



MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO
TOCANTINS *CAMPUS* DIANÓPOLIS
CURSO DE LICENCIATURA EM COMPUTAÇÃO

KARINE DE SOUZA ARAÚJO

ANÁLISE DE REQUISITOS PARA BANCO DE DADOS VOLTADOS A
APLICAÇÕES IOT

DIANÓPOLIS
2023

KARINE DE SOUZA ARAÚJO

**ANÁLISE DE REQUISITOS PARA BANCO DE DADOS VOLTADOS A
APLICAÇÕES IOT**

Trabalho de Conclusão de Curso apresentado à
Coordenação do Curso Superior de Licenciatura
em Computação do Instituto Federal de
Educação, Ciência e Tecnologia do Tocantins
Campus Dianópolis, como exigência à obtenção
do título de Licenciado(a) em
Computação.

Orientador: Prof. Dr. Lucas Arruda Ramalho

**DIANÓPOLIS
2023**

**Dados Internacionais de Catalogação na Publicação (CIP)
Bibliotecas do Instituto Federal do Tocantins**

A658a Araújo, Karine de Souza
ANÁLISE DE REQUISITOS PARA BANCO DE DADOS
VOLTADOS A APLICAÇÕES IOT / Karine de Souza Araújo. – Dianópolis,
TO, 2023.
40 f. : il. color.

Trabalho de Conclusão de Curso (Licenciatura em Computação) –
Instituto Federal de Educação, Ciência e Tecnologia do Tocantins, Campus
Dianópolis, Dianópolis, TO, 2023.

Orientador: Dr. Lucas Arruda Ramalho

1. Automação. 2. Tecnologia. 3. Heterogeneidade. I. Ramalho, Lucas
Arruda. II. Título.

CDD 004

A reprodução total ou parcial, de qualquer forma ou por qualquer meio, deste documento é autorizada para fins de estudo e
pesquisa, desde que citada a fonte.

**Elaborado pelo sistema de geração automática de ficha catalográfica do IFTO com os dados fornecidos pelo(a)
autor(a).**

KARINE DE SOUZA ARAÚJO

**ANÁLISE DE REQUISITOS PARA BANCO DE DADOS VOLTADOS A
APLICAÇÕES IOT**

Trabalho de Conclusão de Curso apresentado à
Coordenação do Curso Superior de Licenciatura
em Computação do Instituto Federal de
Educação, Ciência e Tecnologia do Tocantins
Campus Dianópolis, como exigência à obtenção
do título de Licenciado(a) em
Computação.

Aprovado em: 08/12/2023

BANCA AVALIADORA

Prof. Dr. Lucas Arruda Ramalho
IFTO – *Campus Dianópolis*

Prof. Msc. Tassio Guerreiro Antunes Virginio
IFTO – *Campus Dianópolis*

Prof^a Aparecida de Cássia Rodrigues Ávila
Externo

DIANÓPOLIS-TO

2023

“O sábio que tudo sabe é aquele que sabe que nada sabe.”

Platão.

AGRADECIMENTOS

Em primeiro lugar agradeço a Deus pelo dom da vida e por todas as coisas maravilhosas que só Ele é capaz de me conceder; pela coragem, sabedoria, saúde e por estar a frente de toda minha vida. Agradeço à minha mãe Rosimere Rosa de Souza Araújo, meu pai José Maria de Brito Araújo, à minha irmã Karoline de Souza Araújo e ao meu cunhado Kaique Cardoso por sempre me incentivarem no caminho dos estudos e por não medirem esforços para me ajudar quando precisei; à minha tia Maria Júlia Rosa de Souza e a todos os meus familiares por sempre torcerem e se preocuparem comigo mesmo à distância; aos meus amigos e colegas que acompanharam e fizeram parte dessa jornada de longos anos e momentos diversos, tornando tudo mais leve e divertido. Minha sincera gratidão aos excelentes professores que tive a oportunidade de conhecer no decorrer do curso, principalmente ao meu orientador Professor Doutor Lucas Arruda Ramalho pelo empenho e atenção durante a produção deste trabalho. Os conhecimentos e aprendizados adquiridos nessa jornada acadêmica serão levados com carinho. A paciência e dedicação que tiveram nunca serão esquecidas e espero que o sentimento de missão cumprida se faça presente quando souberem da minha aprovação. <3

RESUMO

A correria da vida cotidiana fez com que o ser humano buscasse cada vez mais por alternativas que auxiliassem na realização de tarefas e na economia. Com o crescente desenvolvimento das ferramentas tecnológicas isso se tornou eficaz e presente em vários ramos da sociedade, tanto dentro de casa quanto no campo, nas indústrias e nos transportes. A necessidade de conhecer essas tecnologias e saber como funciona o Banco de Dados que guarda todas essas informações é de extrema importância dado que em algum momento, devido a modernização, o conhecimento se fará imprescindível. Pensando nisso, conceitos importantes foram abordados neste trabalho, tais como *Automação*, *Internet das Coisas*, *Learn Machine*, *Sistemas Distribuídos*, entre outros. Além disso, ferramentas IoT com interfaces gratuitas foram apresentadas a fim de mostrar que é algo que não está tão longe da realidade. Este trabalho objetivou analisar em artigos da literatura as ferramentas e arquiteturas de armazenamento de dados aplicadas em cenários IoT.

Palavras-chave: Automação. Computação em Névoa. Ferramentas. Framework. Heterogeneidade. Inovação. Protocolos. Sistemas Distribuídos. Tecnologia.

ABSTRACT

The rush of everyday life has made human beings increasingly search for alternatives that help in carrying out tasks and saving money. With the increasing development of technological tools, this has become effective and present in various areas of society, both at home and in the countryside, in industries and in transport. The need to know these technologies and know how the Database that stores all this information works is extremely important given that at some point, due to modernization, knowledge will become essential. With this in mind, important concepts were covered in this work, such as Automation, Internet of Things, Learn Machine, Distributed Systems, among others. Furthermore, IoT tools with free interfaces were presented in order to show that it is something that is not so far from reality. This work aimed to analyze in literature articles the data storage tools and architectures applied in IoT scenarios.

Keywords: Automation. Fog Computing. Tools. Framework. Heterogeneity. Innovation. Protocols. Distributed Systems. Technology.

LISTA DE ILUSTRAÇÕES

Figura 1 - Camadas funcionais de uma solução IoT	12
Figura 2 - Simulação de Banco de Dados heterogêneo de uma aplicação IoT	15
Figura 3 - Computação em Névoa (Fog) - hierarquia	21
Figura 4 - Computação em Nuvem	22
Figura 5 - Computação em Névoa	22
Figura 6 - Tipos de Colaborações	28

LISTA DE TABELAS

Tabela 1 - Redes sem fio	13
Tabela 2 - Principais métricas citadas em artigos e documentações	19

LISTA DE ABREVIATURAS

- ADC - Adaptive Data Cleaning - Limpeza Adaptativa de Dados
- AWS - Amazon Web Service
- BD - Banco de Dados
- DDS - Distributed Data Service - Serviço de Dados Distribuídos
- EHR - Electronic Health Record - Registros Eletrônicos de Saúde
- GB - Giga Byte
- Gbps - Gigabit por segundo
- IA - Inteligência Artificial
- IaaS - Infrastructure as a Service - Infraestrutura como Serviço
- IEEE - Institute of Electrical and Electronic Engineers - Instituto de Engenheiros Elétricos e Eletrônicos
- IIoT - Industrial Internet of Things - Internet Industrial das Coisas
- IoT - Internet of Things - Internet das Coisas
- Kbps - Kilobit por segundo
- LoRa - Long Range - Longo Alcance
- LSTM - Long Short Term Memory - Memória de Longo Prazo
- MB - Mega Byte
- Mbps - Megabits por segundo
- MCU - Microcontroller Unit - Unidades Microcontroladoras
- MQTT - Message Queuing Telemetry Transport - Transporte de Telemetria de Enfileiramento de Mensagens
- M2M - Machine to Machine - Máquina para Máquina
- PaaS - Platform as a Service - Plataforma como Serviço
- RAM - Random Access Memory - Memória de Acesso Aleatório
- SaaS - Software as a Service - Software como Serviço
- SGBD - Sistemas de Gerenciamento de Banco de Dados
- SQL - Structured Query Language - Linguagem Estruturada de Consultas
- SVM - Support Vector Machine - Máquina de Vetor de Suporte
- Wi-Fi - Wireless Fidelity - Fidelidade Sem Fios
- XML - Extensible Markup Language - Linguagem de Marcação Extensível

SUMÁRIO

1 - INTRODUÇÃO	9
2 - EMBASAMENTO TEÓRICO	11
2.1 Conceito IoT e Arquitetura em Camadas	11
2.2 Protocolos de Comunicação Utilizados	13
2.3 Conceitos de Banco de Dados	13
2.3.1 Arquiteturas de Banco de Dados Heterogêneos	14
2.3.2 Arquiteturas de Banco de Dados voltados a grandes volumes de dados	16
2.3.3 Arquiteturas de Bancos de Dados voltados a processamento em tempo real	17
2.4 Sistemas Distribuídos	18
3 - METODOLOGIA DA PESQUISA	19
4 - REVISÃO DA LITERATURA	20
4.1 Artigo 1 - Alshammari et al. 2020.	20
4.2 Artigo 2 - Nandankar et al. 2021.	23
4.3 Artigo 3 - Saqlain et al. 2019	25
4.4 Artigo 4 - Sun et al. 2021.	27
4.5 Discussões	28
5 - CONSIDERAÇÕES FINAIS	31
REFERÊNCIAS	33

1 - INTRODUÇÃO

De acordo com Rosário (2012, p. 18), embora parecidos, os termos “automação” e “automatização” possuem diferentes significados. Uma vez que a automatização esteja diretamente ligada a movimentos repetitivos e mecânicos, a automação, por sua vez, é um conjunto de técnicas capazes de atuar com eficiência pelo uso de informações recebidas do meio sobre o qual atuam, e, com base nas informações recebidas, o sistema calcula e escolhe a ação corretiva apropriada.

Ainda se tratando de automação, para o autor (ROSÁRIO, 2012), presume-se que para que os resultados obtidos sejam precisos, se faz necessária uma auto adaptação dos sistemas de maquinismos a diferentes condições. A automação está ligada à utilização de sistemas automáticos e pode-se definir ainda como sendo um sistema que tende a aumentar a eficiência de um determinado processo. Aliada a Internet das Coisas (IoT - *Internet of Things*), além de nos ajudar na realização de atividades do nosso cotidiano, pode controlar aparelhos eletrônicos, diminuindo o trabalho e consumo de energia, por exemplo, ou até mesmo melhorando a segurança em determinadas situações.

O conceito de IoT é um novo paradigma que surgiu e está rapidamente ganhando terreno no cenário das telecomunicações. A ideia básica desse conceito é a possibilidade de interação, através de uma conexão com a Internet, entre os dispositivos de rede tradicionais - como computador de mesa, notebook, smartphone - e os diversos dispositivos chamados de Coisas - como sensores, eletrodomésticos, etiquetas, lâmpadas, etc. - para alcançar objetivos comuns (ATZORI et al., 2010).

Para Martinhão (2021), Internet das coisas é muito mais que uma geladeira conectada. É a progressiva automatização de setores inteiros da economia e da vida social com base na comunicação máquina-máquina: logística, agricultura, transporte de pessoas, saúde, produção industrial e muitos outros. Para isso, é necessário um ambiente favorável e interoperável ao acesso de um número cada vez maior de dispositivos, além de suporte escalável a grandes quantias de dados.

De acordo com o site Oracle (2023), as aplicações de IoT voltadas a corporações usam algoritmos de *Machine Learning* para analisar grandes quantidades de dados de sensores conectados à nuvem. Usando painéis e alertas da IoT em tempo real, é possível obter visibilidade dos principais indicadores de desempenho, estatísticas do tempo médio entre falhas e outras informações. Os algoritmos baseados em *Machine Learning* podem identificar anomalias na

aplicação, enviar alertas e até mesmo acionar correções automatizadas ou medidas proativas. Para identificar essas anomalias, os algoritmos necessitam realizar o aprendizado baseado em Treinamento Supervisionado (ESCOVEDO, 2020). Nesta etapa, o sistema utiliza bases de dados para treinamento e calibração da predição de padrões. Sendo assim, a base de dados a ser utilizada em aplicações IoT para algoritmos de *Machine Learning* necessitará pré-processar e formatar as informações dos sensores de forma a criar entradas homogêneas no sistema.

Para a aplicação dos algoritmos de Aprendizagem Supervisionada em geral, é importante que os dados estejam em um formato adequado que possa ser interpretado pelo algoritmo. Em muitos casos, é preciso converter informações qualitativas em quantitativas, para que estas possam ser operadas numericamente (FONTANA, 2020).

Os dados coletados dos sensores em aplicações IoT podem ser diversos. Texto, imagens, números, estados (ligado/desligado), áudios, são exemplos de heterogeneidade dos dados IoT. Quais técnicas e ferramentas podem ser utilizadas para pré-formatar dados heterogêneos como forma de criar bancos de dados propícios para aplicações IoT focadas no uso de *Machine Learning*?

Portanto, o objetivo deste trabalho é conceituar, apresentar e analisar algumas técnicas e ferramentas disponíveis para a criação de Banco de Dados voltados a aplicações IoT, abordando a pré-formatação dos dados heterogêneos que são coletados para criação de bases de dados homogêneas. Pretende-se com isso, através da metodologia de Análise Bibliográfica, apresentar algumas das principais arquiteturas utilizadas na atualidade para trabalhar com dispositivos IoT.

Este trabalho está dividido em capítulos. O capítulo 2 traz o Embasamento Teórico, onde estão abordados conceitos muito importantes para o entendimento geral, como o de IoT e Arquitetura em Camadas; Protocolos de Comunicação Utilizados; Banco de Dados (conceito e análise de arquiteturas) e Sistemas Distribuídos. No capítulo 3 é abordada a Metodologia da Pesquisa, contendo também a apresentação da tabela com algumas métricas. No capítulo 4, Revisão de Literatura, são apresentadas sínteses de quatro artigos relacionados ao tema, tendo ao final o tópico Discussões. E, para finalizar, temos o capítulo 5, que é o das Considerações Finais.

2 - EMBASAMENTO TEÓRICO

Para trabalhar com Banco de Dados se faz necessário o conhecimento de alguns conceitos básicos, como métricas, características de entidades e características da natureza dos dados. A primeira seção se trata do Conceito IoT e Arquitetura em Camadas; na segunda seção serão abordados os Protocolos de Comunicação Utilizados; a terceira seção está subdividida em Conceitos de Banco de Dados, Análise de Arquiteturas de Banco de Dados Heterogêneos, Análise de Arquitetura de Banco de Dados voltados a grande volume de dados, Análise de Arquitetura de Banco de Dados voltados a processamento em tempo real e, por fim, na quarta seção é abordado Sistemas Distribuídos.

2.1 Conceito IoT e Arquitetura em Camadas

O termo Internet das Coisas (do inglês *Internet of Things - IoT*) pode ser definido de várias maneiras por autores diferentes. Para Lacerda (2015), desponta como uma evolução da internet e um novo paradigma tecnológico, social, cultural e digital. A Internet das Coisas revolucionará os modelos de negócios e a interação da sociedade com o meio ambiente, por meio de objetos físicos e virtuais, em que esses limites se tornam cada vez mais tênues.

A Internet das Coisas proporciona aos objetos do dia a dia, com capacidade computacional e de comunicação, se conectarem à internet. Essa conexão viabiliza controlar remotamente os objetos e acessá-los como provedores de serviços, fazendo com que se tornem Objetos Inteligentes ou *Smart Objects*. Os Objetos Inteligentes possuem capacidade de comunicação e processamento aliados a sensores (MANCINI, 2020).

Refere-se à rede coletiva de dispositivos conectados e à tecnologia que facilita a comunicação entre os dispositivos e a nuvem, bem como entre os próprios dispositivos. Graças ao advento de chips de computador baratos e telecomunicações de alta largura de banda, agora existem bilhões de dispositivos conectados à Internet. Isso significa que dispositivos do dia a dia, como roupas, eletrodomésticos, veículos e outras máquinas, podem usar sensores para coletar dados e responder de forma inteligente aos usuários (AMAZON, 2023).

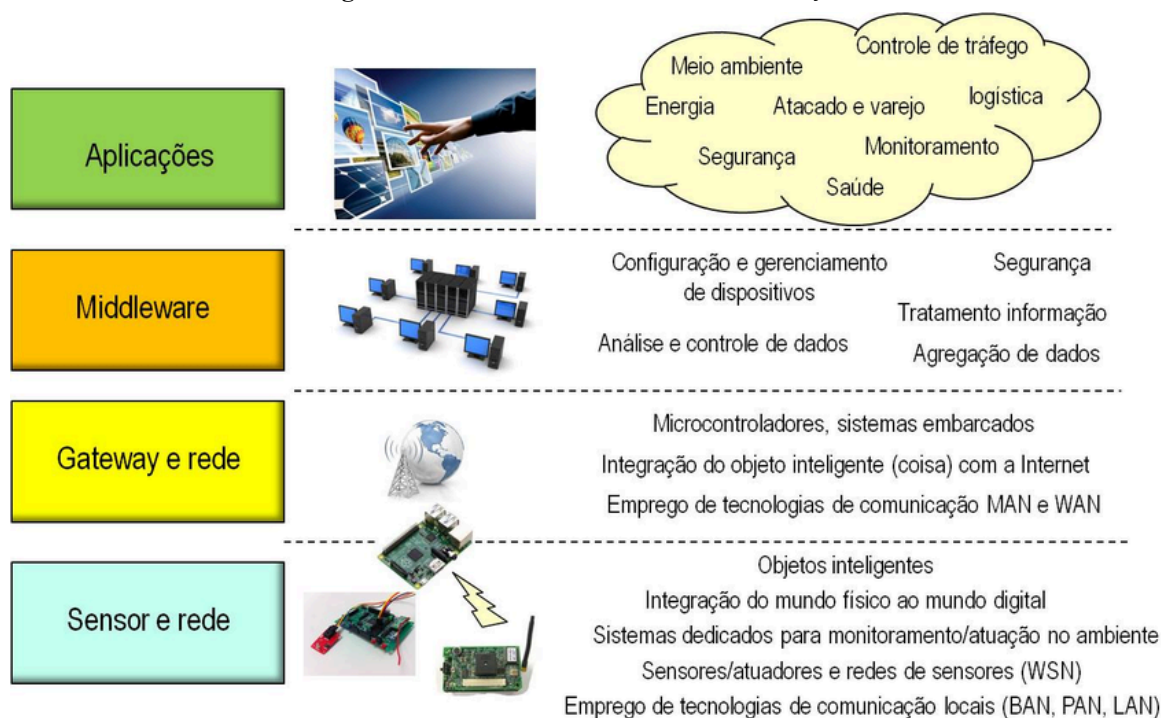
Desde a década de 90 os engenheiros de computação vem adicionando sensores e processadores a equipamentos com o intuito de modernizar ainda mais os objetos do cotidiano. Com o avanço tecnológico esses dispositivos foram aprimorados, diminuídos de tamanho e se tornaram cada vez mais rápidos e inteligentes.

O custo de integrar a capacidade computacional a pequenos objetos diminuiu consideravelmente. Por exemplo, você pode adicionar conectividade com recursos de serviços de

voz Alexa a MCUs com menos de 1 MB de RAM incorporada, como interruptores de luz. Todo um setor surgiu com foco em equipar residências, empresas e escritórios com dispositivos IoT. Esses Objetos Inteligentes podem transmitir dados automaticamente de e para a Internet. Todos esses “dispositivos de computação invisíveis” e a tecnologia associada a eles são coletivamente chamados de Internet das Coisas (AMAZON, 2023).

As camadas de uma arquitetura IoT estão organizadas conforme mostra a Figura 1.

Figura 1 - Camadas funcionais de uma solução IoT



Fonte: ResearchGate (2016).

De acordo com a figura acima, na camada “Sensor e Rede” estão os chamados “Objetos Inteligentes”, nela ocorre a integração do mundo físico ao digital; na camada “Gateway e Rede” ocorre a integração do objeto/coisa com a internet; na camada “Middleware” acontece a configuração e gerenciamento dos dispositivos, além do tratamento da informação e da análise e controle de dados; e na camada “Aplicações” é onde toda a magia se revela, os dados são enviados para os dispositivos dos usuários como foram programados para enviar.

Graças à IoT, muitas tarefas do dia a dia se tornaram mais fáceis, sustentáveis e rápidas de serem realizadas com a Automação. Para (TAGLIARI, 2010), Automação é a ciência que estuda e aplica metodologias, ferramentas e equipamentos, objetivando definir quando e como converter o controle de um processo manual para automático. Com isso, uma coleta metodológica e precisa de dados pode ser empregada para se obter a otimização de um processo (TAGLIARI, 2010).

Nas primeiras implantações, os projetos de automação de instalações industriais visavam apenas ao monitoramento e controle automático dos parâmetros mais importantes, voltados à

garantia de uma operação remota adequada e não assistida. Os avanços tecnológicos, como a instrumentação, as redes de comunicação, o controle e os sistemas de supervisão, foram fundamentais para a obtenção de malhas de controle mais eficientes (TAGLIARI, 2010).

2.2 Protocolos de Comunicação Utilizados

Os Objetos Inteligentes se comunicam por meio de diversos protocolos de redes sem fio que dependem da aplicação IoT final e de outros fatores como por exemplo: alcance e cobertura, consumo/disponibilidade de energia, taxa de transferência de dados, latência, custos, escalabilidade, segurança, ambiente e cenário de aplicação.

Os protocolos de rede Wi-Fi (maior transferência de dados para aplicações de baixo consumo de energia), ZigBee (baixo consumo e padrão avançado de criptografia em canal de comunicação) e LoRa (protocolo wireless que cobre longas distâncias com baixo custo e baixo consumo de energia) são os mais utilizados. A tabela 1 apresenta redes sem fio detalhando suas velocidades e alcance.

Tabela 1 - Redes sem fio

Rede	Alcance	Velocidade
Wi-Fi IEEE 802.11n	50 metros	72 Mbps
Bluetooth IEEE 802.15.1	10 metros	720 Kbps
Zigbee	50 metros	240 Kbps
5G WiMax IEEE 802.16	Dados Móveis	10 Gbps

Fonte: Adaptado de Rivera (2010).

Os protocolos de enlace podem ser encapsulados, o protocolo utilizado para comunicação entre dispositivos IoT é o M2M (Machine - Machine). O site Embarcados (2022) o conceitua como aberto a indústria e criado para gerenciar dispositivos remotamente; são econômicos e usam redes públicas, criam um ambiente onde duas máquinas se comunicam mutuamente e trocam dados, além de reforçar o automonitoramento das máquinas, permite que os sistemas se adaptem de acordo com a necessidade, é usado em casas inteligentes, veículos e caixas eletrônicas.

2.3 Conceitos de Banco de Dados

Para compreender o conceito de Banco de Dados se faz necessário o conhecimento de suas partes, iniciando por “entidade”. Trata-se de um objeto básico com existência física ou conceitual,

que possui características particulares chamadas de atributos. Os valores dos atributos de cada entidade são os dados que preencherão os campos determinados (TAKAI et. al., 2005).

Uma entidade é algo do mundo real que possui uma existência independente. Uma entidade pode ser um objeto com uma existência física - uma pessoa, carro ou empregado - ou pode ser um objeto com existência conceitual - uma companhia, um trabalho ou um curso universitário. Cada entidade tem propriedades particulares, chamadas atributos, que a descrevem. Por exemplo, uma entidade EMPREGADO pode ser descrita pelo seu nome, o trabalho que realiza, idade, endereço e salário. Uma entidade em particular terá um valor para cada um de seus atributos. Os valores de atributos que descrevem cada entidade ocupam a maior parte dos dados armazenados na base de dados (TAKAI et. al., 2005, p. 23).

Os valores ou dados que caracterizam as entidades podem ser classificados como homogêneos ou heterogêneos. Quando heterogêneos, apresentam desafios no armazenamento, na integração das informações e na análise, pois são conjuntos de dados com diferentes formatos e estruturas.

Alguns exemplos descritos deste tipo de dado são:

- **Estrutura:** pode estar relacionada à estrutura dos dados, o que significa que podem estar organizados de maneiras diferentes. Por exemplo, um conjunto de dados pode ser estruturado em formato de grafo enquanto outro pode ser estruturado em formato de tabela. Essa diferença dificulta a análise integrada das informações obtidas.
- **Formato:** os dados podem variar em termos de formato, como imagens, arquivos de texto, vídeo, áudio, planilhas, entre outros. Cada formato de dado requer abordagens específicas para armazenamento, processamento e análise.
- **Tipo:** pode estar relacionada aos diferentes tipos de dados presentes em um conjunto de dados. Por exemplo, pode haver uma combinação de dados numéricos, textuais, categóricos e de data/hora. Isso requer técnicas específicas para tratamento e análise.
- **Semântica:** pode ocorrer em relação à semântica dos dados, ou seja, ao significado atribuído a eles. Dois conjuntos de dados podem ter diferentes interpretações para os mesmos termos ou conceitos. A interpretação correta e integração desses dados requerem uma compreensão adequada de semântica subjacente.

2.3.1 Arquiteturas de Banco de Dados Heterogêneos

Os Sistemas de Gerenciamento de Banco de Dados (SGBD), podem ser homogêneos ou heterogêneos (OZSU et. al., 2001). No SGBD homogêneo, todos os SGBD's locais são semelhantes, ou seja, não existem diferenças estruturais ou semânticas entre os Bancos de Dados,

independentemente do site onde se localizam. Já no caso dos SGBD's heterogêneos, podem existir dois ou mais SGBD's diferentes nos sites locais e são utilizados, de acordo com (CASANOVA, 1985), quando há necessidade de integrar sistemas já existentes.

Diferente dos heterogêneos, quando os dados são homogêneos as informações são mais fáceis de analisar, armazenar e integrar, por não apresentarem características distintas de informações e possuírem a mesma origem.

Segundo (DUARTE, 2004), a heterogeneidade pode ocorrer em qualquer nível do Banco de Dados (BD), envolvendo diferentes elementos que vão desde a implementação do SGBD até o conteúdo do BD. Por exemplo, SGBD's diferentes podem utilizar linguagens de consulta e manipulação de dados distintas, modelos de dados e plataformas diferentes.

Por compreender um conjunto amplamente diversificado de dispositivos e recursos coletados em suas múltiplas camadas, um sistema IoT pode ser definido como heterogêneo. Para que o trabalho seja bem sucedido, é preciso que todas as peças estejam em uma estrutura comum de protocolos e modelos de nível superior para se reconhecerem com segurança e trocar informações. A diversidade de dispositivos e recursos, mesmo em uma única aplicação IoT, requer um conjunto mais extenso de métodos padrão para lidar com as possíveis interações entre as várias partes de um sistema IoT (EMBARCADOS, 2019).

Figura 2 - Simulação de Banco de Dados heterogêneo de uma aplicação IoT



Fonte: Autora (2023).

Na Figura 2 é possível observar o Banco de Dados heterogêneo da aplicação IoT com informações diversas acerca do paciente que utiliza dispositivos vestíveis a fim de coletar dados da sua saúde para realizar seu tratamento a domicílio. Nome, idade, histórico de doenças, estado de saúde, endereço, diagnóstico, recomendações, tudo isso é armazenado no Banco, analisado e tratado, para depois retornar as informações necessárias.

2.3.2 Arquiteturas de Banco de Dados voltados a grandes volumes de dados

Segundo Takai (2005, p. 90), “as métricas são sumarizadas ou detalhadas de acordo com o interesse da análise a ser feita sobre os dados”. A partir daí esses dados são transformados em tabelas e podem ser utilizados para armazenar toda a informação obtida. A seguir ele detalha os três tipos de métricas, são elas:

As métricas mais comuns são as *completamente aditivas*. Diz-se que uma métrica é completamente aditiva quando faz sentido sumarizá-la adicionando seus valores ao longo de qualquer dimensão. No exemplo de pedidos, o valor do pedido, a margem, o custo e a quantidade pedida são todas métricas completamente aditivas. [...] As métricas completamente aditivas são bastante úteis e poderosas, já que não existem limitações em sua utilização. Podem ser facilmente sumarizadas para qualquer combinação de valores das dimensões (TAKAI, 2005, p. 98).

Em resumo, as métricas completamente aditivas são de fácil manipulação caso haja a necessidade de verificar margens de dados com certa frequência, pois podem ser verificadas da forma que for do interesse de quem as manipula.

Já as métricas *não aditivas*, não podem ser adicionadas ao longo dos valores da dimensão, não podem ser sumarizadas de acordo com a dimensão de interesse, como as completamente aditivas. Se tratam de dados percentuais, como porcentagem de vendas mensal em um mercado ou margem de preço para um produto, como veremos a seguir.

Takai cita, “por exemplo, em uma determinada data, um vendedor vende a um cliente 4 tipos diferentes de produtos, cada um deles com uma margem percentual de 25%. Não faz sentido incluir os quatro valores de margem percentual para calcular a margem total para este pedido” (TAKAI, 2005, p. 98).

Ainda para o autor, o terceiro tipo de métrica é a *semi-aditiva*. A métrica semi-aditiva pode ser sumarizada ao longo de determinadas dimensões, porém, não todas. Considerando como exemplo o gerenciamento de saldo das contas de um banco. O saldo é armazenado no final de cada dia, para cada cliente, por conta ao longo do tempo. Em alguns casos este saldo é aditivo. Se um cliente tem uma conta corrente e uma conta poupança, pode-se adicionar os saldos de cada conta no final de um dia e obter resultado significativo. É possível também adicionar os saldos de

uma determinada agência para obter um panorama da situação geral de cada localidade. Entretanto, não faz o menor sentido adicionar o saldo de um cliente ao longo do tempo. Por esta razão, a métrica saldo é considerada semi-aditiva (TAKAI, 2005, p. 98).

Ou seja, a métrica semi-aditiva, como revela o nome, é uma mistura da métrica completamente aditiva com a não aditiva, onde alguns casos podem ser aditivos e outros não.

Nos artigos estudados para elaboração da pesquisa bibliográfica deste trabalho, os autores abordam que os dados gerados decorrentes da utilização da Internet das Coisas vão aumentando ao longo do tempo durante o uso e podem ser de vários tipos, sendo o armazenamento um dos maiores desafios.

2.3.3 Arquiteturas de Bancos de Dados voltados a processamento em tempo real

Plataformas IoT na Nuvem, como a Adafruit IO e Amazon Web Service, fornecem para as aplicações facilidade de coleta de dados e possuem recursos com acesso gratuito aos usuários, além dos tutoriais de fácil acesso nos próprios sites dos desenvolvedores e em canais no YouTube.

De acordo com o site Maker Hero (Maker Hero, 2023), Adafruit IO é uma plataforma em nuvem para IoT, desenvolvida pela empresa Adafruit Industries. Ela permite a criação e gerenciamento de dispositivos conectados à internet, bem como a coleta, armazenamento e visualização de dados gerados por esses dispositivos. Para utilizar, é preciso criar uma conta no próprio site da Adafruit.

Adafruit IO é uma plataforma IOT construída em torno do protocolo MQTT. Que é um protocolo de mensagens leve que fornece aos clientes de rede com recursos limitados uma maneira simples de distribuir informações de telemetria. O protocolo que usa um padrão de comunicação publicar / assinar, é usado para comunicação Máquina a Máquina (M2M) e desempenha um papel importante na Internet das Coisas (THINGS BR, 2022).

É capaz de fazer 30 dados por minuto, com 30 dias de armazenamento de dados na nuvem, 10 alimentações e 5 painéis. Além dos tutoriais que são encontrados facilmente na internet para seguimento passo a passo. A versão paga é capaz de fazer o dobro de dados por minuto, com 60 dias de armazenamento, feeds e painéis ilimitados.

Quanto a Amazon Web Service, também conhecida como AWS, segundo Syozi (TECHNOBLOG, 2021), é um serviço de Computação em Nuvem desenvolvido pela Amazon que oferece mais de 200 serviços completos de data centers por todo o mundo. Com isso, acaba trazendo mais recursos do que outros provedores de nuvem, segundo a própria empresa.

A AWS oferece uma quantidade consideravelmente maior de serviços - e mais recursos com esses serviços - do que qualquer outro provedor de nuvem: de tecnologias de infraestrutura,

como computação, armazenamento e Banco de Dados, a tecnologias emergentes como Machine Learning e Inteligência Artificial, data lakes, análises e Internet das Coisas. Com isso, é mais rápido, mais fácil e mais econômico mover seus aplicativos para a nuvem e construir praticamente qualquer coisa que você possa imaginar. Também tem a funcionalidade mais aprofundada nesses serviços. Por exemplo, a AWS oferece a mais ampla gama de Bancos de Dados especialmente criados para os diversos tipos de aplicativos. Assim, você pode escolher a ferramenta certa para o trabalho, ao melhor custo e com a melhor performance (AMAZON, 2023).

2.4 Sistemas Distribuídos

Um sistema distribuído é aquele no qual os componentes localizados em computadores interligados em rede se comunicam e coordenam suas ações apenas passando mensagens. [...] O compartilhamento de recursos é um forte motivo para a construção de sistemas distribuídos. Os recursos podem ser gerenciados por servidores e acessados por clientes, ou podem ser encapsulados como objetos e acessados por outros objetos clientes (COULOURIS et. al., 2007).

Para Coulouris (COULOURIS et. al., 2007), os desafios advindos da construção de sistemas distribuídos são a heterogeneidade de seus componentes, ser um sistema aberto, o que permite que componentes sejam adicionados ou substituídos, a segurança, a escalabilidade - a capacidade de funcionar bem quando o número de usuários aumenta - , o tratamento de falhas, a concorrência de componentes e a transparência.

Alguns exemplos de sistemas distribuídos são:

- Intranet (um sistema acessível pela web, mas com disponibilidade apenas para usuários de determinado grupo/empresa, com o objetivo de compartilhar informações e recursos de forma interna);
- Extranet (um sistema acessível pela web que pode ser composto de várias intranets; também possui acesso reduzido a determinado grupo/empresa, mas sem que necessariamente estejam em um local só, para acessar é preciso realizar login, o que possibilita ao usuário maior comodidade);
- Internet (acesso de qualquer lugar independentemente da localização);
- Computação Móvel (dispositivos móveis com menor capacidade de memória, armazenamento e desempenho, mas que oferecem disponibilidade de acesso e alta mobilidade aos usuários);
- Computação Ubíqua (presente em locais fechados ou com acesso limitado, são formados por dispositivos móveis, computadores, televisores, assistentes pessoais

digitais e pontos de vendas, estando presentes em supermercados, grandes lojas, residências ou escritórios).

3 - METODOLOGIA DA PESQUISA

A metodologia utilizada para a realização deste Trabalho de Conclusão de Curso será a Pesquisa Bibliográfica, que consiste em estudos mais aprofundados acerca das características dos Bancos de Dados voltados a aplicações IoT, o que precisam ter para funcionar e suas métricas, além dos conceitos relacionados a essa área.

Segundo (PRODANOV; FREITAS, 2013, p. 54) uma Pesquisa Bibliográfica é elaborada a partir de material já publicado, constituído principalmente de: livros, revistas, publicações em periódicos e artigos científicos, jornais, boletins, monografias, dissertações, teses, material cartográfico, internet, com o objetivo de colocar o pesquisador em contato direto com todo material já escrito sobre o assunto da pesquisa. Na Pesquisa Bibliográfica, é importante que o pesquisador verifique a veracidade dos dados obtidos, observando as possíveis incoerências ou contradições que as obras possam apresentar.

Sendo assim, é um estudo com embasamento teórico de autores diversos acerca do tema visando um amplo conhecimento do que será abordado, deixando de ser uma coisa superficial e passando a ser mais confiável com as devidas referências e citações.

A tabela abaixo apresenta as principais métricas citadas em artigos e documentações no Google Acadêmico a partir do ano de 2022 até o presente momento.

Tabela 2: Principais métricas citadas em artigos e documentações

Principais métricas citadas em artigos e documentações	
Métrica	Quantidade de resultados
database	270.000
automation	102.000
internet of things	83.700
dataset	82.000
learn machine	16.500
M2M	15.300
MQTT	15.100
automatization	11.300

Fonte: Autora (2023).

4 - REVISÃO DA LITERATURA

Conforme mencionado anteriormente no Capítulo 3, este capítulo analisa 4 trabalhos acadêmicos que tratam sobre técnicas de utilização, configuração e/ou posicionamento de Banco de Dados voltados a aplicações IoT. Os artigos analisados são:

- Artigo 1 - Big Iot Healthcare Data Analytics Framework Based on Fog and Cloud Computing - Alshammari et al. 2020.
- Artigo 2 - An IoT Based Healthcare Data Analytics Using Fog And Cloud Computing - Nandankar et. al. 2021.
- Artigo 3 - Framework of an IoT-based Industrial Data Management for Smart Manufacturing - Saqlain et al. 2019.
- Artigo 4 - Intelligent Data Collaboration in Heterogeneous-device IoT Platforms - Sun et. al. 2021

4.1 Artigo 1 - Alshammari et al. 2020.

Em seu artigo, partindo do pressuposto da necessidade do trabalho minucioso e eficiente com grande quantidade de dados de diversos tipos no ramo da saúde, Alshammari et al. abordam uma estrutura analítica inovadora para essa heterogeneidade por meio de dispositivos IoT em tecnologias vestíveis.

A análise de dados no campo da saúde está dividida em quatro categorias, são elas:

- Descritiva: identifica o status atual do paciente e gera relatórios, gráficos e histogramas.
- Diagnósticas: compreende e estuda as razões para o reaparecimento de algumas doenças específicas em pacientes individuais por meio de técnicas de agrupamentos e árvores de decisões.
- Preditivas: prevê eventos desconhecidos através de algoritmos e modelos preditivos criados com Aprendizado de Máquina.
- Prescritivas: propõe ações eficazes que levem a tratamentos adequados aos pacientes, a fim de otimizar as decisões.

A estrutura proposta com este trabalho visa criar comunicações mais eficazes e estáveis, prevendo e prevenindo potenciais doenças por meio da integração de Registros Eletrônicos de Saúde (EHRs) e dados sensoriais obtidos no monitoramento 24/7 (todos os dias da semana). Além

de fornecer recomendações de assistências oportunas aos pacientes, isso ajudaria nos cuidados com pessoas de idades mais avançadas e diminuiria a sobrecarga dos serviços de saúde.

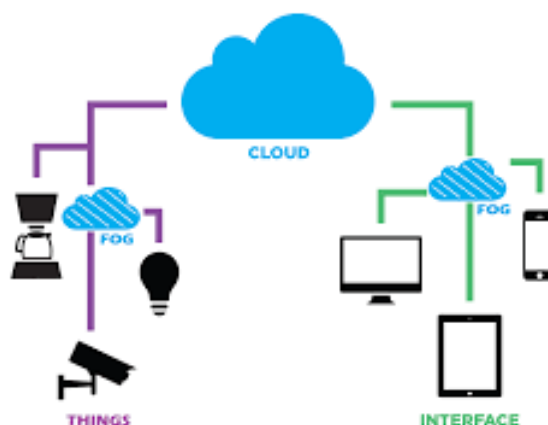
A tecnologia utilizada no processo envolve dispositivos multissensoriais e dispositivos vestíveis inteligentes que identificam os sintomas da doença e enviam para os celulares dos usuários dados e notificações gerais de saúde. Algoritmos de Aprendizado de Máquina e modelos de Aprendizagem Profunda serão utilizados a fim de melhorar o tempo de desempenho e a precisão.

A Computação em Nuvem compreende serviços como IaaS (Infraestrutura como serviço) que oferece acesso a espaço de armazenamento em nuvem ilimitado; PaaS (Plataforma como Serviço) que executa aplicativos com uso intensivo de recursos; SaaS (Software como Serviço) que gerencia o acesso a diferentes softwares; e serviços utilitários, que armazenam grandes quantidades de dados para acesso remoto.

Para o processamento de grande volume de dados provenientes de sensores IoT coletados em tempo real, a Computação em Névoa é essencial. Em aplicações de saúde enormes conjuntos de dados são formados rapidamente, então, se faz necessária a aplicação de nós de Computação em Névoa para obter precisão.

A Figura 3 conceitua de forma mais simples como funciona a Computação em Névoa: a nuvem está no topo da hierarquia, a névoa vem depois e em seguida os dispositivos inteligentes e a interface IoT.

Figura 3 - Computação em Névoa (Fog) - hierarquia

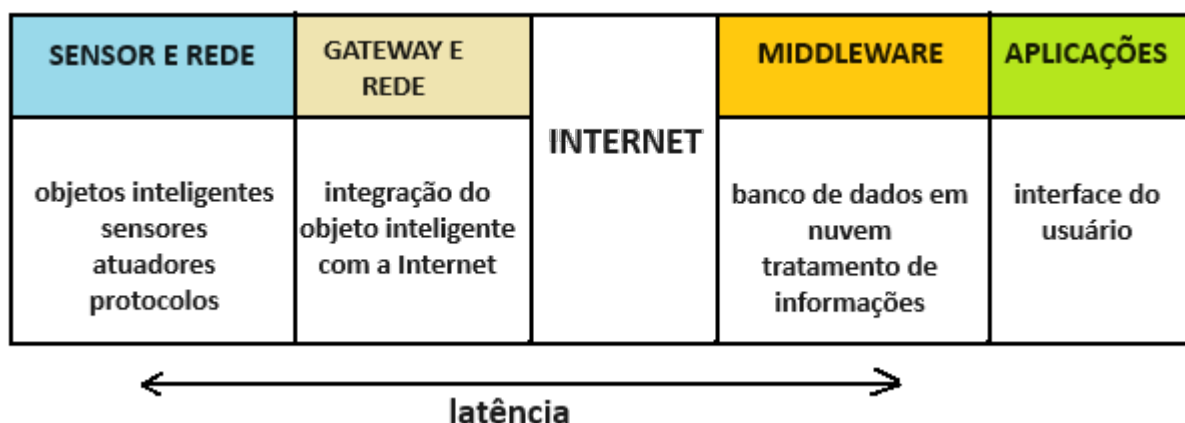


Fonte: HelloWorldSI (2016).

Nas Figuras 4 e 5, respectivamente, temos a representação detalhada das camadas da Computação em Nuvem e da Computação em Névoa, buscando mostrar o diferencial de uma para a outra em termos de latência. Na Computação em Névoa o tempo de latência é relativamente

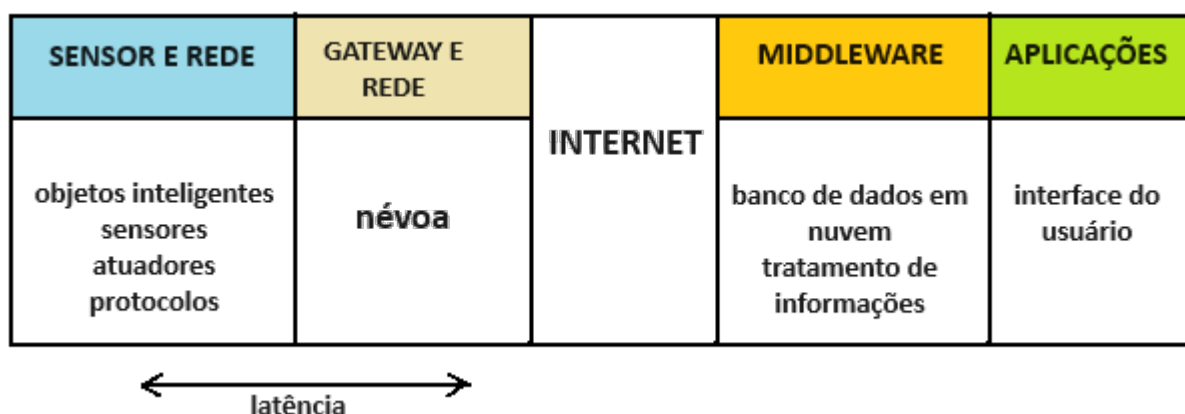
menor quando comparado com o da Computação em Nuvem, devido a diminuição da distância que a informação terá que percorrer para retornar o que o usuário solicitar.

Figura 4 - Computação em Nuvem



Fonte: Autora (2023).

Figura 5 - Computação em Névoa



Fonte: Autora (2023).

Recentemente, a Computação em Névoa (Fog Computing) [Bonomi et al. 2012] vem atraindo interesse pelo seu potencial de satisfazer requisitos que não são atendidos por um modelo centralizado em Nuvem [Khalid et al. 2016]. Este paradigma estende os recursos computacionais disponíveis na Nuvem para a borda da rede visando apoio às soluções em IoT. Dessa forma, possibilita a execução de aplicativos em bilhões de objetos conectados para fornecer dados, processamento, armazenamento e serviços aos usuários (COUTINHO et. al., 2016).

Aplicações em áreas de zona rural, por exemplo, que necessitam do retorno em tempo real, pode-se recorrer a Computação em Névoa a fim de que os recursos da nuvem sejam acessados com mais facilidade e menor latência.

A análise de dados históricos, gerenciamento de armazenamento, autenticação, serviços de autorização, infraestrutura central de gerenciamento de recursos médicos e a mineração de dados em longa escala são processados pela camada de Computação em Nuvem que exige que MapReduce e Apache Spark sejam implementados. Além de depender de uma computação back-end contínua que mantenha à vista os dispositivos médicos vestíveis e atualize os nós de névoa.

O MapReduce é utilizado para processar grandes volumes de dados. O processamento é distribuído em muitas máquinas, todas com conjuntos de dados diferentes, a fim de que as informações sejam processadas em tempo aceitável.

A Apache Spark é uma estrutura poderosa de clustering de big data, baseada em Hadoop e se destaca por sua alta confiabilidade, consistência, velocidade e tolerância a falhas. Dentre seus componentes fundamentais estão as bibliotecas: SQL, streaming, Spark MLlib e GraphX.

Após o comparativo entre algumas estruturas semelhantes, pode-se concluir que a que foi proposta por Alshammari et al. é superior em termos de latência, já que utiliza uma camada de névoa que é encarregada de entregar toda carga pesada de dados para a nuvem e lidar com o processamento computacionalmente intensivo.

4.2 Artigo 2 - Nandankar et al. 2021.

A necessidade de sistemas inteligentes no mundo para cuidados com a saúde se fez devido a escassez de médicos e alguns outros fatores. Com o desenvolvimento da Internet das Coisas, a evolução tecnológica contribuiu para que essas tecnologias se tornassem cada vez mais presentes no nosso cotidiano. Porém, o processamento de dados na área da saúde é problemático por aumentar gradativamente e variar dinamicamente, dificultando todo esse processo.

Em seu artigo, Nandankar sugere uma análise inovadora de um sistema para dados de grande escala obtidos através de dispositivos IoT vestíveis ou dados médicos arquivados, propondo resolver efetivamente o problema da heterogeneidade das informações por meio de middlewares e clusters Map Reduce Hadoop.

Além disso, conscientiza de forma abrangente e segura quanto aos dados médicos para pacientes, permite o uso de Computação em Névoa e sistemas em nuvem para lidar com os problemas enfrentados no meio online e offline, processamento de dados, armazenamento e classificação.

Os dados médicos heterogêneos estão presentes em registros eletrônicos (EHR) e para serem utilizados com garantia de consistência e velocidade é preciso que sejam coletados, integrados, limpos, armazenados, analisados e eficientemente interpretados.

A tomada de decisão exige cautela em seu procedimento, pois, por meio dela, grandes quantidades de dados em tempo real obtidos através dos sensores e dispositivos vestíveis serão revistos e terão grande impacto na saúde do paciente. Esses sistemas de saúde inteligentes, se usados, por exemplo, em uma cidade inteligente, seriam capazes de rastrear e intervir quando pacientes participassem de comportamentos prejudiciais às suas condições de vida. Cientistas utilizam dos benefícios desses dispositivos para detectar sinais que podem distinguir várias doenças. Mas os métodos de Aprendizado de Máquina não deixam de ser cada vez mais comuns, já que também são eficientes.

Com o intuito de reduzir a sobrecarga do médico, ajudar no envelhecimento da população, melhorar o acesso aos dados e diminuir os erros relacionados à área da saúde, o artigo sintetizado também auxilia a melhorar a pressão sobre os orçamentos nacionais para os cuidados de saúde.

Já o que se propõe no trabalho é um sistema de dispositivo vestível inteligente focado em dispositivos multissensoriais capazes de monitorar sintomas identificados da doença e fornecer aos usuários dados gerais de saúde e alertas dos sensores no próprio celular. Para maior precisão serão utilizados algoritmos de aprendizado de máquina como SVM (Support Vector Machine/Vetor de Suporte) e modelos de Aprendizagem Profunda.

A análise de dados em larga escala será dividida em: coleta, limpeza, anotação e extração. A estrutura incorpora os dados médicos para o modelo de análise e a partir daí se tem a visualização dos resultados.

O problema da heterogeneidade será abordado de forma eficiente com o uso de middleware entre as fontes de dados heterogêneas e Map Reduce Clusters Hadoop, que também utiliza algoritmos recentes de Aprendizado de Máquina para obter mais precisão e melhora no tempo de acesso, além de garantir confiabilidade para os pacientes.

Os nós de Computação em Névoa são de suma importância para o armazenamento e processamento de grandes quantidades de dados em tempo real obtidos através dos sensores IoT. A arquitetura proposta evita a latência na Computação em Nuvem devido a camada de Comunicação de Transporte. Os fluxos de dados em aplicações médias formam grandes conjuntos rapidamente e os nós de Computação em Névoa são capazes de gerenciar esses dados de forma confiável em tempo real.

Os dados gerados nos dispositivos médicos vestíveis são coletados e fornecidos às portas IoT, utilizando a camada de Computação em Névoa. A mediação entre os dispositivos vestíveis e

o sistema em nuvem é feita por meio dos nós da Computação em Névoa, que pré-processam os dados (erros, redundância e manuseio excessivo), projetam, classificam, prevêm e visualizam.

Os dados coletados podem ser grandes, ilimitados ou desequilibrados, e sua classificação pode ser cada vez mais complexa. As características dos atributos são o aspecto mais importante na implementação do modelo de classificação. Algoritmos de gerenciamento de dados de alta dimensão como AlexNet, LeNet, perceptron multicamadas e Sistemas Recorrentes de Longo Prazo de Memória (LSTM) podem substituir outras técnicas tradicionais em desempenho, exatidão e precisão.

O sistema proposto enviará alertas aos pacientes em tempo real e apesar da análise da névoa ajudar com que o processamento de grande quantidades de dados seja realizado, há planejamento da equipe para melhorias em trabalhos futuros, seguindo a mesma linha.

4.3 Artigo 3 - Saqlain et al. 2019

As aplicações que utilizam IoT estão se tornando parte essencial da Internet. No ramo industrial, por exemplo, os dispositivos dispõem de informações abundantes nos processos de fabricação. Todavia, a grande quantidade de informação que precisa ser coletada, processada e utilizada para tomadas de decisões em tempo real traz alguns desafios, levando em consideração sua natureza massiva, heterogênea e sensível ao tempo.

O trabalho de Saqlain et al. (2019) apresenta um Sistema de Gerenciamento de Dados Industriais baseado em IIoT (Internet Industrial das Coisas), que gerencia, dá suporte online e controla a fabricação. A estrutura está dividida em cinco camadas básicas: Física, de Rede, de Middleware (gerenciamento de recursos, eventos, dados e recuperação), de Banco de Dados e de Aplicação, tornando possível uma arquitetura orientada a serviços para os usuários finais. Os dados coletados se convertem em informações úteis e melhoram a produtividade e o prognóstico das linhas de produção.

A Internet Industrial das Coisas, também conhecida como Indústria 4.0, nada mais é que uma combinação de aplicações IoT concentrada em dispositivos de fábrica interconectados baratos e de pequeno porte aliada a big data, sistemas ciberfísicos, aprendizado de máquina e simulação, a fim de organizar e melhorar os ambientes industriais. Além disso, permite a transmissão de dados em tempo real através da rede industrial e é econômica.

O escopo da IIoT é cobrir atividades de grande escala como logística ferroviária, geração e distribuição de energia, petróleo e operações de gás e frota de aviação (SQLAIN et. al., 2019).

Os dados obtidos podem ser usados para prever e controlar o sistema de fabricação sem depender 100% de operações humanas, já que sua capacidade de detectar eventos especiais ou falhas se dá instantaneamente. Algumas das desvantagens de adotar a IIoT como solução de produção inteligente são: segurança de dados e serviços; confiança, privacidade de informações e integridade de dados; escalabilidade; e interoperabilidade.

A heterogeneidade e massividade dos dados podem ser tratadas por meio de middlewares IoT e DDS (Serviço de Dados Distribuídos), que processam através de um protocolo de comunicação específico e criam módulos de metadados. Mas o grande desafio ainda está no armazenamento, tanto pelo volume quanto pela variedade de formatos e tipos. Dentre esses desafios estão: heterogeneidade e dinamicidade; visualização de dados para recursos diversos; armazenamento de big data industrial; padronização de protocolos de comunicação; alta produtividade e qualidade do produto; confiabilidade e escalabilidade das linhas de produção.

Para controlar o tráfego e evitar atrasos de comunicação é projetado um servidor de banco de dados distribuído. Os dados industriais brutos são agregados e convertidos em um banco de dados estruturado. Os dados estruturados são entregues ao servidor em nuvem para armazenamento permanente, onde diferentes dados algoritmos de mineração e aprendizado de máquina podem ser executados para extração de conhecimento. (SAQLAIN et. al.).

A estrutura proposta se concentra no ramo industrial e oferece servidor de armazenamento distribuído para os dados serem armazenados antes de irem para a nuvem.

A avaliação do desempenho do framework proposto, por ser no ramo industrial, depende da rede de objetos físicos interconectados que possibilitam a conexão dos demais dispositivos. Por suportar estruturas de rede sem fio, fornece uma plataforma de rede escalável e de alto desempenho. Além disso, os dados da linha de produção podem ser acessados de forma móvel através de smartphones pelos fabricantes e gestores, ajudando a manter o gerente de produção e o gerenciamento de nível superior sempre atualizados, garantindo entrega pontual e alta qualidade de produtos.

Concluiu-se que utilizar a fabricação inteligente melhorou significativamente a escala de produção industrial e impulsionou o desenvolvimento. A estrutura suportava a enorme aquisição de dados e coletava-os em tempo real (8,6GB/h, 207,4GB/dia, 6.220GB/mês e 74.650GB/ano), com isso, a melhora na receita foi de 33,1%, de satisfação do cliente 22,1% e da qualidade do produto 11%, além de tornar os negócios mais seguros, sustentáveis e eficientes.

4.4 Artigo 4 - Sun et al. 2021.

A baixa qualidade dos dados das plataformas IoT heterogêneas é um desafio para o desenvolvimento de aplicações de alta qualidade entre a computação de borda e a Aprendizagem Profunda (Deep Learning). SUN et. al. (2021), em seu trabalho, apresentam uma colaboração inteligente de dados com a capacidade de colaborar de forma adaptativa para a realização de uma tarefa. A fim de que a utilidade e o funcionamento desse sistema fosse demonstrado, uma implementação Adaptativa de Limpeza de Dados (ADC, do inglês “*Adaptive Data Cleaning*”) foi projetada para filtrar os ruídos de leituras de temperatura em uma rede de estação com base IoT.

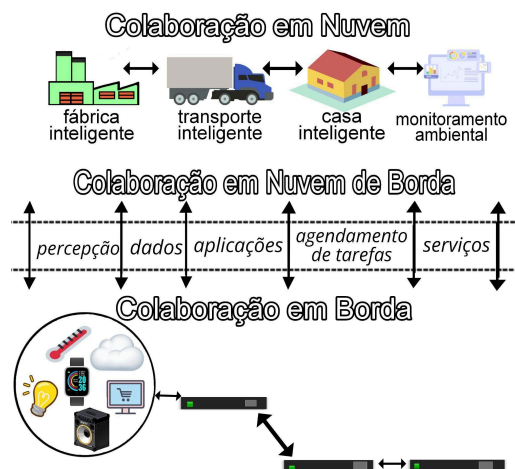
A computação de borda aliada a Internet das Coisas permite que os requisitos em tempo real das aplicações sejam superados devido a capacidade de alocar dinamicamente os recursos, sendo a Computação em Nuvem a espinha dorsal de muitos dos sistemas. Diante disso, as pesquisas acerca desses assuntos, com intuito de aproveitar das vantagens, tiveram crescimento significativo.

Para falar disso, muitos conceitos são importantes, como os três tipos de colaborações, por exemplo:

- Nuvem: diferentes plataformas IoT cooperam, aumentando a flexibilidade da prestação de serviços;
- Borda: máquinas se comunicam entre si na periferia de um sistema, aumentando o desempenho dos aplicativos em Sistemas Distribuídos;
- Nuvem de Borda: vem em diversas formas e a melhoria que traz ao sistema depende da natureza da colaboração.

O tipo utilizado no desenvolvimento do trabalho é o de Colaboração de Dados, mas existem outros, os principais são: **Percepção** (os servidores em nuvem gerenciam os nós de borda e além de monitorar seus estados, fornecem acesso a dispositivos heterogêneos); **Aplicações** (suporte a programas de nós de borda, virtualização de Docker que operam em nós de borda e podem suportar atualizações de programas enquanto o servidor em nuvem fornece o ambiente de desenvolvimento); **Colaboração de Serviços** (implantação de serviços solicitados pelo usuário em nós de borda para satisfazer requisitos específicos); **Agendamento de Tarefas** (orquestração de tarefas que podem abranger vários aplicativos); **Colaboração de Dados** (um servidor em nuvem é responsável por instruir os nós de borda a executarem aspectos específicos de uma tarefa, como filtragem, remoção de ruído, arquivamento, combinação de dados e etc).

Figura 6 - Tipos de colaborações



Fonte: Adaptado de Sun et. al. (2021).

O artigo traz em seu desenvolvimento a integração de plataforma IoT com recursos de Inteligência Artificial (IA) e frisa alguns possíveis problemas em relação aos dados, principalmente quando são heterogêneos. A necessidade de limpeza dos dados antes de serem processados e enviados para a nuvem é muito trabalhosa, já que crescem em escala exorbitante e isso também influencia na latência, trazendo impactos negativos. Para solucionar esses problemas, os autores propõem uma colaboração inteligente de dados que integra os servidores em nuvem e os nós de borda a fim de tirar proveito de suas vantagens.

A colaboração inteligente de dados é um modelo matemático que pode ser independente ou usado em colaboração com outros modelos de IA. [...] formaliza um mecanismo de processamento de dados que fornece dados de alta qualidade (SUN et. al., 2021).

Os resultados obtidos através da comparação feita entre o ADC e o slop máximo do método fixo, com diferentes parâmetros, apontaram que o ADC teve as melhores taxas de filtragem e as menores taxas de erros falsos em uma tarefa de filtragem de ruído com dados de temperatura de uma rede de estação com base IoT do mundo real.

4.5 Discussões

O artigo 1 traz como proposta o uso de dispositivos vestíveis no ramo da saúde com monitoramento 24 horas por dia, todos os dias da semana, a fim de prever e prevenir potenciais doenças a pacientes específicos. A análise de dados coletados dos dispositivos está dividida em quatro categorias: descritiva, diagnóstica, preditiva e prescritiva. Foi utilizada a Computação em Névoa para obter mais precisão ao ter que lidar com a enorme quantidade de dados gerados pelas

aplicações. Comparada às estruturas semelhantes, o autor destacou a superioridade da que propôs (em termos de latência), pois a camada em névoa auxilia no processamento computacional intensivo.

O artigo 2 traz o uso de dispositivos vestíveis no ramo da saúde e a análise de dados médicos já arquivados, propondo resolver o problema da heterogeneidade por meio de middlewares. Assim como o primeiro artigo, recorre à Computação em Névoa e Aprendizado de Máquina e possui resultados significativos em relação à latência, mas mesmo com isso, o autor pretende realizar melhorias em trabalhos futuros.

Diferente dos artigos anteriores, o terceiro traz como proposta um Sistema de Gerenciamento de Dados Industriais baseado em Internet Industrial das Coisas (ou Indústria 4.0), que gerencia, dá suporte e controla a fabricação sem depender completamente de operações humanas. Os dados obtidos são convertidos em banco de dados estruturados para ajudar na manipulação e obtenção de resultados. As melhorias são significativas quanto a escala de produção industrial, qualidade dos produtos, satisfação do cliente e até mesmo receita.

O artigo 4 traz como proposta uma Colaboração Inteligente de Dados, com o objetivo de colaborar de forma adaptativa para a realização de uma tarefa. Para que isso ocorra é utilizada a computação de borda aliada a IoT, que permite que os requisitos em tempo real das aplicações sejam superados. A colaboração de dados se dá onde um servidor em nuvem instrui os nós de borda a executarem aspectos específicos de uma tarefa. Integra Internet da Coisas com recursos de Inteligência Artificial e com isso, torna melhores as taxas de filtragem, diminuindo também as taxas de erros falsos.

Embora todas as ferramentas propostas tenham apresentado resultados positivos, todas elas tiveram dificuldades em relação ao armazenamento dos dados, por serem volumosos e heterogêneos.

Pode se ver que em todos os artigos existe uma grande preocupação com:

- Formatação de dados - Uso de filtros para redução de ruídos, protocolos de formatação de contexto como XML são muito úteis para inibir eventuais problemas de amostras recolhidas ou mesmo o aspecto heterogêneo das fontes de dados.
- Posicionamento do armazenamento - As aplicações apresentam diferentes requisitos de tempo de resposta. Por isso, a posição do banco de dados principal (hora na névoa e hora na nuvem) influencia diretamente na latência da tomada de decisão.

- Redundância de armazenamento - Uso de Computação em Névoa obriga que a aplicação de armazenamento seja distribuída, tendo um conjunto de dados sincronizado com a nuvem por segurança e acesso remoto.
- Utilização inteligente e integrada - A tomada de decisão realizada pela análise de informações isoladas pode prejudicar o resultado final, ao mesmo tempo que uma análise ampla de vários sensores e informações pode tomar tempo para um ser humano. Nesse sentido, em prol da velocidade e aumento da taxa de acerto, o uso de sistemas de análise integradas dos dados IoT é vital para aumento da qualidade da aplicação.

5 - CONSIDERAÇÕES FINAIS

A Internet das Coisas é uma tecnologia cada vez mais emergente, e, embora seja de fácil acesso, é necessário um ambiente favorável e escalável para lidar com o grande volume de dados que são gerados no decorrer de seu uso.

Está presente em diversos ramos da sociedade, como: agricultura, saúde, logística, produção industrial, entre outros. Facilitando a execução de tarefas simples ou complexas, ajudando também na economia e sustentabilidade.

Os objetos inteligentes possuem sensores capazes de coletar dados que são enviados para tomadas de decisões. Os algoritmos utilizados para que isso ocorra são feitos em Aprendizado de Máquina e passam por treinamentos baseados em experiências a fim de reconhecer determinados padrões e prever, ou até mesmo evitar que venham a ocorrer erros, falhas ou anomalias. A comunicação entre esses objetos se dá com a utilização de protocolos de rede sem fio, que dependem de fatores como alcance, latência, disponibilidade de energia, escalabilidade, cenário de aplicação e etc.

Um banco de dados é composto por entidades, que por sua vez, possuem atributos particulares e podem representar objetos existentes no mundo real, sendo uma pessoa, empresa ou até mesmo um curso universitário. Os dados podem ser classificados como homogêneos ou heterogêneos. Quando heterogêneos, podem ser descritos conforme sua estrutura, formato, tipo ou semântica. Já os homogêneos são semelhantes entre si.

Geralmente, o número de dispositivos IoT conectados vai aumentando, e, em decorrência disso, a quantidade de dados aumenta gradativamente. Na maioria das vezes esses dados são heterogêneos e a dificuldade para armazenar e processar também cresce.

Duas plataformas na nuvem desenvolvidas para lidar com grande volume de dados são a Adafruit IO e a Amazon Web Service (AWS), que possuem recursos gratuitos e podem ser acessadas facilmente pelos sites oficiais dos seus desenvolvedores.

Espera-se com esse trabalho entender a utilização e os requisitos das aplicações IoT para o armazenamento de dados em Nuvem. E para isso, alguns trabalhos foram analisados levando em consideração os tipos de tecnologias, protocolos, finalidade e tipos de dados que seriam gerados.

Ainda que as aplicações propostas tenham sido bem sucedidas, houveram preocupações em comum entre elas com relação a formatação dos dados, posicionamento do armazenamento, redundância de armazenamento e utilização inteligente e integrada. Como sugestão para esses desafios, respectivamente, podem ser utilizados protocolos de formatação de contexto, melhorar a

posição do banco de dados principal, armazenamento distribuído e sincronizado com a nuvem e uso de sistemas de análise integradas dos dados IoT.

REFERÊNCIAS

ADAFRUIT. (2023). **Adafruit IO: The Internet of Things for everyone**. Disponível em: <<https://io.adafruit.com>>. Acesso em: 26 jun. 2023.

ALSHAMMARI, Hamoud; EL-GHANY, Sameh Abd; SHEHAB, Abdulaziz. Big IoT healthcare data analytics framework based on fog and cloud computing. **Journal of Information Processing Systems**, v. 16, n. 6, p. 1238-1249, 2020. Disponível em: <https://www.researchgate.net/publication/350213156_Big_IoT_Healthcare_Data_Analytics_Framework_Based_on_Fog_and_Cloud_Computing>. Acesso em: 23 jun. 2023.

AWS. Amazon Web Service. (2023). **Banco de Dados na Nuvem da AWS**. Disponível em: <<https://aws.amazon.com/pt/products/databases/>>. Acesso em: 26 jun. 2023.

AWS. Amazon Web Service. (2023). **O que é IoT**. Disponível em: <<https://aws.amazon.com/pt/what-is/iot/>>. Acesso em: 26 jun. 2023.

CARISSIMI, A. (2016). **Camadas funcionais de uma solução IoT**. Disponível em: <https://www.researchgate.net/figure/Figura-22-Camadas-funcionais-de-uma-solucao-IoT_fig2_301298394>. Acesso em: 22 set. 2023.

COULOURIS, G.; DOLLIMORE, J. KINDBERG, T. (2007). **Sistemas Distribuídos: Conceitos e Projetos**. 4ª e. Porto Alegre, RS: Bookman. Disponível em: <<http://www.inf.ufsc.br/~bosco.sobral/ensino/ine5645/coulouris.pdf>>. Acesso em: 15 nov. 2023.

EMBARCADOS. (2022). **Por que o sucesso da IoT depende da interoperabilidade?**. Disponível em: <<https://embarcados.com.br/sucesso-da-iot-interoperabilidade/>>. Acesso em: 31 out. 2023.

EMBARCADOS. (2022). **Protocolos de Rede sem fio de IoT**. Disponível em: <<https://embarcados.com.br/protocolos-de-rede-sem-fio-de-iot/>>. Acesso em: 01 out. 2023.

ESCOVEDO, T. (2020). **Machine Learning: Conceitos e Modelos - Parte I: Aprendizado Supervisionado**. Disponível em:

<<https://tatianaesc.medium.com/machine-learning-conceitos-e-modelos-f0373bf4f445>>. Acesso em: 01 out. 2023.

FILHO, J. M. (2019). **Monitoramento, com armazenamento de dados em nuvem, das variáveis de ambientes de estoque de peças de aeronáutica**. *Trabalho de Conclusão de Curso* (Graduação em Engenharia de Controle e Automação), Universidade Estadual Paulista (Unesp), Sorocaba, SP. Disponível em: <<https://www.sorocaba.unesp.br/Home/Graduacao/EngenhariadeControleeAutomacao/galdenoro1906/relatoriotg- corrigido.pdf>>. Acesso em: 27 Jun. 2023.

FONTANA, E. (2020). **Introdução aos Algoritmos de Aprendizagem Supervisionada**. *Apostila* (Graduação em Ciência da Computação), Universidade Federal do Paraná (UFPR), Curitiba, PR. Disponível em: <https://fontana.paginas.ufsc.br/files/2018/03/apostila_ML_pt2.pdf>. Acesso em: 01 out. 2023.

KUBOTA, L. C.; ROSA, M. (2023). **Internet das Coisas no Brasil: Análise e recomendação de políticas com ênfase no Agronegócio**. *Nota Técnica (n. 2867)*, Instituto de Pesquisa Econômica Aplicada (IPEA), Rio de Janeiro. Disponível em: <https://repositorio.ipea.gov.br/bitstream/11058/11731/4/TD_2867_Web.pdf>. Acesso em: 22 jun. 2023.

MAGRANI, E. **A internet das coisas**. Editora BOD GmbH, 1 de Janeiro de 2021. *E-book*. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=CyqzEAAAQBAJ&oi=fnd&pg=PA7&dq=conceito+de+internet+das+coisas+&ots=N372akwSMd&sig=wiq8A9MJ_Z3n5SbrwDSEKG1dEeU#v=onepage&q=conceito%20de%20internet%20das%20coisas&f=false>. Acesso em: 21 jun. 2023.

MANCINI, M. **Internet das Coisas: História, Conceitos, Aplicações e Desafios**. *Artigo*. USP, 2020. 10 p. Disponível em: <<http://mmproject.com.br/wp-content/uploads/2020/02/artigo-iot-monicamancini-v1.pdf>>. Acesso em: 01 out. 2023.

NANDANKAR, Praful et al. An iot based healthcare data analytics using fog and cloud computing. **Turk. J. Physiother. Rehabil**, v. 3, p. 32, 2021. Disponível em: <https://www.researchgate.net/publication/352061474_AN_IOT_BASED_HEALTHCARE_DATA_ANALYTICS_USING_FOG_AND_CLOUD_COMPUTING>. Acesso em: 23 jun. 2023.

NICOLETTI, J. L. **Integração de Banco de Dados Heterogêneos através de consultas globais**. *Dissertação* (Mestrado em Ciência da Computação) - Universidade de Caxias do Sul, 2009. Disponível em: <<https://repositorio.ucs.br/xmlui/bitstream/handle/11338/1283/TCC%20Jose%20Luis%20Nicolletti.pdf?sequence=1&isAllowed=y>>. Acesso em: 27 jun. 2023.

NOF, Y. S. **Springer Handbook of Automation**. 2. ed. Cham: Springer Nature, 2023. 1072 p. ISBN 978-3-030-96729-1. DOI: 10.1007/978-3-030-96729-1. Disponível em: <<https://link.springer.com/book/10.1007/978-3-030-96729-1>>. Acesso em: 22 set. de 2023.

OLIVEIRA, R.; KAMIENSKI, C. A. **IoT Redirector: um redirecionador para gerenciamento da heterogeneidade de dados em aplicações IoT**. In: CONGRESSO BRASILEIRO DE ENGENHARIA DE COMPUTAÇÃO, 35., 2022, Natal, Rio Grande do Norte. *Anais [...]*. Natal: SBC, 2022. Disponível em: <<https://sol.sbc.org.br/index.php/courb/article/view/21441/21265>>. Acesso em: 01 out. 2023.

OPENAI. **Aplicações de IoT na Tecnologia**. White paper. 2021. Disponível em: <<https://chat.openai.com/share/b16d2989-4ce6-40ef-bfc0-19aa3df7f9dc>>. Acesso em: 22 jun. 2023.

ORACLE. **O que é IoT?** Artigo. Oracle Brasil, 2023. Disponível em: <<https://www.oracle.com/br/internet-of-things/what-is-iot/#:~:text=As%20aplicações%20de%20IoT%20usam,entre%20falhas%20e%20outras%20informações.>>. Acesso em: 22 jun. 2023.

PRODANOV, C. C.; FREITAS, E. C. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. Novo Hamburgo, RS: Feevale, 2013. 258 p. Disponível em: <<https://revistas.fucamp.edu.br/index.php/cadernos/article/view/2336>>. Acesso em: 27 jun 2023.

RIVERA, Antonio D'Acosta. **Redes de equipamentos sem fio de uso pessoal: comparação de tecnologias emergentes e análise de tendências**. 2010. Dissertação (Mestrado em Sistemas Eletrônicos) - Escola Politécnica, University of São Paulo, São Paulo, 2010. doi:10.11606/D.3.2010.tde-10012011-084232. Acesso em: 27 jan. 2024.

ROSÁRIO, J. M. **Automação Industrial**. *E-book*. Editora Baraúna, 2012. 517 p. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=YsUHLcHdbh4C&oi=fnd&pg=PA9&dq=conceito+de+automação+comercial+e+industrial&ots=3lkWFuSXdi&sig=2aAGXxlHZFLZETQqHPgSIazBOrI#v=onepage&q=conceito%20de%20automação%20comercial%20e%20industrial&f=false>>. Acesso em: 21 jun. 2023.

SAQLAIN, Muhammad et al. Framework of an IoT-based industrial data management for smart manufacturing. **Journal of Sensor and Actuator Networks**, v. 8, n. 2, p. 25, 2019. Disponível em: <<https://www.mdpi.com/2224-2708/8/2/25>>. Acesso em: 24 jun. 2023.

SILVA, R. *Artigo*. **Conceitos de Sistemas Distribuídos**. Dio, 2023. Disponível em: <<https://www.dio.me/articles/conceitos-de-sistemas-distribuidos>>. Acesso em: 15 nov. 2023.

SOFIST. *Post*. **Internet das Coisas: Brasil já tem cerca de 20 milhões de objetos conectados**. One Day Testing, 2019. Disponível em: <<https://blog.onedaytesting.com.br/internet-das-coisas-brasil/>>. Acesso em: 26 jun. 2023.

SUN, Danfeng et al. Intelligent data collaboration in heterogeneous-device iot platforms. **ACM Transactions on Sensor Networks (TOSN)**, v. 17, n. 3, p. 1-17, 2021. Disponível em: <<https://opus.lib.uts.edu.au/handle/10453/158148>>. Acesso em: 25 jun. 2023.

SYOZI, R. *Post*. **O que é AWS [Amazon Web Service]**. Technoblog, 2021. Disponível em: <<https://tecnoblog.net/responde/o-que-e-a-aws-amazon-web-services/>>. Acesso em: 14 nov. de 2023.

TAKAI, O.K.; ITALIANO, I.C.; FERREIRA, J. E. **Introdução a Banco de Dados**. São Paulo, SP: IME-USP, 2005. Disponível em: <<https://www.ime.usp.br/~jef/apostila.pdf>>. Acesso em: 27 jun. 2023.