Online at www.e-iph.co.uk
Indexed in Clarivate Analytics WoS,

E-B
Environment - Behaviour
Proceedings Journal

ICLT2024NakhonPathom
https://iclt.uitm.edu.my/

e-IPH
e-International
Publishing House Ltd.,
United Kingdom

3rd International Conference on Logistics and Transportation
SSRU, Nakhon Pathom, Thailand, 24 - 25 Oct 2024
Organised by: Research Nexus UiTM (ReNeU), Universiti Teknologi MARA

Secure Data Provenance and Semantic Interoperability Model for Big Data in IoT: A healthcare perspective

Nisar Hussain*, Amna Qasim, Muhammad Zain, Grigori Sidorov
**Corresponding Author*

Centro de Investigación en Computación,
Instituto Politécnico Nacional, Mexico City, Mexico

nisar.hussain8400@gmail.com, amnaq2023@cic.ipn.mx, muhammad23@cic.ipn.mx, sidorov@cic.ipn.mx
Tel: +52 55 6176 7696

Abstract
The Internet of Things (IoT) is a heterogeneous network involving diverse communication models. Ensuring trustworthy and reliable data is a major challenge, especially due to the lack of standards for semantic interoperability (SI). This issue is critical in applications like healthcare, where secure data provenance and effective integration are essential. Current tools are often inadequate for achieving SI across diverse IoT devices. This research proposes a novel model to ensure secure SI and robust data provenance in healthcare IoT systems, aiming to improve communication, reliability, and security across connected devices.

Keywords: Data Provenance; Semantic Interoperability; Big Data; Internet of Things (IoT).

1.0 Introduction

The Internet of Things (IoT) refers to an ecosystem of digitally connected physical objects capable of sensing, collecting, processing, and exchanging data over the internet without continuous human intervention. Through the integration of embedded sensors, actuators, radio-frequency identification (RFID), and network connectivity, IoT transforms everyday objects into intelligent, communicative entities that enable the development of smart environments (Ali et al., 2018; Kiljander et al., 2014; Pavithra & Balakrishnan, 2015). Central to IoT functionality are context-aware data processing and intelligent networking mechanisms, which allow devices to interpret environmental conditions and respond dynamically to real-time data streams. In the healthcare domain, IoT has emerged as a transformative technological paradigm, significantly enhancing service delivery, patient safety, and clinical efficiency. IoT-enabled healthcare systems support remote patient monitoring, chronic disease management, and personalised care through interconnected medical sensors, wearable devices, and cloud-based analytics platforms (Ansari et al., 2020; John & Wickramasinghe, 2019; Mehmood et al., 2015). These systems allow continuous observation of patients' physiological conditions, timely detection of anomalies, and proactive clinical interventions, particularly benefiting elderly populations and patients with long-term medical conditions.

Despite the increasing maturity and adoption of IoT technologies in healthcare, critical research challenges remain unresolved, especially in the areas of data provenance and semantic interoperability (Imran & Hlavacs, 2012; Zafar et al., 2017). Healthcare IoT environments generate massive volumes of heterogeneous data originating from diverse sources, devices, and platforms. Ensuring the trustworthiness, traceability, and reliability of this data is essential, as inaccurate or unverifiable information may lead to flawed clinical decisions and compromised patient safety.

Data provenance plays a vital role in addressing these concerns by maintaining comprehensive metadata that describes the origin, creation process, transformation history, ownership, and usage of data throughout its lifecycle (Lomotey et al., 2018; Zafar et al., 2017). Provenance information enables transparency, accountability, and auditability, allowing stakeholders to verify data authenticity, assess data quality, and trace errors or anomalies back to their sources. In healthcare IoT systems, provenance-aware mechanisms are particularly important for regulatory compliance, clinical validation, and secure data sharing across organisational boundaries. Equally critical is semantic interoperability, which ensures that exchanged data is not only transmitted successfully but also understood consistently and unambiguously by different systems, applications, and stakeholders (Kiljander et al., 2014). Semantic interoperability enables heterogeneous healthcare systems, such as electronic health records (EHRs), medical devices, and analytics platforms, to interpret shared data using common meanings and representations. This capability is fundamental for seamless data integration, cross-institutional collaboration, and large-scale health data analytics that support evidence-based decision-making and improved patient outcomes.

Achieving effective semantic interoperability requires the incorporation of explicit semantic descriptions, ontologies, and standardised vocabularies into healthcare data models. Without such semantic enrichment, data may be misinterpreted, incompletely analysed, or incorrectly reused, leading to inaccurate clinical insights and compromised interoperability (Cook, 2020). While data standards play an important role by defining structural agreements and exchange formats, they alone are insufficient to guarantee shared understanding. Semantic technologies complement standards by embedding contextual meaning into data, enabling intelligent reasoning and interoperability across diverse IoT-enabled healthcare systems. In response to these challenges, this study focuses on the integration of secure data provenance and semantic interoperability within healthcare IoT environments, addressing existing gaps in trust, reliability, and meaningful data exchange. By strengthening these foundational components, healthcare IoT systems can better support secure, scalable, and interoperable data-driven services that enhance both individual and population-level health outcomes.

2.0 Literature Review

This section reviews key concepts relevant to this study, including the Internet of Things (IoT), Big Data, data provenance, and semantic interoperability, with a particular focus on their roles and challenges within healthcare environments. This section describes the IoT, Big Data, data provenance, and semantic interoperability in detail.

2.1 Internet of Things (IoT)

The Internet of Things (IoT) refers to an ecosystem of interconnected digital and mechanical devices equipped with unique identifiers and the capability to collect, exchange, and transmit data over network infrastructures without requiring continuous human-to-human or human-to-machine interaction (Ansari et al., 2020; Pavithra & Balakrishnan, 2015). IoT systems enable data communication at any time and from any location, forming the foundation for intelligent and context-aware environments. In the healthcare sector, IoT has significantly reshaped service delivery by enabling remote patient monitoring and continuous health assessment. The adoption of healthcare IoT systems has been accelerated by policy initiatives such as the Health Information Technology for Economic and Clinical Health (HITECH) Act, which promoted the digitalisation and accessibility of electronic health records (Cook, 2020; Mehmood et al., 2015). Through IoT-enabled platforms, patients can access their medical records, improving engagement and facilitating timely clinical decision-making, particularly for elderly and chronically ill patients (John & Wickramasinghe, 2019; Mehmood et al., 2015). Common IoT healthcare devices include wearable sensors, blood pressure monitors, glucose meters, and other medical instruments that collect physiological data, transmit it to cloud-based systems for analysis, and trigger alerts when abnormal conditions are detected.

2.2 Introduction to Big Data

In the digital era, the widespread adoption of IoT and other data-intensive technologies has led to the generation of vast amounts of data across multiple domains. Big Data refers to extremely large datasets that may be structured, semi-structured, or unstructured and cannot be efficiently processed using traditional data management tools. To address these challenges, advanced technologies have been developed to manage the four fundamental characteristics of Big Data, commonly known as the four “V”s: volume, variety, velocity, and veracity (Younas & Qasim, 2019). In healthcare IoT environments, Big Data is generated continuously from heterogeneous sources such as sensors, wearable devices, and clinical information systems. While these datasets offer significant potential for analytics-driven insights and improved healthcare outcomes, their value depends heavily on the reliability, traceability, and interpretability of the data.

2.3 Data Provenance

Data provenance provides essential information regarding the origin, creation process, transformation history, and location of data throughout its lifecycle. Provenance information supports tasks such as debugging, analysis, evaluation, and quality assurance by enabling systems to answer critical questions related to data generation, input parameters, and processing workflows (Imran & Hlavacs, 2012). Provenance metadata typically includes details about data inputs and outputs, platforms used, timestamps, and transformation processes, thereby revealing how, when, and where data enters and moves within a system. By maintaining detailed lineage and ownership records, data provenance enhances data quality by ensuring authenticity, completeness, and verified ownership (Zafar et al., 2017). In healthcare IoT systems, provenance-aware mechanisms are particularly important for ensuring trust, accountability, and compliance when handling sensitive patient data.

2.4 Semantic Interoperability

Interoperability refers to the ability of different computer systems, applications, and devices to exchange data and use the information effectively. In healthcare, interoperability enables Electronic Health Records (EHRs) and other medical information systems to communicate seamlessly, supporting coordinated care and data-driven decision-making (Cook, 2020; Kiljander et al., 2014).

Semantic interoperability extends beyond syntactic data exchange by ensuring that the meaning of the data is preserved and consistently understood across systems. It enables diverse medical applications, tools, and services to collaborate across organisational, regional, and international boundaries (Kiljander et al., 2014; Mehmood et al., 2015). Without semantic alignment, shared healthcare data may be misinterpreted or underutilised, leading to inaccurate analyses and compromised clinical outcomes. Therefore, semantic interoperability relies on ontologies, semantic models, and standardised vocabularies to enrich data with explicit meaning and context.

3.0 Methodology

This section discusses the designed framework and models. Based on the limitations identified in the literature regarding data heterogeneity, lack of secure provenance, and insufficient semantic interoperability in healthcare IoT systems, this study proposes a structured methodology to address these challenges. The proposed approach integrates a layered system architecture with dedicated models for data provenance and semantic interoperability, ensuring secure, reliable, and meaningful healthcare data exchange. The methodology is designed to support heterogeneous IoT devices, manage Big Data characteristics, and enable ontology-driven semantic understanding while maintaining data integrity and confidentiality throughout the data lifecycle.

3.1 Design Framework

The proposed design framework is developed to address the limitations identified in the literature concerning data heterogeneity, secure data provenance, and semantic interoperability in healthcare IoT environments. To manage the complexity and diversity of IoT-generated healthcare data, the framework adopts a layered system architecture that enables systematic data handling, security enforcement, and semantic enrichment throughout the data lifecycle. The framework consists of three primary layers: the IoT Device Layer, the Network Layer, and the Cloud Layer. Each layer is assigned specific functional responsibilities to ensure scalability, security, and interoperability.

The IoT Device Layer comprises heterogeneous healthcare devices such as sensors, wearable medical instruments, and monitoring equipment responsible for collecting real-time physiological data. Given the sensitivity and diversity of healthcare data, this layer focuses on accurate data acquisition and preliminary validation before transmission. Devices are uniquely identifiable and securely connected to the upper layers to support reliable data provenance tracking from the point of origin.

The network layer serves as the communication bridge between IoT devices and cloud-based services. It is responsible for secure data transmission using encrypted channels, ensuring confidentiality and integrity during data exchange. This layer supports authentication, authorisation, and secure routing mechanisms, enabling controlled access and preventing unauthorised interception or modification of healthcare data as it moves across the system. The cloud layer performs centralised data processing, storage, provenance management, and semantic enrichment. Within this layer, incoming data undergoes filtering, validation, and analysis before being stored or made available for further use. The cloud environment hosts the data provenance model, which maintains detailed records of data origin, transmission history, and processing activities, as well as the semantic interoperability model, which transforms raw healthcare data into semantically enriched representations using ontologies and RDF structures.

Data flows sequentially through the three layers, with each stage applying specific security, validation, and processing mechanisms to ensure data usability and trustworthiness. By integrating provenance tracking and semantic interoperability within a unified layered architecture, the proposed framework supports secure, reliable, and meaningful data exchange across heterogeneous healthcare IoT systems. The design framework provides a structured foundation for the proposed models described in Sections 3.2 and 3.3, enabling effective evaluation and comparison as presented in the subsequent findings and discussion.

3.2 Data provenance model

Data provenance alludes to a record trail that accounts for the origin of the data from a database, cloud, or an IoT smart object or device and explains how, when, and from where it got to the current place. This data is used for the analysis of system security.

User and Cloud Service Provider (CSP): In this model, the user or patient first connects their sensor to the protected cloud to send the data securely. The user sends the request to the cloud service provider in order to get the registration key. The security policy of the Cloud Service Provider ensures the confidentiality and integrity of the user data. Private channel: The use of a private channel ensures that data is transmitted securely between the IoT devices and the Cloud Layer.

This private channel is encrypted, making it difficult for unauthorised entities to intercept and decipher the transmitted data. The data provenance model delivers tamper-proof records, ensuring the integrity of the data throughout its lifecycle.

3.3 Semantic interoperability model

The semantic interoperability model described in the context of the Secure Data Provenance and Semantic Interoperability model ensures interoperability in healthcare data exchange by incorporating semantic structures and intelligent services. Semantic interoperability is the capability of data or information exchange with clear and significant meanings. It includes semantics in patient disease symptoms data by aggregating some additional information. It comprises three subsections:

1. Semantic Operation Resources
2. Intelligent Health Cloud services
3. Big-Data Analytics

The proposed methodology directly responds to the limitations identified in the literature by integrating layered architecture, secure data provenance, and ontology-driven semantic interoperability into a unified framework for healthcare IoT systems. This integrated approach ensures secure, reliable, and meaningful data exchange across heterogeneous healthcare environments.

4.0 Findings and Discussion: Testing and Evaluation

The testing of the Data Provenance and Semantic Interoperability model involves two main aspects: testing the data provenance on the Anaconda platform and evaluating the semantic interoperability using the Protégé tool. The evaluation of the proposed Secure Data Provenance and Semantic Interoperability Model was conducted to assess its effectiveness in addressing the challenges identified in the literature, particularly those related to data trustworthiness, provenance tracking, and semantic consistency in healthcare IoT environments. The testing process focused on two main components aligned with the methodology: (i) data provenance validation and (ii) semantic interoperability evaluation. To ensure methodological rigour, the data provenance model was tested using the Anaconda platform, while the semantic interoperability model was evaluated through ontology development and reasoning using the Protégé tool. These tools were selected due to their widespread adoption in data analytics and semantic web research, as well as their suitability for validating the proposed framework components.

4.1 Systematic Comparison

A systematic comparison was conducted between the proposed model and existing approaches to evaluate its relative performance in terms of data provenance extraction, semantic interoperability, and handling of heterogeneous healthcare data. The comparison results, summarised in Table 1, highlight several distinctive advantages of the proposed model.

The findings indicate that many existing models focus either on data-oriented or process-oriented provenance mechanisms, but often fail to integrate semantic structures effectively. In contrast, the proposed model emphasises the application of well-defined ontologies to transform raw and unstructured healthcare sensor data into Resource Description Framework (RDF) representations. This transformation enables structured, machine-interpretable data, thereby enhancing semantic interoperability across heterogeneous systems.

Additionally, unlike several existing approaches that rely on resource-intensive procedures or limited provenance granularity, the proposed model demonstrates improved efficiency by adopting a layered architecture and lightweight provenance mechanisms. By embedding provenance tracking directly into the data flow across IoT Device, Network, and Cloud layers, the model ensures continuous and secure provenance dissemination without introducing excessive computational overhead.

The systematic comparison also reveals that while some existing studies partially address semantic interoperability, they often lack full support for ontology-based reasoning or RDF conversion. The proposed approach overcomes this limitation by explicitly incorporating semantic enrichment processes within the cloud layer, ensuring meaningful data interpretation and interoperability across healthcare platforms.

4.2 Overall Comparison

The overall comparison further demonstrates that many existing healthcare IoT models suffer from resource-intensive processes, which negatively impact system efficiency and scalability. Models that emphasise process-oriented provenance often require extensive navigation through records, increasing latency and raising potential security concerns.

In contrast, the proposed model adopts a data-oriented provenance strategy, focusing on attributes stored within databases and secure query-based access. This approach reduces system complexity while maintaining high levels of data integrity and traceability. The use of encrypted private channels between IoT devices and the cloud layer further strengthens security by preventing unauthorised access and data interception. From a semantic perspective, existing models frequently rely on syntactic interoperability or limited ontology usage, which restricts their ability to support meaningful cross-system data exchange. The proposed semantic interoperability model addresses this gap by integrating semantic operation resources, intelligent health cloud services, and Big Data analytics, enabling enriched data interpretation and scalable analytics-driven insights.

Table 1. Table captions should be placed above the tables.

Authors	Applied Domain	Heterogeneity Addressed	Data/Process Oriented	Provenance Dissemination	Granularity (Data type)	Semantic Web	Ontologies	RDF Format
Proposed	Healthcare	Yes	Data	Queries	Attributes in the database	Yes	Defined ontologies	Yes
[1]	Healthcare	Yes	Data	Queries	Attributes in Database	No	Defined ontology	No
[2]	Wearable IoT devices	Yes	Process	Queries	Abstract data parameters	No	No	No

[5]	Service-oriented system	Yes	Data	Queries	Data	No	Ontological reasoning	No
[6]	Test beds	No	Data	Queries	Triple-Stores	No	No	Yes
[8]	Molecular Engineering	No	Process	Browser	Abstract parameters to workflow	No	Defined ontologies	No
[10]	EHR	No	Data	Queries	Attributes in the database	Yes	Ontology definition	No

4.3 Discussion of the Proposed Model

The findings confirm that the proposed model effectively addresses the key limitations identified in the literature. By combining secure data provenance and semantic interoperability within a unified framework, the model offers a more cost-effective, secure, and scalable solution for healthcare IoT systems. A notable contribution of the proposed approach is its emphasis on transforming raw electronic health data into RDF format using defined ontologies, ensuring both semantic clarity and interoperability. This structured representation supports accurate data exchange, improved reasoning capabilities, and enhanced integration with other healthcare information systems. The evaluation results demonstrate that the proposed model improves data reliability, enhances semantic consistency, and mitigates security risks associated with heterogeneous healthcare IoT environments. These findings validate the effectiveness of the methodology and highlight the model's potential for practical deployment in real-world healthcare scenarios.

5.0 Conclusion and Future Work

This study proposed and evaluated a Secure Data Provenance and Semantic Interoperability Model for healthcare-oriented Internet of Things (IoT) environments, addressing critical challenges related to data trustworthiness, heterogeneity, and meaningful data exchange. Grounded in a layered system architecture comprising IoT Device, Network, and Cloud layers, the proposed framework integrates secure data provenance mechanisms and ontology-driven semantic interoperability to enhance the reliability and interpretability of healthcare data. The findings demonstrate that the proposed model effectively improves data integrity, traceability, and semantic consistency across heterogeneous IoT devices and healthcare information systems. By embedding provenance tracking throughout the data lifecycle and employing encrypted private communication channels, the model ensures tamper-proof records and secure data transmission. Furthermore, the incorporation of semantic structures and RDF-based data transformation enables meaningful interoperability, overcoming limitations observed in existing approaches that rely primarily on syntactic data exchange.

Comparative evaluation results indicate that the proposed model offers a more efficient and scalable alternative to resource-intensive and process-oriented frameworks. Its emphasis on data-oriented provenance extraction and ontology-based semantic enrichment supports accurate data interpretation and facilitates cross-system integration in healthcare IoT ecosystems. These outcomes validate the effectiveness of the proposed methodology in addressing the research gaps identified in the literature. Despite these contributions, several directions for future work remain. Future research will focus on strengthening security at the physical and device levels, including the protection of sensors and edge devices against tampering and unauthorised access. Additionally, the integration of syntactic interoperability standards and protocols will be explored to complement semantic interoperability and further enhance system compatibility. Performance evaluation using large-scale, real-world healthcare datasets and deployment in live clinical environments is also recommended to assess scalability, robustness, and practical feasibility. These enhancements will contribute to the continued development of secure, interoperable, and data-driven healthcare IoT systems.

Acknowledgement

This work was partially supported by the Mexican Government through the grant A1-S-47854 of CONACYT, Mexico, and the grants 20254236, 20253468, and 20254341 provided by the Secretaría de Investigación y Posgrado of the Instituto Politécnico Nacional, Mexico. We gratefully acknowledge the computing resources made available through the Plataforma de Aprendizaje Profundo para Tecnologías del Lenguaje of the Laboratorio de Supercómputo at the INAOE, Mexico, funded by CONACYT. Additionally, we express our sincere gratitude to Microsoft for their support through the Microsoft Latin America PhD Award, which has significantly contributed to the success of this work.

Paper Contribution to the Related Field of Study

This study contributes by proposing a novel model that enhances semantic interoperability and secure data provenance in healthcare IoT systems, addressing communication gaps, reliability issues, and security challenges across heterogeneous devices and networks.

References

Ali, S., Wang, G., Bhuiyan, M. Z. A., & Jiang, H. (2018). Secure data provenance in cloud-centric Internet of Things via blockchain smart contracts. In *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)* (pp. 991-998). IEEE.

Ansari, S., Aslam, T., Poncela, J., Otero, P., & Ansari, A. (2020). Internet of things-based healthcare applications. In *IoT architectures, models, and platforms for smart city applications* (pp. 1-28). IGI Global Scientific Publishing.

Cook, E. L. (2020, March). Could Interoperability Between IoT And Ehr Make Healthcare More Efficient?. In *Proceedings of the Appalachian Research in Business Symposium* (Vol. 1).

Younas, R., & Qasim, A. (2019, April). Streaming State Validation Technique for Textual Big Data Using Apache Flink. In *International Conference on Computational Linguistics and Intelligent Text Processing* (pp. 632-647). Cham: Springer Nature Switzerland.

Hussain, N., Qasim, A., Mehak, G., Kolesnikova, O., Gelbukh, A., & Sidorov, G. (2025). ORUD-Detect: A Comprehensive Approach to Offensive Language Detection in Roman Urdu Using Hybrid Machine Learning–Deep Learning Models with Embedding Techniques. *Information*, *16*(2), 139.

Imran, M., & Hlavacs, H. (2012, October). Applications of provenance data for cloud infrastructure. In *2012 Eighth International Conference on Semantics, Knowledge and Grids* (pp. 16-23). IEEE.

John, B., & Wickramasinghe, N. (2019). A review of mixed reality in health care. *Delivering Superior Health and Wellness Management with IoT and Analytics*, 375-382.

Khan, H., Habib, S., Qasim, A., Hussain, N., Usman, M., Mahmood, A., ... & Zain, M. (2023, December). The Study of Human Action Recognition in Videos with Long Short-Term Memory Model. In *Annual International Conference on Information Management and Big Data* (pp. 217-230). Cham: Springer Nature Switzerland.

Kiljander, J., D'elia, A., Morandi, F., Hyttinen, P., Takalo-Mattila, J., Ylisaukko-Oja, A., ... & Cinotti, T. S. (2014). Semantic interoperability architecture for pervasive computing and internet of things. *IEEE access*, *2*, 856-873.

Hussain, N., Qasim, A., Akhtar, Z. U. D., Qasim, A., Mehak, G., del Socorro Espindola Ulibarri, L., ... & Gelbukh, A. (2023, November). Stock Market Performance Analytics Using XGBoost. In *Mexican International Conference on Artificial Intelligence* (pp. 3-16). Cham: Springer Nature Switzerland.

Lomotey, R. K., Sofranko, K., & Orji, R. (2018). Enhancing privacy in wearable IoT through a provenance architecture. *Multimodal Technologies and interaction*, *2*(2), 18.

Manzoor, M. I., Shaheen, M., Khalid, H., Anum, A., Hussain, N., & Faheem, M. R. (2018). Requirement Elicitation Methods for Cloud Providers in IT Industry. *International Journal of Modern Education & Computer Science*, *10*(10).

Qasim, A., Mehak, G., Hussain, N., Gelbukh, A., & Sidorov, G. (2025). Detection of Depression Severity in Social Media Text Using Transformer-Based Models. *Information*, *16*(2), 114.

Mehmood, R., Faisal, M. A., & Altowaijri, S. (2015). Future networked healthcare systems: a review and case study. In *Handbook of research on redesigning the future of internet architectures* (pp. 531-558). IGI Global Scientific Publishing.

Meque, A. G. M., Hussain, N., Sidorov, G., & Gelbukh, A. (2023). Machine learning-based guilt detection in text. *Scientific Reports*, *13*(1), 11441.

Pavithra, D., & Balakrishnan, R. (2015, April). IoT based monitoring and control system for home automation. In *2015 global conference on communication technologies (GCCT)* (pp. 169-173). IEEE.

Shaheen, M., Anees, T., Hussain, N., & Obaid, I. (2019, April). A Research on SOA in the IT Industry of Pakistan. In *Proceedings of the 2019 5th International Conference on Computer and Technology Applications* (pp. 149-154).

Shaheen, M., Awan, S. M., Hussain, N., & Gondal, Z. A. (2019). Sentiment analysis on mobile phone reviews using supervised learning techniques. *International Journal of Modern Education and Computer Science*, *10*(7), 32.

Tash, M. S., Ahani, Z., Tonja, A., Gemeda, M., Hussain, N., & Kolesnikova, O. (2022, December). Word level language identification in code-mixed kannada-english texts using traditional machine learning algorithms. In *Proceedings of the 19th International Conference on Natural Language Processing (ICON): Shared Task on Word Level Language Identification in Code-mixed Kannada-English Texts* (pp. 25-28).

Hussain, N., Anees, T., Faheem, M. R., Shaheen, M., Manzoor, M. I., & Anum, A. (2018). Development of a novel approach to search resources in IoT. *Development*, *9*(9).

Hussain, N., Qasim, A., Mehak, G., Kolesnikova, O., Gelbukh, A., & Sidorov, G. (2025). Hybrid Machine Learning and Deep Learning Approaches for Insult Detection in Roman Urdu Text. *AI*, *6*(2), 33.

Zafar, F., Khan, A., Suhail, S., Ahmed, I., Hameed, K., Khan, H. M., ... & Anjum, A. (2017). Trustworthy data: A survey, taxonomy and future trends of secure provenance schemes. *Journal of network and computer applications*, *94*, 50-68.