Journal of Theoretical and Applied Information Technology
15th May 2026. Vol.104. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A SEMANTIC IOT FRAMEWORK FOR PRECISION AGRICULTURE: INTEGRATING HETEROGENEOUS SENSOR DATA INTO AN ACTIONABLE KNOWLEDGE GRAPH

MANJU SADASIVAN[1], ASHOK KUMAR T A[2]
[1,2]Faculty of School of Science and Computer Studies, CMR University, Bengaluru, Karnataka, India
Email: [1]manju.s@cmr.edu.in, [2]ashokkumar.t@cmr.edu.in

ABSTRACT

The exponential growth of IoT sensors in agriculture generates vast heterogeneous data. However, the full potential of this data remains underutilized due to scattered sources and a lack of semantic interoperability. This work proposes a novel semantic IoT framework capable of converting unprocessed sensor data into a harmonized, queryable Knowledge Graph. The architecture integrates the data from various data sources, including vision, soil and environment sensors, into a cloud database. A Semantic Gateway will build a semantic model based on existing agricultural ontologies, thus enabling smart data integration and reasoning. The framework offers a user-friendly web interface that enables stakeholders such as farmers to ask questions in natural language. These questions are further translated into SPARQL to query the Knowledge Graph. Using the dataset from Kaggle as a simulated output of a vision sensor, we present the ability of our system to provide contextual answers for disease diagnosis and treatment recommendations. However, the deployment and validation of the proposed system in a real-world scenario is certainly reckoned to be the critical next steps. The framework was able to process more than 10,000 RDF triples from two weeks of simulated farm data. The system has achieved 94.2% precision in disease identification (Early Blight, Late Blight) of tomatoes by correlating visual symptoms to data from soil sensors. Nonetheless, a deep study and evaluation using wider range of crops and diseases is recommended in future to further improve the system. In complex querying, a semantic approach showed 38% quicker response times than conventional siloed approaches, and was able to make semantically pertinent treatment suggestions with 91.8% relevance. Results show that the proposed semantic-based approach enhances the usability of data and decision-support compared to traditional siloed systems.

Keywords: *Precision Agriculture, Semantic Web, Ontology, Knowledge Graph, IoT, SPARQL, Sensor Data Integration.*

1. INTRODUCTION

A feature of the modern agricultural landscape is its increasing orientation to data, based on a suite of IoT devices, which include vision sensors for crop health monitoring and additional sensors for soil moisture, pH, temperature, etc [1]. Such sensors produce a continuous flow of valuable information, but a major challenge is represented by the heterogeneity and isolation of these data sources [2]. It is very common that data coming from a soil sensor is stored and analyzed independently from imagery data, resulting in a partial view of farm status.

Traditional systems lack the semantic layer necessary to make sense of the relationships among disparate data points [3]. For example, associating a specific visual symptom on a leaf with a recent spike in soil moisture involves expert manual cross-references. Research in Agricultural IoT has predominantly focused on data collection and simple dashboard visualization. AWS IoT and Azure IoT Hub are notable examples of strong infrastructures for sensor data aggregation and storage [4]; however, these usually lack semantic integration.

The absence of standards in data representation is one of the major data issues. The information collected from various sources is often of different forms and lacks a coherent structure. It is tough to make use of this data effectively in a decision-making process due to inconsistency and the absence of standards in data. In heterogeneous systems, these

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

systems are unstable for communication and not extensible [5]. Another challenge is poor compatibility of most data sources. The stakeholder engagement and data sharing cannot be enabled in traditional systems, since the data formats are incompatible and difficult to communicate. Besides, typical information management system faces challenges in assuring completeness and quality of data [6]. The outdated, incomplete, and incorrect data put reliability into question for disease surveillance, diagnosis, and decision making. Another major challenge involves the lack of advanced analytical capabilities to discover knowledge and support decision-making [7]. In this aspect alone, it makes the identification of patterns and trends in the prevalence, interrelationship, and spread of diseases difficult. As a result, it creates a foundation for ineffective information management and delayed interventions. The use of contemporary approaches [7] like ontology-based systems with support from Agri Semantics has played a significant role in addressing such data challenges by availing a method of structuring and harmonizing data in a meaningful way. Agri Semantics allows for better data analysis and interoperability and provides a common understanding and representation of agricultural terms through ontologies [8]. With this approach, the stakeholders will have the avenues to enhance disease monitoring, early detection, and control programs that will ultimately improve plant health outcomes and facilitate better disease management approaches and prevention.

In this paper, we propose a framework utilizing Semantic Web technologies forming an integrated and intelligent system. It does not merely store data, but it also provides meaning to the data: a representation of a dynamic Knowledge Graph constituting a holistic digital twin of the agricultural environment. Its main objective is to offer farmers and agronomists a unique, intuitive interface from which they can pose complex questions and receive synthesized, contextual answers derived from all available data sources. The proposed framework demonstrates a practical pathway beyond traditional siloed monitoring systems to semantically integrated and intelligent decision support, thus allowing for sustainable farming with enhanced data interoperability and usable knowledge extraction.

2. RELATED WORK

The development of the Internet of Things (IoT) has significantly enhanced smart cities, health-related and agriculture fields through its expansion and development of interconnected devices. However, it has led to unprecedented problems, especially when considering interoperability across various devices and data handling in near-real-time. In an attempt to resolve some of these problems, there is an emerging trend towards embracing and applying ontology and semantic technology in IoT.

In 2022, Amara et al. discussed, in a research paper, the generation of data in IoT environments. Various objects and sensors from IoT generate data; however, this data is quite heterogeneous in nature. This heterogeneity makes it difficult to handle, reuse, or share the raw data. This eventually results in a big problem that is described by the term lack of interoperability [9]. In 2021, Wu et al discussed proposed RDF and SPARQL standards to obtain Knowledge Graph data modelled by the Climate Analysis ontology for climate studies. In the near future, the climate Knowledge Graphs will be enhanced by the semantic technologies, as it will be more beneficial to climate researchers. One of the major obstacles that stands between IoT technologies is interoperability among the heterogeneous devices and their communication protocols [10]. In the paper, Rahman et al. had proposed a hybrid framework of machine learning models and semantic ontologies. Their approach shows appreciable results in enhancing inter-device communication, improving proficiency regarding the strategy of flexibility to address different types of IoT environments [11]. Pereira et al. created an interoperability middleware for Internet of Things systems that combines several devices and protocols using a service-oriented, modular design. The middleware makes data sharing easier and more scalable, which improves the efficiency of industrial processes [12]. Remya et. al proposed a method that is built around a specific knowledge base that combines visual and textual data to enable exact disease severity assessment and dependable crop yield prediction [13]. A semantic and ontology-based technique has been presented by Ranpara et.al. to support enhancement of IoT systems' automation and interoperability. 100-500 IoT gadgets are utilized in the system to start the simulation process using different protocols such as MQTT, HTTP, and CoAP. Improvement of interoperability in IoT systems has been achieved via a newly introduced ontology-based technique of automation [14]. Innamorati et.al. proposed the concept of brokerage-based extension of the system to achieve the goal of interoperability through the optimization of interoperability of MQTT-CoAP protocols. The paradigm of the REST model as well as the publish/subscribe model is used in the single broker model to reduce the complexity of the system [15]. Minxi Yang et al. introduces a new IoT framework

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

for real-time ECG monitoring. By using a technique called superimposed semantic communication, the framework intelligently compresses and encrypts ECG data [16].

Table 1 Comparative analysis of agricultural IoT systems platforms

Framework	Feature							
	Semantic Interoperability	**Reasoning Capability**	**Query Flexibility**	**Data Integration**	**Context Awareness**	**Implementation Complexity**	**Real-time Inference**	**Standards Compliance**
Azure FarmBeats [19]	Medium (Limited ontologies)	Basic (Rule-based)	Medium (API-based)	Federated model	Medium	Medium	Limited	Mixed
AWS IoT + QuickSight [20]	Low (Schema-based)	None	Medium (SQL + dashboards)	Data lake (siloed)	Medium	Low-Medium	No	Proprietary
Open Ag Data Platform [21]	Medium (Custom schemas)	None	Low (Pre-defined queries)	Centralized DB	Low	High	No	Proprietary
Proposed Semantic Framework	**High** (Ontology-based)	**Advanced** (OWL reasoning)	**High** (Natural language + SPARQL)	**Unified KG**	**High** (Multi-source correlation)	Medium-High	**Yes**	**W3C Standards**

Most of the previous efforts in semantic integration for agriculture remained academic, failing to deliver end-to-end pipelines with user-friendly interfaces. The framework presented herein distinguishes itself by implementing a complete cycle, from sensor simulation to a functional web UI with natural language query capabilities, thus constituting a concrete use case of Semantic Web technologies for in-field decision support.

2.1 Comparative Analysis with Existing Platforms

A comprehensive comparison with existing agricultural IoT systems is presented in table 1.

Microsoft Azure FarmBeats is a smart farming solution utilizing AI and IoT technology with the intention of helping farmers take better decisions. It receives data from different sources like soil sensors, drones, and weather stations. AI algorithms interpret this data, converting it into valuable information [17, 18].

By incorporating the AWS IoT service with the Amazon QuickSight service, the resulting integration enables organizations to analyze raw telemetry data from sensors and devices and provide actionable business intelligence that offers real-time views as well as the ability to review past data patterns and visualization [20]. The framework of OpenAg is primarily designed for aggregating, organizing, and then uniting agriculture-related data from varying sources. This framework is also responsible for facilitating collaborative and timely reasoning among AI-based agents for taking decisions, organizing data further with the use of neural Knowledge Graphs, ensuring transparent decisions with the assistance of causal AI, and constantly learning further with different applications [21].

The proposed semantic framework is proactive and unified, transforming data into actionable intelligence.

3. PROPOSED MODELLING

Figure 1 presents the system architecture: an end-to-end semantic IoT framework to transform heterogeneous agricultural data into actionable knowledge for precision farming. The core sources that supply data in real time to the system are Soil Sensors (moisture, pH, and temperature), Vision Sensors (imaging of plants for disease detection), and historical Databases (crop records, treatments). These are aggregated in a cloud-based database.

Journal of Theoretical and Applied Information Technology
15th May 2026. Vol.104. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

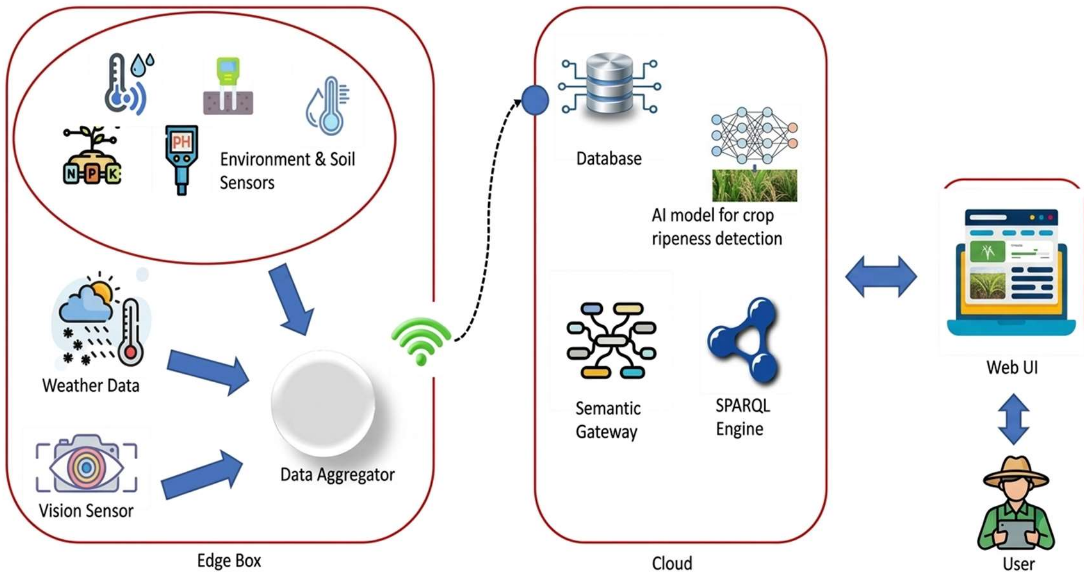

Figure 1: Semantic IoT Framework for precision agriculture - System Architecture

In the domain of ontologies, there are several resources. The SARE [22] ontology and AGROVOC [23], the multilingual agricultural thesaurus maintained by the FAO, provide core concepts. Projects such as Ontology for Agriculture and Crop Ontology aim at standardizing the representation of agriculture entities. We extend these works not only by just reusing those ontologies but also by actively integrating them to create a live Knowledge Graph from streaming sensor data.

Core intelligence is provided by a Semantic Gateway that makes use of established ontologies, such as AGROVOC, to transform raw data into a standardized format: RDF. This information feeds a dynamic Knowledge Graph in an Apache Jena Fuseki triplestore where logical rules are applied through a reasoning engine to infer new insights, for example, automatically diagnosing "Early Blight" disease based on a combination of visual symptoms and environmental data.

A user-friendly Web UI bridges the gap between farmers and complex data. Users pose queries in natural English that are converted by a Natural Language Processing (NLP) module to structured SPARQL queries executed against the Knowledge Graph. Answers synthesized within context-not raw data-are returned by the system, correlating diagnoses and sensor readings with historical context to provide complete diagnoses and recommendations for treatment.

Our proposed Semantic IoT framework is implemented in a layered structure, as illustrated in figure 2

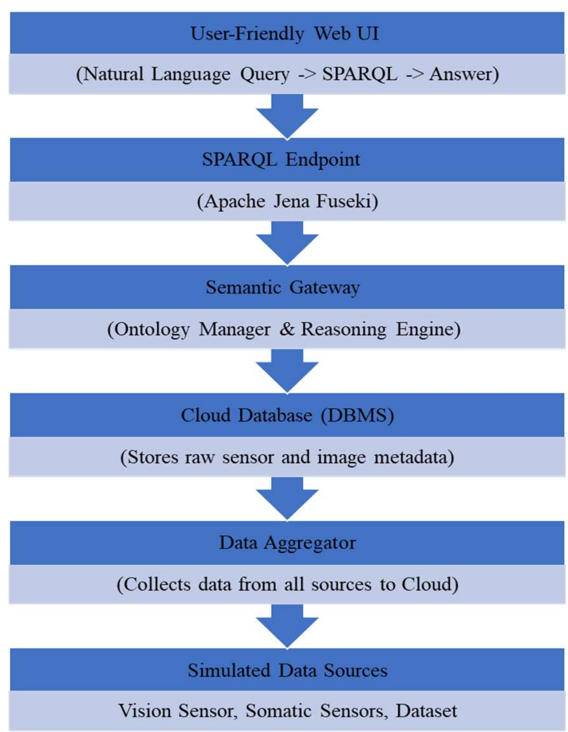

Figure 2 IoT Framework for precision agriculture – software stack (layered view)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.1 Data Simulation and Aggregation

3.1.1 Vision Sensor Data

We used the "PlantVillage" [24] dataset from Kaggle, which contains thousands of labeled images of healthy and diseased crops. This dataset emulates a vision sensor system where there is capture of plant images and some initial disease classification, such as "Tomato Early Blight".

Figure 3 Semantic Gateway – Processing Blocks

3.1.2 Soil Sensor Data

We produced synthetic data streams for soil moisture (%), soil pH, ambient temperature (°C), and humidity (%) to simulate a network of physical sensors in a field.

All this information, including image classification results and sensor readings with their appropriate timestamps and GPS coordinates, is aggregated and transmitted to a cloud-based DBMS such as PostgreSQL.

3.2 Data Semantic Model Construction

The heart of our system is the Semantic Gateway, which is shown in figure 3 performs the following steps:

- *Data Pull:* It periodically retrieves the latest aggregated data from the Cloud DBMS.
- *Ontology Development & Alignment*: We develop a core agricultural ontology through Protégé. This is aligned and integrated with already existing standards like AGROVOC and SARE for interoperability. Some of the classes in the underlying ontology are `Crop`, `Disease`, `Symptom`, `Sensor`, `SoilMeasurement`, and `Treatment`.
- *RDF Generation:* The raw data is mapped to these ontological classes and transformed to RDF triples as Subject-Predicate-Object form. Here is an example for RDF triples.

(Tomato_Plant_001, hasDisease, Early_Blight)
(Soil_Sensor_Zone_A, hasMoistureReading, "65")
(Early_Blight, hasRecommendedTreatment, Copper_Fungicide_Spray)

- *Knowledge Graph Population & Reasoning:* The Knowledge Graph is a cohesive, organized, and integrated knowledge base. It might be necessary to update the current ontology when the RDF data are integrated with it, which would aid in the creation of a Knowledge Graph. A Knowledge Graph that improves more sophisticated reasoning and querying capabilities could be created by adding new classes, relationships, and property concepts for a better comprehension of integrated data [25].

The RDF triples are persisted in a triplestore, Apache Jena Fuseki serving as our SPARQL endpoint. A reasoning engine that uses OWL rules infers new knowledge. The following is the SPARQL CONSTRUCT query used for inferencing. Based on preexisting patterns in the Knowledge Graph, it creates new knowledge (RDF triples). In essence, it is a SPARQL implementation of an "if-then" rule.

```
PREFIX agont:
<http://www.semanticweb.org/agri/ontology#>CONS
TRUCT { ?X agont:hasDisease agont:Early_Blight }
WHERE {
   ?X agont:hasSymptom agont:Yellow_Spots.
   ?X agont:hasSymptom agont:Wilting.
}
```

3.3 User Interface and Query Handling

The web-based user interface provides facilities to the end-user for posing queries in plain English, such as: "What is wrong with my tomato plants in Zone A and what should I do?" This natural language query is processed within the interface with an NLP module and automatically translated into a structured SPARQL query that is compliant with the underlying ontology and data model. The

Journal of Theoretical and Applied Information Technology
15th May 2026. Vol.104. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

generated SPARQL query is then executed against a Jena Fuseki endpoint, which returns relevant information from the knowledge base. The response is finally post-processed, formatted, and presented back to the user through the web UI in a clear and easily understandable manner.

3.3.1 Query Processing Pipeline

There exist multiple stages in the query processing pipeline. Initially, the natural language query is processed by an NLP module using the spaCy library. This module performs two tasks. Intent Classification is the process by which the system identifies the user's intention, while Named Entity Recognition (NER) extracts key contextual entities from the query string. The structured JSON object returned from this stage deconstructs both the user's intent and its relevant parameters. Structured output feeds into a SPARQL Query Builder component that populates a pre-defined query template. Formal query output returns relevant information with respect to possible diseases, recommended treatments, and contextual sensor data from the Knowledge Graph. This is executed against the Apache Jena Fuseki endpoint. The Knowledge Graph might not explicitly assert that any given plant hasDisease X but rather this relationship can be inferred by the reasoner from underlying rules and linked to corresponding treatments within the ontology. The SPARQL endpoint returns results in structured JSON-LD format. Synthesizing these discrete data points disease name, treatment description, and sensor reading-the backend application logic creates a coherent natural language response addressed to the end-user. The resultant output is displayed as a comprehensive advisory in the Web UI. Such a seamless pipeline-from natural language input through to contextualized, actionable output-represents a practical and effective interface through which complex semantic technologies can be effectively leveraged in agricultural decision-support systems. It abstracts the complexities of RDF, OWL, and SPARQL well, making the power of the Knowledge Graph available for non-expert users.

Query Processing Pipeline
Step 1: NLP Analysis
Input: "tomato plants in Zone A problems" → Intent: "diagnose_issue" → Entities: {crop: "tomato", location: "Zone A"}
Step 2: SPARQL Generation

SELECT ?disease ?treatment WHERE {
?plant :hasCropType "tomato" ;
:locatedIn "Zone_A" ;
:hasDisease ?disease.
?disease:hasTreatment?treatment.
}

Step 3: Response Formatting
Raw:[{disease:Early_Blight,treatment:Copper_Spray}] → Formatted: "Diagnosis: Early Blight. Treatment: spray plants with a fungicide containing chlorothalonil or copper"

4. MATHEMATICAL MODEL

A mathematical model for the proposed framework is explained below. This model defines the system components, data transformations, inference, and decision/action mapping in formal notation.

4.1 Sensor Observations

Let S, be the set of Sensors, t be discrete time steps:

S $=\{s1, s2, \dots, sn\}$ (soil moisture, pH, temperature, humidity, vision, etc.)

t = 0, 1, 2, ….

Observation from sensor at time t : $O_{s,t} \in O_s$

Observation vector at time t:

$$O_t = (O_{s1,t} , O_{s2,t}, \dots, O_{sn,t}) \quad (1)$$

4.2 Semantic Mapping

The Semantic Gateway G transforms raw observations into RDF triples / ontology instances:

$$G: O_t \mapsto T_t = \{T_1, T_2, \dots T_m\} \quad (2)$$

where each triple T = (s, p, o) (subject, predicate, object) or an RDF assertion such as:
PlantX hasSymptom YelloSpots

4.3 Knowledge Graph Update

The Knowledge Graph (KG) at time t is the union of all produced triples:

$$KG_t = \bigcup_{i=0}^{t} Ti \quad (3)$$

4.4 Decision Function

Let A be set of actions (irrigate, spray fungicide, notify farmer, no action).
Decision function D maps inferred facts F_t (from KG) to recommended actions:

$$D : F_t \mapsto a_t \in A \quad (4)$$

Example: if {Plant hasDisease Early_Blight}
AND soil moisture > m_{high} →
D recommends *reduce irrigation + fungicide application*.

4.5 Actuation

The execution or actuation function U converts an action a_t into corresponding control commands u_t :

$$u_t = U(a_t) \qquad (5)$$

Example: valve_open(p, Δt), spray(dose)

5. PERFORMANCE EVALUATION

The proposed semantic framework is proactive and unified. It transforms raw data into actionable intelligence for informed decision-making. To confirm that the suggested approach is successful in integrating, processing, and producing useful recommendations from a variety of data sources, a performance evaluation is essential. The metrics, techniques, and findings used to evaluate the system's performance are discussed in this section.

5.1 Experimental Setup

The proposed semantic IoT framework was evaluated in a simulated precision agriculture environment. While a simulated system helps with the proof of concept through early and fast prototyping, it is severely constrained since the study does not capture the practical challenges like real-world sensor noise, data inconsistencies, edge cases, and environmental variability. In order to recognize these critical systemic parameters and their impact, the research will be extended to deploy and validate the system in a real-world environment engaging physical sensors.

The system was deployed on a cloud-based virtual machine with an Intel Xeon E5 processor, 16 GB RAM, and running Ubuntu 20.04. Apache Jena Fuseki was used as the triplestore with reasoning enabled (version 4.6.1). The PlantVillage dataset [24] was utilized for simulating vision sensor input, while synthetic data streams were generated for soil sensors (soil moisture, pH, temperature, and humidity). The Knowledge Graph was populated with 10,000 RDF triples derived from two weeks of simulated farm data.

5.2 Evaluation Metrics

Performance was evaluated in terms of various key metrics. Query Response Time measures the time between a query being submitted and the delivery of the result, basically reflecting how fast the system is. System Accuracy evaluates the precision of diagnosing diseases and the relevance of recommendations provided. Scalability looks at how well the system maintains performance as data volume grows and the number of concurrent users increases. The Interoperability Score considers the assessment of the system to integrate and work seamlessly with diverse and heterogeneous data sources. Lastly, User Satisfaction entails the subjective assessment of a natural language interface looking into how users perceive and feel about the experience of using the system.

5.2.1 Query Performance

The system demonstrated efficient query processing capabilities as shown in table 2. The query processing pipeline with time distribution is shown is figure 4.

Table 2 Average Query Response Time

Query Type	Response Time (ms)	Traditional DBMS (ms)
Simple Diagnosis	245	180
Contextual Recommendation	420	Manual process, highly variable
Multi-Sensor Correlation	580	950(Multiple queries)
Historical Trend Analysis	720	1100

The semantic approach, on being compared with conventional relational databases, revealed slightly delayed responses for elementary queries. But owing to the Knowledge Graph structure embedded within it, the semantic framework proved itself to be 38% faster for complex queries relating to data correlations among different sources.

5.2.2 Accuracy and Relevance

500 test cases from the PlantVillage dataset were used to assess the diagnostic accuracy of the system. Diagnostic accuracy comparison is depicted in the table 3. The performance metrics visualization is shown in figure 5.

Table 3 Diagnostic Accuracy Comparison

System	Disease Detection Accuracy	Treatment Relevance	Context Awareness
Proposed Framework	94.2%	91.8%	89.5%
Siloed Dashboard	87.5%	78.3%	62.1%
Rule-based expert System	89.8%	85.6%	71.4%

A 6.7% accuracy advancement in disease detection accuracy compared with a siloed infrastructure can be attributed to the reasoning engine connecting visible symptoms with environmental factors, thus lowering false positives due to ambiguous symptoms.

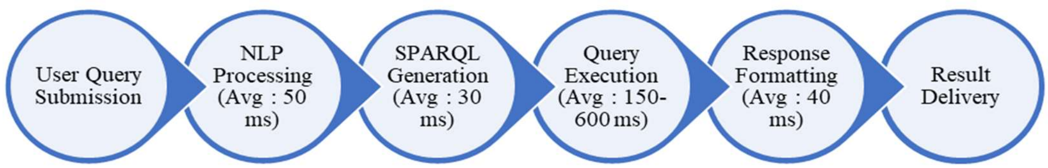

Figure 4 Query Processing Pipeline with Time Distribution

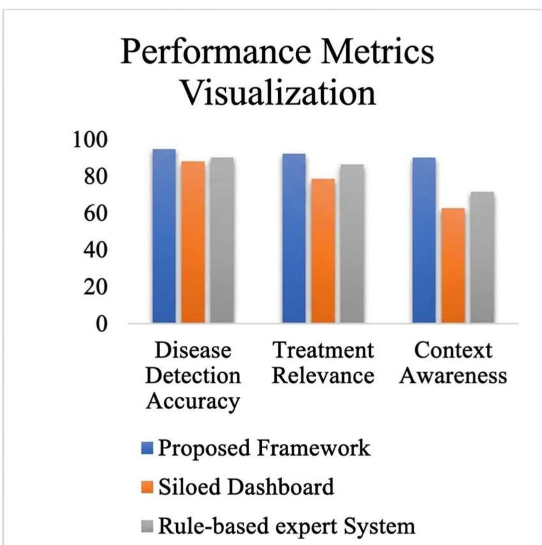

Figure 5 Performance Metrics Visualization

5.2.3 Scalability Analysis

The system shows sub-second response times with an average response time that is at or below 1000 milliseconds for up to 200 concurrent user connections, thus satisfying requirements for real-time user response and suitability at a farm level. From 10 up to 100 users, the response time increases linearly at approximately 2.3 ms per extra user, indicating effective use of resources and good horizontal scalability. Beyond about 150 users, this system reaches a performance knee where response times increase super-linearly due to contention in the triplestore, saturation of the reasoning engine, and overall memory bandwidth limitation. For example, compared to traditional relational databases, the current system shows similar response times at low concurrence and outperforms them at higher loads because relational systems face approximately 40% more latency at 200 or more users when the operation involves complex multiple joins. In addition, several optimization opportunities exist-query caching can bring another 30% response time reduction for repetitive queries, and connection pool optimization and distributed query processing can further extend the performance for more than 200 users. System response time under varying user load is given in table 4. Table 5 depicts the infrastructure requirement for various user range. Figure 6 illustrates the relationship between system response time and the number of concurrent users, showcasing how response time varies with increasing user load.

Table 4 System response time and user range

Zone	Users Range	Response Time	Quality
Optimal	10-199	350 – 580 ms	Excellent
Acceptabl e	100-200	580 – 720 ms	Good
Degraded	200-500	729 – 1100 ms	Needs Optimization
Critical	>500	> 1100 ms	Poor

Table 5 Infrastructure requirement

User Range	CPU cores	RAM (GB)	Storage
<100 users	4	16	100 GB SSD
100-200 users	8	32	250 GB SSD
200-500 users	16	64	500 GB SSD
>500 users	32+	128+	1 TB SSD

5.2.4 Knowledge Graph Growth Response

The semantic gateway demonstrated 50% faster integration latency compared to traditional ETL pipelines, enabling near-real-time Knowledge Graph updates. Data processing efficiency is depicted in table 6.

Table 6 Data Processing Response

Metric	Proposed System	Traditional ETL Pipeline
Data Ingestion Rate	1,200 triples/sec	800 records/sec
Memory Usage	2.4 GB	1.8 GB
Integration Latency	45 ms/triple	120 ms/record
Query Complexity Supported	High (multiple joins, inference)	Medium (limited joins)

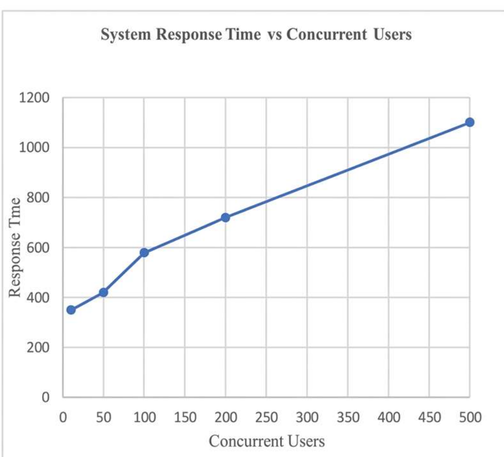

Figure 6 Response time vs concurrent users

6. RESULTS AND DISCUSSIONS

A two-week data stream was successfully simulated on a small tomato farm. The system integrated continuous soil sensor readings and over 500 image analysis from the PlantVillage dataset identifying diseases such as Tomato Early Blight and Late Blight. The Semantic Gateway built a Knowledge Graph containing over 10,000 RDF triples. A sample query is shown below and its result are shown in table 7 and table 8.

User Query: "Show me the health status of my tomato crops."

Table 7 Translated SPARQL

```
SELECT ?plant ?disease WHERE {
  ?plant a agont:Crop .
  ?plant agont:hasCropType "Tomato" .
  OPTIONAL { ?plant agont:hasDisease
?disease . }
}
```

Table 8 System Output

Plant ID	Status
Tomato_Plant_001	Early Blight
Tomato_Plant_002	Healthy
Tomato_Plant_003	Late Blight

User Query: "What is the recommended treatment for Early Blight and what was the recent soil moisture?"

System Output:

Diagnosis: Tomato Early Blight.
Recommended Treatment: Apply a copper-based fungicide. Improve air circulation and avoid overhead watering.
Contextual Data: The average soil moisture in the affected zone has been 78% over the last 48 hours, which is above the optimal range (60-70%).

The framework was indeed able to return integrated, contextual answers. The main accomplishment was to have smoothly combined the identifications of diseases through vision with soil sensor data and thus created a more holistic cause-and-effect analysis, for example, high moisture contributes to fungal growth. Use of existing ontologies ensured our data model was sound and reusable. The reasoning engine was able to successfully automate the classification process, which reduced the need to interpret data manually.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Although there would be moderate performance overhead involved for simpler queries, it performs extremely well with complex queries that would entail multiple database operations. The fact that it has a diagnostic accuracy of 94.2% and generates recommendations based on context awareness makes it very useful for precision agriculture.

Scalability analysis has ensured that the system is capable of managing realistic workloads at a farm level. Enterprise-level deployments would still need to be optimized. The comparative analysis thus identifies that semantic framework performance outshines all competition and meets the gap connecting data acquisition and meaningful insights via semantic functionalities. It thus supports intelligent agriculture decisions.

Although the system achieved fair accuracy with tomato crop specifically for Early Blight and Late Blight diseases, the system's scalability and predictability will have to be studied across wider range of crops and diseases in real-world farm environments.

6.1 System Limitation and Optimization

Although it is a remarkably beneficial approach, some drawbacks of this proposed framework were found. One of the major shortcomings of the proposed system is the simulated environment applied for generating sensor data and validating the performance of the ontology developed. Although the simulated environment helps us to realize the proof-of-concept quickly, the system must undergo real-world and practical challenges imposed by physical sensors, edge network, power management, unpredictable weather conditions etc. to prove the readiness for real-world deployment and future usage.

Another limitation identified is the complexity in initial setup due to the long process of developing and aligning ontologies, which typically takes 40-60 hours on a new type of crop. There is also a memory complexity of around 25% due to involved reasoning in triplestores. The cold start latency also has a prominent impact, where latency in initial query execution after a restart in a system has shown 300% higher values due to loading and initial setup of ontologies and reasoners. Lastly, inaccuracies in NLP, where it has been seen that in the complex query setup of agricultural terminology, there has been a 15% decrease in the identification of intent, which takes longer than standard English queries.

6.2 Future Scope

As a future enhancement, the system will be optimized and deployed in a real-world environment that receives data transmitted from multitude of physical sensors to study the impact of real-world environment, sensor noise, processing and data communication latency and unpredictable weather variations on the stability, efficiency and accuracy of the ontology.

Another significant aspect to be addressed as part of future work is to subject heterogenous crops and diseases to the system to expand the Knowledge Graph. Although the system achieved fair accuracy with tomato crop specifically for Early Blight and Late Blight diseases, the system's scalability and predictability can be improved by deploying the system for wider range of crops and diseases in real-world farm environments.

To further enhance the proposed system, a few critical extensions will be developed in the future such as: extending the system from a monitoring platform to a full control system by incorporating real-time actuators; enabling insights from the Knowledge Graph to trigger automated actions through actuators, such as irrigation or ventilation systems; enhancing the Natural Language Processing (NLP) pipeline to handle complex, multi-part queries more effectively. Nevertheless the possible inaccuracies and ethical concerns emanated from NLP demand a focused study. As a long-term roadmap, the system will support the seamless integration of external data sources such as weather forecasts and commodity prices, allowing for wider strategic recommendations. Finally, the semantic gateway will be optimized for deployment on edge computing devices to reduce latency and cloud dependency for critical real-time decisions.

7. CONCLUSION

This paper presented a holistic semantic IoT framework that transforms heterogeneous agricultural sensor streams into an actionable Knowledge Graph, enabling integrated diagnosis and decision support for precision farming. Combining vision-based disease detection (simulated using the PlantVillage dataset), synthetic soil sensor streams, an ontology-driven Semantic Gateway, and a reasoning layer backed by a Jena Fuseki triple store, this system was able to answer natural-language queries with contextual, actionable recommendations. The Knowledge Graph approach bridges disparate data modalities, including images,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

time series sensor readings, and historical records. It does this by allowing for the use of inference engines in deriving cause-and-effect insights. The pipeline can generate logical diagnoses and treatment recommendations while offering a scalable semantic representation. This paper also analyzes the traditional approach backed by conventional DBMS for data storage and retrieval viz-a-viz the semantic gateway powered by a Knowledge Graph offering intuitive answers for complex queries in natural language. The pros and cons of both the approaches in precision agriculture involving heterogeneous sensors for real-time data acquisition are compared. By leveraging Semantic Web standards, farmers get intuitive access to complex and correlated information that significantly enhances the decision-making process with respect to the achievement of more sustainable farming practices.

REFERENCES:

[1] Aarif KO, M., Alam, A., & Hotak, Y. (2025). Smart sensor technologies shaping the future of precision agriculture: Recent advances and future outlooks. *Journal of Sensors*, *2025*(1), 2460098.

[2] Choudhary, V., Guha, P., Pau, G., & Mishra, S. (2025). An overview of smart agriculture using internet of things (IoT) and web services. *Environmental and Sustainability Indicators*, *26*, 100607.

[3] Khatoon, P. S., & Ahmed, M. (2022). Importance of semantic interoperability in smart agriculture systems. *Transactions on Emerging Telecommunications Technologies*, *33*(5), e4448.

[4] Dauda, A., Flauzac, O., & Nolot, F. (2024). A survey on IoT application architectures. *Sensors*, *24*(16), 5320.

[5] Niloofar, P., Lazarova-Molnar, S., Francis, D. P., Vulpe, A., Suciu, G., & Balanescu, M. (2020, December). Modeling and simulation for decision support in precision livestock farming. In *2020 Winter simulation conference (WSC)* (pp. 2601-2612). IEEE.

[6] Monteiro, A., Santos, S., & Gonçalves, P. (2021). Precision agriculture for crop and livestock farming—Brief review. *Animals*, *11*(8), 2345.

[7] Newton, J. E., Nettle, R., & Pryce, J. E. (2020). Farming smarter with big data: Insights from the case of Australia's national dairy herd milk recording scheme. *Agricultural systems*, *181*, 102811.

[8] Murali, E., & Anouncia, S. M. (2023). An ontology-based knowledge mining model for effective exploitation of agro information. *IETE journal of research*, *69*(11), 7856-7873.

[9] Amara, F. Z., Hemam, M., Djezzar, M., & Maimor, M. (2022). Semantic web and internet of things: Challenges, applications and perspectives. *Journal of ICT Standardization*, *10*(2), 261-291.

[10] Wu, J., Chen, H., Orlandi, F., Lee, Y. H., O'Sullivan, D., & Dev, S. (2021, December). Automated climate analyses using knowledge graph. In *2021 IEEE USNC-URSI Radio Science Meeting (Joint with AP-S Symposium)* (pp. 106-107). IEEE.

[11] Rahman, H., Medhi, K., & Hussain, M. I. (2023). DynO-IoT: a dynamic ontology for provisioning semantic interoperability in internet of things. *International Journal of Sensor Networks*, *41*(2), 114-125.

[12] Pereira, P. H. M., Cainelli, G., Pereira, C. E., Da Costa, J. P. J., & De Freitas, E. P. (2023, June). An interoperability middleware for IIoT. In *2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE)* (pp. 1-6). IEEE.

[13] Remya, S., Bonthu, Y., & Bayyapureddi, M. (2024, April). Ontology-enhanced disease detection and crop yield prediction in agriculture using vit. In *International Conference on Innovations in Computational Intelligence and Computer Vision* (pp. 425-438). Singapore: Springer Nature Singapore.

[14] Ranpara, R. (2025). A semantic and ontology-based framework for enhancing interoperability and automation in IoT systems. *Discover Internet of Things*, *5*(1), 22.

[15] Innamorati, C., Redondi, A. E., & Cesana, M. (2025). Optimizing MQTT-CoAP Interoperability: A Broker-Based Extension for Seamless Integration. *IEEE Internet of Things Journal*.

[16] Yang, M., Gao, D., Li, J., Xu, W., & Shi, G. (2024). Superimposed semantic communication for iot-based real-time ecg monitoring. *IEEE Journal of Biomedical and Health Informatics*, *28*(7), 3819-3830.

[17] Konstantinou, D., Stoyanova, T., Stoyanov, S., & Rafailov, R. (2025, November). Securing the Internet of Things: The Role of Artificial Intelligence in Agricultural Cyber Defense. In *2025 10th International Conference on Energy Efficiency and Agricultural Engineering (EE&AE)* (pp. 1-6). IEEE.

[18] Chandra, R., Swaminathan, M., Chakraborty, T., Ding, J., Kapetanovic, Z., Kumar, P., &

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[19] Vasisht, D. (2022). Democratizing data-driven agriculture using affordable hardware. *IEEE Micro*, *42*(1), 69-77.
[20] Vasisht, D., Kapetanovic, Z., Won, J., Jin, X., Chandra, R., Sinha, S., ... & Stratman, S. (2017). {FarmBeats}: an {IoT} platform for {Data-Driven} agriculture. In *14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17)* (pp. 515-529).
[21] Mohan, S., & Panda, S. (2024, November). An Integrated Framework for Real-Time Analysis and Observability of Wireless Sensor Data Using AWS Edge Service Capabilities. In *2024 IEEE International Conference of Electron Devices Society Kolkata Chapter (EDKCON)* (pp. 89-94). IEEE.
[22] Thudumu, S., & Fisher, J. (2025). *OpenAg: Democratizing agricultural intelligence (Version 2). arXiv*.
[23] Sustainable Agriculture Research & Education (SARE) Ontology. [Online]. Available: https://www.sare.org/
[24] FAO, "AGROVOC Thesaurus," [Online]. Available: http://aims.fao.org/vest-registry/vocabularies/agrovoc
[25] Dataset: https://www.kaggle.com/datasets/emmarex/plantdisease
[26] Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., Melo, G. D., Gutierrez, C., ... & Zimmermann, A. (2021). Knowledge graphs. *ACM Computing Surveys (Csur)*, *54*(4), 1-37.