Proceedings of the 20th Conference on Computer Science and Intelligence Systems (FedCSIS) pp. 297–302
DOI: 10.15439/2025F5425
ISSN 2300-5963 ACSIS, Vol. 43

Data Interoperability Using Smart Data Models and NGSI-LD for the Norwegian Agrifood Sector

Rustem Dautov
0000-0002-0260-6343
SINTEF Digital
Oslo, Norway
Email: rustem.dautov@sintef.no

Simeon Tverdal
0000-0003-1660-4127
SINTEF Digital
Oslo, Norway
Email: simeon.tverdal@sintef.no

André Skoog Bondevik
Agdir Drift AS
Arendal, Norway
Email: asb@agdir.no

Svein Arild Frøshaug
Agdir Drift AS
Arendal, Norway
Email: saf@agdir.no

Vera Szabo
Aersea AS
Kristiansand, Norway
Email: vera.szabo@aersea.com

Jan Robert Fiksdal
Aersea AS
Kristiansand, Norway
Email: jan.robert.fiksdal@aersea.com

***Abstract*—The growing adoption of digital technologies in agriculture has led to a proliferation of heterogeneous data from sources such as drones, robotic platforms, and IoT sensors. However, the lack of interoperability across these data streams poses major challenges for integration into decision support systems. This paper presents an approach to harmonising such data using NGSI-LD and Smart Data Models, developed within the Norwegian research project *SMARAGD*. We demonstrate how domain-specific semantic models and linked data principles can be applied to standardise and enrich geospatial and temporal metadata across three key agritech domains: aerial imagery, robotic sensing, and environmental monitoring. The resulting information assets are integrated into a shared, FIWARE-compatible data space, enabling cross-platform visualisation, querying, and reuse. This work contributes to the development of an open, standards-based digital infrastructure for interoperable, data-driven agriculture in Norway and beyond.**

***Index Terms*—Data interoperability, NGSI-LD, Smart Data Models, FIWARE, precision agriculture, semantic data integration, linked data**

I. Introduction

THE digital transformation of agriculture is producing an increasing volume of data from diverse sources such as drones, robotic platforms, and IoT sensors. Realising the full potential of these data streams requires seamless integration across different formats, systems, and stakeholders – a challenge that is especially pressing in the fragmented agrifood sector. Interoperability is essential for enabling advanced decision support tools, improving data reuse, and fostering sustainable, data-driven practices.

This paper addresses the challenge of data interoperability by applying Smart Data Models (SDMs) and the NGSI-LD standard to create a unified, semantically enriched representation of heterogeneous agritech data. Developed within the Norwegian research project *SMARAGD*[1] (Smart Agriculture Data Fusion for Decision Support), our approach focuses on transforming raw data, such as imagery metadata and sensor measurements, into structured and interoperable information assets. The emphasis lies on semantic modelling, rather than on the technical details of the underlying data pipelines or system integration components.

The remainder of this paper is organised as follows: Section II introduces the research context and outlines key interoperability challenges in the Norwegian agrifood domain. Section III summarises related research work. Section IV presents our modelling approach using SDMs and NGSI-LD. Section V discusses the outcomes and implications of our approach, while Section VI concludes the paper and outlines directions for future work.

II. Research Context and Motivation

A. Sector Overview and Interoperability Vision

Norway's agrifood sector is undergoing rapid digitalisation, driven by the deployment of advanced technologies such as UAVs, robotic platforms, and IoT sensors. These tools offer significant potential to enhance productivity, sustainability, and data-informed decision making across the entire agricultural value chain. However, they also introduce a critical barrier – *data heterogeneity*. In practice, each technology stack typically operates in isolation, producing data in incompatible formats and lacking a shared semantic framework.

This fragmentation results in data silos that hinder integration, reuse, and meaningful analytics. For instance, when agronomists or technology providers attempt to combine aerial imagery with in-soil sensor data, or align robotic observations with external weather inputs, they face tedious manual harmonisation processes. Without a common data foundation, these efforts are error-prone and labour-intensive, ultimately undermining the value of digital tools in smart farming.

To address this issue, the research project SMARAGD has adopted a data-centric strategy centred on semantic interoperability. The project's architecture is built on the open-source

[1] https://www.sintef.no/en/projects/2022/smaragd-smart-agriculture-data-fusion-for-decision-support/

FIWARE ecosystem[2] [1], enabling the creation of so-called Smart Data Spaces where information from heterogeneous systems can be fused, queried, and analysed in a unified way.

The cornerstone of this approach is not only the technical integration of data flows, but also the harmonisation of data semantics using shared, standardised data models. Specifically, the NGSI-LD specification and domain-specific SDMs are used to semantically enrich the data so that it can be understood and reused across platforms. FIWARE's modular components, such as the Context Broker and its standard APIs, enable interoperability across data sources and stakeholders, laying the foundation for collaborative, data-driven agriculture.

B. Real-World Interoperability Challenges

Despite the availability of digital infrastructure and advanced sensing technologies, data interoperability remains a major obstacle in real-world agritech applications. The following examples illustrate the nature and diversity of these challenges across three representative domains in SMARAGD.

Aerial imagery. High-resolution images captured by UAVs provide valuable spatial insights for crop monitoring and yield estimation. However, these images are typically offloaded manually from SD cards after field missions, and their embedded metadata (stored in formats like EXIF or proprietary DJI tags) lacks semantic structure. Different camera types (RGB, thermal, multispectral) introduce further variability. As a result, integrating UAV imagery into downstream systems requires specialised extraction and reformatting steps, often with little reuse potential beyond the initial use case.

Robotic imagery. Mobile agricultural robots are increasingly used for close-range inspection, weeding, and precision spraying. These platforms often rely on the Robot Operating System (ROS), which records multi-sensor data into so-called bag files. While ROS is highly effective within robotic environments, it is not designed for external data interoperability. Each data stream (e.g., from GPS, IMU, or cameras) follows a different structure, and the data is typically not enriched with semantic or geospatial context. Converting this raw content into usable, interoperable formats requires extensive processing and domain knowledge.

Environmental sensors. IoT systems deployed in greenhouses, tunnels, and open fields continuously monitor soil, climate, and air parameters. However, these devices are typically vendor-specific, each exposing data through bespoke APIs, measurement units, and timestamp conventions. The absence of shared ontologies or data models results in a proliferation of disconnected systems. For example, two moisture sensors from different providers may use different naming conventions, measurement intervals, and units, hindering aggregation and comparative analysis.

These examples underscore the need for structured, semantically aligned representations that facilitate consistent integration across space, time, and source. By adopting NGSI-LD-based SDMs, the SMARAGD project enables the transformation of heterogeneous sensor outputs into harmonised entities, thereby making them accessible for advanced analytics, visualisation, and long-term reuse.

[2]https://www.fiware.org/

III. Related Work

Interoperability remains a key challenge for smart agriculture, where heterogeneous systems often lack shared standards [2]. Existing modelling vocabularies for data representation and exchange [3] are often either too generic or tailored to narrow, non-reusable use cases. Similarly, ontologies for IoT, sensor networks, and cyber-physical systems [4], [5] need further refinement to address the specific needs of the agritech domain. Although various competing agritech data standards exist,[3] none has achieved wide adoption.

Equally important is the design of data pipelines – automated workflows that perform extraction, transformation, and loading (ETL) of data into usable formats [6]. In agriculture, traditional pipelines often operate in silos, processing drone imagery, sensor streams, or robotic data separately. More recent efforts promote unified, context-aware pipelines based on open standards and semantic models. Initiatives such as FIWARE[4] and NGSI-LD SDMs provide standardised representations to enable cross-platform integration [7], [8], [9], [10]. Another influential reference architecture is Microsoft's FarmBeats [11], [12], [13], which combines edge computing, cloud storage, and AI to process large-scale agritech data. Many pipeline architectures adopt containerised microservices [14] and event-driven components for scalability and flexibility. However, challenges persist, including the harmonisation of proprietary formats, temporal-spatial alignment of multimodal data, and generalisability beyond isolated deployments.

IV. Semantic Interoperability Approach

While numerous promising solutions exist, much of the current work remains experimental or domain-specific. This paper builds on best practices from these efforts, advancing semantic interoperability through SDMs and open, standards-based tooling. Achieving semantic interoperability across diverse agritech data sources is essential for enabling data sharing, fusion, and value-added analytics. This section outlines the key enabling technologies and modelling principles that underpin our approach, developed in the SMARAGD project.

A. SDMs and NGSI-LD for Interoperable Agritech Data

SDMs[5] are at the core of the FIWARE ecosystem, designed to standardise the structure and semantics of data exchanged across systems. They are based on the NGSI-LD specification [15], which represents context data as entities composed of properties and relationships, enriched with geospatial and temporal context. NGSI-LD also supports linked data principles, facilitating richer semantics and integration across domains.

The SDM initiative maintains a large and evolving repository of domain-specific models. In agriculture, these include

[3]https://www.aspexit.com/standards-and-data-exchange-in-agriculture/
[4]https://www.fiware.org/community/smart-agrifood/
[5]https://smartdatamodels.org/

templates for weather, soil, crop, and sensor data. Reusing existing models accelerates development and fosters interoperability. Where necessary, custom models can be contributed to fill domain gaps, as in our work with drone/robotic imagery.

NGSI-LD enables interoperable representation of core attributes such as geolocation (GeoJSON), time (ISO 8601), and sensor metadata. Its machine-readable and extensible structure allows for integration with third-party data sources, including weather and satellite data. Together, NGSI-LD and SDMs provide a robust foundation for harmonising data across heterogeneous agritech systems.

B. Modelling Strategy for Data Harmonisation

Our modelling approach targets the integration of siloed datasets, such as drone imagery, robotic data, and IoT sensor streams, into a unified data space based on NGSI-LD. All data points are modelled as NGSI-LD entities with standardised attributes, timestamps, and geospatial references, supporting spatio-temporal alignment across sources [16], [17]. In particular, GeoJSON-based GPS coordinates and ISO 8601 timestamps serve as key enablers of data interoperability in precision agriculture, allowing observations from different systems to be fused and interpreted in context. Our approach is guided by three core principles:

- **Heterogeneity awareness**: We account for the fragmented nature of agritech systems and the diversity of their data formats.
- **Reuse of open standards**: Wherever possible, we leverage the FIWARE stack, NGSI-LD format, and existing SDMs to avoid redundant development.
- **Future-proof interoperability**: All modelling aligns with EU data space principles and emerging best practices to ensure reusability and policy compliance.

This modelling layer forms the semantic backbone of our architecture, enabling integration with third-party datasets (e.g., weather forecasts or satellite data) and supporting scalable, interoperable decision support services. The resulting harmonised data space serves as the foundation for downstream decision support services, ranging from visual analytics to predictive models.

1) Aerial Imagery Metadata Transformation: Aerial imagery captured by UAVs is a valuable data source in precision agriculture, often used for monitoring crop health or planning interventions. These images, typically stored in the raw DNG format, include embedded EXIF metadata describing camera properties, geolocation, and capture time. However, this metadata is not natively interoperable.

Several challenges emerge when using this data directly: (1) metadata is fragmented across namespaces like `EXIF`, `XMP`, and `Composite`; (2) fields such as `EXIF:GPSLatitude` or `EXIF:FocalLength` lack semantic context; and (3) proprietary tags from manufacturers like DJI hinder automated processing. While rich in content, such raw metadata is difficult to integrate across systems.

To address this, we developed a semantically enriched representation using a custom NGSI-LD SDM, `DroneImage`, which standardises key attributes such as GPS coordinates in GeoJSON, capture date, and camera settings. GPS coordinates and timestamps serve as primary anchors for interoperability, enabling spatial and temporal alignment with data from other sources. The transformation maps each image into an NGSI-LD entity with human-readable, structured attributes, suitable for integration into a shared data space. An example transformation is illustrated in Table I, comparing raw metadata fields to their harmonised NGSI-LD representation.

2) Robotic Imagery Metadata Transformation: Agricultural robots perform close-up imaging and sensor-based monitoring during tasks such as spraying, weeding, or harvesting. These systems typically use ROS to synchronise multiple sensor streams, including imagery, GPS, and orientation data, into ROS bag files. Although ROS provides a robust internal structure, its data format is not directly interoperable with external systems. Three main challenges limit reusability: (1) sensor data is fragmented across asynchronous topics, (2) message types are diverse and lack semantic annotations, and (3) timestamps and geolocation are scattered and inconsistently structured. Manual extraction and interpretation are therefore required to fuse imagery with context metadata such as position and orientation.

To address this, we developed a custom NGSI-LD SDM called `RoboticFrame`, which unifies data from ROS bag files into semantically enriched entities. The model aligns image frames with spatial, temporal, and camera-specific attributes, using ISO 8601 timestamps and GeoJSON-formatted GPS coordinates as primary anchors for interoperability. An example transformation is shown in Table II, where fragmented ROS data is merged into a single structured and linked entity.

3) Environmental Sensor Data Harmonisation: Environmental sensors deployed in fields, tunnels, or greenhouses generate continuous readings of soil moisture, temperature, and other parameters. These are often transmitted in fragmented batches via NB-IoT gateways, with metadata scattered across configuration files or encoded in vendor-specific formats. The lack of semantic structure, inconsistent timestamp formats, and loosely coupled geolocation data make it difficult to integrate such measurements into a unified system.

To address these challenges, we adopt the existing SDM `DeviceMeasurement`[6] to encode individual readings as semantically structured NGSI-LD entities. This model enables consistent representation of core attributes such as measurement type, value, timestamp, elevation, and device identity. Where applicable, we enrich these entities using additional SDMs from the `SmartAgrifood` domain. As with the other two domains, GPS coordinates (GeoJSON) and ISO 8601 timestamps are used as primary anchors for geospatial and temporal alignment across sources – key enablers for precision agriculture and longitudinal analysis. Each entity also establishes NGSI-LD relationships to its originating device, location, and provider. Table III illustrates a representative example of this transformation.

[6]https://github.com/smart-data-models/Smart-Sensoring

TABLE I
SAMPLE TRANSFORMATION OF THE AERIAL IMAGERY METADATA.

Original format: EXIF metadata

```
{
  "SourceFile": "DJI_0007.DNG",
  "File:FileName": "DJI_0007.DNG",
  "File:FileSize": 40920003,
  "EXIF:Make": "DJI",
  "EXIF:Model": "FC6540",
  "EXIF:ModifyDate": "2024:09:25 17:33:10",
  "EXIF:ImageWidth": 6016,
  "EXIF:ImageHeight": 3376,
  "EXIF:GPSLatitude": -48.875485,
  "EXIF:GPSLongitude": -123.391860,
  "EXIF:GPSAltitude": 189.172,
  "EXIF:LensModel": "DJI DL 24mm F2.8 LS ASPH",
  "EXIF:FocalLength": 24,
  "EXIF:ApertureValue": 2.8,
  "EXIF:ShutterSpeedValue": 0.002,
  "EXIF:ISO": 200,
  "XMP:AbsoluteAltitude": "+189.17",
  "XMP:RelativeAltitude": "+122.01",
  "XMP:GimbalYawDegree": "+90.30",
  "XMP:GimbalPitchDegree": -45.0,
  "Composite:GPSPosition": "-48.875485 -123.391860",
  "Composite:FOV": 53.13
}
```

NGSI-LD format

```
{
  "id": "urn:ngsi-ld:DroneImage:DJI_0007",
  "type": "DroneImage",
  "fileName": {
    "type": "Property",
    "value": "DJI_0007.DNG"
  },
  "fileSize": {
    "type": "Property",
    "value": 40920003
  },
  "manufacturer": {
    "type": "Property",
    "value": "DJI"
  },
  "model": {
    "type": "Property",
    "value": "FC6540"
  },
  "imageDimensions": {
    "type": "Property",
    "value": {
      "width": 6016,
      "height": 3376
    }
  },
  "captureDate": {
    "type": "Property",
    "value": "2024-09-25T17:33:10Z"
  },
  "geoLocation": {
    "type": "GeoProperty",
    "value": {
      "type": "Point",
      "coordinates": [-48.875485, -123.391860]
    }
  },
  "altitude": {
    "type": "Property",
    "value": 189.172
  },
  "lensModel": {
    "type": "Property",
    "value": "DJI DL 24mm F2.8 LS ASPH"
  },
  "focalLength": {
    "type": "Property",
    "value": 24
  },
  "aperture": {
    "type": "Property",
    "value": 2.8
  },
  "shutterSpeed": {
    "type": "Property",
    "value": 0.002
  },
  "iso": {
    "type": "Property",
    "value": 200
  },
  "fieldOfView": {
    "type": "Property",
    "value": 53.13
  },
  "gimbalOrientation": {
    "type": "Property",
    "value": {
      "yaw": 90.3,
      "pitch": -45.0
    }
  },
  "imageUrl": "..." // Truncated for brevity
}
```

V. DISCUSSION: BENEFITS AND IMPLICATIONS OF INTEROPERABLE AGRITECH DATA

Based on the described data harmonisation strategy, in the context of the SMARAGD project, we developed data transformation pipelines that enabled the integration of drone imagery, robotic observations, and environmental sensor measurements into a unified, standards-based information space. Each dataset is harmonised using SDMs and encoded in NGSI-LD format, allowing them to be jointly queried, visualised, or analysed. Despite differences in data origin, structure, and operational constraints, all entities share a common representation of key attributes such as geolocation (via GeoJSON) and timestamps (ISO 8601). This alignment enables integration across domains and supports spatial-temporal data fusion. NGSI-LD's linked data structure provides machine-readable context by connecting each observation to its corresponding device, location, and provider. Although NGSI-LD representations are more verbose than raw formats and require additional modelling effort, they bring long-term benefits in maintainability, cross-platform compatibility, and ease of reuse, effectively mitigating future technical debt.

TABLE II
SAMPLE TRANSFORMATION OF THE ROS BAG METADATA.

Original format: ROS bag

```
{
  "timestamp": "2025-01-21T12:00:00Z",
  "px4/position_world": {
    "Point": {
      "X": 47.11528015136719,
      "Y": 8.670736312866211,
      "Z": 1040.333984375
    }
  },
  "px4/orientation": {
    "Quaternion": {
      "X": -0.04053421614832881,
      "Y": 0.1062795290984786,
      "Z": -0.9375996305227535,
      "W": 0.3285857146308133
    }
  },
  "cam0/image_raw": {
    "sensor_msgs/Image": {
      "Height": 480,
      "Width": 752,
      "Encoding": "mono8",
      "Data": "[87, 92, 93, ...]" // Truncated for brevity
    }
  }
}
```

NGSI-LD format

```
{
  "id": "urn:ngsi-ld:RoboticFrame: 2025-01-21T12:00:00Z",
  "type": "RoboticFrame",
  "timestamp": {
    "type": "Property",
    "value": "2025-01-21T12:00:00Z"
  },
  "position": {
    "type": "GeoProperty",
    "value": {
      "type": "Point",
      "coordinates": [-48.875485, -123.391860],
      "altitude": 1040.333984375
    }
  },
  "orientation": {
    "type": "Property",
    "value": {
      "quaternion": {
        "x": -0.04053421614832881,
        "y": 0.1062795290984786,
        "z": -0.9375996305227535,
        "w": 0.3285857146308133
      }
    }
  },
  "camera": {
    "type": "Property",
    "value": {
      "id": "cam0",
      "resolution": {
        "height": 480,
        "width": 752
      },
      "encoding": "mono8",
      "imageUrl": "..." // Truncated for brevity
    }
  }
}
```

As a result, this semantically enriched and standardised representation supports a wide range of decision support use

TABLE III
SAMPLE TRANSFORMATION OF SOIL SENSOR DATA.

Original format: vendor specific JSON documents

```json
{
  "_id": {
    "$oid": "634dc15fe4132c6a5cff7372"
  },
  "time": {
    "$date": "2024-07-06T18:41:18.531Z"
  },
  "value": 21,
  "name": 20,
  "elevation": 0,
  "companyId": {
    "$oid": "634db5fd036971b6bf5aba89"
  },
  "location": {
    "_id": {
      "$oid": "634db600036971b6bf5abaa1"
    }
  },
  "device": {
    "_id": {
      "$oid": "634db601036971b6bf5abab7"
    },
    "vendor": 2,
    "type": "hydra",
    "soilConfiguration": {
      "soilType": "organic",
      "depth": 10,
      "soilDryBulkDensity": 0.5,
      "fieldCapacity": 60,
      "wiltingPoint": 15,
      "irrigationThreshold": 0.5
    },
    "serialNumber": "19344",
    "externalIdentity": "11003"
  }
}
```

NGSI-LD format

```json
{
  "id": "urn:ngsi-ld:DeviceMeasurement: 634dc15fe4132c6a5cff7372",
  "type": "DeviceMeasurement",
  "measurementType": {
    "type": "Property",
    "value": "soilTemperature"
  },
  "measurementValue": {
    "type": "Property",
    "value": 21
  },
  "elevation": {
    "type": "Property",
    "value": 0
  },
  "timestamp": {
    "type": "Property",
    "value": "2024-07-06T18:41:18.531Z"
  },
  "refCompany": {
    "type": "Relationship",
    "object": "urn:ngsi-ld:Company: 634db5fd036971b6bf5aba89"
  },
  "refLocation": {
    "type": "Relationship",
    "object": "urn:ngsi-ld:Location: 634db600036971b6bf5abaa1"
  },
  "refDevice": {
    "type": "Relationship",
    "object": "urn:ngsi-ld:Device: 634db601036971b6bf5abab7"
  },
  "deviceDetails": {
    "type": "Property",
    "value": {
      "vendor": 2,
      "type": "hydra",
      "soilConfiguration": {
        "soilType": "organic",
        "depth": 10,
        "soilDryBulkDensity": 0.5,
        "fieldCapacity": 60,
        "wiltingPoint": 15,
        "irrigationThreshold": 0.5
      },
      "serialNumber": "19344",
      "externalIdentity": "11003"
    }
  }
}
```

cases. For instance, drone-based vegetation indices can be correlated with robotic imagery and soil moisture readings to inform irrigation or treatment decisions. Integration with weather forecasts further enables optimal scheduling of agricultural operations. Even non-expert users can benefit from basic time-series visualisations, while domain specialists can perform detailed queries and analytics.

Beyond technical integration, this approach offers value for both end users and data providers. Farmers and agronomists benefit from timely, context-aware insights, while data providers can contribute semantically aligned datasets to shared platforms, enabling monetisation, collaboration, and extended reuse. For example, combining ground-based observations with satellite data or treatment recommendations from suppliers creates synergy effects that increase the value of individual datasets. This interoperable infrastructure also opens the door to new business models, such as real-time advisory services, predictive analytics, or autonomous workflows. As interoperability improves, stakeholders across the value chain, from growers to regulators, can engage with trustworthy, machine-readable information that supports sustainable and data-driven agriculture.

Contributions in a European Context. The presented approach aligns with ongoing European efforts to foster data sharing and interoperability in agriculture. Initiatives like AgriDataSpace[7] and policy frameworks such as the EU Data Act [18] emphasise common standards and linked data principles to facilitate innovation and reuse. By adhering to NGSI-LD and contributing domain-specific SDMs, the SMARAGD project lays the groundwork for scalable, semantically interoperable agritech systems. The resulting information assets are reusable in broader European data ecosystems, promoting sustainable agriculture and cross-sector innovation in both national and international contexts.

VI. CONCLUSION

This paper presented a standards-based approach to agritech data interoperability through the application of NGSI-LD and SDMs. Focusing on three heterogeneous data sources – UAV imagery, robotic sensing, and environmental IoT measurements – we demonstrated how semantically enriched and temporally aligned NGSI-LD entities can be used to bring previously siloed datasets into a unified, interoperable information space. This transformation facilitates cross-domain integration, supports spatial-temporal analysis, and lays the groundwork for advanced decision-support services in the Norwegian agrifood sector.

The resulting information assets are designed to be reusable and interoperable across platforms, aligning with the principles and goals of European initiatives such as AgriDataSpace and the EU Data Act. By embracing shared ontologies, linked data principles, and modular, open-source tooling, this work contributes to a scalable, forward-compatible digital infrastructure for sustainable agriculture. In addition to supporting smarter farm-level decision-making, it enables new forms of data collaboration and monetisation, unlocking synergy effects among different data providers and applications.

Future Work. Future work will focus on the broader validation and evolution of the domain-specific SDMs introduced for drone and robotic imagery. While initially developed to meet specific needs in the SMARAGD project, these models will benefit from wider community review and iterative refinement within the open SDMs ecosystem.

[7]https://agridataspace-csa.eu/

TABLE IV
KEY CHALLENGES IN ORIGINAL DATA FORMATS VS. BENEFITS OF NGSI-LD TRANSFORMATION.

Key challenges of the original formats	**Key benefits of the transformed NGSI-LD format**
• **Scattered and Disjoint Metadata:** Contextual data (e.g., timestamps, geolocation, sensor or device metadata) is spread across multiple namespaces, files, or topics (e.g., `EXIF`, `ROS bags`) and separate sensor metadata files. • **Proprietary and Inconsistent Structures:** Each data source uses its own format (EXIF for UAV imagery, ROS message types for robots) and custom enumerations for IoT sensors, requiring domain-specific logic to interpret or align. • **Limited Interoperability:** Original data structures are tightly coupled to specific vendors or platforms (e.g., DJI cameras, ROS message types, vendor-specific IoT APIs), making cross-system integration difficult. • **No Semantic Clarity or Linked Context:** Field names lack semantic meaning and are not explicitly linked (e.g., sensor measurements not connected to geolocation or device metadata in machine-readable form).	• **Unified Entity Representation:** Each dataset is transformed into a single NGSI-LD entity (e.g., `DroneImage`, `RoboticFrame`, `DeviceMeasurement`) representing relevant context in one structure. • **Semantic Interoperability and Enrichment:** All properties and relationships are semantically defined using SDMs (e.g., `measurementType`, `deviceDetails`, `captureDate`, `geoLocation`). • **Spatial and Temporal Alignment:** Geo-coordinates are consistently represented using GeoJSON; timestamps are standardised to ISO 8601 across all data types. • **Cross-Platform Integration:** NGSI-LD enables seamless ingestion into context-aware systems (e.g., FIWARE's Orion-LD), supporting integration with third-party data sources like satellite imagery or weather forecasts. • **Machine-Readable Relationships:** Explicit links are established between entities (e.g., sensor → device → location → company), enabling scalable querying and analytics.

Moreover, we plan to quantitatively evaluate the performance and scalability of the underlying data transformation pipelines under realistic, large-scale workloads [19]. This includes stress-testing the ingestion and processing capabilities of the system across different deployment environments (edge, fog, cloud) and assessing its responsiveness and throughput when handling high-frequency sensor streams or large image collections. Such benchmarks will inform optimisation strategies and ensure that the system remains robust, efficient, and suitable for operational use in demanding agritech scenarios.

ACKNOWLEDGMENT

This work is funded by the Research Council of Norway under grant agreement no. 337012 (SMARAGD) and the European Union's Horizon Europe Research and Innovation Programme under grant agreement No. 101135576 (INTEND).

REFERENCES

[1] R. Dautov, S. Tverdal, A. S. Bondevik, S. A. Frøshaug, V. Szabo, J. R. Fiksdal, and M. F. Stølen, "SMARAGD: Data Interoperability for Decision Support in the Norwegian Agrifood Sector," in *Companion Proceedings of the 8th International Joint Conference on Rules and Reasoning (RuleML+RR-Companion 2024)*, 2024.

[2] O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow, and M. N. Hindia, "An overview of Internet of Things (IoT) and data analytics in agriculture: Benefits and challenges," *IEEE Internet of things Journal*, vol. 5, no. 5, pp. 3758–3773, 2018. doi: 10.1109/JIOT.2018.2844296

[3] V. Pesce, G.-W. Kayumbi, J. Tennison, L. Mey, and P. Zervas, "Agrifood data standards: a gap exploration report," *F1000Research*, vol. 7, no. 176, p. 176, 2018. doi: 10.7490/f1000research.1115261.1

[4] S. Mishra and S. Jain, "Ontologies as a semantic model in IoT," *International Journal of Computers and Applications*, vol. 42, no. 3, pp. 233–243, 2020. doi: 10.1080/1206212X.2018.1504461

[5] I. Szilagyi and P. Wira, "Ontologies and Semantic Web for the Internet of Things-a survey," in *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2016. doi: 10.1109/IECON.2016.7793744 pp. 6949–6954.

[6] A. Raj, J. Bosch, H. H. Olsson, and T. J. Wang, "Modelling data pipelines," in *2020 46th Euromicro conference on software engineering and advanced applications (SEAA)*. IEEE, 2020. doi: 10.1109/SEAA51224.2020.00014 pp. 13–20.

[7] J. López-Riquelme, N. Pavón-Pulido, H. Navarro-Hellín, F. Soto-Valles, and R. Torres-Sánchez, "A software architecture based on FIWARE cloud for Precision Agriculture," *Agricultural water management*, vol. 183, pp. 123–135, 2017. doi: 10.1016/j.agwat.2016.10.020

[8] P. Corista, D. Ferreira, J. Gião, J. Sarraipa, and R. J. Gonçalves, "An IoT agriculture system using FIWARE," in *2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*. IEEE, 2018. doi: 10.1109/ICE.2018.8436381 pp. 1–6.

[9] M. A. Rodriguez, L. Cuenca, and A. Ortiz, "FIWARE open source standard platform in smart farming-a review," in *19th Working Conference on Virtual Enterprises (PRO-VE 2018)*. Springer, 2018. doi: 10.1007/978-3-319-99127-6_50 pp. 581–589.

[10] I. Roussaki, K. Doolin, A. Skarmeta, G. Routis, J. A. Lopez-Morales, E. Claffey, M. Mora, and J. A. Martinez, "Building an interoperable space for smart agriculture," *Digital Communications and Networks*, vol. 9, no. 1, pp. 183–193, 2023. doi: 10.1016/j.dcan.2022.02.004

[11] D. Vasisht, Z. Kapetanovic, J. Won, X. Jin, R. Chandra, S. Sinha, A. Kapoor, M. Sudarshan, and S. Stratman, "FarmBeats: an IoT platform for Data-Driven agriculture," in *14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17)*. USENIX Association, 2017, pp. 515–529.

[12] S. Malvar, A. Badam, and R. Chandra, "FarmBeats: Digital Water for Agriculture," *Resource Magazine*, vol. 29, no. 4, pp. 40–42, 2022.

[13] Z. Kapetanovic, D. Vasisht, J. Won, R. Chandra, and M. Kimball, "Experiences deploying an always-on farm network," *GetMobile: Mobile Computing and Communications*, vol. 21, no. 2, pp. 16–21, 2017. doi: 10.1145/3131214.3131220

[14] D. Jaramillo, D. V. Nguyen, and R. Smart, "Leveraging microservices architecture by using Docker technology," in *SoutheastCon 2016*. IEEE, 2016. doi: 10.1109/SECON.2016.7506647 pp. 1–5.

[15] ETSI Industry Specification Group (ISG), "Context Information Management (CIM); NGSI-LD API," ETSI, Tech. Rep. RGS/CIM-009v161, 2019.

[16] R. Dautov and S. Distefano, "Distributed data fusion for the Internet of Things," in *Proceedings of 14th International Conference on Parallel Computing Technologies (PaCT 2017)*. Springer, 2017. doi: 10.1007/978-3-319-62932-2_41 pp. 427–432.

[17] R. Dautov and S. Distefano, "Three-level hierarchical data fusion through the IoT, edge, and cloud computing," in *Proceedings of the 1st International Conference on Internet of Things and Machine Learning*, 2017. doi: 10.1145/3109761.3158388 pp. 1–5.

[18] European Commission and Directorate-General for Communication, *Data Act – The path to the digital decade*. Publications Office of the European Union, 2022.

[19] D. Firmani, F. Leotta, J. G. Mathew, J. Rossi, L. Balzotti, H. Song, D. Roman, R. Dautov, E. J. Husom, S. Sen, V. Balionyte-Merle, A. Morichetta, S. Dustdar, T. Metsch, V. Frascolla, A. Khalid, G. Landi, J. Brenes, I. Toma, R. Szabó, C. Schaefer, C. Udroiu, A. Ulisses, V. Pietsch, S. Akselsen, A. Munch-Ellingsen, I. Pavlova, H.-G. Kim, C. Kim, B. Allen, S. Kim, and E. Paulson, "INTEND: Intent-Based Data Operation in the Computing Continuum," in *CEUR Workshop Proceedings*, vol. 3692, 2024, pp. 43–50.