**ITUPublications**

**International Telecommunication Union**
Standardization Sector

# ITU-T Technical Report

**(07/2024)**

# YSTR.DataModelling-Agri

## Data processing, management and analytics with artificial intelligence for digital agriculture

# Technical Report ITU-T YSTR.DataModelling-Agri

## Data processing, management and analytics with artificial intelligence for digital agriculture


### Summary

Technical Report ITU-T YSTR.DataModelling-Agri on data modelling for digital agriculture delves into the integration of data technologies and artificial intelligence (AI) modelling for optimizing farming practices. It explores diverse data sources, techniques, pre-processing methods, and modelling algorithms used in digital agriculture. It illuminates how digital agriculture is revolutionizing crop management, resource utilization, and sustainability practices, ultimately paving the way for a more efficient and resilient agricultural sector.



### Keywords

AI, data collection, digital agriculture, IoT, modelling.


### Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

## Table of Contents

# Technical Report ITU-T YSTR.DataModelling-Agri

## Data processing, management and analytics with artificial intelligence for digital agriculture

## 1 Scope

This Technical Report provides a comprehensive overview of how data modelling technologies are transforming agricultural practices, enhancing productivity, sustainability, and efficiency in the digital age.

## 2 References

None.

## 3 Definitions

### 3.1 Terms defined elsewhere

This Technical Report uses the following terms defined elsewhere:

**3.1.1 artificial intelligence (AI)** [b-ITU-T M.3080]: Computerized system that uses cognition to understand information and solve problems.

NOTE 1 – ISO/IEC 2382-28 defines AI as "an interdisciplinary field, usually regarded as a branch of computer science, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning".

NOTE 2 – In computer science AI research is defined as the study of "intelligent agents": any device that perceives its environment and takes actions to achieve its goals.

NOTE 3 – This includes pattern recognition, the application of machine learning and related techniques.

NOTE 4 – Artificial-intelligence is the whole idea and concept of machines being able to carry out tasks in a way that mimics human intelligence and would be considered "smart".

**3.1.2 Internet of things (IoT)** [b-ITU-T Y.2060]: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

### 3.2 Terms defined in this Technical Report

None.

## 4 Abbreviations and acronyms

This Technical Report uses the following abbreviations and acronyms:

AI Artificial Intelligence

API Application Programming Interface

FMIS Farm Management Information System

IAM Identity and Access Management

IoT Internet of Things

ML Machine Learning

MQTT Message Queuing Telemetry Transport

REST Representational State Transfer

| | |
|---|---|
| ROI | Region of Interest |
| SDK | Software Development Kit |
| USN | Ubiquitous Sensor Network |

## 5 Conventions

None.

## 6 Introduction

Artificial intelligence (AI) and the Internet of things (IoT) play pivotal roles in shaping the landscape of digital agriculture, offering innovative solutions for enhancing productivity, sustainability, and efficiency in farming practices. IoT devices, such as sensors, drones, and smart farming equipment, enable the collection of real-time data on various agricultural parameters like soil moisture levels, temperature, humidity, and crop health. This data is then processed and analyzed using AI algorithms to extract valuable insights, patterns, and trends that empower farmers to make data-driven decisions.

Leveraging AI for agriculture revolutionizes traditional farming practices by integrating modern technologies to optimize crop production, enhance sustainability, and improve overall efficiency.

This innovative approach would utilize an array of complementary technologies including AI, machine learning (ML), IoT sensors, and data analytics to transform every stage of the agricultural production process.

At the core of this would lie an AI-centric architecture oriented towards robust data collection mechanisms that gather crucial information on factors like soil quality, weather patterns, crop health, and resource usage. This data is then processed using advanced AI algorithms to derive actionable insights and patterns, empowering farmers to make informed decisions in real time. Decision-support systems based on AI recommendations enable farmers to implement precision agricultural techniques, such as targeted irrigation, personalized fertilization, and customized pest control strategies. By continuously monitoring crop health and environmental conditions, AI systems facilitate early detection of issues like diseases, pests, or nutrient deficiencies, allowing for prompt intervention and mitigation. Automation technologies, driven by AI algorithms, streamline farming operations by autonomously performing tasks like seeding, weeding, and harvesting with precision and efficiency. Predictive maintenance tools anticipate equipment failure, thereby reducing downtime and optimizing operational workflows. Moreover, AI-enabled market analysis tools provide farmers with valuable insights into market trends, consumer preferences, and pricing data, enabling strategic decision-making regarding crop selection and sales timing.

This transformative fusion of AI and agriculture not only improves operational efficiency but also fosters sustainable practices. By optimizing resource utilization through data-driven insights, AI architecture minimizes waste, reduces environmental impact, and enhances crop yields. The seamless integration of AI into agricultural processes enables farmers to adapt quickly to changing conditions, mitigate risks, and maximize productivity. Real-time monitoring of crop health and environmental parameters empowers farmers to take proactive measures, ensuring the well-being of their crops while increasing overall yield quality. Automation technologies driven by AI algorithms not only streamline labor-intensive tasks but also enhance precision and consistency in farming operations.

AI systems allow forecasting the future. This forecasting allows for predicting the impact of environmental changes such as weather, as well as fertilization and harvest time, and facilitates earlier intervention on a data-driven basis. This enables greater efficiency, increased food security and maximized harvests. Furthermore, the predictive maintenance capabilities of AI systems preemptively identify equipment issues, enabling timely repairs and preventing costly downtime. The market intelligence provided by AI aids farmers in making strategic decisions that optimize profitability and market competitiveness. Ultimately, AI architecture for agriculture represents a paradigm shift towards a more sustainable, efficient, and technologically advanced agricultural sector that meets the demands of a rapidly evolving global food system.

## 7 Challenges in AI architecture for agriculture: overcoming data and modelling hurdles

Data modelling in agriculture faces numerous challenges, including issues with data quality and consistency, where inconsistent formats and incomplete data can lead to unreliable models. The sheer volume of data from various sources, such as sensors and satellites, demands significant storage and processing capabilities, while the integration of diverse data types presents additional complexity. Accessibility is another hurdle, with limited access to relevant data due to proprietary restrictions or lack of digitization in certain regions. Ensuring data privacy and security, especially for proprietary agricultural practices and personal information, adds another layer of difficulty [b-MDPI].

The high variability in environmental conditions requires robust models that can generalize well across different agro-ecological zones. A shortage of skilled personnel in data science and AI within the agricultural sector exacerbates these issues, compounded by resource constraints, particularly in developing regions. Additionally, balancing the need for high-resolution temporal and spatial data with practical collection and processing capabilities is challenging. Ensuring model accuracy and reliability under varying conditions, validating models with ground-truth data, and developing adaptable, scalable solutions are critical. Interdisciplinary collaboration between agronomists, data scientists, engineers, and policymakers is essential for bridging these gaps and creating effective, comprehensive models for the agricultural sector.

Additionally, one of the primary challenges has been ensuring interoperability and compatibility with existing data systems and standards across different regions and industries. Addressing these challenges has required close collaboration with stakeholders, and ongoing refinement of data integration capabilities, to ensure seamless data exchange and compatibility with diverse data sources within the agricultural domain [b-PeerJ].

The main challenges in this context are associated with:

a) Data quality and consistency:
  - Inconsistent data formats and quality can hinder analysis and modelling efforts
  - Incomplete or inaccurate data can lead to unreliable models.

b) Data volume and storage
  - Large volumes of data generated from various sources (e.g., sensors and satellites) require significant storage and processing capabilities
  - Efficient management and retrieval of this data is critical.

c) Data integration
  - Integrating data from diverse sources (e.g., weather data, soil data, and crop data) can be complex
  - Ensuring compatibility and seamless integration between different data sets and formats is challenging.

d) Data accessibility
  - Limited access to relevant data due to proprietary restrictions or lack of digitization in some regions
  - Difficulty in obtaining real-time data for timely decision-making.

e) Data privacy and security
  - Protecting sensitive data, such as proprietary agricultural practices and the personal information of farmers, is essential
  - Ensuring compliance with data privacy regulations adds another layer of complexity.

f) Environmental variability

- High variability in environmental conditions (e.g., weather patterns and soil types) requires robust models that can generalize well
- Models must account for diverse agro-ecological zones and microclimates.

g) Technical expertise

- Lack of skilled personnel in data science and AI within the agricultural sector.
- Bridging the gap between agricultural domain knowledge and technical expertise in data modelling.
- Fostering collaboration between agronomists, data scientists, engineers, and policymakers.
- Bridging the gap between different disciplines, to create comprehensive and effective models.

h) Resource constraints

- Limited financial and technological resources, especially in developing regions
- High costs associated with advanced data collection tools and infrastructure.

i) Temporal and spatial resolution

- Balancing the need for high-resolution data with the practicalities of data collection and processing
- Ensuring temporal resolution is adequate for capturing relevant agricultural phenomena.

j) Model validation and accuracy

- Ensuring models are accurate and reliable under varying conditions and for different crops
- Validating models with ground-truth data, to ensure their practical applicability.

k) Adaptability and scalability

- Developing models that can adapt to changing agricultural practices and technologies
- Ensuring scalability of models to accommodate large and diverse agricultural landscapes.

l) Data and model interoperability

- Ensuring models can be used within different systems without needing to be retrained/rewritten
- Ensuring data from different systems can be used with models that are not specifically built for the system.

# 8 Essential features of AI architecture: A comprehensive overview

AI architecture for agriculture involves implementing AI technologies to enhance various aspects of agriculture, such as crop management, monitoring, and decision-making. Here are some key components of AI architecture in agriculture:

a) **Data collection:** Sensors, drones, satellites, and other IoT devices collect data such as weather conditions, soil moisture levels, crop health, and more.

b) **Data processing**: AI algorithms process the collected data to extract meaningful insights and patterns. ML models are often used for tasks like image recognition, predictive analytics, and anomaly detection.

c) **Decision support systems**: AI systems provide farmers with recommendations and insights for optimizing crop yields, reducing resource wastage, and improving overall efficiency. This helps in making informed decisions about irrigation, fertilization, pest control, and harvesting.

d) **Monitoring and automation**: AI-powered monitoring systems track crop health in real time, enabling early detection of diseases, pests, or nutrient deficiencies. Automation technologies like robotic farming equipment can be controlled by AI algorithms to perform tasks like planting, weeding, and harvesting.

e) **Forecasting**: Forecasting in agriculture requires AI systems that model complex environmental interactions between soil, weather, crop genetics and management practices. These systems generate accurate predictions for optimal planting windows, irrigation scheduling, fertilizer applications, and harvest timing based on real-time sensor data and historical outcomes. By modelling agricultural states and making data-driven predictions, AI enables farmers to optimize resource use, reduce risks, and maximize sustainable yields through precise timing of critical farming operations.

f) **Predictive maintenance**: AI can predict equipment failure and maintenance needs, reducing downtime and optimizing operational efficiency on the farm.

g) **Market analysis**: AI tools can analyse market trends, pricing data, and consumer preferences to help farmers make strategic decisions about what crops to grow and when to sell them.

## 9 AI system architecture: core aspects for digital agriculture

In response to the increasingly complex challenges facing agriculture, particularly concerning sustainability and efficiency, this section presents a comprehensive solution aimed at revolutionizing farming practices predicated on an AI-based architecture.

Agriculture stands at a critical juncture, tasked with the formidable challenge of feeding a growing global population while minimizing its environmental footprint and adapting to changing climatic conditions. In light of these challenges, the integration of cutting-edge technologies such as AI and semantic standards offers immense potential to enhance agricultural productivity, optimize resource utilization, and promote environmental stewardship. The digital platform that serves as a nexus for data aggregation, analysis, and decision-making is vital within the agricultural domain. At its core lies the concept of the digital twin, a virtual counterpart of farms and fields that captures a wealth of metadata crucial for optimizing agricultural processes. This digital twin acts as a central repository, aggregating data from various sources such as sensors, drones, and farm management systems, and facilitating seamless interaction with AI-driven services.

Central to the framework of an AI-based architecture for agriculture would be the development of semantic data models tailored specifically for agricultural information. These models enable the standardized representation and exchange of data, fostering interoperability among diverse AI services, data platforms, and farm management systems. By adopting semantic standards, we aim to overcome the inherent challenges associated with data heterogeneity, enabling seamless integration and analysis of disparate datasets.

In addition to semantic data models, the architectural framework incorporates self-descriptive mechanisms for AI services, enabling automated registration, discovery and invocation of services by the digital twin. This self-descriptive capability streamlines the integration of AI-driven solutions into agricultural workflows, empowering farmers with the tools and insights needed to optimize their operations.

Key stakeholders within the digital agricultural ecosystem would include farmers, agricultural researchers, AI service providers, data platform developers, and policymakers (see Figure 1).

Furthermore, the architectural framework would be designed to cater to the diverse needs of companies, customers (farmers), research institutions and, potentially regulatory authorities. It serves as a comprehensive ecosystem that not only aggregates data from various sources, including other farm management information systems (FMIS) and IoT devices but also operates according to principles aligned with Gaia-X and International Dataspace Association (IDSA) standards.

At its core, AI-based architecture facilitates the seamless integration and exchange of data, enabling stakeholders to harness the full potential of agricultural information. By adhering to Gaia-X and IDSA principles, the platform ensures data sovereignty, interoperability, and compliance with regulatory frameworks, thereby instilling trust among users and fostering collaboration across the agricultural ecosystem. Such an architectural framework goes beyond data aggregation by providing sophisticated semantic descriptions for all types of data. This semantic layer enables scientific researchers to leverage the platform for knowledge discovery and hypothesis testing, empowering them to derive valuable insights from heterogeneous datasets. Simultaneously, farmers benefit from the semantic representation of data, as it facilitates the interoperable use of independent AI services for data analysis. By leveraging these semantic descriptions, farmers can unlock new insights and optimize their agricultural practices, thereby enhancing productivity and sustainability.

It is essential that the framework should be scalable and adaptable for global application. The standards and semantic descriptions employed within the platform should also be crafted with a global perspective, ensuring that users from different countries can effectively enhance their agricultural practices in line with existing international standards [b-NALAMKI].

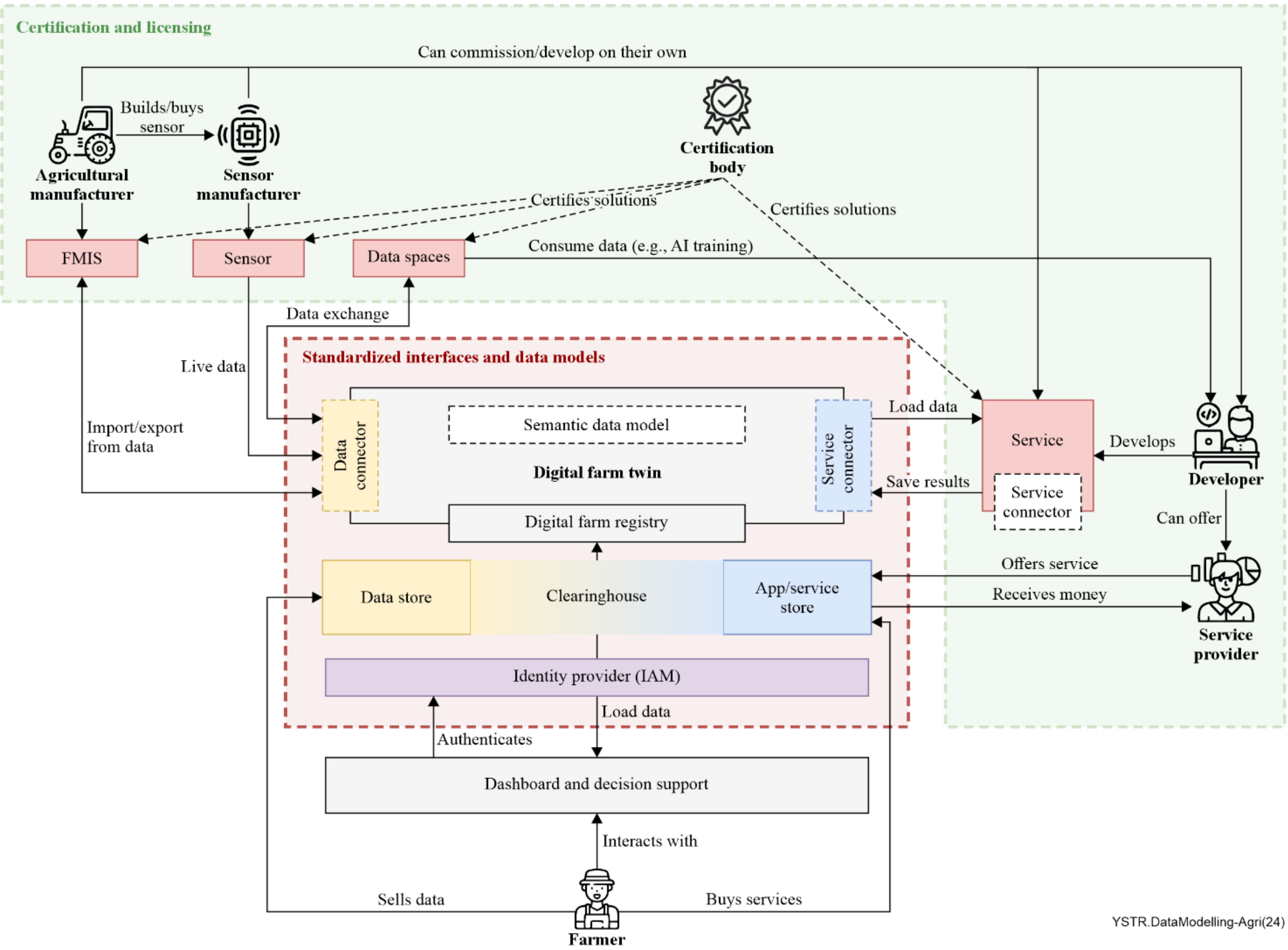


**Figure 1 – Facilitating data exchanges between stakeholders [b-NALAMKI]**

## 10 Architectural considerations

### 10.1 Data management

The data management strategy should be designed to accommodate both structured, semantically enriched information about agricultural entities (such as fields, farms, and cultivation methods) and diverse data types, including IoT data, images (such as drone and satellite imagery), and other geolocated agricultural data. This approach enables the association of various data sources with specific agricultural objects, facilitating comprehensive analysis and insights generation, particularly through AI services.

The captured data, along with agricultural information, is stored holistically, ensuring seamless integration and semantic alignment. To achieve this, a combination of database technology and storage is utilized, providing scalability to the platform to handle the vast amounts of agricultural data effectively. This hybrid approach enables to scale up or down based on the demands of data processing and storage, ensuring optimal performance and resource utilization.

Figure 2 illustrates a conceptual data model for integrating AI services with traditional farm management systems. The model shows the relationships between key agricultural entities, from company/product level down to individual fields and their regions of interest (ROIs). It demonstrates how farm data synchronizes with management information systems while connecting to AI services through a service catalogue. This structured approach enables systematic collection and evaluation of agricultural information, supporting data-driven decision-making at various levels of farm operations.

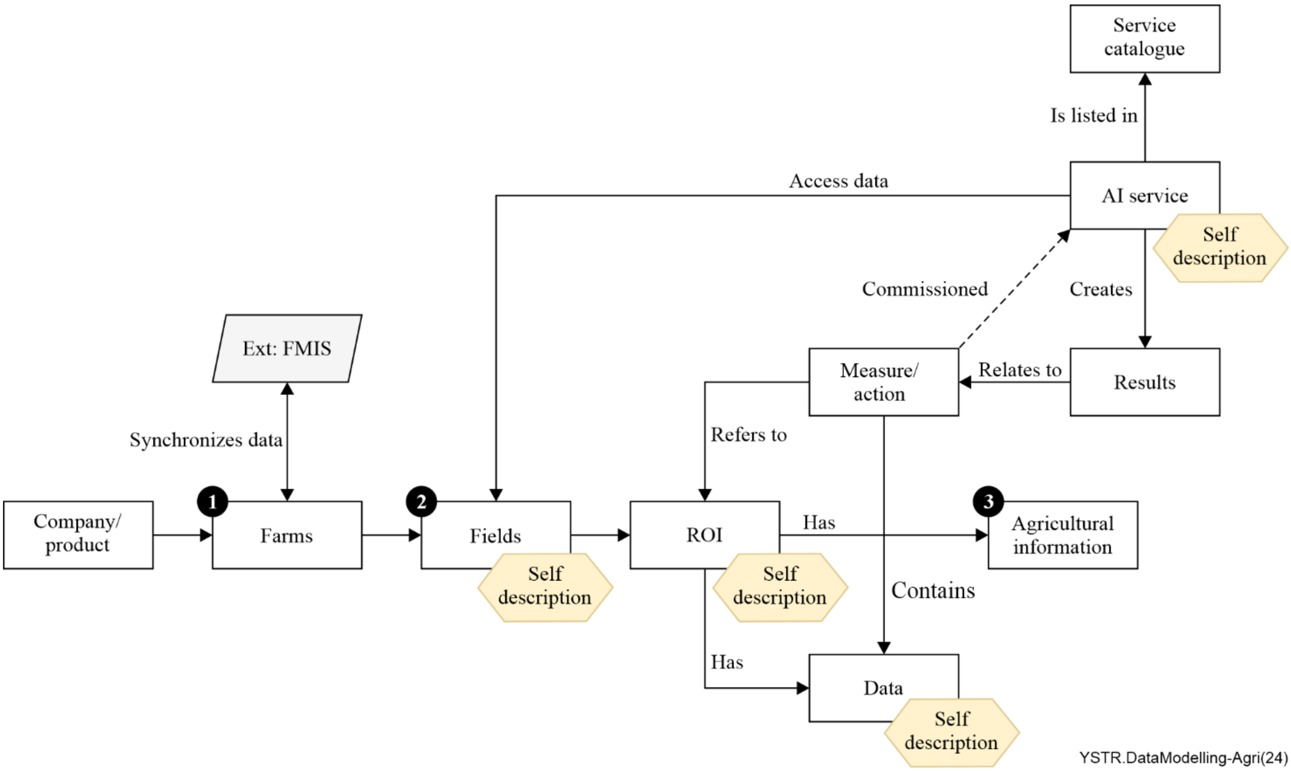


**Figure 2 – Data flow and access optimization for digital agriculture**

An integral aspect of an adequate data management framework is its adherence to principles of ownership, security, and privacy, in line with Gaia-X and IDSA principles. Clear ownership and access control mechanisms are implemented to safeguard the integrity and confidentiality of data, ensuring that only authorized users can access and manipulate the information stored within the platform. These measures not only bolster data security but also instil trust among users, encouraging collaboration and data sharing while maintaining strict compliance with regulatory requirements.

## 10.2 System architecture

The core services constitute the core of the platform and encompass various key functions. These include the digital farm as a central data repository, storing structured information about agricultural operations, fields, and cultivation methods. The service store enables users to explore and book (AI) services to gain insights from the collected data. Identity and access management (IAM) (see Figure 3) ensures secure authentication and authorization of users, while the dashboard provides a user-friendly interface through which end users can access all functionalities.

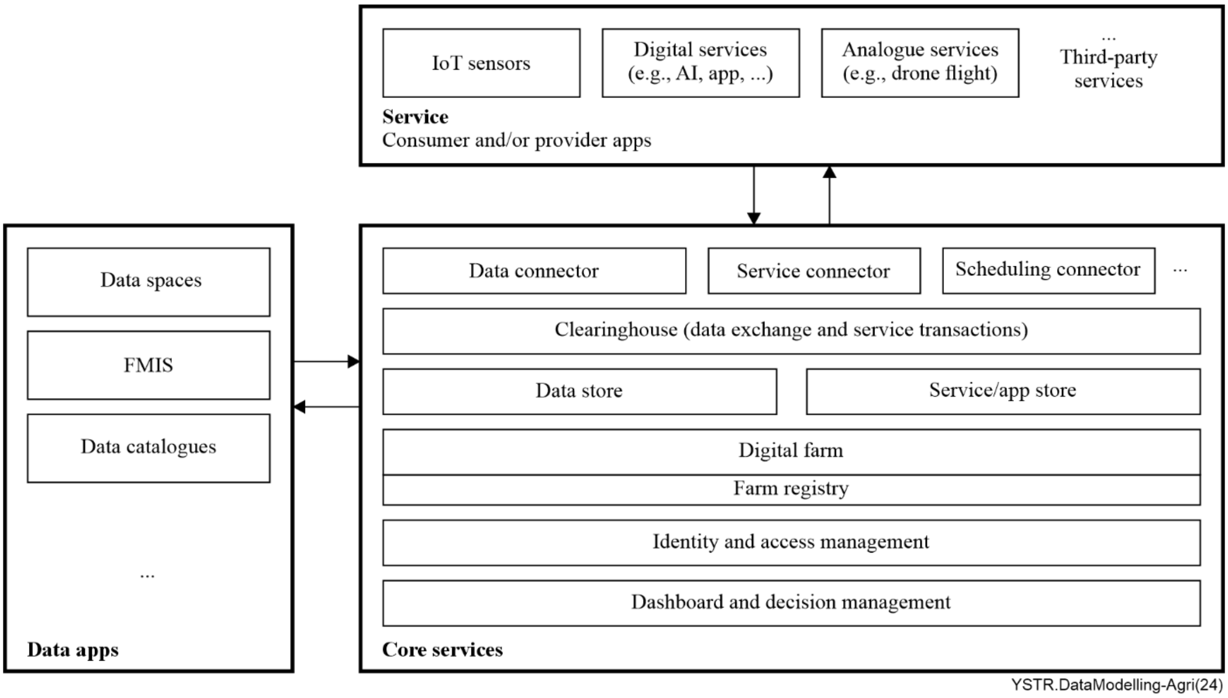


**Figure 3 – Overview of technical system architecture**

The extensible services and applications provide a wide range of functions that allow users to gain insights from the collected data. These services can be extended as needed and include tools for analysis, prediction models, and optimization algorithms.

The data apps serve as an interface to external applications and platforms relevant for data and sharing. A central use case is FMIS, from which a wealth of data can be extracted.

The architecture of the (AI) services is designed to run anywhere and adhere only to the specified protocols and semantics, thus covering edge use cases as well. A Python software development kit (SDK) facilitates the development of services, while other languages and frameworks can also be used.

The core services form the standardized environment, while extensions have no technological requirements and only need to adhere to a selection of international standards. This enables a flexible and scalable platform that meets the evolving needs of agriculture. The identified international standards are as listed in Table 1.

**Table 1 – List of complementary international standards**

| Source | Title of standard |
|---|---|
| Y.4800 | Requirements and functional architecture of an automatic location identification system for ubiquitous sensor network (USN) applications and services |
| Y.4804 | Multimedia information access triggered by tag-based identification – Identification scheme |
| Y.4805 | Identifier service requirements for the interoperability of smart city applications |
| Y.4806 | Security capabilities supporting safety of the Internet of things |
| Y.4810 | Requirements for data security of heterogeneous Internet of things devices |

**Table 1 – List of complementary international standards**

| Source | Title of standard |
|---|---|
| Y.4811 | Reference framework of converged service for identification and authentication for IoT devices in a decentralized environment |
| Y.3601 | Big data – Framework and requirements for data exchange |

## 10.3 Communication infrastructure

The architecture would employ a robust set of interfaces and protocols for seamless communication and interaction within its ecosystem. All functionalities and operations within this framework are exposed and controlled via a representational state transfer (REST) API, providing users with a standardized and accessible interface for managing and utilizing the platform's features. This RESTful API enables users to perform various tasks, such as querying data, initiating processes, and accessing platform functionalities, in a simple and intuitive manner [b-Elsevier].

Furthermore, the architecture would support the uploading of raw and result data directly to its storage, allowing users to efficiently manage and store large volumes of data with ease. This capability enables users to seamlessly integrate their data into the platform, facilitating data-driven decision-making and analysis. [b-NALAMKI]

When it comes to services, such as AI, data, IoT, or external service providers, the architecture would adopt a slightly different approach for communication. When a service is commissioned, the trigger is initiated via the REST API. The communication between the service and the architecture would occur through a message queuing telemetry transport (MQTT) broker. This setup enables real-time communication between services and the architectural framework, facilitating both complex interactions and simple task executions.

Figure 4 presents a system architecture for agricultural model deployment that emphasizes standardized communication protocols. While the implementation uses Docker containers and Python due to SDK requirements, the key architectural feature is the use of S3 storage and message brokers (like RabbitMQ) to ensure consistent communication between components. The system enables farmers to access AI models through a service catalogue, with numbered steps (1-5) indicating the flow from initial data selection to result storage. Though alternative communication pathways are possible, this streamlined approach maintains system simplicity while ensuring reliable data exchange.

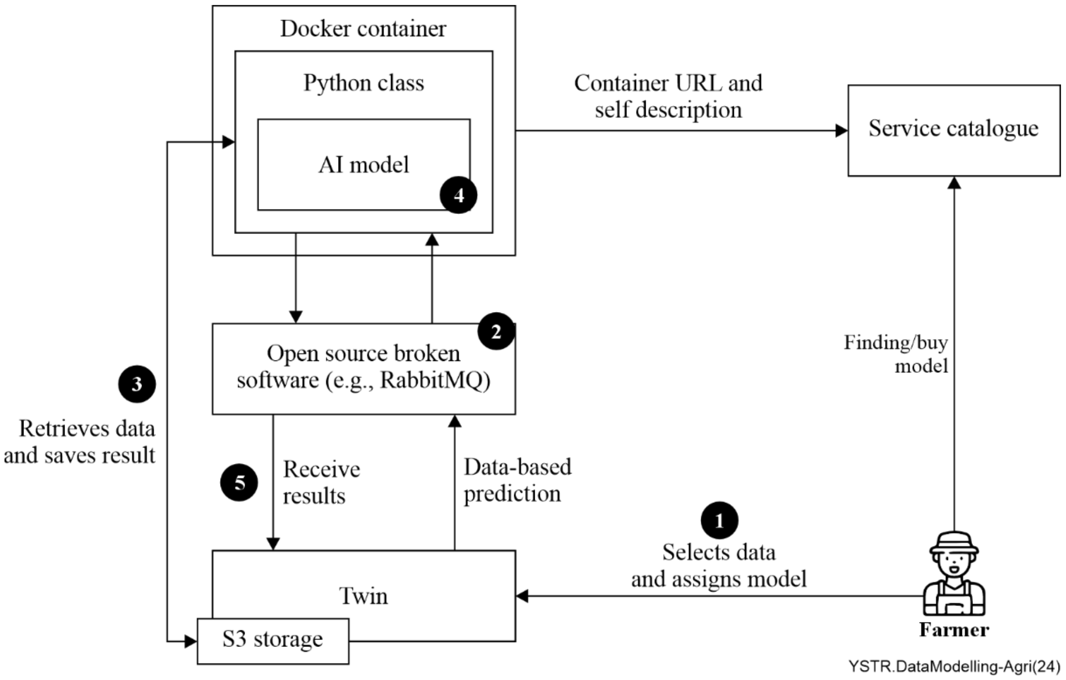


**Figure 4 – AI service container architecture**

In addition, if a service requires data from the platform for processing, it can retrieve this data directly via a storage sharing link. This streamlined communication mechanism ensures efficient data exchange between services and the core platform, minimizing latency and enhancing overall performance. The communication between services and core services follows a straightforward handshake protocol, ensuring smooth and reliable interaction between different components of the platform. This communication framework enables seamless integration and interoperability of services, empowering users to leverage a wide range of functionalities and capabilities for agricultural innovation and decision-making.

## 11 AI-based modelling for agriculture

In the realm of digital agriculture, AI modelling holds immense potential for revolutionizing farming practices, optimizing crop yields, and enhancing sustainability. By leveraging AI techniques, farmers can make data-driven decisions, predict outcomes and streamline agricultural operations. Here is how AI modelling can be applied in digital agriculture. First, AI modelling in digital agriculture involves collecting and analysing vast amounts of data from sources like sensors, satellite imagery, weather forecasts, and historical crop information. This data is crucial for understanding crop health, soil conditions, pest infestations, and other factors affecting agricultural productivity. Next, AI algorithms are used to process and interpret this data, uncovering hidden patterns, trends and insights that can guide farmers in making informed decisions. ML models, such as regression, classification, clustering, and deep learning, play a vital role in predicting outcomes like crop yields, disease outbreaks, and optimal planting times. Moreover, AI modelling enables precision farming techniques, where resources like water, fertilizers, and pesticides are applied more efficiently based on real-time data and predictive analytics. This not only maximizes crop yields but also minimizes environmental impact and reduces production costs.

Furthermore, AI-powered decision-support systems can provide personalized recommendations to farmers, helping them optimize planting schedules, crop rotations and pest management strategies. This proactive approach enhances productivity while reducing the risks associated with unpredictable weather patterns and market fluctuations. The main steps related to leveraging AI for modelling for agriculture include:

1) *Define the problem*: Clearly articulate the specific agricultural problem or task that you aim to solve using AI.
2) *Data collection:* Gather relevant data such as crop types, soil properties, weather conditions, and historical yield information.
3) *Data pre-processing:* Clean the data, handle missing values, normalize or standardize the features, and prepare the dataset for modelling.
4) *Feature selection:* Choose the most relevant features that will contribute to the accuracy of the model.
5) *Model selection:* Select an appropriate AI model such as ML algorithms (e.g., random forests and support vector machines) or deep learning models (e.g., neural networks).
6) *Model training:* Train the selected model on the pre-processed data to learn patterns and relationships in the dataset.
7) *Model evaluation:* Evaluate the model's performance using metrics like accuracy, precision, recall, F1 score, or others, to assess how well it generalizes to new data.
8) *Model optimization:* Fine-tune the model's parameters to improve its performance further if necessary.
9) *Deployment:* Implement the trained model into a user-friendly interface or system that can be used by farmers or agricultural experts.
10) *Monitoring and maintenance:* Continuously monitor the model's performance, retrain it with new data periodically, and make the necessary updates to ensure its reliability and accuracy over time.

Figure 5 demonstrates a practical implementation of an agricultural monitoring system, translating theoretical concepts into a tangible field solution. The system combines readily available components – standard soil sensors, a Raspberry Pi 3 model B, and cloud infrastructure – to collect, process, and analyse critical field data. Through this hands-on approach, farmers can access real-time soil parameters and receive data-driven recommendations through ML models, enabling practical decision-making in day-to-day farm operations.

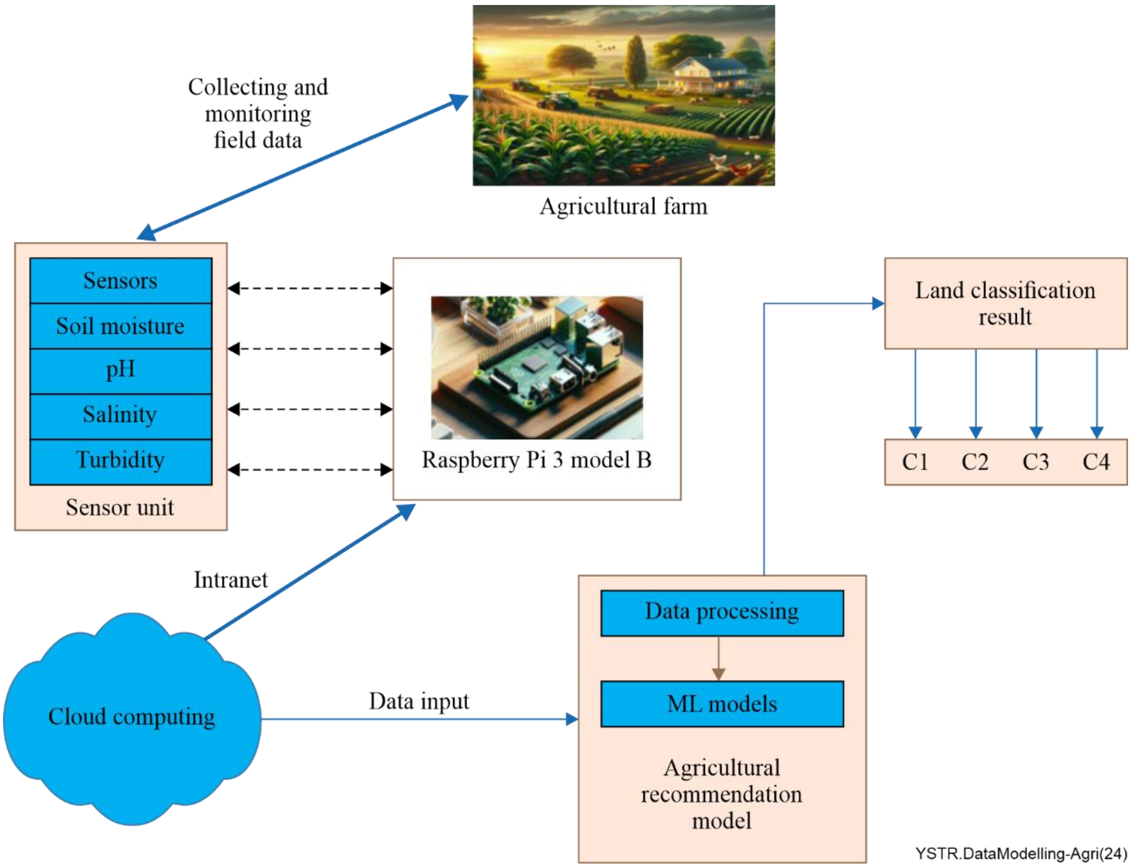


**Figure 5 – AI modelling for digital agriculture**

## 12 Conclusion

The fusion of data modelling techniques presents a transformative opportunity for the agricultural sector. By effectively collecting, processing and analysing vast amounts of agricultural data, farmers can make informed decisions that optimize crop production, resource utilization, and overall farm management practices.

Through advanced technologies such as sensors, drones, satellite imagery, and IoT devices, farmers can gather real-time data on soil conditions, weather patterns, crop health, and more. This data serves as the backbone for developing AI and ML models that can predict crop yields, identify potential risks, and guide precision farming practices.

The integration of data modelling not only enhances operational efficiency but also promotes sustainability by reducing waste, minimizing environmental impact, and improving resource management. Farmers can leverage predictive analytics to mitigate risks, adapt to changing conditions, and maximize productivity in a dynamic agricultural landscape.

As the world moves towards a more data-driven agricultural future, it is imperative for stakeholders to embrace and invest in data modelling technologies. By doing so, it is possible to unlock new insights, drive innovation, and foster a resilient agricultural ecosystem that meets the challenges of a rapidly evolving world.

Data modelling technologies, when integrated with traditional farming practices, are transforming agriculture, by enabling precise resource management, predictive decision-making, and sustainable farming methods that balance productivity with environmental stewardship.

# Bibliography


[b-ITU-T M.3080] Recommendation ITU-T M.3080 (2021), *Framework of artificial intelligence enhanced telecom operation and management (AITOM)*.

[b-ITU-T Y.2060] Recommendation ITU-T Y.2060 (2012), *Overview of the Internet of things*.

[b-Elsevier] Elsevier (2024), *A systematic review of open data in agriculture*.

[b-MDPI] MDPI (2022), *Data Type and Data Sources for Agricultural Big Data and Machine Learning*.

[b-NALAMKI] NALAMKI: https://nalamki.de/

[b-PeerJ] PeerJ, *Agricultural big data and methods and models for food security analysis – a mini-review*.