



QUALIDADE DOS DADOS NA COLETA AUTOMATIZADA DE PLACAS DE VEÍCULOS EM TRÂNSITO

Alexandre Augusto Biz¹
Daniel Henrique Rodrigues²

Resumo: *A qualidade dos dados em grandes corporações é fundamental para o gerenciamento eficiente das informações, essenciais à tomada de decisões. Na gestão de segurança pública em Santa Catarina, introduziu-se o uso de dispositivos de reconhecimento automático de placas veiculares, integrados para gerar alertas às forças policiais. Analisa-se, então, a qualidade dos dados capturados, com foco em dimensões relevantes e nos desafios de precisão nas leituras. Após revisão da literatura e seleção das dimensões estudadas, realiza-se a análise de uma amostra para avaliar sua qualidade.*

Palavras-chave: qualidade; dados; dimensões; reconhecimento; placa veicular.

Abstract: *The quality of data in large corporations is essential for efficient information management, crucial for decision-making. In public security management in Santa Catarina, automatic license plate recognition devices were introduced and integrated to generate alerts for police forces. This study analyzes the quality of captured data, focusing on relevant dimensions and accuracy challenges in readings. After a literature review and selection of study dimensions, a sample analysis is conducted to assess its quality.*

Keywords: quality; data; dimensions; recognition; license plate.

Resumen: *La calidad de los datos en grandes corporaciones es fundamental para la gestión eficiente de la información, esencial para la toma de decisiones. En la gestión de seguridad pública en Santa Catarina, se introdujeron dispositivos de reconocimiento automático de matrículas, integrados para generar alertas para las fuerzas policiales. Este estudio analiza la calidad de los datos capturados, enfocándose en dimensiones relevantes y en los desafíos de precisión en las lecturas. Tras una revisión de la literatura y la selección de las dimensiones estudiadas, se realiza un análisis de una muestra para evaluar su calidad.*

Palabras clave: calidad; datos; dimensiones; reconocimiento; matrícula.

1. INTRODUÇÃO

O estado de Santa Catarina registra em agosto de 2024 uma frota de 6.096.825 (seis milhões noventa e seis mil oitocentos e vinte e cinco) de veículos automotores efetivamente

¹ PPGEGC – UFSC Florianópolis – Brasil. ORCID: <https://orcid.org/0000-0003-3235-9328> email: alexandre.biz@ufsc.br

² PPGEGC – UFSC, Florianópolis – Brasil. ORCID <https://orcid.org/0009-0009-8962-7638> e-mail: danielh6640@gmail.com



em circula o por sua malha de infraestrutura rodovi ria, seja em n vel municipal, estadual ou federal, conforme dados dispon veis no *site* do Departamento Estadual de Tr nsito de Santa Catarina (DETRAN/SC) atrav s do endere o www.detran.sc.gov.br/estatisticas-veiculos-transparencia/ e toda a gest o de emplacamentos, transfer ncias, licenciamentos e d bitos   efetivado atrav s do sistema de informa es denominado DetranNet, sustentado pelo Centro de Inform tica e Automa o do Estado de Santa Catarina:

O Portal de Servi os DetranNet proporciona: seguran a, operadores, autoriza es, auditoria de opera es do sistema, estat sticas da frota em tempo real, escolha de placas, interc mbio de dados, Integra o com o SERPRO (integra o entre o banco de dados do DETRAN e a base nacional do sistema RENAVAM), integra o com o Sistema Nacional Gravames, consulta consolidada de ve culos, consultas anal ticas com refer ncia a um determinado ve culo de pessoa f sica ou jur dica, dossi  do ve culo, dados cadastrais, caracter sticas cadastrais mais restritas, d bitos, multas, infra es em autua o,  ltimo processo, hist rico de cobran a, simula o dos lan amentos do exerc cio seguinte, pagamentos inconsistentes, hist rico de infra es notificadas, hist rico de multas, hist rico de publica o em editais, avisos de recebimento, recurso de infra o, hist rico de processos, cadeia sucess ria, hist rico de impedimentos e hist rico de CRLV. (CIASC, 2024)³

Assim, sempre que h  uma consulta em uma abordagem policial, por exemplo, um operador pode verificar uma placa e constatar a condi o de um determinado ve culo quanto a estas restri es. Entretanto, diante da clara impossibilidade de um agente policial proceder a m ltiplas consultas veiculares simultaneamente frente   grande massa em circula o e tamb m diante da import ncia de obter consci ncia situacional quanto   movimentaa o dos ve culos em circula o, adotam-se tecnologias para automatizar em larga escala a leitura e verifica o da condi o dos ve culos em tr nsito nas vias urbanas e rodovi rias.

A maneira mais usual para esta coleta e posterior verifica o de ve culos   atrav s do recurso tecnol gico denominado ALPR - *Automatic License Plate Recognition* (Reconhecimento Autom tico de Placas de Licen a), o qual se mostrou uma ferramenta importante em diferentes aplica es, uma vez que   um dispositivo com um sensor digital que reconhece os caracteres e n meros da placa de um ve culo a partir da imagem gerada (Du, Ibrahim, Shehata & Badawy, 2013). A partir desta imagem, um algoritmo executa quatro etapas: 1) captura da imagem de um ve culo, 2) detec o da placa, 3) segmentaa o dos caracteres e 4) reconhecimento dos caracteres, de modo que existem diferentes abordagens para efetivamente se extrair a placa de um ve culo a partir de uma imagem digital, como Rede Neural Artificial (ANN - *Artificial Neural Network*), Rede Neural Probabil stica (PNN - *Probabilistic Neural*

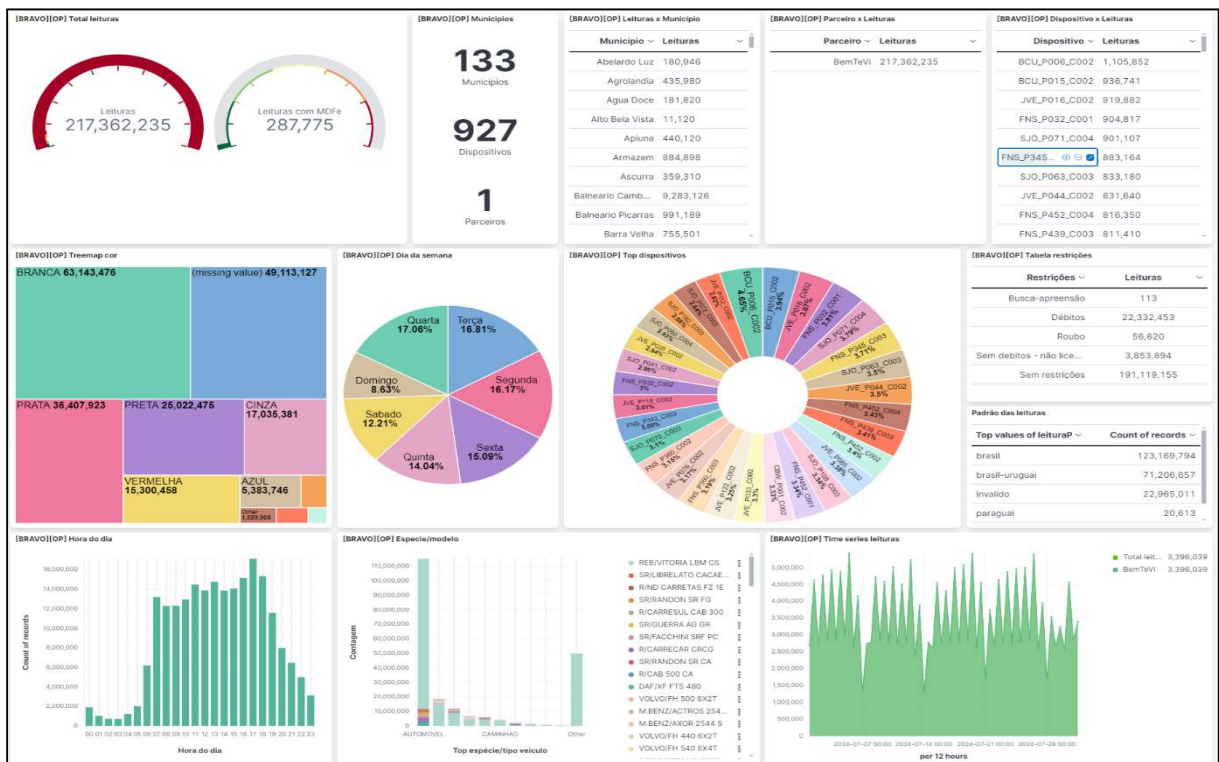
³ <https://www.ciasc.sc.gov.br/detrannet/> - Acesso em 03 de agosto de 2024.

Network), Reconhecimento Ótico de Caracteres (OCR – *Optical Character Recognition*), *Feature salient*, MATLAB, Janela Deslizante de Concentração (SCW – *Sliding Concentrating Window*), Rede Neural de Retropropagação (BPNN – *Back Propagation Neural Network*), dentre várias outras (Patel, Shah & Patel, 2013).

Patel et al. (2013) destacam que o ALPR pode ser explorado também para identificação do proprietário do veículo, identificação do seu modelo, controle de tráfego, controle de velocidade, rastreamento e localização, além de vários benefícios como fiscalização de atividades suspeitas com disponibilidade imediata de informações, sendo incomparável à busca destes detalhes manualmente frente ao benefício proporcionado pelo ALPR.

Neste norte, no ano de 2022 a Secretaria de Estado da Segurança Pública de Santa Catarina (SSP/SC) adquiriu um quantitativo de 1.625 (mil seiscentos e vinte e cinco) equipamentos ALPR para o estabelecimento de uma rede de coleta de informações de veículos em circulação a nível estadual para integração ao Sistema BRAVO, uma plataforma para o recebimento e gestão dos bilhetes de leitura de placas de veículos em circulação, cuja denominação decorre de um acrônimo de “*Bot de Recepção de Alertas e Visualização Operacional*” e aglutina basicamente um conjunto de dados referentes à data, hora, geolocalização (latitude e longitude) e os caracteres da placa identificada na imagem capturada.

Figura 1 – Dashboard operacional do Sistema BRAVO



Fonte: Sistema BRAVO, acessado pelo autor em 02 de agosto de 2024.



Conforme se extrai da Figura 1, na qual está disposta o *dashboard* do Sistema BRAVO, no mês de agosto estavam implantados um total de 927 dispositivos ALPR, os quais promoveram a captura de 217.362.235 (duzentos e dezessete milhões trezentos e sessenta e dois mil duzentos e trinta e cinco) imagens de veículos em circulação com a respectiva identificação da placa, de modo que a Administração Pública, em especial a Secretaria de Estado da Segurança Pública – mantenedora da plataforma – deve promover a avaliação da qualidade dos dados e informações gerados através do emprego da tecnologia ALPR, uma vez que incorreções podem ocasionar em tomadas de decisões não assertivas por parte das forças de segurança, como por exemplo a abordagem policial de um veículo com registro de furto por uma leitura inadequada, sem que tal veículo tenha tal restrição.

Embora não haja consenso sobre a distinção entre “qualidade dos dados” e “qualidade da informação”, existe uma tendência para utilizar a “qualidade dos dados” para se referir a questões técnicas e a “qualidade da informação” para se referir a questões não técnicas. (Madnick, Wang, Lee & Zhu, 2009)

Desta forma, o presente trabalho se propõe a analisar a qualidade dos dados gerados decorrentes das capturas efetuadas pelos dispositivos ALPR adquiridos pela Administração Pública catarinense em consonância com o indicado em bibliografia referencial na área de abrangência, identificando eventuais lacunas passíveis de melhoria para otimização dos processos e procedimentos, através do estudo de uma amostra das capturas relacionadas às restrições de furto e/ou roubo em determinada localidade de Florianópolis, capital de Santa Catarina, com foco à correta identificação da placa capturada.

2. METODOLOGIA

Para o presente caso, promoveu-se uma revisão de escopo da literatura disponível sobre qualidade de dados no âmbito da internet das coisas (IoT – *Internet of Things*), partindo da premissa que um dispositivo ALPR, na condição de componente integrante de uma rede de sensores de imagem com inteligência embarcada, conectados em um sistema centralizado para processamento e análise dos dados coletados, é um dispositivo que se enquadra como IoT.

A revisão de escopo é definida por Costa, Fontanari e Zoltowski (2022) como a melhor escolha para mapear a abrangência de um determinado assunto ou um levantamento amplo da literatura, bem como de conceitos e características específicas. As revisões de escopo também



forneem informações úteis para tomada de decisão acerca de um conceito e seu estudo ao longo do tempo, inclusive para futuras revisões sistemáticas (PETERS et al., 2020).

Para identificar as mais recentes abordagens envolvendo o tema qualidade do dado (DQ – *Data Quality*), o enfoque se concentrará em quais as principais dimensões são as comumente analisadas especificamente em plataformas envolvendo IoT, de modo que promover-se-á uma listagem envolvendo 1) autores, 2) ano, 3) dimensões de destaque e 4) contatações. Não serão abordados estudos que não envolvam as dimensões da qualidade do dado dentro do contexto de IoT. Desta forma, a presente revisão de escopo tem como pergunta de pesquisa a seguinte indagação: quais as dimensões da qualidade da informação mais importantes no contexto de Internet das coisas (IoT – *Internet of Things*) de acordo com os estudos promovidos nos últimos 5 anos?

Utilizar-se-á como norteadora da presente pesquisa as orientações do guia *Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews* (PRISMA-ScR) disponível em <http://www.prisma-statement.org/> e fruto do trabalho de um grupo de especialistas internacionais que identificou critérios mínimos para serem abordados em uma publicação científica de qualidade, de modo que o *checklist* PRISMA é um dos instrumentos com maior reconhecimento global (Lockwood, Dos Santos & Pap, 2019).

A identificação de produção científica utilizou-se do emprego das seguintes palavras-chave e operadores lógicos para composição da estratégia de busca: *TITLE-ABS-KEY ("DATA QUALITY") AND ("DIMENSION*") AND ("IOT" OR "INTERNET OF THINGS") AND ("REVIEW")* na base de dados da Scopus e da *Web of Science* (WoS) em busca de estudos que objetivamente possam delinear as dimensões relacionadas à qualidade dos dados dentro do contexto de um ecossistema de IoT. Adotou-se o intervalo dos últimos cinco anos, em razão da implementação do conceito de IoT ser algo recente, bem como a intenção de identificar as considerações mais recentes no contexto de qualidade de dados, uma vez que os estudos iniciais remontam à virada da década de 80 para 90. As consultas às bases de dados da Scopus e *Web of Science* foram realizadas em 05 de fevereiro de 2024.

O processo de seleção levou em conta na etapa de triagem, após a verificação de artigos duplicados, a aplicação dos filtros e análise do título e palavras-chave. Na fase de elegibilidade, foram realizadas as leituras de resumos, introdução e conclusão e, posteriormente, a leitura integral de cada estudo. Como requisito para seleção levou-se em conta (1) aderência do tema ao contexto de qualidade dos dados em ambientes IoT, (2) possuir alguma abordagem às dimensões desejadas quanto à qualidade dos dados no contexto de IoT; (3) expressar

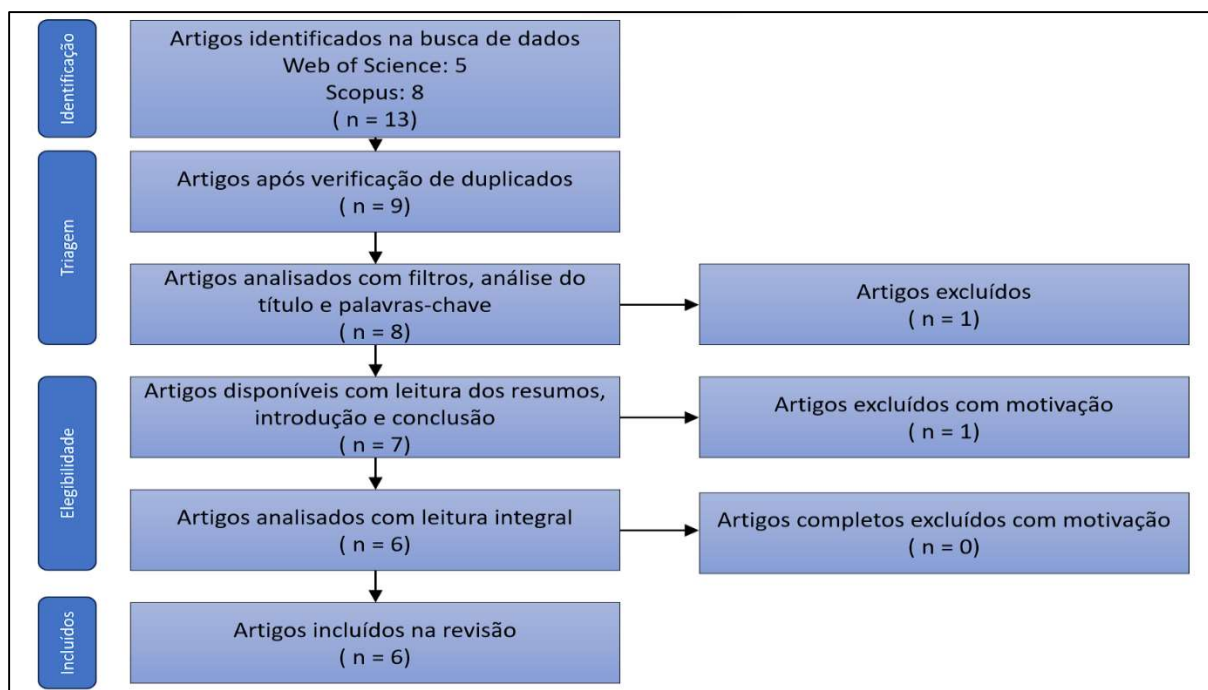
objetivamente quais os requisitos, preceitos e aplicabilidade das dimensões relacionadas (4) os estudos decorrentes da *string* de pesquisa serem de acesso aberto, com o decorrente descarte daqueles com acesso restrito ou indisponível.

Finalmente, definido os principais conceitos no que concerne às dimensões da qualidade de dados em ambientes de IoT, proceder-se-á análise da pertinência e emprego destas aos dados coletados pelos dispositivos ALPR que subsidiam a operação do Sistema BRAVO.

3. RESULTADOS

Dos 13 resultados inicialmente obtidos através da estratégia de busca, ao fim, foram selecionados 6 trabalhos, todos em inglês – provenientes das bases Scopus e *Web of Science*, cuja seleção compreendeu as etapas constantes no fluxograma apresentado a seguir:

Figura 2 – Diagrama do processo de inclusão e exclusão dos estudos.



Fonte: Elaborado pelo autor conforme protocolo PRISMA.

Dos artigos incluídos para revisão foram extraídos os seguintes dados para tabulação: autores, ano, dimensões de qualidade de dados abordadas e as constatações da revisão pesquisada. Embora apenas 6 estudos tenham sido selecionados, alguns englobam a revisão de diversos estudos no mesmo eixo temático, conforme pode ser sintetizado na Tabela 1:



Tabela 1 – Constatações dos estudos sobre revisões sobre qualidade dos dados em *Internet of Things* (n = 6)

Autor(es)	Ano	Dimensões de <i>Data Quality</i>	Constatações
ROMERO, Marcos Lima; SUYAMA, Ricardo	2023	Precisão Completude Consistência Credibilidade Atualidade Acessibilidade Conformidade Confidencialidade Eficiência Precisão Rastreabilidade Compreensibilidade Disponibilidade Portabilidade Recuperabilidade	<p>Os autores analisam a qualidade dos dados em aplicações da Internet das Coisas (IoT) de 2016 a 2022, examinando métodos, conjuntos de dados, padrões e segmentos de casos de uso com base na normativa ISO 25012 e dividindo as aplicações de IoT em duas categorias: aplicações com dispositivos em massa e aplicações críticas à missão, destacando diferentes demandas de qualidade. Após revisar mais de 3200 artigos, 124 foram selecionados para análise detalhada, mostrando um foco crescente na dimensão da credibilidade dos dados: o nível de confiança que se pode ter na precisão e confiabilidade dos dados coletados e processados. Em contextos onde a tomada de decisão é crítica, a credibilidade dos dados é fundamental para garantir que as decisões baseadas nestes dados sejam válidas e seguras. Isso envolve aspectos como a autenticidade, precisão e integridade dos dados, bem como a confiança nos métodos de coleta e processamento desses dados. Ao fim, aventa a necessidade de pesquisas futuras e identificar oportunidades em dimensões de qualidade de dados não mencionadas, novos métodos usando tecnologias emergentes e casos de uso subexplorados.</p>
ALWAN, Ahmed Abdulhasan; CIUPALA, Mihaela Anca; BRIMICOMBE, Allan J.; GHORASHI, Seyed Ali; BARAVALLE, Andres; FALCARIN, Paolo	2022	Precisão Consistência Completude Pontualidade	<p>Os autores abordam os desafios relacionados à qualidade dos dados em sistemas ciber-físicos (CPS - <i>Cyber-Physical Systems</i>) de grande escala, especialmente em aplicações de cidades inteligentes, através de uma revisão sistemática da literatura para investigar os principais desafios da qualidade dos dados e as técnicas mais comuns usadas para enfrentá-los após a identificação de 542 estudos e a seleção final de 60 para revisão completa. Constatou-se que há um histórico de esforços significativos para o gerenciamento da qualidade dos dados em aplicações de CPS de grande escala em cidades inteligentes, mas ainda há necessidade de uma solução abrangente de gerenciamento da qualidade dos dados que possa detectar erros em medições de nós de sensor e lidar com as principais dimensões da qualidade dos dados, como precisão, pontualidade, completude e consistência.</p>
ZHANG, Lina; JEONG, Dongwon; LEE, Sukhoon	2021	Precisão Consistência Completude Pontualidade Confiança Volume de Dados	<p>Os autores discutem a importância da qualidade dos dados no contexto do IoT (Internet das Coisas), destacando que as exigências de qualidade variam significativamente dependendo da aplicação ou da organização. Abordam uma ampla gama de metodologias e <i>frameworks</i> destinados à definição, avaliação e melhoria da qualidade dos dados, considerando a diversidade dos tipos de dados, definições de qualidade, dimensões e métricas, e como escolher as dimensões de avaliação adequadas. O objetivo é fornecer um guia que auxilie na escolha da técnica de gestão de qualidade de dados de IoT mais apropriada, mediante a comparação de <i>frameworks</i> e metodologias existentes, bem como padrões internacionais relacionados.</p>



MANSOURI, Taha; MOGHADAM, Mohammad Reza Sadeghi; MONSHIZADEH, Fatemeh; ZARERAVASAN, Ahad	2021	Precisão Compleitude Pontualidade Utilidade Volume de dados Concordância	Os autores abordam questões de qualidade de dados na Internet das Coisas (IoT) e propõe soluções para esses problemas promovendo uma revisão sistemática da literatura de 2000 a 2020 mediante a análise de 58 artigos para identificar dimensões e problemas de qualidade de dados no IoT e suas categorizações. O estudo oferece uma classificação de características de dados do IoT, examina a segurança dos dados como um problema importante de qualidade de dados e sugere soluções potenciais para superar problemas de segurança no IoT, propondo um novo campo de pesquisa para examinar a qualidade dos dados no contexto do IoT. Destaca também a importância da qualidade dos dados para o sucesso das aplicações IoT, identificando desafios como a integração de dados heterogêneos e a gestão da veracidade dos dados, propondo a adoção de técnicas avançadas de processamento de dados e normas rigorosas para garantir a qualidade dos dados. O estudo também sugere a criação de uma estrutura abrangente para avaliação e melhoria contínua da qualidade dos dados no ecossistema IoT, visando otimizar a eficiência operacional e a tomada de decisões baseada em dados.
BYABAZAIRE, John; O'HARE, Gregory; DELANEY, Declan	2020	Precisão Consistência Compleitude Pontualidade	O conceito deste artigo é examinar os desafios e oportunidades relacionados à qualidade dos dados e à confiança no compartilhamento de dados dentro do contexto da Internet das Coisas (IoT). Destaca a importância crítica dos dados de qualidade para resultados de modelos e decisões de negócios na era do <i>Big Data</i> e IoT. Com o aumento do número de dispositivos IoT conectados, torna-se desafiador avaliar a qualidade dos dados usando métodos de avaliação tradicionais. O artigo introduz uma nova perspectiva de aplicativos IoT que se beneficiam do compartilhamento de dados entre diferentes domínios para criar novas aplicações, enfatizando que a qualidade aceitável dos dados compartilhados é essencial para realizar os benefícios dessas aplicações. Enfatiza, ainda, a necessidade de construir confiança no compartilhamento de dados entre dispositivos IoT, apontando que a confiança é um componente essencial para garantir a cooperação e o intercâmbio eficaz de informações entre diferentes entidades e sistemas. Destaca-se a ideia de que melhorar a qualidade dos dados e estabelecer um alto nível de confiança pode impulsionar significativamente o desenvolvimento e a adoção de aplicações IoT, oferecendo melhores insights, otimizando operações e permitindo decisões mais assertivas.
LIU, Caihua; NITSCHKE, Patrick; WILLIAMS, Susan P.; ZOWGHI, Didar	2019	Precisão Tempestividade Compleitude Utilidade Volume de dados Concordância	Os autores promovem uma revisão sistemática da literatura sobre qualidade dos dados no contexto da Internet das Coisas (IoT) desde o surgimento do termo IoT em 1999 até 2018. São examinados 45 estudos empíricos para identificar temas de pesquisa sobre qualidade de dados em IoT, estabelecendo ligações entre dimensões de qualidade dos dados, manifestações de problemas de qualidade dos dados e métodos utilizados para medir a qualidade dos dados. A revisão sugere novas áreas para investigação e destaca implicações para praticantes na definição e medição da qualidade dos dados em IoT. Aborda também como a qualidade dos dados é crucial para o desenvolvimento sustentável de aplicações de IoT, considerando a variedade e a quantidade de dados gerados por dispositivos conectados, bem como a importância de métodos eficazes para a medição da qualidade dos dados, de modo a garantir a confiabilidade, a precisão e a utilidade dos dados coletados em ambientes de IoT, o que se é essencial para a tomada de decisão, otimização de processos e geração de insights valiosos em diversos campos de aplicação da IoT.



4. DISCUSSÃO

Preliminarmente é importante destacar uma breve conceituação e no que consiste o termo “qualidade”, conforme destacado por Alwan et al. (2022), e como esta questão de qualidade se relaciona quando falamos de dados, abordados anteriormente:

De acordo com a Organização Internacional de Normalização (ISO), qualidade, em geral, é a "totalidade das características de uma entidade que se referem à sua capacidade de satisfazer necessidades declaradas e implícitas". Em comparação, a Qualidade dos Dados é definida como a adequação dos dados para o propósito do uso pretendido ou sua conformidade com os requisitos. Esta definição destaca que os dados que atendem muito bem a algumas expectativas, especificações ou padrões pré-definidos são considerados dados de alta qualidade adequados para uso em uma aplicação específica. (Alwan et al., 2022)

É recordado por Zhang et al. (2021) que Richard Y. Wang, um dos precursores da temática, define a qualidade de dados como “dados que são adequados para uso pelos consumidores de dados”. Contudo, a qualidade dos dados deve considerar dois eixos: primeiramente a nuance prática decorrente da perspectiva do usuário, enfatizando a satisfação deste e também dos produtores e gestores de dados; e, em segundo lugar, a partir de avaliações com perspectiva orientada a sistemas, considerando a qualidade dos dados como um conceito abrangente e multidimensional. Ainda assim, Byabazaire, O’hare e Delaney (2020) alertam que é algo subjetivo e dependente do caso de uso e área de domínio da aplicação.

Romero e Suyama (2023) destacam que o primeiro uso do termo *Internet of Things* (IoT) – Internet das Coisas – é atribuído à Kevin Ashton para se referir a objetos do cotidiano que poderiam ser conectados em rede. Desde então, com a evolução da infraestrutura de conectividade, processadores embarcados e protocolos de comunicação, as aplicações de IoT tiveram um salto, de modo que passou a ser tópico multidisciplinar envolvendo áreas do conhecimento como redes sem fio, sensores, sistemas embarcados, eletrônicos de baixo consumo energético, análise de dados, aprendizado de máquina, computação em borda e na nuvem, dentre outros. Zhang et al. (2021) conceituam a qualidade de dados em IoT como dados coletados por dispositivos – que geralmente monitoram uma variável de interesse no mundo físico – apropriados para os usuários de IoT fornecerem serviços ubíquos.

Após revisar mais de 3200 artigos e selecionar 124 artigos para análise, Romero e Suyama (2023) destacam um interesse crescente nas dimensões da qualidade de dados (DQ – *Data Quality*) em diversos segmentos da Internet das Coisas (IoT) e nos métodos usados para



abordar essas questões, de modo que as dimensões de DQ mais citadas incluem credibilidade e confidencialidade, das dimensões que adaptam da norma ISO 25012.

Alwan et al. (2022), ao definir especificamente o que é Qualidade dos Dados (DQ – *Data Quality*) menciona que mais de 200 dimensões foram criadas ao longo dos anos, de modo que elas se concentram em quatro dimensões principais:

A qualidade dos dados pode ser quantificada, medida e monitorada usando um conjunto de parâmetros ou indicadores dependentes de contexto conhecidos como Características ou Dimensões da Qualidade dos Dados. Mais de 200 dimensões de qualidade de dados foram introduzidas desde os anos oitenta. No entanto, essas dimensões podem ser categorizadas em quatro dimensões principais de qualidade de dados: **precisão, completude, pontualidade e consistência**, que são, tipicamente, associadas aos requisitos de qualidade de dados e mapeadas para definir critérios de avaliação da qualidade dos dados. (Alwan et al., 2022, grifo nosso)

Neste norte, dos 60 artigos revisados por Alwan et al. (2022), 53% tem como destaque a dimensão da precisão (*accuracy*), 14% a dimensão da completude (*completeness*) e da pontualidade (*timeliness*), 12% da consistência (*consistency*) e 7% outras dimensões diversificadas. As mesmas dimensões – precisão, completude, consistência e pontualidade – também são as destacadas por Byabazaire et al. (2020), embora elenquem um total de 15 diferentes em seu trabalho e constatando que a qualidade dos dados é essencial para tomadas de decisão nos negócios, sendo que sua má qualidade pode ter impactos negativos significativos. Fatores que podem degradar a qualidade dos dados em contextos de IoT incluem escala de implantação, restrições de recursos, falhas de segurança, vulnerabilidade de segurança, preservação de privacidade e processamento (Byabazaire et al., 2020).

Também são destacadas estas dimensões – precisão, completude, consistência e pontualidade – por Zhang et al. (2021) como elementos básicos da qualidade da informação, de modo que os autores adicionam outras duas dimensões quando se trata de IoT, vindo a serem relatados os problemas havidos com estas dimensões. Romero e Suyama (2023) sugerem que pesquisas futuras explorem dimensões e segmentos de DQ que receberam menos atenção em pesquisas anteriores, além de investigar novos métodos e técnicas que possam abordar tal desiderato e estudos focados no desenvolvimento e implementação de *frameworks* abrangentes que abordem holisticamente as questões de DQ em diferentes domínios de aplicações de IoT.

Desta forma, para análise do contexto de qualidade de dados em relação aos dispositivos ALPR que subsidiam o Sistema BRAVO, serão abordadas as quatro dimensões que englobam precisão, completude, pontualidade e consistência, as quais serão preliminarmente definidas:



Precisão refere-se ao nível em que os dados têm a propriedade de representar corretamente o valor verdadeiro das propriedades esperadas de um conceito ou evento em um determinado contexto (Zhang et al., 2021) e definida como o grau em que os dados estão livres de erros, corretos ou sem defeitos (Byabazaire et al., 2020). Tal definição é corroborada por Mansouri et al. (2021), que definem a precisão como o grau em que os dados possuem atributos que representam corretamente o verdadeiro valor do atributo pretendido de um conceito ou evento em um contexto específico de uso", constituindo-se de medida pela proximidade dos valores capturados dos pontos de dados com seus valores originais, estando intimamente relacionada à correção dos dados.

Completude refere-se ao grau em que os dados apresentam valores para todos os atributos esperados (Romero & Suyama, 2023) ou o grau em que os dados não estão ausentes e possuem uma amplitude e profundidade suficientes para a tarefa pretendida, sendo a dimensão usada para descrever o problema de valores faltantes devido a falhas ou defeitos nos sensores, revestindo-se do nível em que os dados têm valores de todas as propriedades pretendidas e instâncias de entidades associadas no ambiente de uso especificado (Byabazaire et al., 2020).

Os problemas associados à completude e mencionados por Liu et al. (2019) incluem dados faltantes, que podem ser causados por ineficiências de sensores, problemas de comunicação, ou interceptação e manipulação de dados em decorrência de ataques. A falta de atualizações pode afetar a obtenção dos dados necessários, limitando a sua disponibilidade.

Pontualidade consiste na tempestividade dos dados em relação ao contexto da aplicação e discrepâncias entre o carimbo de data/hora (*timestamp*) de origem dos dados e o carimbo de data/hora atual (Zhang et al., 2021), no grau em que os dados são suficientemente atualizados para a tarefa em questão ou grau em que uma observação para um objeto é atualizada em um momento desejado de interesse (Byabazaire et al., 2020). Os problemas de qualidade dos dados (DQ – *Data Quality*) manifestados nessa dimensão foram a falta de atualizações (*missing updates*) e a baixa taxa de dados (*low data rate*) (Liu et al., 2019).

Consistência refere-se ao grau em que os dados estão livres de contradições e coerentes (Alwan et al., 2022), ou ao nível em que valores de um atributo obedecem às restrições definidas, como estar em um intervalo específico ou evitar valores negativos em sensores de umidade. A verificação de regras e restrições pode servir como métricas de consistência, descrevendo dados sem contradições. Zhang et al. (2021) exemplifica com dois sensores de temperatura no mesmo local, mas com leituras divergentes, o que indica baixa consistência.



Para análise dos quesitos propostos de Qualidade dos Dados (DQ – *Data Quality*), foram extraídos dados do Sistema Bravo entre os dias 20 de dezembro de 2023 e 20 de janeiro de 2024, o que compreende um total de 245.848.605 (duzentos e quarenta e cinco milhões oitocentas e quarenta e oito mil seiscentos e cinco) de coletas de placas de veículos em trânsito provenientes de 1.878 dispositivos ALPR implantados em todo o estado de Santa Catarina.

Diante da clara impossibilidade de análise do total de leituras, optou-se por obter como amostra as coletas procedidas na Rodovia João Carlos Daux (SC-401), reconhecidamente a rodovia estadual mais movimentada de Santa Catarina, de oito dispositivos dispostos em frente à base da Polícia Militar Rodoviária, conforme Tabela 2:

Tabela 2 – Alertas decorrentes das capturas realizadas na SC-401

LOCAL	ID DISPOSITIVO	CAPTURAS	ALERTAS	PERCENTUAL
SC-401 Sentido Norte - Faixa 1	FNS_P438_C003	191.267	24	1,25%
SC-401 Sentido Norte - Faixa 4	FNS_P438_C004	537.167	49	0,91%
SC-401 Sentido Norte - Faixa 2	FNS_P438_C005	66.946	9	1,34%
SC-401 Sentido Norte - Faixa 3	FNS_P438_C006	514.471	92	1,79%
SC-401 Sentido Centro - Faixa 4	FNS_P439_C003	747.213	86	1,15%
SC-401 Sentido Centro - Faixa 3	FNS_P439_C004	480.275	57	1,19%
SC-401 Sentido Centro - Faixa 2	FNS_P439_C005	2.531	0	0,00%
SC-401 Sentido Centro - Faixa 1	FNS_P439_C006	3.217	0	0,00%
TOTAL		2.543.087	317	1,25%

Fonte: Elaborado pelo autor.

Para fins de verificação das capturas realizadas pelos dispositivos ALPR, a amostra analisada concentrou-se nas leituras que resultaram em alerta de registro de furto ou roubo do veículo referente à placa identificada pelo dispositivo, resultando em 317 elementos a serem analisados de um universo de 2.543.087 de veículos. Tem-se, portanto, de que, do quantitativo total em circulação, apenas um montante de 1,25% apresentou inconformidade de interesse da segurança pública – vinculação de furto ou roubo do veículo capturado –, com os seguintes resultados:

Tabela 3 – Resultado da análise das leituras dos dispositivos ALPR

SITUAÇÃO	QUANTIDADE	PERCENTUAL
Leituras inviáveis	18	5,68%
Leituras corretas	74	23,34%
Leituras incorretas	225	70,98%

Fonte: Elaborado pelo autor.



As leituras inviáveis são aquelas que as condições ambientais de iluminação, inclinação, obstrução ou deterioração da placa veicular impedem naturalmente que o dispositivo ALPR proceda com a leitura adequada dos caracteres, identificando-os incorretamente.

Uma segunda problemática se relaciona à constatação de que das 225 leituras incorretas, representando quase 71% do montante registrado como alerta de interesse da segurança pública, todas foram por erro de reconhecimento dos caracteres das placas dos veículos pelo processamento embarcado dos dispositivos ALPR. Realizada análise das principais imprecisões, destacam-se 52,44% das leituras incorretas por erro de identificação, com substituição do caractere “I” pelo número “1”, dentre outros dispostos na Tabela 4:

Tabela 4 – Incidência de imprecisão no reconhecimento de caracteres pelos dispositivos ALPR.

CARACTERE CORRETO	CARACTERE IDENTIFICADO	QUANTIDADE	PERCENTUAL
I	1	118	52,44%
A	8	29	12,89%
B	8	8	3,56%
I	A	6	2,67%
I	U	4	1,78%
M	J	4	1,78%
7	2	3	1,33%
A	4	3	1,33%
E	5	3	1,33%
J	3	3	1,33%
Q	D	3	1,33%
Outros (35)		41	18,22%

Fonte: Elaborado pelo autor.

Denota-se uma incidência, no tocante à qualidade dos dados (DQ – *Data Quality*) de problemas na dimensão de precisão dos dispositivos ALPR, visto que o processamento embarcado automaticamente atribui um valor àquele que não tem condições de ser efetivamente identificado, resultando em imprecisão da leitura. O desejável seria o dispositivo classificar aquela captura como irreconhecível e proceder ao seu descarte.

5. CONCLUSÃO

É cristalina a importância crítica da qualidade dos dados no contexto de operações corporativas em geral e, neste caso em específico, na coleta de dados relacionados ao tráfego



de veículos para gestão da segurança pública, utilizando dispositivos ALPR (*Automatic License Plate Recognition*) – que atuam como componentes IoT – integrados ao Sistema BRAVO em Santa Catarina. Percebe-se como dados de alta qualidade, de acordo com as dimensões usualmente estipuladas, são fundamentais para a eficiência e eficácia das decisões organizacionais e operacionais, neste caso em particular, em aplicações que envolvem a segurança pública e a gestão de infraestruturas críticas.

Percebeu-se ao longo dos levantamentos que os dispositivos ALPR, que constituem componentes de borda do Sistema BRAVO, possuem um problema de precisão que precisa ser analisado com especial atenção, principalmente pela validação da leitura por elemento humano, antes de proceder-se a intervenção policial para abordagem do veículo destacado.

Nas demais dimensões aqui destacadas, não foram verificadas inadequações: as placas coletadas com a precisão adequada atendem à dimensão de completude, uma vez que é dada a amplitude com a fusão de múltiplos dados decorrentes da consulta e ingestão da geolocalização do dispositivo, resultando em informações qualificadas para gerar valor e subsidiar ações.

No tocante à pontualidade, todos os dispositivos apresentaram alimentação contínua e transmissão em tempo real por rede de comunicação baseada em fibra ótica, não verificando-se qualquer perda de pacotes ou indisponibilidade por solução de continuidade do serviço implantado e, quanto à consistência, esta sofre com o problema de precisão, uma vez que as capturas incorretas levam à consulta e registro de tráfego de veículo em determinado local que não corresponde com a realidade dos fatos.

A qualidade dos dados capturados pelos dispositivos ALPR é um fator determinante para a garantia da consistência da operação, pois são o item basilar no fluxo da informação, de modo que estudos futuros podem envolver a aplicação de um segundo processo de validação dos caracteres capturados através de inteligência artificial mediante o emprego de recursos de Visão Computacional, verificando a conformidade das placas capturadas pelos dispositivos IoT de borda e validando se os caracteres inicialmente identificados condizem com a imagem de referência, atuando como um segundo fator de validação e incrementando a precisão.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Código de Financiamento 001.



REFERÊNCIAS

- Alwan, A. A., Ciupala, M. A., Brimicombe, A. J., Ghorashi, S. A., Baravalle, A., & Falcarin, P. (2022). Data quality challenges in large-scale cyber-physical systems: A systematic review. *Information Systems*, *105*, 101951. <https://doi.org/10.1016/j.is.2021.101951>
- Byabazaire, J., O'Hare, G., & Delaney, D. (2020). Data Quality and Trust: Review of Challenges and Opportunities for Data Sharing in IoT. *Electronics*, *9*(12), 2083. mdpi. <https://doi.org/10.3390/electronics9122083>
- CIASC - Centro de Informática e Automação de Santa Catarina. (2018, October 11). *DetranNet*. CIASC. <https://www.ciasc.sc.gov.br/detrannet/>
- Costa, A.B., Fontanari, A. M. & Zoltowski, A. P. (2022). Como escrever um artigo de revisão sistemática: um guia atualizado. In Sampaio, M. I. C., Sabadini, A. A. Z. P., & Koller, S. H. (2022). *Produção Científica: um Guia Prático*. Universidade de São Paulo. Instituto de Psicologia. DOI: [10.11606/9786587596280](https://doi.org/10.11606/9786587596280) Acesso em 18 Fev. 2024.
- Du, S., Ibrahim, M., Shehata, M., & Badawy, W. (2013). Automatic License Plate Recognition (ALPR): A State-of-the-Art Review. *IEEE Transactions on Circuits and Systems for Video Technology*, *23*(2), 311–325. <https://doi.org/10.1109/tcsvt.2012.2203741>
- Liu, C., Nitschke, P., Williams, S. P., & Zowghi, D. (2019). Data quality and the Internet of Things. *Computing*, *102*(2), 573–599. <https://doi.org/10.1007/s00607-019-00746-z>
- Lockwood, C., dos Santos, K. B., & Pap, R. (2019). Practical Guidance for Knowledge Synthesis: Scoping Review Methods. *Asian Nursing Research*, *13*(5), 287–294. <https://doi.org/10.1016/j.anr.2019.11.002>
- Madnick, S. E., Wang, R. Y., Lee, Y. W., & Zhu, H. (2009). Overview and Framework for Data and Information Quality Research. *Journal of Data and Information Quality*, *1*(1), 1–22. <https://doi.org/10.1145/1515693.1516680>
- Mansouri, T., Sadeghi Moghadam, M. R., Monshizadeh, F., & Zareravasan, A. (2021). IoT Data Quality Issues and Potential Solutions: A Literature Review. *The Computer Journal*, *66*(3), 615–625. <https://doi.org/10.1093/comjnl/bxab183>
- Patel, C., Shah, D., & Patel, A. (2013). Automatic Number Plate Recognition System (ANPR): A Survey. *International Journal of Computer Applications*, *69*(9), 21–33. <https://doi.org/10.5120/11871-7665>
- Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated Methodological Guidance for the Conduct of Scoping Reviews. *JBIM Evidence Synthesis*, *18*(10), 2119–2126.
- Romero, M. L., & Suyama, R. (2023). Data Quality in IoT Applications: A Scoping Review. *2023 15Th Ieee International Conference on Industry Applications (Induscon)*. <https://doi.org/10.1109/induscon58041.2023.10374893>
- Zhang, L., Jeong, D., & Lee, S. (2021). Data Quality Management in the Internet of Things. *Sensors*, *21*(17), 5834. <https://doi.org/10.3390/s21175834>