as novas tecnologias podem facilitar a análise de grandes coleções de dados, mas também trazem desafios relacionados à validação e integridade das informações. A aplicação de métodos automatizados, segundo os autores, pode ser um aliado no aumento da eficiência dos processos de avaliação de qualidade.

No Brasil, o uso de dados geoespaciais no Censo Demográfico 2022 trouxe à tona a importância da geoinformação no planejamento operacional e na coleta de dados censitários. Strauch e Azevedo (2024) ressaltam que as inovações tecnológicas no campo da geoinformação têm contribuído para o aumento da qualidade e

confiabilidade dos dados, especialmente na integração de dados geoespaciais com informações censitárias.

Damiani (2018), por sua vez, discute os desafios relacionados à qualidade de dados de localização, especialmente em aplicações móveis e ambientes internos. Segundo o autor, a especificação de métricas de qualidade e procedimentos padronizados de avaliação é essencial para lidar com erros posicionais, que podem comprometer a utilidade dos dados em análises mais precisas.

Essas contribuições contextualizadas pelos autores citados, demonstram a complexidade envolvida na garantia da qualidade dos dados geoespaciais, que abrange desde a completude, acurácia posicional, acurácia temática e a consistência lógica, esta última tratada nesta pesquisa.

Dessa forma, a avaliação da qualidade de dados geoespaciais é um elemento essencial para garantir sua confiabilidade e efetividade em diversas aplicações. A discussão e aplicação de normas, como a ET-CQDG e os métodos recentes sugeridos pela literatura, são fundamentais para aprimorar os processos de controle de qualidade e assegurar a utilidade prática dos dados.

2.2 NORMAS E ESPECIFICAÇÕES TÉCNICAS PARA CONTROLE DE QUALIDADE DE DADOS GEOESPACIAIS

A norma ISO 19157:2013 estabelece diretrizes para a descrição, medição e avaliação da qualidade de dados geoespaciais, sendo uma referência internacional amplamente reconhecida para a padronização da avaliação de qualidade em bases de dados geoespaciais. Ela define conceitos, categorias e elementos de qualidade que permitem uma avaliação estruturada e consistente das informações geográficas.

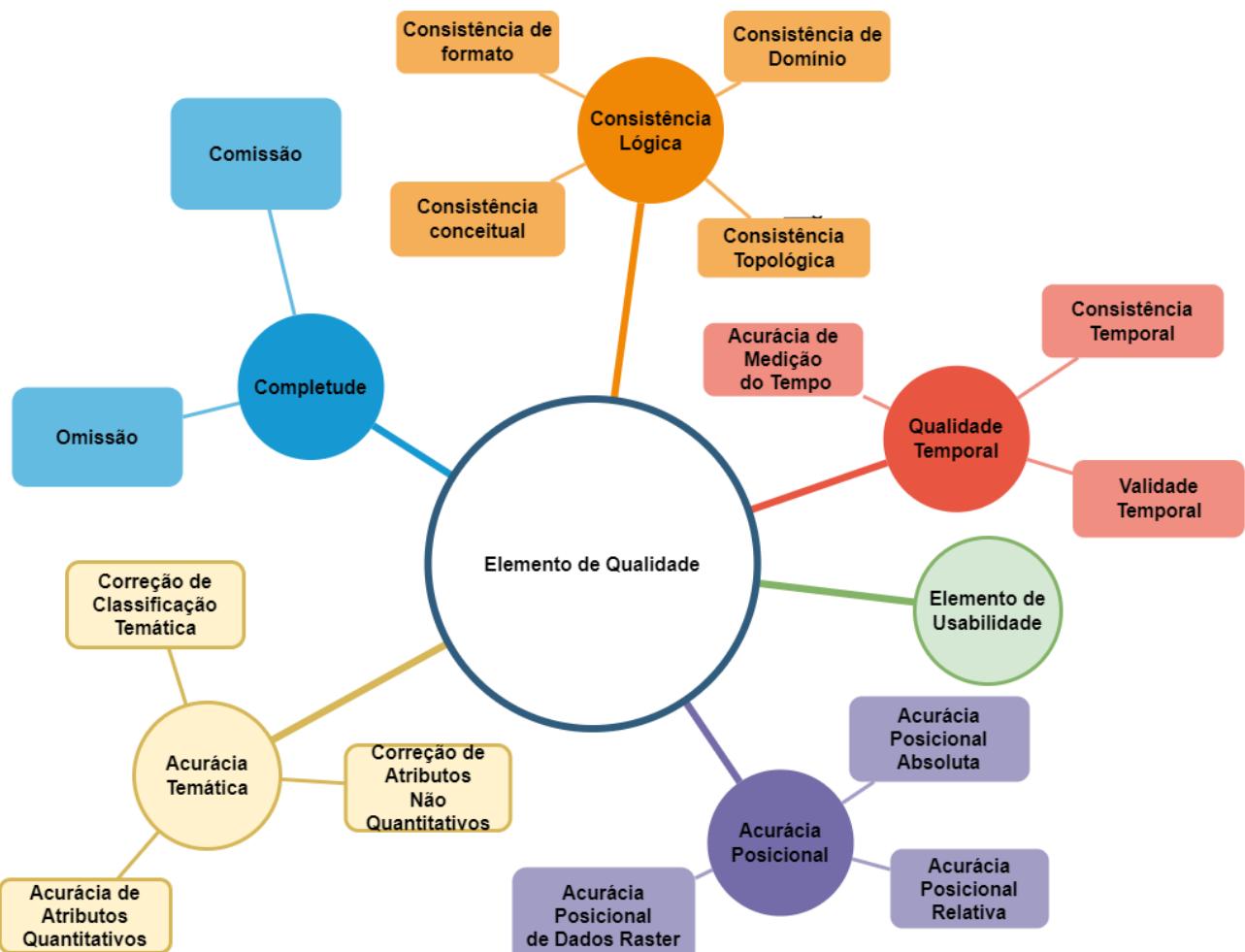
Segundo Devillers et al. (2007), critérios amplamente aceitos para avaliar a qualidade de dados espaciais incluem completude, consistência lógica, acurácia posicional, precisão temporal e precisão temática (ou acurácia temática). Esses critérios são mencionados como fundamentais para a mensuração de qualidade e foram posteriormente consolidados na norma ISO 19157:2013, cujo objetivo é garantir que os dados geoespaciais sejam confiáveis e adequados às suas finalidades, promovendo maior transparência.

Um dos aspectos mais destacados da norma é a introdução de um modelo de metadados de qualidade, que permite documentar informações detalhadas sobre

a avaliação da qualidade de um conjunto de dados. Segundo João e Smith (2019), o modelo de qualidade da ISO 19157 facilita a interoperabilidade entre sistemas e usuários, garantindo que os dados sejam compreendidos e utilizados de forma eficiente.

A Figura 1 apresenta uma síntese das categorias e elementos de qualidade definidos pela ISO 19157:2013. Para facilitar a leitura, adota-se a seguinte convenção: a circunferência indica uma categoria de qualidade e o retângulo indica um elemento de qualidade.

Figura 1 – Categorias e Elementos de Qualidade segundo a ISO 19157:2013.



Fonte: Adaptado ISO 19157 (2013).

A norma estabelece métodos para medir cada elemento de qualidade, como a utilização de amostras estatísticas para verificar a acurácia posicional e a aplicação de regras lógicas para validar a consistência de atributos e relações espaciais. Tais

procedimentos mostram-se eficazes em diferentes aplicações geoespaciais, como demonstrado por Ariza-López et al. (2022) no controle da qualidade de nuvens de pontos obtidas por varredura a laser terrestre, utilizadas na documentação e preservação de patrimônios ambientais e culturais.

Considerando a proposta deste estudo, os conceitos da ISO 19157 orientam a aplicação das medidas de qualidade 201 e 204 da ET-CQDG, que compartilham diretamente os elementos de consistência conceitual e de domínio definidos pela norma. Embora a ET-CQDG seja uma especificação nacional, ela reflete os fundamentos estabelecidos pela ISO, adotando as mesmas categorias de qualidade e adaptando-as ao contexto brasileiro.

A ET-CQDG, publicada pelo Diretório de Serviço Geográfico do Exército Brasileiro (DSG) em 2016, é uma norma nacional desenvolvida para avaliar e garantir a qualidade de dados geoespaciais no Brasil. Inspirada nos conceitos da ISO 19157:2013, a ET-CQDG adapta elementos internacionais às demandas operacionais e legais brasileiras, proporcionando diretrizes práticas para o controle de qualidade em bases de dados geoespaciais. A especificação tem sido amplamente utilizada em estudos recentes como base metodológica para avaliação de acurácia posicional e consistência dos dados geoespaciais (LIMA et al., 2022; TEIXEIRA; SCHMIDT, 2023), indicando o seu uso para fins público e privado.

Assim como a ISO 19157, a ET-CQDG organiza os elementos de qualidade em categorias amplas, apresentadas na Figura 1. Para fins de clareza, essas categorias podem ser descritas da seguinte forma: completude, consistência lógica, acurácia posicional, acurácia temática e qualidade temporal. A norma também detalha como as medidas de qualidade devem ser documentadas e aplicadas, sendo a Tabela 1 dedicada aos componentes que descrevem uma medida de qualidade e a Tabela 2 a um exemplo prático de sua implementação.

Tabela 1 – Componentes que descrevem uma medida de qualidade.

COMPONENTE	DESCRIÇÃO
Nome da medida	Identifica a medida de qualidade.

Objetivo	Descreve a finalidade da medida e o que ela avalia.
Parâmetro	Detalha os parâmetros necessários para a aplicação da medida.
Método de avaliação	Explica como a medida será aplicada e quais os critérios de análise.
Tipo de resultado	Indica o formato do resultado esperado (quantitativo ou qualitativo).

Fonte: Adaptado DSG (2016, p. 3-1).

Tabela 2 – Exemplo de medida de qualidade adaptada.

ELEMENTO DE QUALIDADE	NOME	OBJETIVO	MÉTODO DE AVALIAÇÃO	RESULTADO ESPERADO
Medida 201 Consistência Conceitual	Conformidade com o modelo de dados	Avaliar a conformidade com o esquema conceitual.	Verificação lógica baseada em regras semânticas.	Identificação de erros conceituais.
Medida 204 Consistência de Domínio	Porcentagem de não conformidade com o domínio	Validar valores atribuídos aos atributos.	Comparação com domínios válidos definidos previamente.	Relatório com valores inconsistentes.

Fonte: Adaptado DSG (2016).

A documentação detalhada e estruturada das medidas de qualidade na ET-CQDG permite a aplicação padronizada e a rastreabilidade dos processos de avaliação. As Tabelas apresentadas demonstram como as medidas são especificadas

e exemplificadas, auxiliando na implementação prática em sistemas de informações geográficas (SIG) e em bancos de dados geoespaciais.

3. CONSISTÊNCIA LÓGICA

A Consistência Lógica é uma das categorias de qualidade descritas na ISO 191157:2013 e adotadas pela ET-CQDG (2016). Essa categoria avalia a conformidade dos dados geoespaciais com regras lógicas, abordando aspectos essenciais relacionados à integridade estrutural e semântica. Garantir a consistência lógica é fundamental para assegurar que os dados representem adequadamente o mundo real e possam ser utilizados de forma confiável em SIG.

De acordo com Maranhão e Carneiro (2016), a consistência lógica desempenha um papel essencial na validação da integridade dos dados espaciais, evitando erros que possam resultar em interpretações equivocadas ou incompatibilidades entre diferentes bases de dados. A falta de consistência pode gerar problemas na interoperabilidade dos dados, dificultando sua integração em SIGs. Além disso, o aumento da disponibilidade de dados geoespaciais, muitas vezes provenientes de fontes colaborativas, reforça a necessidade de mecanismos eficazes para garantir a qualidade e a padronização das informações.

Complementando essa perspectiva, França e Portugal (2022) enfatizam que a consistência lógica, incluindo a consistência topológica, deve ser tratada como um dos pilares fundamentais na estruturação e validação de bases de dados geoespaciais. A ausência de consistência pode comprometer análises espaciais, principalmente em aplicações críticas como zoneamentos urbanos e cadastro técnico multifinalitário.

A validação da consistência lógica pode ser realizada de forma manual ou automatizada. Métodos manuais, apesar de precisos, são geralmente ineficientes quando aplicados a grandes volumes de dados. Nesse contexto, abordagens automatizadas têm se destacado como soluções viáveis para assegurar a integridade dos dados de maneira eficiente e escalável. Conforme proposto nesta pesquisa, a automação da validação da consistência lógica por meio de scripts desenvolvidos em PyQGIS permite reduzir significativamente o tempo de análise e aumentar a precisão na detecção de inconsistências estruturais e conceituais.

Garantir a consistência lógica nos conjuntos de dados geoespaciais é fundamental para assegurar que a representação do mundo real seja fidedigna e que os dados possam ser utilizados com segurança em aplicações críticas. A conformidade com normas e especificações técnicas, como a ISO 191157 e a ET-

CQDG, estabelece diretrizes claras para esse processo, promovendo a robustez, a interoperabilidade e a padronização dos dados espaciais.

De acordo com a ET-CQDG, a consistência lógica é composta por quatro elementos principais:

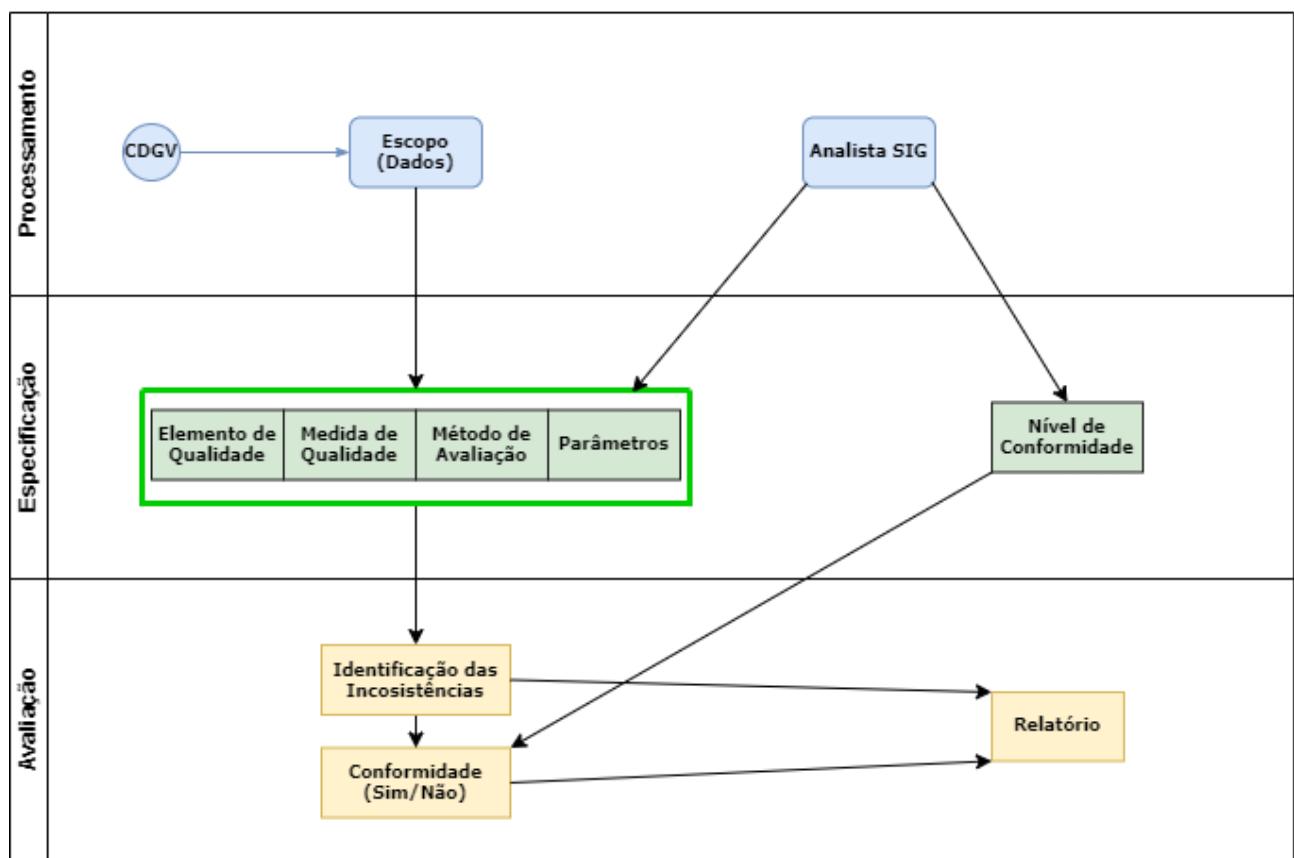
1. Consistência Conceitual: Verifica a aderência dos dados ao modelo conceitual definido, analisando a conformidade de classes, atributos e relacionamentos;
2. Consistência de Domínio: Avalia se os valores atribuídos aos atributos dos dados estão dentro dos limites ou domínios estabelecidos;
3. Consistência de Formato: Examina a compatibilidade entre o formato de armazenamento e o modelo físico do conjunto de dados; e
4. Consistência Topológica: Analisa as relações espaciais entre feições, identificando erros como sobreposições, lacunas e cruzamentos inadequados.

Os elementos da categoria de consistência lógica são avaliados por meio de medidas específicas de qualidade, como a CQDG:201 (Consistência Conceitual) e a CQDG:204 (Consistência de Domínio), previstas na ET-CQDG (2016) e fundamentadas nos princípios da ISO 19157:2013. Essas medidas visam assegurar que os dados estejam em conformidade com o modelo conceitual estabelecido e que os valores atribuídos aos atributos respeitem os domínios definidos. Conforme Meneses et al. (2023), a aplicação de indicadores de qualidade com base na ISO 19157 permite identificar lacunas e inconsistências que podem comprometer a integridade semântica dos dados geoespaciais.

De acordo com Silva (2023), a avaliação da consistência lógica em Infraestrutura de Dados Espaciais é fundamental para garantir que os padrões, normas e tecnologias empregados na infraestrutura possibilitem a interoperabilidade entre os dados e sua reutilização em diferentes aplicações

A Figura 2 ilustra o fluxo de avaliação da consistência lógica, destacando as etapas principais, como a definição de parâmetros, identificação de inconsistências e elaboração do relatório final. Este processo é fundamental para assegurar que os dados geoespaciais atendam aos requisitos de qualidade estabelecidos.

Figura 2 – Fluxo de avaliação da consistência lógica para os dados de um CDGV.



Fonte: Adaptado de França (2021, p. 90).

3.1 CONSISTÊNCIA TOPOLOGICA

A consistência topológica é um dos elementos da categoria Consistência Lógica, descrita na ET-CQDG (2016). Esse elemento avalia a conformidade das relações espaciais entre feições de um conjunto de dados geoespaciais, assegurando que regras como conectividade, adjacência e ausência de sobreposições ou lacunas indevidas sejam respeitadas.

Erros topológicos recorrentes em dados geoespaciais incluem sobreposição inadequada entre polígonos, lacunas entre feições contíguas, cruzamento de linhas sem conexão e conectividade imprecisa em redes geométricas. Segundo Lazorenko-Hevel et al. (2021), a consistência topológica pode ser verificada por meio da codificação explícita das relações espaciais entre feições, tanto do mesmo tipo quanto de tipos diferentes, sendo esta uma etapa fundamental na validação automatizada da

consistência lógica em bancos de dados topográficos digitais. Esses erros são ilustrados na Figura 3, que apresenta exemplos visuais de cada situação.

- a) Sobreposição inadequada entre polígonos: Quando polígonos que deveriam ser contíguos apresentam áreas sobrepostas. Esse erro é frequentemente causado por inconsistências no delineamento de fronteiras compartilhadas.
- b) Lacunas entre feições contíguas: Espaços vazios entre polígonos que deveriam compartilhar limites, resultando em falhas na cobertura espacial.
- c) Conectividade incorreta em redes geométricas: Segmentos que deveriam estar conectados, mas permanecem desconexos, afetando a integridade de redes viárias ou sistemas de drenagem.
- d) Cruzamento de linhas sem conexão: Linhas que se cruzam sem um nó de junção, comprometendo a conectividade lógica, particularmente em redes como estradas ou dutos.

A ET-CQDG especifica diversas medidas de qualidade para avaliar a consistência topológica, descrevendo métodos de avaliação e resultados esperados para cada tipo de análise. Por exemplo, a medida CQDG:215 verifica sobreposições entre polígonos que não deveriam existir, enquanto a medida CQDG:217 identifica lacunas em dados poligonais, como áreas não mapeadas entre feições ou desconexões em linhas.

Figura 3 – Exemplos de inconsistências topológicas: sobreposição (a), lacunas (b) e conectividade (c).



Fonte: Autor (2025).

A Figura 3 ilustra esses erros:

- Figura A: Representa uma sobreposição inadequada entre dois polígonos, erro identificado pela medida CQDG:215. Nesse caso, a avaliação busca identificar áreas em que feições da mesma classe se sobrepõem indevidamente.
- Figuras B e C: Mostram exemplos relacionados à medida CQDG:217, como uma lacuna em um conjunto de polígonos contíguos (Figura B) e desconexões em uma linha representando um rio (Figura C).

As Tabelas 3 e 4 exemplificam as medidas de qualidade CQDG:215 e CQDG:217 são apresentadas a seguir, descrevendo os parâmetros, métodos de avaliação e resultados esperados para cada caso.

Tabela 3 – Medida porcentagem de objetos com sobreposição inválida (CQDG: 215).

Linha	Componente	Descrição
1	Nome	Porcentagem de objetos com sobreposição inválida
2	Nome alternativo	-
3	Elemento de qualidade	Consistência topológica
4	Medida básica	Taxa de erro
5	Definição	Porcentagem de objetos do tipo polígono que possuem uma sobreposição inválida com outros objetos da mesma classe onde a sobreposição seja inválida
6	Descrição	Na região avaliada procura-se por qualquer sobreposição de áreas entre objetos da mesma classe onde tal fenômeno não seja previsto no modelo de dados. Em cada área inválida encontrada, soma-se um erro para cada objeto que participa da sobreposição. Ao final, divide-se o número de erros pelo número total de objetos na classe considerada. Esse procedimento pode ser aplicado a mais de uma classe por vez, somando-se o total de erros encontrados e dividindo pelo total de objetos nas classes avaliadas. O resultado é expresso como uma porcentagem.
7	Parâmetro	-

8	Tipo de valor	Real (porcentagem)
9	Estrutura do valor	-
10	Referência da medida	-

Fonte: Adaptado DSG (2016, p. 3-18).

Tabela 4 – Medida porcentagem de objetos com sobreposição inválida (CQDG: 217).

Linha	Componente	Descrição
1	Nome	Porcentagem de objetos que desrespeitam o raio de topologia
2	Nome alternativo	-
3	Elemento de qualidade	Consistência topológica
4	Medida básica	Taxa de erro
5	Definição	Porcentagem de objetos que desrespeitam o raio de topologia, ou seja, deveriam ser adjacentes, mas não são, em relação ao número total de objetos
6	Descrição	Considerando as regras do modelo de dados, conta-se como erro todo objeto que possui vértices a uma distância menor que o raio de topologia, mas não estão conectados
7	Parâmetro	Raio de topologia (real, em metros)
8	Tipo de valor	Real (porcentagem)
9	Estrutura do valor	-
10	Referência da medida	-

Fonte: Adaptado DSG (2016, p. 3-20).

Erros de consistência topológica podem comprometer análises geoespaciais, gerando resultados imprecisos. Em redes viárias, por exemplo, linhas desconectadas podem impedir a correta análise de rotas. Em dados poligonais, sobreposições ou lacunas afetam a delimitação de áreas e a atribuição de atributos.

Esse elemento reforça a importância de metodologias e ferramentas automatizadas para validação de dados, sendo uma etapa relevante no controle de qualidade.

3.2 CONSISTÊNCIA DE FORMATO

A consistência de formato é um elemento da categoria consistência lógica, conforme descrito pela ET-CQDG (2016). Esse elemento avalia se os dados geoespaciais estão armazenados no formato especificado pelo modelo físico, garantindo que a estrutura de armazenamento esteja alinhada às normas técnicas definidas.

De acordo com Ariza-López et al. (2019), a verificação da consistência de formato é parte integrante do controle de qualidade de dados geoespaciais e pode ser conduzida com base na ISO 19157. Os autores apontam que a ausência de atributos obrigatórios, falhas no preenchimento de campos e o descumprimento das estruturas previstas podem comprometer a utilização e integração dos dados, sendo essencial aplicar métodos sistemáticos para identificar e quantificar essas inconformidades.

França (2021) reforça essa perspectiva ao destacar diversas situações comuns de inconsistências de formato, como o fornecimento de arquivos com extensões diferentes das especificadas, presença de arquivos corrompidos, uso de sistemas de referência incompatíveis e codificações de caracteres distintas das previstas. O autor também chama atenção para as limitações do formato Shapefile, amplamente adotado, mas com restrições significativas, como o limite de tamanho de arquivos, número e tipo de atributos, e a dependência de múltiplos arquivos complementares para funcionamento adequado.

Esses problemas, quando não tratados, podem comprometer a compatibilidade com o modelo físico e inviabilizar a análise e interoperabilidade dos dados. Um exemplo recorrente é a ausência do arquivo .prj nos shapefiles, essencial para definir o sistema de referência de coordenadas. Também é comum a omissão ou configuração incorreta de campos obrigatórios nos atributos.

A ET-CQDG define medidas específicas para avaliar a consistência de formato, como a CQDG:206, que verifica a conformidade dos dados com o modelo físico. Essa medida prevê inspeções internas para garantir que os dados estejam devidamente organizados e que todos os componentes requeridos estejam presentes.

A Tabela 5, apresenta os parâmetros e métodos de avaliação relacionados à medida CQDG:206.

Tabela 5 – Medida conflito de estrutura física.

Linha	Componente	Descrição
1	Nome	Conflito de estrutura física
2	Nome alternativo	-
3	Elemento de qualidade	Consistência de formato
4	Medida básica	Indicador de erro
5	Definição	Instrumento que indica se existe algum erro de conflito na estrutura física de armazenamento do dado
6	Descrição	Deve ser verificado se existe algum erro de conflito na estrutura física do arquivo de acordo com o formato indicado no parâmetro. O arquivo que contém os dados geoespaciais terá sua estrutura avaliada de acordo com o formato indicado no parâmetro desta medida. Se algum conflito for encontrado entre o arquivo e o formato usado, o resultado desta medida é verdadeiro, caso contrário, falso é retornado.
7	Parâmetro	Definição do formato do arquivo (texto)
8	Tipo de valor	Booleano
9	Estrutura do valor	-
10	Referência da medida	-
11	Exemplo	Exemplos de conflitos no formato (retorna verdadeiro): 1) Arquivo GML é inválido se informa que usa UTF-8 no código de caracteres, mas na verdade usa o ISO-8859-1 (LATIN1); 2) Arquivo Shapefile é inválido se é entregue o SHP e o DBF, mas não o SHX; 3) Arquivo GeoTIFF é inválido se não possui o georreferenciamento no arquivo, mas sim num arquivo anexo (TFW).
12	Identificador	CQDG:206

Fonte: Adaptado DSG (2016, p. 3-12).

A consistência de formato desempenha um papel importante no controle de qualidade, pois assegura que os dados possam ser manipulados corretamente nos sistemas e processos planejados. Conforme Ariza-López et al. (2019), a aplicação de regras formais para validação automática é uma abordagem eficaz para minimizar a ocorrência de erros e promover a padronização das estruturas de armazenamento.

Com base na abordagem proposta, a análise da consistência de formato complementa os elementos de qualidade apresentados nos itens anteriores, contribuindo para a construção de bases de dados geoespaciais que atendam às especificações técnicas e aos objetivos do estudo.

3.3 CONSISTÊNCIA CONCEITUAL

A consistência conceitual é um dos componentes da categoria consistência lógica, conforme estabelecido pela ET-CQDG (2016). Esse elemento examina a aderência dos dados geoespaciais ao modelo conceitual predefinido, considerando a adequação das classes, atributos e relacionamentos às regras semânticas e estruturais estabelecidas. Tal verificação é imprescindível para assegurar uma representação coerente da realidade no banco de dados.

Conforme observado por França (2021), a consistência conceitual contribui para a confiabilidade da estrutura semântica dos dados, evitando desvios de classificação que possam dificultar sua utilização em SIGs. A correta categorização das feições, segundo o autor, favorece a reutilização dos dados em diferentes contextos e aplicações.

As inconsistências conceituais podem manifestar-se por meio da utilização indevida de classes, omissão de atributos obrigatórios ou atribuição de valores fora dos limites esperados. Tais falhas afetam diretamente a integração e a análise dos dados, podendo comprometer decisões baseadas nessas informações.

Para apoiar a identificação desses problemas, a Tabela 6 resume os principais tipos de domínio associados aos diferentes tipos de dados geoespaciais. Esses domínios ajudam a definir as regras esperadas para atributos do tipo numérico, textual, temporal, booleano e geométrico, sendo fundamentais na validação automatizada das estruturas conceituais e semânticas dos bancos analisados nesta pesquisa.

Por exemplo, ao se trabalhar com o atributo *situacao fisica* de uma classe vetorial, espera-se que seus valores estejam restritos a uma lista codificada como {1 = existente, 2 = em construção, 3 = demolido}. A ausência desse controle pode resultar na inserção de valores livres como "em obras" ou "antiga", comprometendo análises posteriores e dificultando a integração com outras bases. Esse tipo de erro evidencia

a relevância dos domínios predefinidos para assegurar a uniformidade da base de dados e dos resultados produzidos a partir dela.

Tabela 6 – Resumo dos principais tipos de domínios.

Tipo de dado	Tipo de domínio	Descrição	Exemplo
Numérico - inteiro	Código (Codelist)	Valor específico associado a um código.	$x \in (0, 1, 3, 7)$
	Intervalo (Range)	Espaço definido por valores mínimo e/ou máximo.	$1 < x < 10 \text{ e } y > 0$
Numérico – ponto flutuante	Intervalo (Range)	Espaço definido por valores mínimo e/ou máximo.	$0 < x < 3.14 \text{ e } y > 0.01$
Textual	Ortografia	Atributo textual que atende a norma ortográfica	‘eucalípito’
	Digitação	Observância a regras de digitação.	‘vila São João’
	Sintaxe	Disposição lógica entre palavras	‘ruas isolada’
Temporal	Intervalo de tempo	Faixa contínua de tempo para data ou hora definida por valores máximo e/ou mínimo	1999 a 2021 08:00 às 17:00
	Sazonal	Períodos descontínuos na linha de tempo	Verão Segunda a Quinta
Booleano	Não nulo (not null)	Atributo de preenchimento obrigatório	True False
Geometria	Região 2D	Geometria (G) dentro de um Espaço 2D definido (S2D)	G within S2D
	Região 3D e 4D	Idem ao anterior, acrescentando a definição de limites para as coordenadas Z e M	G within S2D e min < Z < max e/ou min < M < max

Fonte: Adaptado de França (2021, p. 38).

A ET-CQDG propõe medidas específicas para avaliar esse tipo de qualidade. A medida CQDG:201, por exemplo, verifica a correspondência entre as feições e o modelo conceitual, enquanto a CQDG:202 examina a completude dos atributos e a CQDG:203 avalia a coerência entre os relacionamentos.

A Tabela 7 resume os elementos que compõem a medida CQDG:201, utilizada para avaliar a consistência conceitual em conjuntos de dados geoespaciais. Essa medida verifica se as classes, os atributos e seus respectivos tipos de dados estão de acordo com o modelo conceitual previamente definido, como a EDGV. Um

exemplo ilustrativo dessa avaliação pode ser observado na linha 6, componente “Descrição”, onde se afirma que “se uma classe possui atributos que não existem no modelo (inclusive a geometria), isso configura um erro conceitual.”

Tabela 7 – Detalhes da medida CQDG:201 (Consistência Conceitual).

Linha	Componente	Descrição
1	Nome	Conformidade com o modelo de dados
2	Nome alternativo	-
3	Elemento de qualidade	Consistência conceitual
4	Medida básica	Indicador de acerto
5	Definição	Essa medida indica que o produto está em conformidade com o modelo de dados. A conformidade é em relação aos conceitos, que englobam classes, atributos e tipos de atributos.
6	Descrição	Partindo do conjunto de dados a ser avaliado, cada classe presente é comparada com sua classe correspondente no modelo de dados: a) Se a classe não possui correspondente no modelo de dados, é um erro conceitual; b) Verifica-se cada atributo de cada classe presente. Se um atributo não existe no modelo (inclusive a geometria), é um erro conceitual; c) Para cada atributo, se o tipo de dado no conjunto avaliado é incompatível com o modelo de dados, é um erro conceitual. Após concluir a avaliação de todas as classes e seus atributos presentes no conjunto avaliado, se houver algum erro, o resultado da medida é falso, que significa não conformidade com o modelo. Se nenhum erro é encontrado, então a medida retorna verdadeiro e o produto é conforme.
7	Parâmetro	-
8	Tipo de valor	Booleano
9	Estrutura do valor	-
10	Referência da medida	-
11	Exemplo	Considerando a EDGV 2.1.3 (CONCAR/DSG, 2010), os exemplos a seguir representam não conformidade com o modelo de dados, ou seja, o resultado da medida é falso: 1) Uma classe com o nome 'Trecho de Rodovia', quando o modelo prevê 'Trecho_Rodoviario';

		<p>2) Classe 'Identificador_Trecho_Rodoviario' com um atributo chamado 'abreviatura', quando o modelo prevê 'sigla';</p> <p>3) Classe 'Túnel' faltando o atributo 'altura'. É um atributo que pode assumir o valor nulo, mas deve constar do modelo mesmo que não esteja preenchido;</p> <p>4) Atributo 'tipo Ponte' na classe 'Ponte' com um valor numérico, quando o modelo prevê uma lista controlada de valores texto.</p>
12	Identificador	CQDG:201

Fonte: Adaptado DSG (2016, p. 3-7).

O modelo conceitual desempenha um papel central nesse processo, servindo como referência para estruturar os dados e definir as regras semânticas e relacionais. Segundo França (2021), a aplicação dessas medidas de qualidade permite identificar e corrigir inconsistências que poderiam passar despercebidas em análises manuais. A Tabela 8 ilustra exemplos de erros conceituais, de algumas feições foram atribuídas a classes inadequadas, destacando a importância de um modelo bem definido.

Por exemplo, no primeiro caso apresentado na Tabela 8 — classe com nome incorreto — o modelo de dados estabelece a nomenclatura padrão *Trecho_Rodoviario*. Contudo, no conjunto de dados avaliado, foi identificada a classe nomeada como *Trecho_de_Rodovia*. Essa divergência, embora pareça sutil, representa uma inconsistência conceitual segundo a medida CQDG:201, pois o nome da classe não está conforme o modelo conceitual estabelecido. Como resultado, essa ocorrência é considerada não conforme, e a medida retorna o valor *false*, evidenciando a necessidade de revisão e padronização dos dados para garantir a interoperabilidade e a qualidade da informação.

Tabela 8 – Exemplos de erros de consistência conceitual.

Exemplo	Modelo de Dados	Dado Avaliado	Resultado da Medida
Classe com nome incorreto	Trecho_Rodoviario	Trecho_de_Rodovia	<i>False</i>

Atributo com nome incorreto	sigla	abreviatura	False
Atributo ausente	altura presente (pode ser nulo)	altura ausente	False
Atributo fora do domínio	tipo Ponte com lista de valores texto	tipo Ponte com valor numérico	False

Fonte: Autor (2025).

Nesta pesquisa, a consistência conceitual é um dos focos principais, pois está diretamente relacionada à integridade lógica e semântica dos dados geoespaciais. As medidas descritas e exemplificadas não apenas avaliam a qualidade dos dados, mas também oferecem subsídios para a construção de uma base sólida que atenda aos padrões técnicos e às demandas de usuários.

3.4 CONSISTÊNCIA DE DOMÍNIO

A Consistência de Domínio é um elemento fundamental da qualidade de dados geoespaciais, abordando a conformidade dos valores atribuídos aos atributos em relação aos domínios estabelecidos no modelo de dados. Segundo a ET-CQDG (2016), esse elemento avalia se os valores registrados para atributos seguem os padrões, intervalos ou listas controladas definidos pelo modelo, garantindo a integridade semântica e a coerência das informações.

França (2021) destaca que a consistência de domínio é essencial para evitar erros que comprometem a confiabilidade de análises geoespaciais, como valores fora de intervalo ou atributos incompatíveis com suas definições. Esses erros podem ocorrer devido a falhas no processo de coleta ou integração de dados de diferentes fontes.

A avaliação da consistência de domínio envolve a aplicação de medidas de qualidade específicas, como a CQDG:204, que verifica a conformidade de valores em relação ao modelo. Essa medida é aplicada por meio de inspeções internas e comparações automáticas, assegurando que os valores estejam dentro dos limites esperados.

Os erros mais comuns incluem:

- a) Valores fora do intervalo: Atributos numéricos que excedem ou não alcançam os limites definidos pelo domínio.
- b) Valores não listados: Atributos textuais com valores que não pertencem à lista de valores válidos.
- c) Ausência de valores obrigatórios: Campos obrigatórios que estão vazios ou preenchidos inadequadamente.

A Tabela 9 apresenta a medida CQDG:204, exemplificando como a avaliação da consistência de domínio é realizada:

Tabela 9 - Medida porcentagem de não conformidade com o domínio.

Linha	Componente	Descrição
1	Nome	Porcentagem de não conformidade com o domínio
2	Nome alternativo	-
3	Elemento de qualidade	Consistência de domínio
4	Medida básica	Taxa de erro
5	Definição	Proporção da quantidade de valores de propriedades de objetos que não estão em conformidade com o domínio estabelecido no modelo de dados em relação ao número total de objetos no conjunto de dados avaliado

6	Descrição	<p>Para o conjunto de dados de testes, são verificados todos os objetos presentes no conjunto. Cada objeto possui suas propriedades, ou seja, um atributo e seu respectivo valor, excluindo-se aqui as propriedades geométricas. Cada propriedade deve ser verificada junto ao modelo de dados, a ver se cumpre com o estabelecido. Qualquer valor discrepante é considerado um erro. Um valor nulo em propriedade não-nula também configura um erro.</p> <p>Ao final, o resultado é uma proporção da quantidade de erros encontrados sobre o número total de objetos no conjunto de dados, expresso como porcentagem.</p>
7	Parâmetro	-
8	Tipo de valor	Real (porcentagem)
9	Estrutura do valor	-
10	Referência da medida	-

Fonte: Adaptado DSG (2016, p. 3-10).

O impacto da consistência de domínio está diretamente relacionado à qualidade das análises derivadas dos dados. Valores inconsistentes podem levar a interpretações equivocadas e decisões inadequadas, especialmente em contextos críticos, como planejamento urbano e gestão de recursos naturais. Como enfatizado por França (2021), a automatização de verificações utilizando ferramentas de SIG é indispensável para assegurar que os dados atendam aos padrões definidos.

Um exemplo prático dessa avaliação pode ser observado na análise do atributo "situação física" de um banco de dados geoespacial, no qual um modelo predefinido lista nove valores válidos. No processo de avaliação, foi identificado que, dos nove valores predefinidos no modelo, cinco dos valores presentes no banco verificado não coincidiam com os padrões estabelecidos. Isso representa uma taxa de erro de 55,56% (5/9), indicando a necessidade de ajustes para que o banco analisado esteja em conformidade com o modelo.

A Tabela 10 resume os valores esperados e encontrados no atributo "situação física", permitindo uma visualização dos desvios detectados:

Tabela 10 – Exemplos de inconsistências de domínio em atributos textuais e numéricos.

Classes	edgv_300	edgv_300_ortho	Coincide?
situação física	(0, 'Desconhecida (0)')	(0, 'Desconhecida (0)')	Sim
situação física	(1, 'Abandonada (1)')	(1, 'Abandonada (1)')	Sim
situação física	(2, 'Destruída (2)')	(2, 'Destruída (2)')	Sim
situação física	(3, 'Em construção (3)')	Não coincide	Não
situação física	(4, 'Planejada (4)')	Não coincide	Não
situação física	(5, 'Construída (5)')	Não coincide	Não
situação física	(6, 'Construída, mas em obras (6)')	Não coincide	Não
situação física	(97, 'Não aplicável (97)')	Não coincide	Não
situação física	(9999, 'A SER PREENCHIDO (9999)')	(9999, 'A SER PREENCHIDO (9999)')	Sim
situação física	Não presente	(3, 'Construída (3)')	Não
situação física	Não presente	(4, 'Em construção (4)')	Não

Fonte: O autor (2025).

Neste estudo, a consistência de domínio é avaliada como parte do controle de qualidade de dados geoespaciais, utilizando as especificações da ET-CQDG (2016) e metodologias automatizadas em ambientes SIG. O próximo capítulo detalha o processo de aplicação dessas medidas e a geração de relatórios de qualidade.

4. PROCEDIMENTOS METODOLÓGICOS

A metodologia adotada neste trabalho tem como objetivo estruturar e sistematizar as etapas necessárias para a avaliação da qualidade de dados geoespaciais, com ênfase nos elementos de consistência lógica conceitual e consistência de domínio. Segundo as diretrizes estabelecidas pela ISO 19157:2013 e pela ET-CQDG, foram elaborados procedimentos que integram análise teórica e aplicação prática para validar os dados conforme os padrões estabelecidos.

O capítulo apresenta uma descrição dos procedimentos realizados, que incluem a seleção de bases de dados, a definição de critérios e medidas de qualidade, a execução de ferramentas específicas para a verificação de inconsistências e a interpretação dos resultados obtidos. Cada etapa é fundamentada em técnicas amplamente aceitas na literatura, como a validação contra esquemas de dados conceituais e físicos (Ariza-López et al., 2019), a verificação de integridade de atributos (França, 2021), e a aplicação de scripts automatizados por meio do PyQGIS e do PostgreSQL/PostGIS, conforme sugerido por Meneses et al. (2023). Tais práticas favorecem a reproduzibilidade e a confiabilidade da avaliação, especialmente no contexto de aplicação das medidas da ET-CQDG (2016). Ademais, testes foram conduzidos em cada etapa com o intuito de verificar a eficiência dos procedimentos adotados.

A abordagem metodológica contempla dois eixos principais:

1. Análise de Dados Geoespaciais: Aplicação de medidas específicas da ET-CQDG para identificar erros relacionados à consistência lógica conceitual e de domínio.
2. Desenvolvimento de Ferramentas Automatizadas: Implementação de scripts para o controle de qualidade, otimizando o processo de detecção e correção de inconsistências em bases de dados geoespaciais.

4.1 ANÁLISE DE DADOS GEOESPACIAIS

O objeto de estudo desta pesquisa é a análise da consistência lógica, com ênfase na consistência conceitual e de domínio, de dados geoespaciais estruturados segundo diferentes modelagens. Esta abordagem visa não apenas garantir a

aderência dos dados a normas técnicas e especificações (como a ISO 19157:2013 e a ET-CQDG, 2016), mas também fortalecer as bases para a conversão entre essas modelagens, promovendo a interoperabilidade e a consistência no intercâmbio de dados entre bancos de dados distintos.

Para atender às necessidades de interoperabilidade, racionalização de recursos e integração de dados geoespaciais no âmbito da Infraestrutura Nacional de Dados Espaciais (INDE), foi adotada a ET-EDGV para padronizar a organização, o armazenamento e o compartilhamento de dados geoespaciais vetoriais. Com base na modelagem conceitual orientada a objetos, a EDGV é amplamente utilizada para a produção de dados cartográficos e para a integração de bases em diferentes escalas, promovendo a padronização e o compartilhamento eficiente de informações.

Para este trabalho, foram utilizadas bases de dados geoespaciais estruturadas conforme as seguintes especificações:

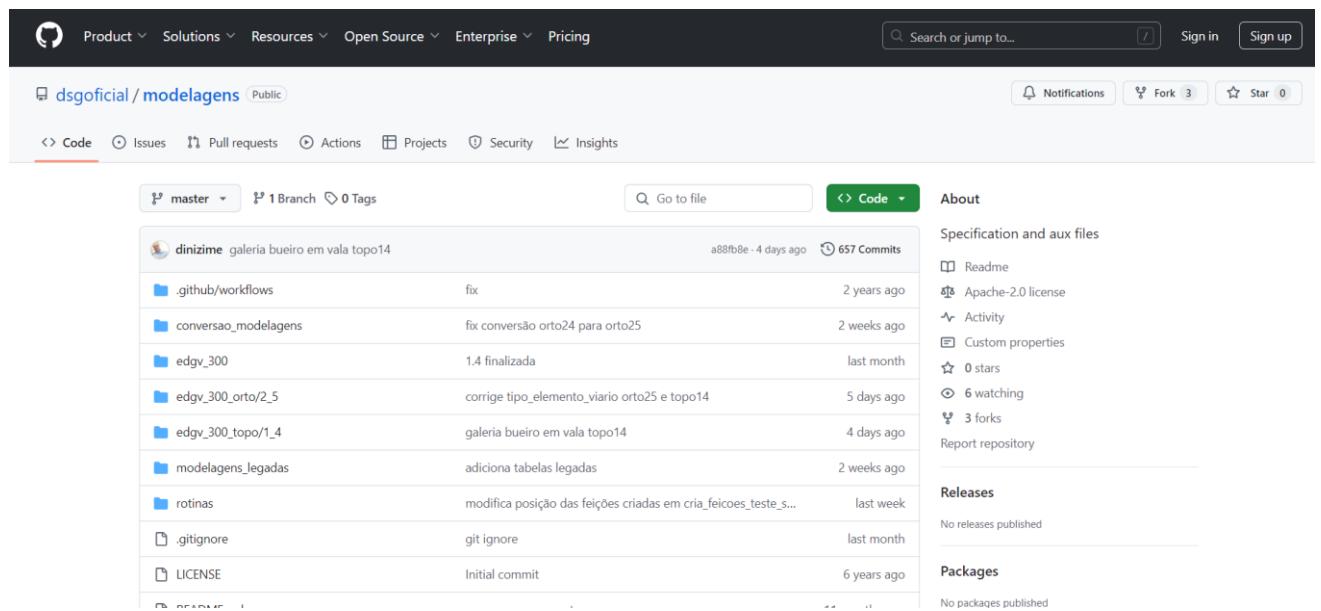
1. EDGV 3.0 - Atual padrão de modelo de CDVG, utilizado também para construção de cartas topográficas e modelos de dados de grandes escalas.
2. EDGV 3.0 Orto (Carta Ortoimagem) – Modelo de CDVG para construção da Cartas Ortoimagem.
3. EDGV Aracaju – Personalização da EDGV 3.0 que integra peculiaridades regionais para atender às demandas específicas da cartografia de Aracaju.
4. EDGV 2.1.3 – Versão anterior amplamente utilizada, escolhida para permitir a comparação de avanços entre diferentes versões da modelagem.

As especificações EDGV 3.0, EDGV 3.0 Orto e EDGV 2.1.3 utilizadas neste estudo estão disponíveis no repositório oficial da Diretoria de Serviço Geográfico (DSG) no GitHub, conforme ilustra a Figura 4. A estrutura do repositório foi organizada para facilitar o acesso às modelagens e sua aplicação em projetos que demandem estruturação padronizada de dados geoespaciais. Já a modelagem utilizada em Aracaju, por se tratar de uma adaptação local, foi obtida a partir de fontes regionais específicas.

A seleção dessas modelagens fundamenta-se na relevância da EDGV como principal referência para a organização de dados geoespaciais no Brasil. A comparação entre versões oficiais e adaptações regionais visa identificar padrões estruturais e inconsistências conceituais, contribuindo para o aprimoramento de práticas de controle de qualidade de dados no país. O uso de repositórios abertos,

como o da DSG, também favorece a reproduzibilidade das análises e estimula a adoção de práticas interoperáveis entre instituições públicas e privadas.

Figura 4 – Repositório oficial da Diretoria de Serviço Geográfico (DSG) no GitHub contendo as especificações EDGV.



The screenshot shows the GitHub repository page for 'dsgoficial/modelagens'. The repository is public and has 657 commits. The code tab is selected, showing a list of commits. The commits are as follows:

- dinizime galeria bueiro em vala topo14 (a88fb8e - 4 days ago)
- .github/workflows fix (2 years ago)
- conversao_modelagens fix conversão orto24 para orto25 (2 weeks ago)
- edgv_300 1.4 finalizada (last month)
- edgv_300.orto/2_5 corrige tipo_elemento_viajario orto25 e topo14 (5 days ago)
- edgv_300.topo/1_4 galeria bueiro em vala topo14 (4 days ago)
- modelagens_legadas adiciona tabelas legadas (2 weeks ago)
- rotinas modifica posição das feições criadas em cria_feicoes_teste_s... (last week)
- .gitignore git ignore (last month)
- LICENSE Initial commit (6 years ago)

The repository has 3 forks and 0 stars. The 'About' section includes links to 'Specification and aux files', 'Readme', 'Apache-2.0 license', 'Activity', 'Custom properties', '0 stars', '6 watching', '3 forks', and 'Report repository'. The 'Releases' section shows 'No releases published'. The 'Packages' section shows 'No packages published'.

Fonte: DSG. Repositório de modelagens de banco de dados EDGV.

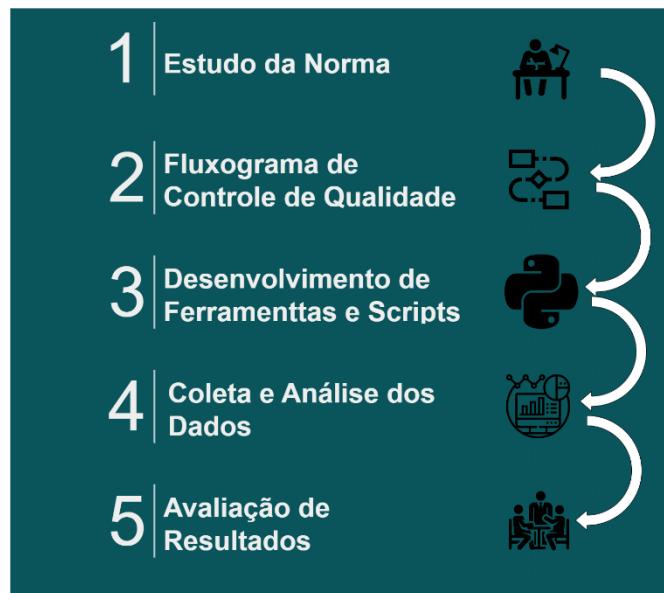
Disponível em: <https://github.com/dsgoficial/modelagens>. Acesso em: 13/01/2025.

4.2 DESENVOLVIMENTO E IMPLANTAÇÃO DE PROCESSOS AUTOMATIZADOS E PROCEDIMENTOS DE VERIFICAÇÃO

O fluxo de trabalho desta pesquisa foi organizado em cinco etapas principais, que refletem o procedimento metodológico adotado para a análise e validação dos dados geoespaciais. Essas etapas estão representadas no fluxograma apresentado na Figura 5.

O processo inicia com o estudo das normas, que abrangeu a análise detalhada de documentos como as diretrizes da ISO para qualidade de dados geoespaciais, ET-CQDG e a ET-EDGV, para compreender os requisitos e especificações necessários a verificação da consistência lógica, conceitual e de domínio.

Figura 5 – Fluxograma do Procedimento Metodológico.



Fonte: O autor (2025).

Na etapa seguinte, ocorreu o desenvolvimento de ferramentas e scripts, subitem 4.2.1, com a criação de rotinas automatizadas em PyQGIS para a validação dos dados geoespaciais. Esses scripts foram desenvolvidos para verificar a conformidade dos dados com os padrões estabelecidos, realizando análises como a validação de atributos obrigatórios, tipos de dados e estruturas geométricas.

Com as ferramentas prontas, foi realizada a coleta e análise dos dados, organizando as bases em um banco de dados PostgreSQL/PostGIS e submetendo-as aos scripts automatizados. Essa etapa foi responsável por identificar inconsistências estruturais e conceituais, que foram registradas e corrigidas conforme necessário.

Por fim, a avaliação de resultados, consolidou as análises realizadas, gerando relatórios sobre as conformidades e não conformidades identificadas. Esses relatórios auxiliaram na interpretação dos resultados e na proposição de ajustes para garantir que os dados estivessem alinhados às normas aplicáveis.

4.2.1 Desenvolvimento do Processo de Automatização das Verificações

A abordagem metodológica integrou o uso do QGIS e da biblioteca PyQGIS para manipulação e processamento de dados geoespaciais, juntamente com o

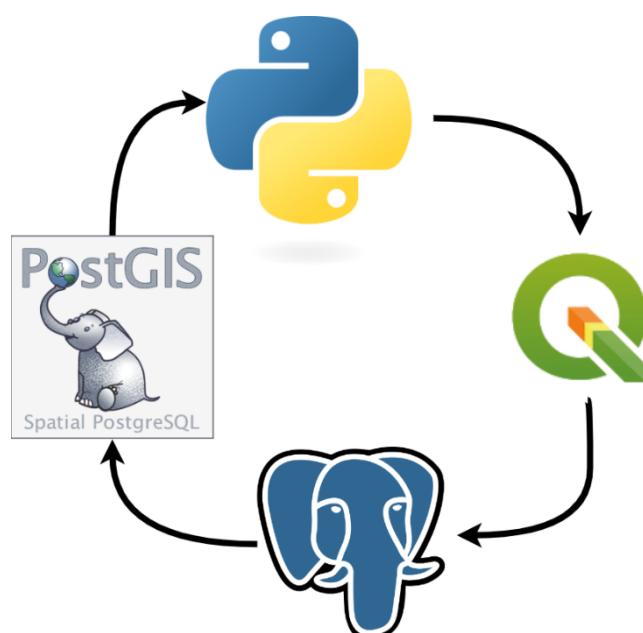
PostgreSQL e sua extensão PostGIS para armazenamento e análise em banco de dados geoespacial. Essa escolha garantiu acesso às funcionalidades necessárias para o trabalho, alinhando-se às práticas recomendadas no uso de tecnologias abertas.

O QGIS é um software livre de SIG, amplamente utilizado para análises geoespaciais, que permite a manipulação de dados vetoriais e raster, além de oferecer suporte para automação de tarefas por meio de sua biblioteca integrada, o PyQGIS. O uso da biblioteca PyQGIS foi essencial para a implementação de scripts automatizados que viabilizaram a análise e validação dos dados geoespaciais.

O PostgreSQL, em conjunto com sua extensão espacial PostGIS, foi empregado para armazenamento e consulta de dados espaciais. O PostGIS possibilita a realização de operações espaciais avançadas, como interseções e validações topológicas, sendo uma escolha amplamente documentada e reconhecida para análises geoespaciais.

Essas ferramentas são integradas, facilitando o desenvolvimento do script, a Figura 6, ilustra a integração entre ferramentas de software livre utilizadas nesta pesquisa: QGIS, PostgreSQL, PostGIS e Python (PyQGIS). Essa integração possibilita o armazenamento, processamento e análise automatizada de dados geoespaciais, garantindo interoperabilidade e alinhamento com os padrões estabelecidos pela ET-EDGV e ET-CQDG.

Figura 6 – Integração de Ferramentas Geoespaciais.



Fonte: O autor (2025).

As especificações técnicas da ET-CQDG, em especial as medidas de qualidade CQDG:201 (consistência conceitual) e CQDG:204 (consistência de domínio), bem como a modelagem definida pela ET-EDGV, forneceram as diretrizes metodológicas para o desenvolvimento de um conjunto de scripts em Python voltados à validação da qualidade dos dados geoespaciais.

O conjunto de scripts foi denominado GeoQuali CQDG e abrange rotinas específicas de verificação automatizada com base nas medidas mencionadas. Cada script foi implementado para atuar de forma modular, mantendo o nome da medida correspondente no ambiente do QGIS (como CQDG 201 e CQDG 204), facilitando sua aplicação e compreensão.

O fluxo metodológico do GeoQuali CQDG foi estruturado em etapas que integram o uso do QGIS, da API PyQGIS, e dos bancos de dados PostgreSQL/PostGIS, abrangendo desde a preparação das bases até a análise e documentação dos resultados obtidos. As etapas principais incluem:

:

1. Preparação das Bases de Dados: Os dados geoespaciais foram carregados e organizados no PostgreSQL com a extensão PostGIS. A modelagem de banco de dados seguiu os padrões estabelecidos pelas ET-EDGV para garantir conformidade com a estrutura conceitual e topológica.
2. Análise e Processamento Automatizado: Os scripts desenvolvidos em PyQGIS foram utilizados para implementar a análise de consistência conceitual e de domínio com base nas diretrizes do ET-CQDG.
3. Validação dos Resultados: O QGIS foi empregado para visualizar e validar os resultados das análises. Camadas resultantes das operações foram verificadas visualmente e comparadas com os padrões esperados para identificar possíveis inconsistências.
4. Geração de Relatórios: Dados sobre inconsistências e conformidade foram exportados para relatórios detalhados, que auxiliam na interpretação dos resultados e na identificação de padrões que possam ser ajustados.

4.2.2 Coleta e Procedimentos de Análise

Os procedimentos metodológicos desta etapa foram direcionados à coleta e análise de dados geoespaciais, com foco na verificação da consistência lógica, especialmente nos aspectos conceitual e de domínio. As análises foram conduzidas com base nos referenciais normativos já descritos anteriormente, notadamente as medidas CQDG 201 e CQDG 204 da ET-CQDG (2016) e a ET-EDGV. Conforme discutido por França (2021), a aplicação de mecanismos sistemáticos de verificação é essencial para assegurar a coerência entre o modelo conceitual e sua materialização física.

A etapa de coleta consistiu na obtenção das modelagens físicas disponibilizadas em fontes oficiais, como o repositório GitHub da Diretoria de Serviço Geográfico (DSG), contemplando versões como EDGV 3.0, EDGV Aracaju e EDGV 2.1.3. Esses dados foram organizados em um banco PostgreSQL com a extensão PostGIS, estruturando o ambiente para análise.

A EDGV 3.0 foi utilizada como modelo de referência, por ser amplamente adotada na DSG e representar a padronização oficial. Já as modelagens EDGV Aracaju e EDGV 2.1.3 foram incluídas de forma comparativa, a fim de demonstrar como a falta de uniformidade pode comprometer a consistência conceitual e de domínio dos dados geoespaciais.

A Tabela 11 ilustra um exemplo do processo de análise comparativa realizado na classe *enc_torre_energia_p*, pertencente a duas bases de dados distintas, denominadas Banco de Dados X e Banco de Dados Y. Essa comparação permitiu verificar a presença e o tipo de cada atributo, evidenciando diferenças que podem comprometer a interoperabilidade e a consistência conceitual dos dados.

Tabela 11 - Comparaçāo de Atributos entre BancoX e BancoY.

Classe	Atributo	Tipo BD X	Tipo BD Y
<i>enc_torre_energia_p</i>	<i>id</i>	<i>integer</i>	<i>uuid</i>
<i>enc_torre_energia_p</i>	<i>nome</i>	<i>character varying</i>	<i>character varying</i>
<i>enc_torre_energia_p</i>	<i>geometriaaproximada</i>	<i>boolean</i>	<i>boolean</i>
<i>enc_torre_energia_p</i>	<i>operacional</i>	<i>text</i>	<i>smallint</i>
<i>enc_torre_energia_p</i>	<i>situacaofisica</i>	<i>text</i>	<i>smallint</i>
<i>enc_torre_energia_p</i>	<i>ovgd</i>	<i>text</i>	<i>smallint</i>
<i>enc_torre_energia_p</i>	<i>alturaestimada</i>	<i>double precision</i>	<i>real</i>

enc_torre_energia_p	observacao	Não presente	character varying
enc_torre_energia_p	Geometria	Point	Multipoint

Fonte: O autor (2025).

As análises foram realizadas por meio de scripts automatizados desenvolvidos em PyQGIS, que operacionalizam as verificações propostas nas medidas CQDG 201 e 204. Foram validados atributos obrigatórios, tipos de dados e possíveis inconformidades estruturais, conforme detalhado a seguir.

4.2.2.1 Validação de Estruturas Conceituais

Verificação da correspondência entre os atributos das classes armazenadas no banco de dados e o modelo conceitual da EDGV, assegurando conformidade em relação à presença, nomenclatura e tipos de dados dos atributos. Na tabela 11, é apresentado um extrato do relatório e nele, é verificado a correspondência entre os atributos da classe torre de energia.

4.2.2.2 Análise de Consistência de Domínio

A verificação da consistência de domínio compõe uma das etapas centrais desta pesquisa, conforme previsto pela medida CQDG:204 da ET-CQDG (2016). Essa análise busca identificar se os valores atribuídos aos atributos estão em conformidade com os domínios definidos nos próprios bancos de dados utilizados, assegurando coerência semântica e reduzindo a ocorrência de erros interpretativos ou estruturais.

É importante destacar que, embora os bancos de dados analisados sigam a modelagem conceitual da EDGV, não há uma padronização normativa externa única que defina os domínios válidos para todos os atributos. Os domínios utilizados foram aqueles previamente configurados nos bancos de dados disponibilizados publicamente pela Diretoria de Serviço Geográfico (DSG), sendo observadas, inclusive, variações de domínio para um mesmo atributo entre diferentes versões da base. Nesse contexto, os scripts desenvolvidos permitiram verificar automaticamente a aderência dos valores registrados aos domínios existentes em cada base, facilitando a identificação de possíveis inconsistências ou divergências internas.

De acordo com França (2021), a análise sistemática dos domínios presentes nos dados é uma estratégia importante para garantir a qualidade semântica da base, especialmente em contextos em que a interoperabilidade entre diferentes versões ou implementações é necessária.

4.2.2.3 Automatização das Verificações

O uso de scripts permite a automatização das verificações, reduzindo significativamente o tempo necessário para identificar inconsistências em um grande volume de classes e atributos. Os scripts foram configurados para gerar relatórios, facilitando a interpretação dos resultados.

A aplicação desses procedimentos permite avaliar a eficiência do modelo conceitual da EDGV e sua implementação em diferentes contextos. Conforme observado por França (2021), a consistência conceitual é fundamental para assegurar que os dados geoespaciais atendam aos requisitos de qualidade exigidos por projetos de grande escala, como os relacionados à gestão urbana e ambiental.

5 RESULTADOS E DISCUSSÕES

Os scripts implementados realizam a análise da estrutura conceitual dos dados geoespaciais vetoriais presentes em bancos PostgreSQL/PostGIS, avaliando aspectos fundamentais da conformidade com os modelos oficiais da EDGV. A automatização abrange a extração de informações da estrutura física dos bancos de dados e a execução de rotinas de comparação com os parâmetros conceituais e de domínio definidos nas especificações.

Foram avaliados, de forma sistemática, os seguintes critérios:

- Existência ou ausência de tabelas exigidas pelo modelo;
- Presença e nomenclatura dos atributos obrigatórios;
- Tipagem dos atributos conforme definição da EDGV;
- Geometria declarada (tipo e dimensionalidade);
- Associação de atributos a domínios;
- Existência de valores divergentes nos campos codificados;
- Diferenças entre modelagens regionais e o modelo nacional de referência;
- Estrutura de implementação dos domínios (por CREATE DOMAIN com cláusula CHECK ou por tabelas auxiliares);
- Casos de sensibilidade a maiúsculas e minúsculas que impactam consultas e armazenamento.

A apresentação dos resultados está organizada por categorias de análise, conforme a estrutura das medidas de qualidade adotadas. Em 5.1, são descritos os resultados da verificação da consistência conceitual (CQDG 201), com destaque para as comparações entre diferentes modelagens da EDGV, como EDGV 3.0, EDGV 2.1.3, EDGV 3.0 Orto e uma modelagem municipal adaptada (referida aqui como "modelo exemplo"). O subitem 5.1.1 detalha a quantidade de atributos analisados e divergências encontradas. Já o subitem 5.1.2 apresenta um panorama das diferenças estruturais entre os modelos avaliados.

Em 5.2, são apresentados os resultados da consistência de domínio (CQDG 204), com verificação da aderência dos valores armazenados aos domínios predefinidos. O subitem 5.2.1 mostra exemplos de inconsistências nos valores codificados, enquanto o subitem 5.2.2 compara a estruturação dos domínios nos diferentes bancos analisados.

Por fim, em 5.3, são descritas as funcionalidades da aplicação desenvolvida, com destaque para as vantagens do processo automatizado em termos de eficiência, padronização e replicabilidade.

5.1. QUANTO A MEDIDA DE QUALIDADE CQDG 201

A consistência conceitual avaliada com base na medida CQDG 201 deve verificar se os atributos presentes nas tabelas físicas dos bancos de dados geoespaciais estão em conformidade com o modelo conceitual da EDGV 3.0. No critério estabelecido no script, se considerou a: presença, nomenclatura e tipo dos atributos definidos para cada classe.

As comparações foram realizadas de forma automatizada que verificou os atributos obrigatórios previstos no modelo conceitual. Os resultados foram organizados conforme os seguintes tópicos:

5.1.1 Comparação entre Modelagens (EDGV 3.0 versus outras versões)

Foram avaliadas três comparações principais:

- a) EDGV 3.0 versus EDGV 2.1.3;
- b) EDGV 3.0 versus EDGV 3.0 Orto; e
- c) EDGV 3.0 versus EDGV Aracaju

Os resultados completos dessas comparações estão disponíveis no repositório do projeto no GitHub, conforme indicado no Apêndice C, com as saídas geradas automaticamente por cada um dos scripts aplicados.

A Tabela 12 apresenta um recorte dos resultados obtidos na comparação entre a EDGV 3.0 e o modelo exemplo. O foco está na quantidade de atributos verificados e no número de divergências detectadas em cinco classes distintas.

Tabela 12 – Comparação de Atributos: EDGV 3.0 x Aracaju (CQDG 201).

Classe/Tabela	Atributos Comparados	Diferenças Encontradas
hdv_trecho_hidroviario_l	11	10

hdv_atracadouro_terminal_a	13	11
fer_trecho_ferroviario_I	18	13
eco_plataforma_a	6	5
edf_posto_policia_militar_a	19	17

Fonte: O autor (2025).

Esses resultados ilustram divergência estrutural entre o modelo oficial EDGV 3.0 e a modelagem adaptada, mesmo em classes recorrentes nos bancos de dados. A classe *edf_posto_policia_militar_a*, por exemplo, apresentou 17 divergências em um total de 19 atributos verificados, o que demonstra um alto grau de desacordo conceitual. As discrepâncias observadas envolvem principalmente nomes de campos divergentes, atributos ausentes ou com tipos incorretos, o que reforça a importância da padronização dos modelos utilizados.

O detalhamento da quantidade de classes por modelagem e suas implicações na estrutura conceitual está apresentado no item 5.1.2, que discute as diferenças estruturais entre as versões analisadas.

5.1.2 Diferenças Estruturais e Conceituais

A Tabela 13 apresenta um resumo das classes avaliadas nas diferentes modelagens utilizadas, destacando aspectos quantitativos que orientaram a análise comparativa. Observa-se uma variação significativa na quantidade de classes entre as versões da EDGV.

Tabela 13 – Resumo das Classes Avaliadas.

Modelagem	Número de Classes	Descrição Geral
EDGV 3.0	287	Modelo conceitual atualizado, atualmente utilizado no mapeamento cadastral.

EDGV 3.0 Orto	67	Modelo para produção de cartas ortoimagem.
EDGV Aracaju	466	Adaptação regional da EDGV 3.0 com peculiaridades locais.
EDGV 2.1.3	330	Versão anterior amplamente utilizada para comparação.

Fonte: O autor (2025).

A EDGV 3.0, atualmente adotada como padrão nacional para estruturação de dados geoespaciais vetoriais, apresenta 287 classes, enquanto a sua adaptação voltada à produção de cartas ortoimagem (EDGV 3.0 Orto) possui 67 classes. A EDGV 2.1.3, por sua vez, possui 330 classes e foi incluída neste estudo por ainda estar sendo amplamente utilizada, a exemplo do Estado da Bahia que seu mapeamento recente, foi realizado utilizando a versão do banco de dados EDGV 2.1.3.

Destaque especial deve ser dado à modelagem de Aracaju, que apresenta 466 classes, número consideravelmente superior às demais modelagens. Esse volume decorre da adaptação local da EDGV 3.0 realizada pela gestão municipal, que incorporou subclasses específicas para atender peculiaridades urbanas, administrativas e cartográficas. A análise do banco de dados evidenciou a presença de diversas classes com sufixos, subdivisões e detalhamentos inexistentes na estrutura padrão da EDGV 3.0.

Por exemplo, classes como edf_marquise_a, edfRepresentacao_diplomatica e lml_bairro_a não integram o modelo oficial, mas foram adicionadas à modelagem municipal para contemplar demandas específicas do contexto urbano e administrativo de Aracaju. Esse tipo de detalhamento reforça a flexibilidade do modelo local, ao mesmo tempo em que evidencia os impactos negativos da ausência de padronização. Ainda que essa granularidade ampliada represente necessidades locais legítimas, ela impõe desafios à interoperabilidade e à padronização nacional, sobretudo quando se busca integrar ou comparar dados entre diferentes entes federativos. A inexistência de normas que regulem de forma detalhada a criação e a padronização de subclasses, por exemplo, mediante codificação unificada ou domínio controlado nacional, contribui diretamente para as divergências estruturais observadas entre os modelos.

5.2 QUANTO A MEDIDA DE QUALIDADE CQDG 204

Conforme definida pela ET-CQDG, 2016, a medida de qualidade 204 tem como finalidade verificar se os valores registrados nos atributos dos dados geoespaciais estão de acordo com os domínios predefinidos, ou seja, se respeitam os conjuntos válidos de valores esperados para cada campo, conforme a estrutura conceitual da EDGV.

No script automatizado implementado em PyQGIS, o processo consistiu em cruzar os valores efetivamente armazenados com os domínios válidos definidos nos bancos de dados disponibilizados no repositório oficial da DSG, os quais, embora não estejam formalmente padronizados por uma norma nacional, foram considerados como referência por estarem amplamente utilizados.

A verificação concentrou-se em atributos com domínios do tipo codelist, como por exemplo: *situacao_fisica*, *tipopista*, *revestimento*, entre outros, que apresentaram variações significativas entre os diferentes modelos analisados. A Tabela 14 apresenta um exemplo da comparação entre os valores esperados e os valores observados em determinadas classes, destacando inconsistências encontradas.

Tabela 14 – Exemplo de Verificação de Domínio (CQDG 204).

Classes	edgv_300	edgv_300_ortho	Coincide?
situacao_fisica	(0, 'Desconhecida (0)')	(0, 'Desconhecida (0)')	Sim
situacao_fisica	(1, 'Abandonada (1)')	(1, 'Abandonada (1)')	Sim
situacao_fisica	(2, 'Destruída (2)')	(2, 'Destruída (2)')	Sim
situacao_fisica	(3, 'Em construção (3)')	Não coincide	Não
situacao_fisica	(4, 'Planejada (4)')	Não coincide	Não
situacao_fisica	(5, 'Construída (5)')	Não coincide	Não
situacao_fisica	(6, 'Construída, mas em obras (6)')	Não coincide	Não
situacao_fisica	(97, 'Não aplicável (97)')	Não coincide	Não
situacao_fisica	(9999, 'A SER PREENCHIDO (9999)')	(9999, 'A SER PREENCHIDO (9999)')	Sim
situacao_fisica	Não presente	(3, 'Construída (3)')	Não
situacao_fisica	Não presente	(4, 'Em construção (4)')	Não

Fonte: O autor (2025).

Os resultados evidenciaram que diversas inconsistências de domínio estão associadas à ausência de padronização, especialmente quando os domínios utilizados foram modificados localmente ou definidos sem referência direta ao modelo conceitual oficial.

Embora a Tabela 14 não apresente esse tipo de ocorrência relacionada a variações tipográficas, é importante destacar que, em situações em que os domínios são definidos como texto livre — em vez de por códigos controlados (codelists) —, podem ocorrer discrepâncias sutis, como diferenças de capitalização (por exemplo, “Construída” versus “construída”), inserção de espaços em branco ou pequenas inconsistências ortográficas. Esses desvios comprometem a uniformidade dos dados e dificultam a execução de consultas automatizadas, impactando negativamente a eficiência dos filtros e a geração de relatórios consistentes.

Adicionalmente, a Tabela 14 revela um tipo diferente de inconsistência crítica: a divergência na associação entre códigos e descrições entre diferentes modelagens. Por exemplo, no atributo situacao_fisica, o código (3) no banco edgv_300_orto está associado à descrição 'Construída (3)', enquanto, no banco edgv_300, o mesmo código (3) refere-se à descrição 'Em construção (3)'. Nesse caso, o valor correspondente a 'Construída' aparece no banco edgv_300 como código (5). Isso significa que, se um usuário realizar uma operação direta de transferência de dados entre os bancos, como exportar vetores de um banco e importar em outro, sem verificar previamente a correspondência dos domínios, poderá alterar indevidamente a classificação do atributo. Na prática, um objeto originalmente registrado como *construído* pode ser interpretado como *em construção* no banco de destino, comprometendo a fidedignidade da informação geoespacial.

Essa verificação é essencial para garantir que os dados geoespaciais possam ser corretamente interpretados por sistemas automatizados, que dependem da uniformidade e previsibilidade dos valores para a geração de representações visuais, relatórios e análises confiáveis. Como também destaca Silva (2023), a consistência de domínio está diretamente relacionada à usabilidade e à reusabilidade dos dados geoespaciais em ambientes interoperáveis, especialmente no contexto das Infraestruturas de Dados Espaciais (IDE).

5.2.2 Verificação da consistência de domínio

Na análise dos modelos de dados, observou-se que ambos utilizam a estrutura CREATE DOMAIN, mas com abordagens distintas na definição dos valores permitidos para os atributos controlados por domínio.

O modelo adotado (Figura 7) utiliza domínios associados a tabelas auxiliares, onde cada valor possível é representado por um código e uma descrição. Essa organização permite o reaproveitamento dos domínios por diversas tabelas e facilita o controle centralizado dos valores utilizados. Essa abordagem é especialmente útil para gerar interfaces com listas suspensas e para manter coerência na entrada de dados.

Figura 7 – Representação alternativa de domínio implementada por tabela de códigos e descrições.

```

CREATE TABLE dominios.administracao (
    code smallint NOT NULL,
    code_name text NOT NULL,
    CONSTRAINT administracao_pk PRIMARY KEY (code)
) #

INSERT INTO dominios.administracao (code,code_name) VALUES (0,'Desconhecida (0)')#
INSERT INTO dominios.administracao (code,code_name) VALUES (1,'Federal (1)')#
INSERT INTO dominios.administracao (code,code_name) VALUES (2,'Estadual/Distrital (2)')#
INSERT INTO dominios.administracao (code,code_name) VALUES (3,'Municipal (3)')#
INSERT INTO dominios.administracao (code,code_name) VALUES (7,'Concessionada (7)')#
INSERT INTO dominios.administracao (code,code_name) VALUES (15,'Privada (15)')#
INSERT INTO dominios.administracao (code,code_name) VALUES (97,'Não aplicável (97)')#
INSERT INTO dominios.administracao (code,code_name) VALUES (9999,'A SER PREENCHIDO (9999)')#

ALTER TABLE dominios.administracao OWNER TO postgres#

```

Fonte: Diretoria de Serviço Geográfico – DSG (2025).

Já o modelo baseado na Especificação Técnica da EDGV (Figura 8) define os valores diretamente no próprio domínio, por meio da cláusula CHECK com uma lista fixa (ARRAY[...]). Embora essa estrutura seja funcional e eficiente para situações com poucas opções, ela torna mais trabalhosa a atualização dos valores permitidos, pois exige alterações diretas no domínio e recompilação do banco. Além disso, dificulta a associação de descrições complementares aos valores, o que pode limitar a clareza para o usuário final em ferramentas de visualização ou edição.

Esse tipo de estrutura define explicitamente os valores aceitáveis como strings, o que pode dificultar a padronização e a realização de consultas no banco de dados. Por exemplo, ao tentar filtrar registros cuja administração seja “Estadual”, é

necessário conhecer exatamente como o valor foi escrito, por exemplo ('Estadual / Distrital', 'ESTADUAL / DISTRITAL', etc.), já que o banco diferencia letras maiúsculas de minúsculas.

Figura 8 – Definição de domínio do tipo enumerado diretamente no modelo de dados.

```

CREATE DOMAIN public.administracao AS text
  CONSTRAINT administracao_chk CHECK ((VALUE = ANY (ARRAY['Concessionada'::text,
  'Desconhecida'::text, 'Estadual / Distrital'::text, 'Federal'::text,
  'Municipal'::text, 'Não aplicável'::text, 'Privada'::text])));
  
ALTER DOMAIN public.administracao OWNER TO postgres;

-- 
-- Name: administracao_v; Type: DOMAIN; Schema: public; Owner: postgres
-- 

CREATE DOMAIN public.administracao_v AS text[]
  CONSTRAINT administracao_v_chk CHECK ((VALUE <@ ARRAY['Concessionada'::text,
  'Desconhecida'::text, 'Estadual / Distrital'::text, 'Federal'::text,
  'Municipal'::text, 'Não aplicável'::text, 'Privada'::text]));
  
ALTER DOMAIN public.administracao_v OWNER TO postgres;

```

Fonte: O autor (2025).

Ambas as abordagens da estruturação de dados são válidas, porém a estrutura com tabelas auxiliares proporciona maior flexibilidade para organização dos dados, atualização de valores e integração com ferramentas de controle ou edição de atributos.

5.3 CRIAÇÃO E VALIDAÇÃO DA APLICAÇÃO

Além dos extratos de parte dos resultados apresentados e descritos anteriormente, os scripts desenvolvidos automatizam e asseguram a verificação da conformidade da consistência conceitual e da consistência de domínio de funcionalidades específicas, como:

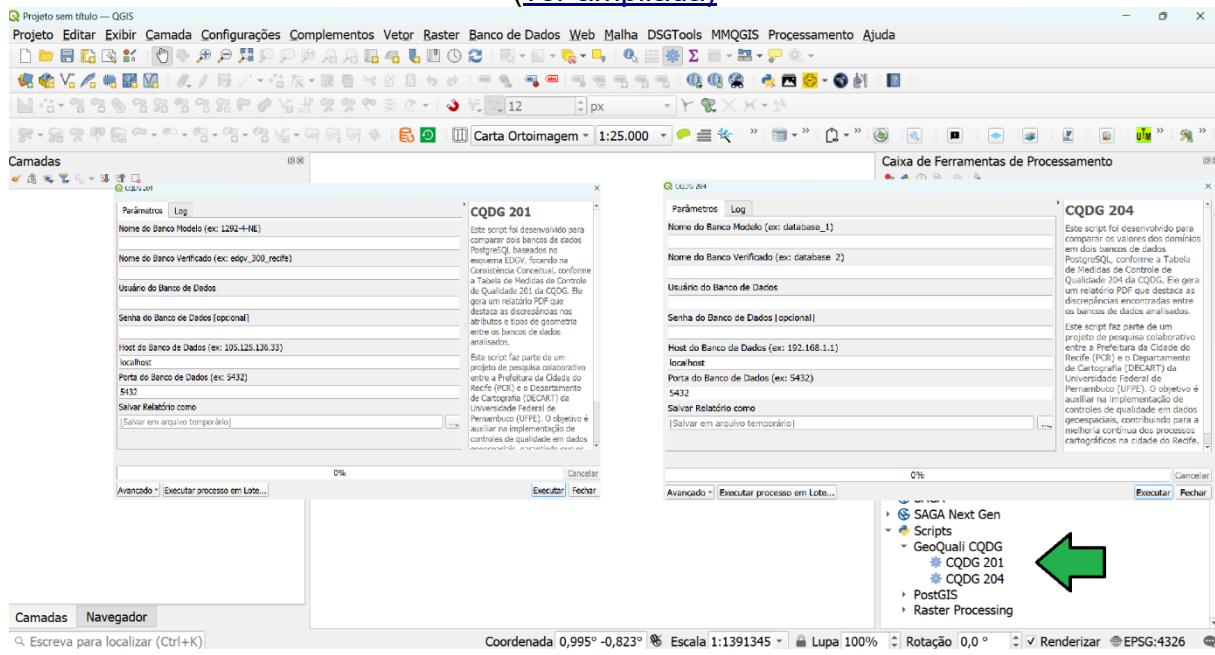
- Comparação de atributos com os definidos na EDGV 3.0;
- Validação de valores conforme domínios predefinidos (codelists);
- Geração de relatórios em tabelas de inconsistências e indicadores binários (conforme/não conforme);

O uso desses scripts proporcionou:

- a) Redução significativa do tempo de análise;
- b) Padronização do processo de verificação; e
- c) Reprodutibilidade metodológica, uma vez que os mesmos parâmetros e funções podem ser reutilizados em outras modelagens.

Todos os códigos estão disponíveis publicamente no GitHub: https://github.com/thiagoarruda/Projeto_PCR/tree/main/scripts para que demais usuários possam usufruir e continuar o desenvolvimento de novas funcionalidades. A Figura 9 apresenta a janela do ambiente do QGIS com a interface do painel Python, onde os scripts podem ser testados e executados pelo usuário.

Figura 9 – Ambiente QGIS com os scripts desenvolvidos para análise automatizada. (ver ampliada)



Fonte: O autor (2025).

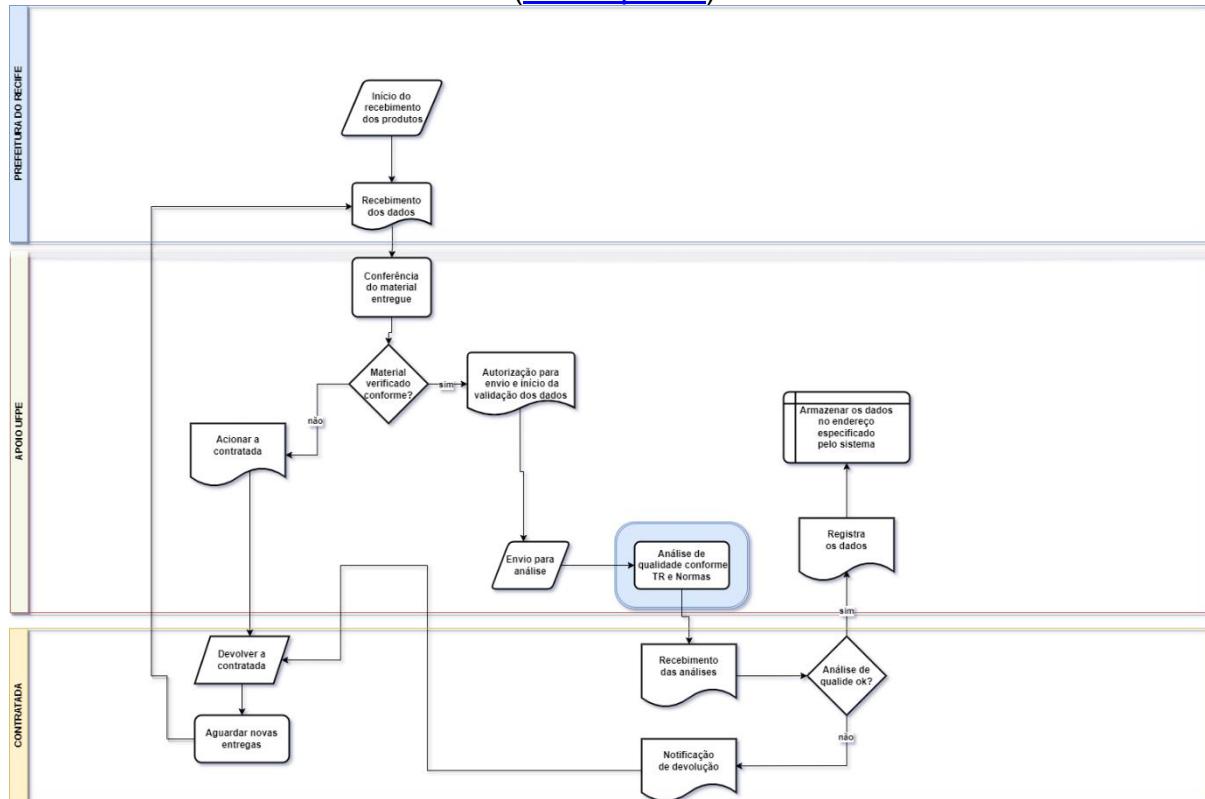
5.3.1 Vantagens da processos automatizados

A adoção de processos automatizados para o controle de qualidade dos dados geoespaciais, como os scripts GeoQuali CQDG desenvolvidos nesta pesquisa auxilia na validação de dados cartográficos utilizados em instituições públicas e privadas. Normalmente o fluxo de trabalho institucional de conferência de produtos e insumos cartográficos referente a uma base de dados geoespacial estruturada em conformidade com a ET-EDGV e ET-CQDG, se apresenta de acordo com a Figura 10.

O fluxograma foi desenhado para ser aplicado em situações de contextos contratuais com empresas fornecedoras de BDG a uma instituição contratante.

Figura 10 – Fluxo institucional de conferência e controle de qualidade de dados geoespaciais.

[\(ver ampliada\)](#)



Fonte: O autor (2025).

Neste modelo “tradicional”, as etapas de recebimento de materiais para verificação de qualidade, validação em conformidade com as especificações aqui discutidas, são desenvolvidas manualmente, isso significa que serão utilizados recursos humanos especializados para verificação de cada etapa, além de elaboração de relatórios em caso de aprovação se tudo estiver em conformidade ou no caso de divergências, todos os apontamentos e descrição das inconsistências deverão ser descritas e apresentadas finalmente no relatório a ser entregue à contratada. Esse tipo de processo é mais suscetível a erros humanos, e obviamente, exige maior tempo de finalização do fluxo de trabalho, além disso, dificulta a rastreabilidade sistemática dos desvios detectados.

Em contraste, a metodologia proposta de verificação da qualidade da consistência lógica automatizada tem as seguintes vantagens: permitir a identificação

direta e padronizada de inconsistências conceituais e de domínio, reduz o tempo de trabalho das operações, reduz a quantidade de operadores altamente qualificados para a validação dos dados, reduz o retrabalho, torna os processos mais transparentes e auditáveis

A análise dos resultados obtidos ao longo desta pesquisa evidencia a importância da organização sistemática dos processos de controle de qualidade em dados geoespaciais. A estruturação metodológica com base na ET-CQDG, aliada à automatização via scripts, demonstrou ser uma abordagem eficaz para identificar inconsistências conceituais e de domínio em diferentes modelagens do padrão EDGV.

A adoção de processos automatizados foi fundamental para lidar com a elevada complexidade das bases de dados analisadas. A existência de centenas de classes e atributos em diferentes modelagens torna inviável a verificação manual, reforçando a necessidade de ferramentas que possibilitem um controle de qualidade escalável, transparente e padronizado.

Um dos principais pontos críticos identificados foi a falta de padronização dos domínios (codelists) entre os bancos de dados analisados. Embora as modelagens estejam estruturadas sobre a EDGV, as adaptações locais, como no caso da modelagem de Aracaju, apresentaram divergências relevantes, tanto em nomenclatura quanto nos valores atribuídos aos atributos codificados. Em muitos casos, domínios foram alterados, ampliados ou utilizados de forma não padronizada, o que dificulta a interoperabilidade entre bases e compromete a reusabilidade dos dados em diferentes sistemas.

A ausência de um repositório nacional unificado que defina e mantenha os domínios controlados (codelists) para cada versão da EDGV contribui para a heterogeneidade entre implementações. Tal lacuna pode ser minimizada por meio de iniciativas voltadas à padronização normativa e tecnológica, possibilitando que entes públicos e privados compartilhem e integrem informações com maior segurança e fidelidade.

Nesse sentido, os resultados alcançados não apenas evidenciam os problemas recorrentes na gestão de dados geoespaciais, como também apontam caminhos para sua mitigação, reafirmando a relevância da qualidade semântica e conceitual como alicerce para sistemas geográficos interoperáveis e confiáveis.

6 CONCLUSÕES E RECOMENDAÇÕES

A verificação da consistência dos dados geoespaciais é uma etapa essencial no processo de produção cartográfica, especialmente em contextos que exigem interoperabilidade, reusabilidade e integração entre diferentes fontes de dados. Este trabalho teve como objetivo avaliar, de forma automatizada, os subelementos de qualidade consistência lógica conceitual e de domínio, em conformidade com a ET-CQDG, com foco em dados vetoriais estruturados segundo a ET-EDGV.

A implementação de scripts automatizados no ambiente QGIS, utilizando a linguagem PyQGIS, permitiu uma análise sistemática e reproduzível de diferentes bases de dados geoespaciais. As comparações entre modelagens revelaram diferenças significativas na estrutura e nos atributos das classes, com destaque para as adaptações locais, como o modelo utilizado como exemplo. A ausência de padronização dos domínios de atributos, em muitos casos, comprometeu a validade dos dados frente aos *codelists* definidos no modelo conceitual da EDGV.

A avaliação da medida CQDG 201 revelou divergências na nomenclatura, presença e tipo dos atributos, especialmente em bancos que implementaram variações ou extensões do modelo original. Por sua vez, a aplicação da medida CQDG 204 demonstrou a relevância da verificação de domínio como etapa essencial para garantir a conformidade dos dados com padrões preestabelecidos, sendo observadas inconsistências em valores não previstos, variações textuais e ausência de uniformidade nos domínios utilizados.

A partir da consolidação dos procedimentos automatizados de verificação, esta pesquisa contribuiu com um fluxo metodológico alinhado às normativas nacionais, promovendo a padronização das etapas de controle de qualidade e a rastreabilidade dos resultados. Esse processo mostrou-se eficiente para apoiar a gestão e a melhoria da qualidade dos dados, sendo aplicável em contextos institucionais diversos.

Dentre os principais achados, destaca-se a necessidade de padronização nacional dos domínios utilizados nos atributos dos dados geoespaciais. Embora a ET-CQDG e a ET-EDGV forneçam diretrizes gerais para o controle de qualidade, observa-se a ausência de uma norma que estabeleça *codelists* unificados para uso em todas as instâncias da Infraestrutura Nacional de Dados Espaciais (INDE). Tal lacuna

compromete a interoperabilidade entre entes federativos e limita o potencial de reuso dos dados.

Nesse sentido, recomenda-se que as instituições responsáveis pela normatização cartográfica no Brasil promovam a ampliação das especificações técnicas vigentes, incluindo a definição de um repositório centralizado de domínios e codelists padronizados.

Além disso, sugere-se o fortalecimento de ferramentas e mecanismos de validação automatizada como parte integrante do fluxo de produção cartográfica, com incentivo ao uso de softwares livres e iniciativas de código aberto.

Como trabalhos futuros, recomenda-se a ampliação da metodologia para contemplar outras medidas de qualidade definidas na ET-CQDG, além da consistência conceitual e de domínio. Sugere-se também o desenvolvimento de ferramentas que tornem o processo mais acessível a diferentes usuários, bem como estudos comparativos com outros modelos e especificações internacionais. Tais iniciativas podem contribuir para o fortalecimento da padronização e para a melhoria contínua da qualidade dos dados geoespaciais.

REFERÊNCIAS

- ARIZA-LÓPEZ, F. J. **Calidad en la Información Geográfica (Introducción)**. En: Experto Universitario en Evaluación de la Información Geográfica (4^a Edición). Máster Universitario en Evaluación y Gestión de la Calidad de la Información Geográfica (1^a Edición). Jaén: Universidad de Jaén, 2011.
- ARIZA-LÓPEZ, F. J.; MORALES, J.; MAESTRE, D.; IGLESIAS, R. **Quality Control of “As Built” BIM Datasets Using the ISO 19157 Framework and a Multiple Hypothesis Testing Method Based on Proportions**. *ISPRS International Journal of Geo-Information*, v. 8, n. 12, p. 569, 2019. <https://doi.org/10.3390/ijgi8120569>.
- ARIZA-LOPEZ, FRANCISCO & REINOSO, J. F. & GARCÍA-BALBOA, JOSÉ & ARIZA-LÓPEZ, ÍÑIGO. (2022). **Quality specification and control of a point cloud from a TLS survey using ISO 19157 standard**. *Automation in Construction*. <https://doi.org/10.1016/j.autcon.2022.104353>.
- ATTAH, R. U.; GARBA, B. M. P.; GIL-OZOUDEH, I.; IWUANYANWU, O. **Leveraging geographic information systems and data analytics for enhanced public-sector decision-making and urban planning**. *Magna Scientia Advanced Research and Reviews*, v. 12, n. 2, p. 152–163, 2024. Disponível em: <https://doi.org/10.30574/msarr.2024.12.2.0191>.
- BATISTA, P.; SILVA, F.; ALMEIDA, C. **Uso de dados geoespaciais na gestão urbana**. *Revista Brasileira de Geografia*, v. 35, n. 2, p. 123-137, 2019.
- 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, jan.-mar. 2015. Disponível em: <http://dx.doi.org/10.1590/S1982-21702015000100005>.
- COMISSÃO NACIONAL DE CARTOGRAFIA - CONCAR. **Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais (ET-EDGV)**. Versão 3.0. Brasília-DF, 2017.
- DAMIANI, M. L. (2018). **Location data quality: A holistic view and a key enabling technology for mobile applications**. Disponível em: <https://doi.org/10.48550/arXiv.1804.00524>.

DEVENDRAN, A. A., & LAKSHMANAN, G. (2014). **A Review On Accuracy and Uncertainty of Spatial Data and Analyses with special reference to Urban and Hydrological Modelling.** ISPRS Annalss of the Photogrammetry, Remote Sensing and Spatial Information Sciences, II-8, 171–178. Disponível em: <https://doi.org/10.5194/isprsaannals-ii-8-171-2014>.

DEVILLERS, R.; BÉDARD, Y.; JEANSOULIN, R.; MOULIN, B. **Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data.** International Journal of Geographical Information Science, v. 21, n. 3, p. 261-282, 2007. Disponível em: <https://doi.org/10.1080/13658810600911879>.

DIRETORIA DE SERVIÇO GEOGRÁFICO - DSG. **Norma da Especificação Técnica para Aquisição de Dados Geoespaciais Vetoriais (ET-ADGV)** – 1^a Edição. Brasília-DF, 2018

DIRETORIA DE SERVIÇO GEOGRÁFICO - DSG. **Norma da Especificação Técnica para Controle de Qualidade de Dados Geoespaciais (ET-CQDG)** – 1^a Edição. Brasília-DF, 2016.

DIRETORIA DE SERVIÇO GEOGRÁFICO - DSG. **Norma da Especificação Técnica para Produtos de Conjunto de Dados Geoespaciais (ET-PCDG)** – 1^a Edição. Brasília-DF, 2014.

ELIAS, E. N. N. **A qualidade da informação geográfica colaborativa no OpenStreetMap: avaliação de feições urbanas em bairros de Salvador (BA).** Geografia (Londrina), Londrina, v. 30, n. 2, p. 255–275, jul. 2021. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/geografia/article/view/41768>. Acesso em: 3 mar. 2025.

FERREIRA, M. B. P.; PAZ, O. L. de S. da. **Qualidade de dados geoespaciais: análise da consistência topológica no mapeamento de cobertura da terra do município de Curitiba - PR.** Revista Cerrados, Montes Claros, v. 16, n. 1, p. 123–147, 2018.

FRANÇA, L. L. S. **Processo de validação da consistência lógica de conjunto de dados geoespaciais vetoriais com abordagem da consistência semântica.**

Dissertação (Mestrado em Ciências Geodésicas e Tecnologias da Geoinformação) – Universidade Federal de Pernambuco, Recife, 2021.

FRANÇA, L. L. S.; PORTUGAL, J. L. **Consistência topológica de dados geoespaciais**. Revista Brasileira de Cartografia, v. 74, n. 2, p. 378–395, 2022.

GALO, M.; CAMARGO, P. O. **O uso do GPS no controle de qualidade de cartas**. 1º Congresso Brasileiro de Cadastro Técnico Multifinalitário, Florianópolis, Santa Catarina, Brasil. p.41- 48. 1994.

GUTIÉRREZ, E.; MARTÍNEZ, M.; RAMOS, J. **Ferramentas de geoprocessamento para usuários não especialistas**. Geoinformática Aplicada, v. 22, n. 1, p. 89-103, 2020.

HAHMANN, S.; BURGHARDT, D.; WEBER, B. **“80% of All Information is Geospatially Referenced”??? Towards a Research Framework: Using the Semantic Web for (In)Validating this Famous Geo Assertion**. Institute for Cartography, Institute for English and American Studies, Dresden University of Technology, 2011.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Base Cartográfica Contínua do Brasil, escala 1:250.000 – BC250, versão 2023**. Rio de Janeiro: IBGE, 2023.

ISO – International Organization for Standardization. ISO 19157:2013. **Geographic Information – Data quality**. Geneva: ISO, 2013.

LAZORENKO-HEVEL, N.; KARPINSKYI, Y.; KIN, D.; LETS, O. **Automation of quality control of digital topographic maps at the scale 1:50 000 of the Main State Topographic Map in Ukraine**. Proceedings of the International Cartographic Association, v. 4, p. 65, 2021. <https://doi.org/10.5194/ica-proc-4-65-2021>.

MARANHÃO, V. C.; CARNEIRO, A. F. T. **Análise de Consistência Lógica conforme padrões de qualidade da norma ISO 19133**. Revista Cartográfica, n. 92, p. 149-171, jan.-jun. 2016. <https://doi.org/10.35424/rcarto.i92.442>.

MENESES, P. R.; GUIMARÃES, A. T. B.; SANTOS, R. R. **Evaluation of the Consistency and Completeness of Geospatial Data Using Quality Indicators**

Based on ISO 19157. ISPRS International Journal of Geo-Information, v. 12, n. 407, 2023.

SATO, Simone Sayuri. **Sistema de controle de qualidade dos processos fotogramétricos digitais para produção de dados espaciais.** 2003. Tese (Doutorado em Engenharia de Transportes) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2003.

SANTOS, A. P. **Avaliação da Acurácia Posicional em Dados Espaciais com o uso de Estatística Espacial.** 110 p. (Magister Scientiae). Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil. 2010.

SERVIGNE, S.; LESAGE, N.; LIBOUREL, T. **Quality components, standards, and metadata.** In: DEVILLERS, R.; JEANSOULIN, R. (ed.). Fundamentals of spatial data quality. London: ISTE, 2006. p. 179-210.

SILVA, P. de L. e. **Metodologia para Avaliação da Usabilidade e Consistência em Infraestruturas de Dados Espaciais.** 2023. Tese (Doutorado em Engenharia Civil) – Universidade Federal de Viçosa, Viçosa, 2023.

APÊNDICE A – CÓDIGO PYTHON: VERIFICAÇÃO CQDG 201

```

from qgis.core import QgsProcessing
from qgis.core import QgsProcessingAlgorithm
from qgis.core import QgsProcessingMultiStepFeedback
from qgis.core import QgsProcessingParameterString
from qgis.core import QgsProcessingParameterFileDestination
from PyQt5.QtCore import QCoreApplication
import psycopg2
from fpdf import FPDF

class CompareDatabases(QgsProcessingAlgorithm):

    def initAlgorithm(self, config=None):
        self.addParameter(QgsProcessingParameterString(
            'dbname1',
            'Nome do Banco Modelo (ex: 1292-4-NE)'
        ))
        self.addParameter(QgsProcessingParameterString(
            'dbname2',
            'Nome do Banco Verificado (ex: edgv_300_recife)'
        ))
        self.addParameter(QgsProcessingParameterString(
            'usuario',
            'Usuário do Banco de Dados'
        ))
        self.addParameter(QgsProcessingParameterString(
            'password',
            'Senha do Banco de Dados',
            optional=True
        ))
        self.addParameter(QgsProcessingParameterString(
            'host',
            'Host do Banco de Dados (ex: 105.125.136.33)',
            defaultValue='localhost'
        ))
        self.addParameter(QgsProcessingParameterString(
            'port',
            'Porta do Banco de Dados (ex: 5432)',
            defaultValue='5432'
        ))
        self.addParameter(QgsProcessingParameterFileDestination(
            'output_file',
            'Salvar Relatório como',
            fileFilter='PDF files (*.pdf)'
        ))

    def processAlgorithm(self, parameters, context, model_feedback):
        feedback = QgsProcessingMultiStepFeedback(1, model_feedback)
        results = {}
        outputs = {}

        dbname1 = self.parameterAsString(parameters, 'dbname1', context)
        dbname2 = self.parameterAsString(parameters, 'dbname2', context)
        usuario = self.parameterAsString(parameters, 'usuario', context)
        password = self.parameterAsString(parameters, 'password', context)
        host = self.parameterAsString(parameters, 'host', context)
        port = self.parameterAsString(parameters, 'port', context)
        output_file = self.parameterAsString(parameters, 'output_file',
        context)

        def connect_to_database(dbname):

```

```

        return psycopg2.connect(
            dbname=dbname,
            user=usuario,
            password=password,
            host=host,
            port=port
        )

feedback.pushInfo('Conectando aos bancos de dados...')
conn1 = connect_to_database(dbname1)
conn2 = connect_to_database(dbname2)

def get_table_structure(cursor, schema='edgv'):
    cursor.execute(f"""
        SELECT table_name, column_name, data_type
        FROM information_schema.columns
        WHERE table_schema = '{schema}'
        AND column_name <> 'geom'
        ORDER BY table_name, ordinal_position
    """)
    return cursor.fetchall()

def get_geometry_type(cursor, schema='edgv'):
    cursor.execute(f"""
        SELECT f_table_name, type
        FROM geometry_columns
        WHERE f_table_schema = '{schema}'
    """)
    return cursor.fetchall()

feedback.pushInfo('Obtendo estruturas dos bancos de dados...')
cur1 = conn1.cursor()
cur2 = conn2.cursor()

estrutura1 = get_table_structure(cur1)
estrutura2 = get_table_structure(cur2)

geometry1 = get_geometry_type(cur1)
geometry2 = get_geometry_type(cur2)

cur1.close()
cur2.close()
conn1.close()
conn2.close()

def organize_structure(estrutura):
    estrutura_dict = {}
    for tabela, coluna, tipo in estrutura:
        if tabela not in estrutura_dict:
            estrutura_dict[tabela] = {}
        estrutura_dict[tabela][coluna] = tipo
    return estrutura_dict

estrutura1_dict = organize_structure(estrutura1)
estrutura2_dict = organize_structure(estrutura2)

geometry1_dict = {table: geom_type for table, geom_type in
geometry1}
geometry2_dict = {table: geom_type for table, geom_type in
geometry2}

```

```

        todas_tabelas =
set(estrutura1_dict.keys()).union(set(estrutura2_dict.keys()))

    def compare_table_attributes(modelo, verificado, todas_tabelas,
geom1, geom2):
        inconsistencias = []
        for tabela in todas_tabelas:
            colunas_modelo = modelo.get(tabela, {})
            colunas_verificado = verificado.get(tabela, {})
            for coluna, tipo_verificado in colunas_verificado.items():
                tipo_modelo = colunas_modelo.get(coluna)
                if tipo_modelo:
                    inconsistencias.append([
                        tabela, coluna,
                        f'{tipo_modelo}', f'{tipo_verificado}',
                        'Sim' if tipo_modelo == tipo_verificado else
'Não'
                    ])
                else:
                    inconsistencias.append([
                        tabela, coluna,
                        'Não presente', f'{tipo_verificado}',
                        'Não'
                    ])
            for coluna, tipo_modelo in colunas_modelo.items():
                if coluna not in colunas_verificado:
                    inconsistencias.append([
                        tabela, coluna,
                        f'{tipo_modelo}', 'Não presente',
                        'Não'
                    ])
            geom_tipo1 = geom1.get(tabela, 'Não presente')
            geom_tipo2 = geom2.get(tabela, 'Não presente')
            inconsistencias.append([tabela, 'Geometria', geom_tipo1,
geom_tipo2, 'Sim' if geom_tipo1 == geom_tipo2 else 'Não'])
        return inconsistencias

    inconsistencias = compare_table_attributes(estrutura1_dict,
estrutura2_dict, todas_tabelas, geometry1_dict, geometry2_dict)

    class PDF(FPDF):
        def header(self):
            self.set_font('Arial', 'B', 12)
            self.cell(0, 10, 'Relatório de Comparação de Atributos -
CQDG 201', 0, 1, 'C')

        def footer(self):
            self.set_y(-15)
            self.set_font('Arial', 'I', 8)
            self.cell(0, 10, f'Page {self.page_no()}', 0, 0, 'C')

        def add_table(self, data, dbname1, dbname2):
            self.set_font('Arial', 'B', 10)
            col_widths = [70, 50, 50, 50, 30]
            row_height = self.font_size * 1.5

            self.cell(col_widths[0], row_height*2, 'Tabela', border=1,
align='C')
            self.cell(col_widths[1], row_height*2, 'Atributo',
border=1, align='C')

```

```

        self.cell(col_widths[2], row_height*2, dbname1, border=1,
align='C')
        self.cell(col_widths[3], row_height*2, dbname2, border=1,
align='C')
        self.cell(col_widths[4], row_height*2, 'Coincide?', border=1, align='C')
        self.ln(row_height * 2)

        self.set_font('Arial', '', 10)
        for row in data:
            for i, item in enumerate(row):
                self.cell(col_widths[i], row_height, str(item),
border=1, align='C')
            self.ln(row_height)

feedback.pushInfo('Gerando PDF de relatório...')
pdf = PDF(orientation='L')
pdf.add_page()
pdf.add_table(inconsistencias, dbname1, dbname2)
pdf.output(output_file)

feedback.pushInfo(f"Relatório PDF gerado com sucesso:
{output_file}")
return results

def name(self):
    return 'compare_databases'

def displayName(self):
    return 'CQDG 201'

def group(self):
    return 'MEDIDAS DE QUALIDADE DOS DADOS'

def groupId(self):
    return 'cqdg_comparisons'

def createInstance(self):
    return CompareDatabases()

def shortHelpString(self):
    return QCoreApplication.translate(
        "CompareDatabases",
        "Este script foi desenvolvido para comparar dois bancos de
dados PostgreSQL baseados no esquema EDGV, "
        "focando na Consistência Conceitual, conforme a Tabela 7 -
Medida conformidade com o modelo de dados. "
        "Ele gera um relatório PDF que destaca as discrepâncias nos
atributos e tipos de geometria entre os bancos de dados analisados.\n\n"
        "Este script faz parte de um projeto de pesquisa colaborativo
entre a Prefeitura da Cidade do Recife (PCR) "
        "e o Departamento de Cartografia (DECART) da Universidade
Federal de Pernambuco (UFPE). O objetivo é auxiliar "
        "na implementação de controles de qualidade em dados
geoespaciais, garantindo que os atributos e geometrias estejam "
        "em conformidade com o modelo de dados.\n\n"
    )

```

APÊNDICE B – CÓDIGO PYTHON: VERIFICAÇÃO CQDG 204

```

from qgis.core import QgsProcessing
from qgis.core import QgsProcessingAlgorithm
from qgis.core import QgsProcessingMultiStepFeedback
from qgis.core import QgsProcessingParameterString
from qgis.core import QgsProcessingParameterFileDestination
from PyQt5.QtCore import QCoreApplication
import psycopg2
from fpdf import FPDF
import re

class CQDG204CompareDomains(QgsProcessingAlgorithm):

    def initAlgorithm(self, config=None):
        self.addParameter(QgsProcessingParameterString(
            'dbname1',
            'Nome do Banco Modelo (ex: database_1)')
        )
        self.addParameter(QgsProcessingParameterString(
            'dbname2',
            'Nome do Banco Verificado (ex: database_2)')
        )
        self.addParameter(QgsProcessingParameterString(
            'usuario',
            'Usuário do Banco de Dados')
        )
        self.addParameter(QgsProcessingParameterString(
            'password',
            'Senha do Banco de Dados',
            optional=True)
        )
        self.addParameter(QgsProcessingParameterString(
            'host',
            'Host do Banco de Dados (ex: 192.168.1.1)',
            defaultValue='localhost')
        )
        self.addParameter(QgsProcessingParameterString(
            'port',
            'Porta do Banco de Dados (ex: 5432)',
            defaultValue='5432')
        )
        self.addParameter(QgsProcessingParameterFileDestination(
            'output_file',
            'Salvar Relatório como',
            fileFilter='PDF files (*.pdf)')
        )

    def processAlgorithm(self, parameters, context, model_feedback):
        feedback = QgsProcessingMultiStepFeedback(1, model_feedback)
        results = {}
        outputs = {}

        # Capturar as entradas do usuário
        dbname1 = self.parameterAsString(parameters, 'dbname1', context)
        dbname2 = self.parameterAsString(parameters, 'dbname2', context)
        usuario = self.parameterAsString(parameters, 'usuario', context)
        password = self.parameterAsString(parameters, 'password', context)
        host = self.parameterAsString(parameters, 'host', context)
        port = self.parameterAsString(parameters, 'port', context)

```

```

        output_file = self.parameterAsString(parameters, 'output_file',
context)

    def connect_to_database(dbname):
        return psycopg2.connect(
            dbname=dbname,
            user=usuario,
            password=password,
            host=host,
            port=port
        )

        feedback.pushInfo('Conectando aos bancos de dados...')
conn1 = connect_to_database(dbname1)
conn2 = connect_to_database(dbname2)

    def fetch_table_names(cursor):
        cursor.execute("SELECT table_name FROM
information_schema.tables WHERE table_schema = 'dominios'")
        return cursor.fetchall()

    def fetch_table_data(cursor, table_name):
        cursor.execute(f"SELECT * FROM dominios.{table_name}")
        return cursor.fetchall()

    def normalize_table_name(name):
        return re.sub(r'^(\w+)_(\w+)$', '\1_\2', name).lower()

    def normalize_value(value):
        return re.sub(r'\s*(\d+)', '\1', value)

cur1 = conn1.cursor()
cur2 = conn2.cursor()

dominios1 = fetch_table_names(cur1)
dominios2 = fetch_table_names(cur2)

    normalized_dominios1 = {normalize_table_name(item[0]): item[0] for
item in dominios1}
    normalized_dominios2 = {normalize_table_name(item[0]): item[0] for
item in dominios2}

    table_data = []

    for norm_name, table_name1 in normalized_dominios1.items():
        table_name2 = normalized_dominios2.get(norm_name)
        if table_name2:
            lista1 = fetch_table_data(cur1, table_name1)
            lista2 = fetch_table_data(cur2, table_name2)

            normalized_lista1 = [(item[0], normalize_value(item[1]))]
            for item in lista1:
                normalized_lista2 = [(item[0], normalize_value(item[1]))]
                for item in lista2:
                    for row1 in lista1:
                        normalized_row1 = (row1[0], normalize_value(row1[1]))
                        if normalized_row1 in normalized_lista2:
                            table_data.append([table_name1, str(row1),
str(row1), 'Sim'])

```

```

        else:
            table_data.append([table_name1, str(row1), 'Não
coincide', 'Não'])
        for row2 in lista2:
            normalized_row2 = (row2[0], normalize_value(row2[1]))
            if normalized_row2 not in normalized_list1:
                table_data.append([table_name2, 'Não presente',
str(row2), 'Não'])
            else:
                possible_match = None
                for other_norm_name, other_table_name2 in
normalized_dominios2.items():
                    if norm_name in other_norm_name or other_norm_name in
norm_name:
                        possible_match = other_table_name2
                        break
                if possible_match:
                    lista1 = fetch_table_data(cur1, table_name1)
                    lista2 = fetch_table_data(cur2, possible_match)

                    normalized_list1 = [(item[0],
normalize_value(item[1])) for item in lista1]
                    normalized_list2 = [(item[0],
normalize_value(item[1])) for item in lista2]

                    table_data.append([table_name1, 'Nome não coincide mas
relacionado', possible_match, 'Possível relação'])

                    for row1 in lista1:
                        normalized_row1 = (row1[0],
normalize_value(row1[1]))
                        if normalized_row1 in normalized_list2:
                            table_data.append([table_name1, str(row1),
str(row1), 'Sim'])
                        else:
                            table_data.append([table_name1, str(row1),
'Não
coincide', 'Não'])
                    for row2 in lista2:
                        normalized_row2 = (row2[0],
normalize_value(row2[1]))
                        if normalized_row2 not in normalized_list1:
                            table_data.append([possible_match, 'Não
presente', str(row2), 'Não'])
                        else:
                            table_data.append([table_name1, 'Não presente no banco
2', 'N/A', 'Não'])

                    for norm_name, table_name2 in normalized_dominios2.items():
                        if norm_name not in normalized_dominios1:
                            table_data.append([table_name2, 'N/A', 'Não presente no
banco 1', 'Não'])

        cur1.close()
        conn1.close()
        cur2.close()
        conn2.close()

    class PDF(FPDF):
        def header(self):
            self.set_font('Arial', 'B', 12)

```

```

        self.cell(0, 10, 'Relatório de Comparação de Domínios - 
CQDG 204', 0, 1, 'C')

    def footer(self):
        self.set_y(-15)
        self.set_font('Arial', 'I', 8)
        self.cell(0, 10, f'Page {self.page_no()}', 0, 0, 'C')

    def add_table(self, data, dbname1, dbname2):
        self.set_font('Arial', 'B', 12)
        col_widths = [60, 90, 90, 40] # Ajustar larguras das
        colunas
        row_height = self.font_size * 1.5
        headers = ["Classes", dbname1, dbname2, "Coincide?"]

        for i, header in enumerate(headers):
            self.cell(col_widths[i], row_height,
header.encode('latin1', 'replace').decode('latin1'), border=1)
            self.ln(row_height)

        self.set_font('Arial', '', 12)
        for row in data:
            for i, item in enumerate(row):
                self.cell(col_widths[i], row_height,
str(item).encode('latin1', 'replace').decode('latin1'), border=1)
                self.ln(row_height)

    pdf = PDF(orientation='L')
    pdf.add_page()
    pdf.add_table(table_data, dbname1, dbname2)
    pdf.output(output_file)

    feedback.pushInfo(f"Relatório PDF gerado com sucesso:
{output_file}")
    return results

    def name(self):
        return 'cqdg_204_compare_domains'

    def displayName(self):
        return 'CQDG 204'

    def group(self):
        return 'GeoQuali CQDG'

    def groupId(self):
        return 'cqdg_comparisons'

    def createInstance(self):
        return CQDG204CompareDomains()

    def shortHelpString(self):
        return QCoreApplication.translate(
            "CQDG204",
            "Este script foi desenvolvido para comparar os valores dos
domínios em dois bancos de dados PostgreSQL, "
            "conforme a Tabela de Medidas de Controle de Qualidade 204 da
CQDG. Ele gera um relatório PDF que destaca "
            "as discrepâncias encontradas entre os bancos de dados
analizados.\n\n"
)

```

"Este script faz parte de um projeto de pesquisa colaborativo entre a Prefeitura da Cidade do Recife (PCR) "
"e o Departamento de Cartografia (DECART) da Universidade Federal de Pernambuco (UFPE). O objetivo é auxiliar "
"na implementação de controles de qualidade em dados geoespaciais, contribuindo para a melhoria contínua dos "
"processos cartográficos na cidade do Recife.\n\n"
"Desenvolvido por Thiago Arruda."

)

APÊNDICE C – RELATÓRIO DE VERIFICAÇÃO CQDG 201

Este apêndice refere-se ao relatório completo de verificação da consistência conceitual (medida CQDG 201), gerado automaticamente pelo script descrito na Seção 5.1 desta dissertação. O relatório contém os resultados da análise do banco de dados EDGV 30 comparado ao banco de dados personalizado EDGV 30 Aracaju.

O relatório completo, com 181 páginas, pode ser acessado no seguinte endereço:

https://github.com/thiagoarruda/Projeto_PCR_Scripts/blob/main/examples/aracaju_edgv300_cqdg201.pdf

Por questões de otimização do volume, o relatório não foi incluído integralmente neste trabalho. Contudo, seu conteúdo está integralmente disponível no repositório público da pesquisa.

Relatório de Comparação de Atributos - CQDG 201

Tabela	Atributo	aracaju	edgv_300	Coincide?
emu_passag em_p	id	integer	Não presente	Não
emu_passag em_p	geometriaapro ximada	boolean	Não presente	Não
emu_passag em_p	matconstr	text	Não presente	Não
emu_passag em_p	nome	character varying	Não presente	Não
emu_passag em_p	operacional	text	Não presente	Não
emu_passag em_p	situacaoespaci al	text	Não presente	Não
emu_passag em_p	situacaofisica	text	Não presente	Não
emu_passag em_p	id_tlograd	integer	Não presente	Não
emu_passag em_p	tipopassagem	text	Não presente	Não
emu_passag em_p	Geometria	POINT	Não presente	Não
enc_torre_en ergia_p	id	integer	uuid	Não
enc_torre_en ergia_p	nome	character varying	character varying	Sim
enc_torre_en ergia_p	geometriaapro ximada	boolean	boolean	Sim
enc_torre_en ergia_p	operacional	text	smallint	Não
enc_torre_en ergia_p	situacaofisica	text	smallint	Não
enc_torre_en ergia_p	ovgd	text	smallint	Não

enc_torre_en ergia_p	alturaestimada	double precision	real	Não
enc_torre_en ergia_p	observacao	Não presente	character varying	Não
enc_torre_en ergia_p	Geometria	POINT	MULTIPOINT	Não

APÊNDICE D – RELATÓRIO DE VERIFICAÇÃO CQDG 204

Este apêndice refere-se ao relatório completo de verificação da consistência de domínio (medida CQDG 204), gerado automaticamente pelo script descrito na Seção 5.2 desta dissertação. O relatório contém os resultados da análise do banco de dados EDGV 2.1.3 comparado ao banco de dados EDGV 3.0.

O relatório completo, com 61 páginas, pode ser acessado no seguinte endereço:

https://github.com/thiagoarruda/Projeto_PCR_Scripts/blob/main/examples/edgv213_x_edgv300_cqdg204.pdf

Por questões de otimização do volume, o relatório não foi incluído integralmente neste trabalho. Contudo, seu conteúdo está integralmente disponível no repositório público da pesquisa.

Relatório de Comparação de Domínios - CQDG 204

Classes	edgv_213	edgv_300	Coincide?
tipotransporte	(0, 'Desconhecido')	(0, 'Desconhecido')	Sim
tipotransporte	(21, 'Passageiros')	Não coincide	Não
tipotransporte	(22, 'Cargas')	Não coincide	Não
tipotransporte	(98, 'Misto')	(98, 'Misto')	Sim
tipotransporte	(999, 'A SER PREENCHIDO')	Não coincide	Não
tipo_transporte	Não presente	(21, 'Passageiro (21)')	Não
tipo_transporte	Não presente	(22, 'Carga (22)')	Não
tipo_transporte	Não presente	(9999, 'A SER PREENCHIDO (9999)')	Não
tipoestrut	Não presente no banco 2	N/A	Não
tipooperativo	Não presente no banco 2	N/A	Não
matconstr	(0, 'Desconhecido')	(0, 'Desconhecido')	Sim
matconstr	(1, 'Alvenaria')	(1, 'Alvenaria')	Sim
matconstr	(2, 'Concreto')	(2, 'Concreto')	Sim
matconstr	(3, 'Metal')	(3, 'Metal')	Sim
matconstr	(4, 'Rocha')	(4, 'Rocha')	Sim
matconstr	(5, 'Madeira')	(5, 'Madeira')	Sim
matconstr	(6, 'Arame')	Não coincide	Não
matconstr	(7, 'Tela ou Alambrado')	Não coincide	Não
matconstr	(8, 'Cerca viva')	Não coincide	Não