International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
e-ISSN: 2455-2585
Volume 10, Issue 10, October-2024

# The Impact of Data Quality Assurance Practices in Internet of Things (IoT) Technology

Kishan Patel[*1],
[*1]Independent Researcher, kpatelscholarnetwork@gmail.com
Sandeep Gupta[*2]
[*2]SATI, Vidisha, M.P. India, Sandeepguptabashu@gmail.com

**ABSTRACT-A rapid growth of the Internet of Things has led to its widespread use in several fields, including precision agriculture and smart cities. At the moment, much of the research and development in the field of the IoT is going into creating integrated platforms that will allow it to fulfil its potential in large-scale commercial applications. IoT data quality is critical in real-world applications. We rely on data to generate goods, make informed choices, and learn about people, places, and things in the world, thus it's important to think about the aspects and problems related to the quality of IoT data. We cannot get useful findings from low-quality data, and that is an important point to make. However, ensuring the quality of this data presents significant challenges due to the diverse sources, formats, and velocities involved. This review paper examines the critical role of Data Quality Assurance (DQA) practices in mitigating these challenges and enhancing the reliability, integrity, and usability of IoT data. Such topics are significant DQA frameworks, techniques for data validation, and data cleaning, data integrity management, and implications of DQA on sectors such as smart city, healthcare, and industrial automation. Thus, this study goal to offer an understanding of the prominent aspects of effective DQA practices in enhancing IoT operations and supporting decision-making across different fields based on the existing academic works and industrial trends. Further research opportunities are growing effective methods of anomaly detection, data recovery and counteracting privacy and network reliability concerns to enhance the IoT data quality.**

**Keywords:** Internet of Things (IoT), Data Quality Assurance (DQA), Data Quality (DQ), Open issues, Taxonomy, Factors affecting QD.

## I. INTRODUCTION

The new internet, known as the IoT, is constantly creating massive volumes of data as it links various entities, things, and items from all over the world. The data generated by the IoT is more suited for consumption by the linked items than by humans. You can tell how many internet-connected devices there are by looking at the relatively high number of servers required to store data for user access[1].

A primary element in determining the foundation that permits intelligent decision-making and the exploration and utilisation of additional services is Internet of Things (IoT) data[2],. Data comes from intelligent objects that can link and exchange vast amounts of data with other objects in the IoT ecosystem. The reliability of the data, however, determines both a quality of the data gathered and a service offered. A most reliable information is high-quality data because it allows us to draw useful conclusions from events and helps us serve people well. Therefore, actions supported by reliable and high-quality data may be beneficial, whereas those supported by inaccurate or unreliable data are detrimental and may lead to serious outcomes. In order for businesses to reap the benefits of their IoT projects, data integrity is paramount. Because information can be utilised to make new choices, create new items, and expand markets, data is one of the most valuable assets in the IoT. The importance of high-quality data for data mining operations and how poor data quality affects the trustworthiness of results have been the subject of several research. Appropriateness of data for analytical purposes degree to which a collection of inherent qualities fulfils needs [3] [4][5], and "how well data satisfies the expectations of consumer" are all definitions of data quality[6]. Additionally, data quality refers to an appropriateness of the collected data from smart devices for the Internet of devices [7]. The data quality area faces new challenges due to the variety of sources and the amount of data. In order to determine if data sets are "fit for use," researchers and practitioners must first determine this[8][9]. Poor judgements are a direct result of a lack of knowledge about data quality, which is why data quality has emerged as a key component of the IoT. Every user of the data relies on the used data meeting their own requirements in order to complete critical work[10]. Accuracy, timeliness, completeness, and dependability are some of the Data Quality Dimensions that include these requirements. Data quality, volume, accuracy, security, privacy preservation, and trust have all been highlighted in various publications as important considerations for the IoT. There are four key ways data quality is categorised in literature: accessibility, representation, contextual, and intrinsic[11][12]. From the many perspectives highlighted in this paper, a number of papers have investigated data quality aspects, problems, and methods for enhancing data quality within the framework of the IoT.

**Motivation and Significance**

A rapid proliferation of the IoT has revolutionised a way entities, objects, and systems interact, generating vast amounts of data in real-time. However, the efficacy and reliability of IoT applications are fundamentally tied to a quality of this data. Poor data quality can lead to erroneous decisions, potentially causing severe consequences in critical applications such as healthcare, smart cities, and industrial automation. Therefore, ensuring high data quality is paramount. This study

of data quality assurance practices in IoT Technology is motivated by the necessity to address the diverse and complex challenges inherent in IoT data, such as resource constraints, heterogeneity, and environmental factors. By exploring and enhancing data quality assurance practices, this research aims to improve the reliability, accuracy, and overall utility of IoT data, thus enabling more intelligent decision-making processes and fostering the development of robust, scalable, and trustworthy IoT ecosystems. This review paper contributes as:

- **Enhanced Reliability:** Effective handling of data quality creates confidence in IoT data with processes of decision-making and business operations benefiting from quality data.
- **Improved Data Integrity:** By applying proper QA techniques, it is possible to protect the IoT data from the low quality meaning its lack of consistency, incompleteness, or inaccuracies, which are especially likely to occur in the context of massive and heterogeneous IoT systems.
- **Optimized Resource Utilization:** Effective QA procedures enable efficient use of the available resources in the IoT systems since the QA overhead includes the time needed for the cleaning, validation, and correction of the data.
- **Mitigation of Operational Risks:** Implementing preventive tools and strategies to counter the issues connected with data quality enhances performance and mitigates operational risks in terms of wrong decisions making, failures within IoT systems, and vulnerable data protection.
- **Facilitation of Innovation:** Accurate and high-quality IoT data promotes innovation as predictive analysis, machine learning models, and insights in real-time increase the chances of fresh applications and opportunities in the business world.
- **Promotion of Trust and Adoption:** Ensuring data quality enhances stakeholder trust in IoT technologies, promoting wider adoption across industries and sectors, thereby contributing to the advancement of smart cities, healthcare systems, industrial automation, and other IoT-driven domains.

These contributions underscore the critical importance of robust data quality assurance practices in realising the full potential and benefits of IoT technology across diverse applications and industries.

The following paper organised as: Section II provide the aspect of data quality in IoT, then provide the DQA techniques within IoT in Section III, And Section IV give the taxonomy for QD of IoT, Next Section V and VI discuss a factor affecting and open issues of data quality in IoT. Section VII provide the literature review on DQA in IoT area. Last Section VIII summarise the work with conclusion and future work.

## II. THE ASPECT OF DATA QUALITY IN IOT

Researchers and practitioners alike have come to appreciate the significance of high-quality IoT data as the

network expands into more and more domains. Both the quantity and quality of data needed by various applications and organisations are context and application-specific. Previous study [13][14] has made it clear that the DQ management approach often involves an essential step at the beginning when the data and DQ are defined. This covers the kind of data as well as its context. Additionally, DQ evaluation is impacted by the analysis of the systemic components that may have an influence on DQ. This chapter explains the definition of DQ, offers an overview of how general and IoT data are classified, and talks about some possible influences on IoT DQ.

**Types of Data**

The provided definition states that data are "selected aspects of real-world objects, events, and ideas represented and understood by specifically stated standards related to their meaning, collection, and storage." In some research, the terms "information" and "data" are used interchangeably with no obvious difference. In [15], the term "data" is used by the authors to refer to structured data found in databases; other, more general forms of data, such as linked open data and big data, are characterised by "information." In some studies [16], instead of naming a particular kind of data, the phrase "information" is used to emphasise that any sort of data may be included in the research. Depending on its use in various contexts, data may be classified into several categories. As seen in Table 1, researchers suggested several categories for the data. The data categorisation method based on data structure is the most often employed in the research that are currently available.

 

**Table 1.** Classifications for data[17].

| Basis | Data Types | Description |
|---|---|---|
| Structure | Structured data | Data that has a defined formal schema; (for instance, relational tables) |
| | Unstructured data | Generic sequence of symbols (e.g., video) |
| | Semi-structured data | Data that lacks a schema or is only partially organised (e.g., XML file) |
| Change frequency | Stable data | Data impossible to change |
| | Long-term changing data | Data that changes very infrequently |
| Product | Frequently changing data | Dramatically changing data, (e.g., real-time traffic information) |
| | Raw data items | Data that have not been processed |
| | Information products | Results of manufacturing activities |
| Nature | Component data items | Semi-processed information |
| | Federated data | Data from different heterogeneous sources |
| | Web data | Data from the Web |
| | High-dimensional data | Big data |
| | Descriptive data | consists of several tables with intricate interactions between them. |
| | Longitudinal data | Time series data |
| | Streaming data | Data produced consecutively at a greater pace in a single source. |

### Data Quality

DQ has been defined in a variety of ways across different disciplines and historical periods. The concept of DQ can be interpreted from two distinct perspectives: practical measurement, which emphasises user satisfaction, and system-oriented evaluation, which acknowledges DQ as a comprehensive and multidimensional concept. It is critical to assess its core attributes from many angles, including timeliness, accuracy, completeness, and consistency.

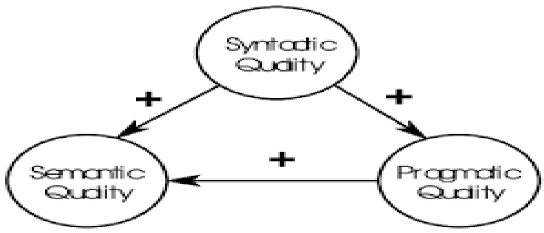


**Figure 1:** Categories of Data Quality

Figure 1 displays Data syntactic quality, including metadata consistency, is a degree to which data follows a defined syntax. A data's semantic quality is a measure of how well it represents the meaning it represents. Data are considered to be of pragmatic quality if they are suitable and useful for a certain goal [18]. Specifically, DQ incorporates the following concepts as outlined in ISO 8000:

- The data is suitable for its intended use;
- Being at the correct location with the correct information at the right moment;
- Complete the data needs as agreed upon by the customer;
- The elimination of waste and duplication throughout enhancement stages and the prevention of data defect recurrence.

## III. QUALITY OF DATA ASSURANCE TECHNIQUES IN AN IOT PARADIGM

Never trust things" is a more appropriate guideline to utilise in the IoT than the old adage "Never trust user input" in traditional programming. This is shown by the fact that there are uncertainties and inconsistencies in the sensor data. To lessen an impact of low QoD and its expensive consequences, data preparation and improvement are essential [19][20]. The five

most promising methods for achieving QoD in an IoT environment are as follows:

### Outlier detection

The objective is to identify the variables that do not follow the expected pattern of data or that do not fit the normal distribution. Showing anomalies is the end aim. Enabling outlier identification in a model enhances its overall efficiency and dependability. Moreover, the first stage in handling all instances of discrepancies is to identify outliers. The trustworthiness and precision of data processing follows. QoD will be guaranteed and more informed decisions will be rendered if these issues are resolved [21].

### Interpolation

Interpolation is a way to generate data that may increase the quality of data size (i.e., include all the accessible data pieces), thereby improving the QoD dimension. One way to look at completeness is as a ratio of accessible data to non-interpolated items in the stream window. Interpolation, on the other hand, is the inverse of completeness in effect. Meeting user-defined quality-of-life standards is the only objective here, yet finding the sweet spot between the two might be seen as an optimisation effect [22].

**Data Integration**

To go above structural disparities and inconsistencies and really help the universal service, all the diverse data from various landscapes has to be combined. Standardised data description methods for more direct processing, retrieval, and search are provided by frameworks for DQ approaches like the RDF and the OWL, 2009. Once again, a paradigm known as the Service Architecture Paradigm (SOA) is created in order to enhance interoperability within the context of e-health systems by extracting heterogeneity from intelligent objects.

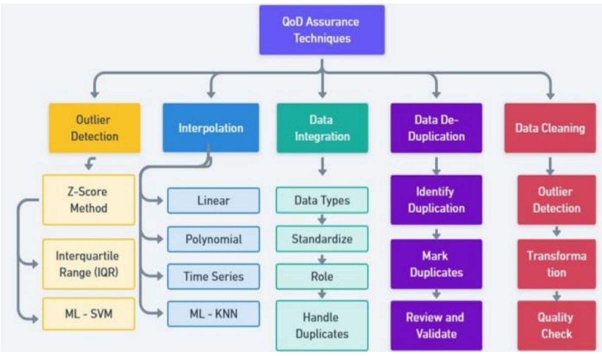


**Figure 2:** QoD Assurance techniques[1].

**Data Deduplication**

compresses data by removing redundant information and substituting it with distinct but unaltered references, therefore reducing the amount of data saved. Reducing data redundancy is what data deduplication is all about. A decrease in data quality is associated with a decrease in data quantity.

**Data Cleaning**

Data cleaning is a first step in a data life cycle, which also includes error selection, error rectification, and error identification and potential detection. In the field of big data analysis, data cleansing has been extensively studied. There are primarily 3 steps to it: (i) figuring out what kind of mistake it is, (ii) finding other possible mistakes, and (iii) fixing those mistakes. Data storage is another typical setting for managing company data. Fig. 2 shows a few methods for quality of data.

## IV. TAXONOMY FOR QUALITY OF DATA IN IOT

An IoT and related areas, including WSN[23], share a common set of deployment-and implementation-friendly traits, features, characteristics, protocols, and technologies. Data trustworthiness measurement approaches, key performance indicators for quality of service assurance, domain integration in the IoT data trustworthiness ecosystem, and previous and future work in the field may all be better understood with the help of a taxonomy.

**Data Source**

The reliability of the data is determined by its source, which may originate from one or more sources. The data could come from people, especially if the model or the environment requires interactive input, or it might come from sensors that are placed to capture readings of occurrences.

**Data Processing**

Data is often provided in the first form, known as a stream or batch, at a predetermined period. The majority of the data inside a batch is same or similar. The device's temporary memory is the first data collection point and storage location until data is sent to the server or another base station over the network. The sensor may provide batches of data for processing. Time must also be taken into consideration, since such data cannot be processed instantly.

**Data Type**

The user may be presented with numerical statistics, symbols, or both based on the data retrieved from the sensor. The parameter being measured and the way the results are represented determine the kind of data.

**Trust Type**

There are two main kinds of trust: direct and indirect. The term "direct trust" describes data that comes straight from a sensor, while "indirect trust" describes data that travels from one sensor to another through a network, where it is more likely to be intercepted or otherwise changed due to factors outside of the network's control.

**Trust Computation Location**

A reliability of a node may be ascertained by computing its location. There are two possible computing models: distributed and centralised. If the value is less than a certain threshold, the cluster head in a WSN confirms the node's integrity; if not, the data is sent to a gateway and then an application layer. This is thought of as a computation that is centralised. In contrast, each node in a distributed trust computation system determines its own reliability and communicates that information to the rest of the system via the gateway and cluster.

**Data QoS**

The maximum number of packets received should then be obtained by adding up all of the packets in the transmission that have the maximum delay (400 ms standard) and a consistent jitter interval (1 ms) between their successive packets.

**Node Quality**

There are two varieties of this: one is a resistant node that can withstand side-channel assaults, is unclonable, and is immune to memory extraction of any kind. For further security, most of these nodes have many ICs stacked on top of each other, making it almost impossible for an attacker to steal any data from them. An unresistant node is one that cannot defend itself against the assault stated.

**Data Accuracy**

As long as the node's data is exact, correct, and dependable, it is deemed accurate. Accurate data serves as the foundation for wise decision-making, but incomplete data leads to harsh decisions with unfavourable outcomes. Accurate data is stated to accurately depict the course of an event as it really occurs.

**Measurable Parameters**

The purpose of configuring these nodes to measure parameters is to determine various amounts. The characteristics can include things like the item's volume, the room's or location's temperature, the equipment's pressure, the humidity, the ground's slope, the object's fitness level, the pipes' sleeping habits, the area's condition, the liquids' movement positions, and so on.

**Data Consistency**

Standards and integrity must be met by the data in order for it to be considered consistent, and the codes that are burnt into the device must consistently provide the correct output each time the program is executed.

## V. FACTORS AFFECTING QUALITY OF DATA IN IOT DEVICES

The IoT's data is its own weakest link since there are several variables that might compromise its quality. Poor quality data makes it impossible for it to accurately depict the situation it is meant to track and may have further detrimental repercussions on the choice being made as well as the operational levels of any company or organisation [24][25][26]. Some possible IoT issues must be resolved for there to be a phenomenon of interest. The following are only a few of the issues with maintaining QoD in the IoT:

- **Resource constraints:** The IoT has long been defined as having limited resources. Data flowing from IoT devices isn't always reliable due to their low processing power, little memory, and minuscule battery consumption. Particularly when the data produced by the gadget exceeds its storage capacity, which is an inevitable consequence of the ever-increasing data production pace of these devices.
- **Scalability:** Global deployment of IoT is already taking place, originating from households, businesses, cities, and now the whole world. Big volumes of data are produced by every IoT deployment option, and when settings are combined or applications are integrated, the quantity of data increases even more, raising the possibility of errors in the data.
- **Heterogeneity:** The data collected by IoT devices varies depending on their origin, just as the devices themselves originate from diverse environments. Data of the same kind is inherently easier to handle than data of a different type. Heterogeneous data enables IoT devices to function at their peak efficiency. Consequently, ideal resolution of the heterogeneity problem is required.
- **Sensors:** It is possible for sensors to lose calibration or give inaccurate values when they are deployed. As a result of malfunction, certain sensors may provide false readings. In a large-scale deployment, this is a major obstacle that makes identifying the defective sensor very difficult.
- **Environment:** The majority of the deployment takes place in an exposed area that is vulnerable to storms, earthquakes, erosion, the top of the mountain, strong winds, and even human aggression.
- **Network:** The connection is sometimes dropped and then regained due to limited resources, bad weather, interference from infrastructure, and a weak signal. An IP network, the IoT has lower packet loss than conventional networks.
- **Vandalism:** Physical attacks, such as data theft, alteration, destruction, and forced extraction, are commonplace in the environment due to its lack of protection. Animals who go out into the wild in search of food or a new home are also vulnerable to vandalism. Consequently, the QoD is impacted by this component.
- **Dead node:** Numerous circumstances may result in a dead node, although the node is still receiving data. The quality of the data has become untrustworthy as a result.
- **Privacy:** This plays a significant role in the IoT's widespread adoption. There is no assurance that the data of individuals will be safe, and when data breaches occur (such as with medical data), the harm is too great.
- **Data stream:** The back-end ubiquitous apps that employ the smart IoT devices continually receive and send data from them.

Additional issues include incompleteness, unauthorised access, certain nodes' insomniac habits, changing the source code and properties, etc. Constraints prevented the memory devices from reporting events often or sending big packets; as a result, only small messages could be delivered, which is inadequate to record every occurrence. Additionally, due to resource constraint, items will enter a sleep state in order to save energy. Nevertheless, due to Internet Protocols (IP) being inappropriate for sleep modes, the smart objects must be operational at all times. This poses a number of challenges, such as inaccurate readings, data duplication, data leakage, inconsistencies and misalignment among several sources, etc.

## VI. OPEN ISSUES IN QUALITY OF DATA FOR IOT

Here are a few topics that need further investigation, divided into four main groups [1]:

**Scalability**

At now, there is a noticeable increase in the deployment of the IoT, surpassing even the scope of the conventional Internet on an unprecedented scale. Instead of offering enough flexibility and scalability for large-scale deployment, the majority of solutions are concentrated, and this sets them apart from distributed systems.

**Heterogeneity of Data Sources**

Data is created by an IoT ecosystem by many different places, such as sensors, RFID tags, objects, entities, and more. The variety of data sources must be accommodated by the architecture designed for the IoT. The suggested technologies need to address a lot of things in order to be suitable for IoT applications. To meet the needs of IoT applications, which may provide complex services based on criteria such as user actions, energy consumption, and the relative indoor and outdoor temperatures [27].

**Domain-agnostic/automated verification**

According to their setup, objects in IoT visions automatically communicate data with nearby nodes. Domain-agnostic data cleaning techniques verify that information is sent between "things" in an uninterrupted, human-free, and minimally controlled manner—a crucial component of a flawless IoT service [28]v.

**Distributed Architectures**

When it comes to the IoT, distributed designs solve problems with scalability and robustness in the face of errors. From an IoT point of view, these tasks are critical since data cleaning infrastructure is always available and continues to provide comprehensive services, even when ecosystems fail [29].

## VII. LITERATURE REVIEW

This section summarises prior literature reviews on DQ in IoT. Reza et al. [30] established an adaptable and widely applicable " OODA" methodology for identifying and improving DQ across sectors, kinds of organisations, and scales of operation. Observe, orient, decide, and act are the four phases that make up the OODA framework cycle. Following the principle of using existing DQ metrics and measurement tools, the OODA DQ method just stresses the need for a measuring algorithm for each DQ dimension. While OODA DQ

methodology lacks a structured approach to analysis and improvement, tools like dashboards and regular reporting may help identify DQ problems while they are still in observe phase.

Carretero et al., (2016) developed a versatile framework known as the "MADM frame-work" for process reference modelling, assessment, and improvement. At last, it was tested using a real-life hospital scenario. A MADM Framework cycle is comprised of the two-part Process Reference Model according to ISO 8000-61 and the Assessment and Improvement Model compliant with ISO/IEC 33000. With its twenty-one procedures, the MAMD Process Reference Model might be useful in three areas: data quality management, data governance, and data management[31].

Sebastian-Coleman, (2013) introduced a "DQAF" that provides 48 standard measurement types based on integrity, timeliness, validity, consistency, and completeness. This framework gives DQ assessment companies a broad range of objective DQ metrics to choose from. A "measurement type" is proposed by the authors of DQAF as a generic form that is suitable for a specific metric. They then come up with ways to describe the six characteristics of every measure type: support processes, business concerns, definition, programming, measurement methodology, and a logical model for both measures and skills[32].

Carlo et al., (2011) introduced a " HDQM" that has been validated via the use of examples and may be used to assess and enhance the DQ. Additional data kinds may be included into the HDQM measurement and improvement phase by using the DQ dimensions. A significant advancement in HDQM is that it offers a more qualitative method for guiding the selection of suitable improvement strategies by drawing on cost-benefit analysis methods used in TIQM, COLDQ, and CDQ[33].

Angeles and García-Ugalde, (2009) detailed "DQPA" that provided an overview of a DQ framework and demonstrated its application with a use case in a heterogeneous multi-database setting. The developers of DQPA offer the Measurement Model by extending DQ assessment metrics to include metrics for evaluating both the original data sources and created data[34].

Despite these advancements, research gaps remain, such as the lack of standardised formal processes for DQ improvement in some methodologies, limited integration of dynamic and real-time data quality assessment, and the need for more comprehensive cross-industry applicability and adaptability to rapidly evolving IoT environments.

**Table 2.** Summary of Data Quality Methodologies in IoT Technology

| Authors | Methodology | Framework Cycle | Key Features |
|---|---|---|---|
| Reza et al. [30] | Observe–Orient–Decide–Act (OODA) | Observe, Orient, Decide, Act | • Adaptive, employ of current DQ metrics and tools<br>• Issues identified through routine reports, dashboards, alerts, and feedback |
| Carretero et al. [31] | Alarcos Data Improvement Model (MADM) | Process Reference Model (21 processes), Assessment and Improvement Model | • Based on ISO 8000-61 and ISO/IEC 33000<br>• Applicable to data governance, data management, and DQ management quality |
| Laura et al. [32] | Data Quality Measurement Framework (DQAF) | Define, Measure, Analyze, Improve, Control | • 48 universal DQ metrics focusing on validity, completeness, consistency, timeliness, integrity<br>• Continuous monitoring |
| Batini et al. [33] | Heterogeneous Data Quality Methodology (HDQM) | State Reconstruction, Assessment, Improvement | • Applicable to different data types<br>• Cost-benefit analysis, qualitative approach for improvement techniques |
| Maria et al. [34] | Data Quality Practical Approach (DQPA) | Identification of DQ issues, Relevant Data Identification, Evaluation, Business Impact Determination, Data Cleansing, Monitoring, Assessment | • Metrics for primary and derived data<br>• Applicable to multi-database environments, different levels of granularity |

## VIII. CONCLUSION AND FUTURE WORK

The impact of Data Quality Assurance (DQA) practices in IoT technology is profound, addressing the critical issues of data availability and veracity which are pivotal for the accurate functioning of IoT systems. This review emphasises that DQA practices play a critical in guaranteeing the success and dependability of IoT implementations. Within this paper, starting with the analysis of the issues concerning IoT data quality, the role of DQA for the upgrade of IoT environments has been described regarding data reliability, integrity, and pragmatic utility. The outcomes have stressed that the prospective DQA actions do not only reduce risks connected with the insufficient quality of formal data, but also develop

innovation and confidence in IoT solutions all through the spheres. Further research and development of DQA together with the consolidation of globally unified requirements will remain as the key drivers of the worldwide IoT market development and the progress towards a more connected, smart future. Moreover, essential themes of privacy, stream management, and dependable networks shall be essential to attaining further DQA in IoTs meaning conducive to enhancing the recognition of reliable IoTs around the world.

The future of IoT data quality research comprises several promising avenues in the following areas. Firstly, developing new and improved methods of using machine learning to detect anomalies and reconstruct lost data will improve the actuality and realism of data. Integrating various and dispersed data in an efficient and sophisticated way is an answer to scalability and heterogeneity challenges. Also, advancements in Lite IoT devices that are powered by improved computation and systematic power control will help overcome resource limitations. Stressing on aspects of security and protection from hacking and vandalism plays the crucial role of safeguarding the data. Likewise, the refinement of network protocols to cope with the loss of connection and the steady control of the stream data will add to the overall quality of the IoT data. All these endeavours will in a combined fashion help in building more dependable and efficient IoT systems that will effectively assist in better decision-making processes and thereby facilitate improved IoT application throughout different domains.

## IX. REFERENCES

[1] J. S. Yalli, M. H. Hasan, N. S. Haron, M. U. R. Shaikh, N. Y. Murad, and A. L. Bako, "Quality of Data (QoD) in Internet of Things (IOT): An Overview, State-of-the-Art, Taxonomy and Future Directions," *Int. J. Adv. Comput. Sci. Appl.*, 2023, doi: 10.14569/IJACSA.2023.01412110.

[2] J. Merino, I. Caballero, B. Rivas, M. Serrano, and M. Piattini, "A Data Quality in Use model for Big Data," *Futur. Gener. Comput. Syst.*, 2016, doi: 10.1016/j.future.2015.11.024.

[3] T. Banerjee and A. Sheth, "IoT Quality Control for Data and Application Needs," *IEEE Intell. Syst.*, 2017, doi: 10.1109/MIS.2017.35.

[4] V. Rohilla, M. S. S. Kumar, S. Chakraborty, and M. S. Singh, "Data Clustering using Bisecting K-Means," in *Proceedings - 2019 International Conference on Computing, Communication, and Intelligent Systems, ICCCIS 2019*, 2019. doi: 10.1109/ICCCIS48478.2019.8974537.

[5] V. V. Kumar and F. T. S. Chan, "A superiority search and optimisation algorithm to solve RFID and an environmental factor embedded closed loop logistics model," *Int. J. Prod. Res.*, 2011, doi: 10.1080/00207543.2010.503201.

[6] K. Patel, "Quality Assurance In The Age Of Data Analytics: Innovations And Challenges," *Int. J. Creat. Res. Thoughts*, vol. 9, no. 12, pp. f573–f578, 2021.

[7] A. Karkouch, H. Mousannif, H. Al Moatassime, and T. Noel, "A model-driven framework for data quality management in the Internet of Things," *J. Ambient Intell. Humaniz. Comput.*, 2018, doi: 10.1007/s12652-017-0498-0.

[8] R. Fries, "Software Verification and Validation," *Reliab. Des. Med. Devices, Third Ed.*, pp. 401–410, 2012, doi: 10.1201/b12511-35.

[9] P. Khare, "Enhancing Security with Voice: A Comprehensive Review of AI-Based Biometric Authentication Systems," *Int. J. Res. Anal. Rev.*, vol. 10, no. 2, pp. 398–403, 2023.

[10] V. V. Kumar, F. W. Liou, S. N. Balakrishnan, and V. Kumar, "Economical impact of RFID implementation in remanufacturing: a Chaos-based Interactive Artificial Bee Colony approach," *J. Intell. Manuf.*, 2015, doi: 10.1007/s10845-013-0836-9.

[11] Y. W. Lee, D. M. Strong, B. K. Kahn, and R. Y. Wang, "AIMQ: A methodology for information quality assessment," *Inf. Manag.*, 2002, doi: 10.1016/S0378-7206(02)00043-5.

[12] T. Mansouri, M. R. S. Moghadam, F. Monshizadeh, and A. Zareravasan, "IoT Data Quality Issues and Potential Solutions: A Literature Review," *Comput. J.*, vol. 66, no. 3, pp. 615–625, 2023, doi: 10.1093/comjnl/bxab183.

[13] C. Batini, C. Cappiello, C. Francalanci, and A. Maurino, "Methodologies for data quality assessment and improvement," *ACM Comput. Surv.*,

2009, doi: 10.1145/1541880.1541883.
[14] S. Kumar, V. K. Solanki, S. K. Choudhary, A. Selamat, and R. G. Crespo, "Comparative study on ant colony optimization (ACO) and k-means clustering approaches for jobs scheduling and energy optimization model in internet of things (IoT)," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 6, no. 1, pp. 107–116, 2020, doi: 10.9781/ijimai.2020.01.003.
[15] S. E. Madnick, R. Y. Wang, Y. W. Lee, and H. Zhu, "Overview and framework for Data and information quality research," *J. Data Inf. Qual.*, 2009, doi: 10.1145/1515693.1516680.
[16] L. P. English, "Warehouse and Business Information Quality Methods for Reducing Costs and Increasing Profits," *Time*, 1999.
[17] L. Zhang, D. Jeong, and S. Lee, "Data quality management in the internet of things," 2021. doi: 10.3390/s21175834.
[18] ISO, "8000:8 Information and data quality - Concepts and measuring," *Int. Stand. ISO 8000*, 2015.
[19] A. P. A. Singh, "STRATEGIC APPROACHES TO MATERIALS DATA COLLECTION AND INVENTORY MANAGEMENT," *Int. J. Bus. Quant. Econ. Appl. Manag. Res.*, vol. 7, no. 5, 2022.
[20] R. Togneri, G. Camponogara, J. P. Soininen, and C. Kamienski, "Foundations of Data Quality Assurance for IoT-based Smart Applications," in *Proceedings - 2019 IEEE Latin-American Conference on Communications, LATINCOM 2019*, 2019. doi: 10.1109/LATINCOM48065.2019.8937930.
[21] M. Bures, T. Cerny, and B. S. Ahmed, "Internet of things: Current challenges in the quality assurance and testing methods," in *Lecture Notes in Electrical Engineering*, 2019. doi: 10.1007/978-981-13-1056-0_61.
[22] J. Buelvas, D. Múnera, D. P. Tobón V, J. Aguirre, and N. Gaviria, "Data Quality in IoT-Based Air Quality Monitoring Systems: a Systematic Mapping Study," 2023. doi: 10.1007/s11270-023-06127-9.
[23] S. Mathur and S. Gupta, "An Energy-Efficient Cluster-Based Routing Protocol Techniques for Extending the Lifetime of Wireless Sensor Network," in *2023 International Conference on the Confluence of Advancements in Robotics, Vision and Interdisciplinary Technology Management, IC-RVITM 2023*, 2023. doi: 10.1109/IC-RVITM60032.2023.10434975.
[24] C. Liu, P. Nitschke, S. P. Williams, and D. Zowghi, "Data quality and the Internet of Things," *Computing*, 2020, doi: 10.1007/s00607-019-00746-z.
[25] S. Kumar and Z. Raza, "A Priority-Based Message Response Time Aware Job Scheduling Model for the Internet of Things (IoT)," *Int. J. Cyber-Physical Syst.*, vol. 1, no. 1, pp. 1–14, 2019, doi: 10.4018/ijcps.2019010101.
[26] J. Thomas, "The Effect and Challenges of the Internet of Things (IoT) on the Management of Supply Chains," *Int. J. Res. Anal. Rev.*, vol. 8, no. 3, pp. 874–878, 2021.
[27] I. Bedhief, M. Kassar, and T. Aguili, "Empowering SDN-Docker Based Architecture for Internet of Things Heterogeneity," *J. Netw. Syst. Manag.*, 2023, doi: 10.1007/s10922-022-09702-3.
[28] J. Gaskin, A. Elmaghbub, B. Hamdaoui, and W. K. Wong, "Deep Learning Model Portability for Domain-Agnostic Device Fingerprinting," *IEEE Access*, 2023, doi: 10.1109/ACCESS.2023.3305257.
[29] F. Azzedin and T. Alhazmi, "Secure Data Distribution Architecture in IoT Using MQTT," *Appl. Sci.*, 2023, doi: 10.3390/app13042515.
[30] A. Sundararaman and S. K. Venkatesan, "Data quality improvement through OODA methodology," in *Proceedings of the 22nd MIT International Conference on Information Quality, ICIQ 2017*, 2017.
[31] A. G. Carretero, A. Freitas, R. Correia, and M. Piattini, "A case study on assessing the organizational maturity of data management, data quality management and data governance by means of MAMD," in *Proceedings of the 21st International Conference on Information Quality, ICIQ 2016*, 2016.
[32] L. Sebastian-Coleman, *Measuring data quality for ongoing improvement: A data quality assessment framework*. 2013. doi: 10.1016/C2011-0-07321-0.
[33] B. Carlo, B. Daniele, C. Federico, and G. Simone, "A Data Quality Methodology for Heterogeneous Data," *Int. J. Database Manag. Syst.*, 2011, doi: 10.5121/ijdms.2011.3105.
[34] M. del P. Angeles and F. García-Ugalde, "A Data Quality Practical Approach," *Int. J. Adv. Softw.*, 2009.