Ecological Informatics 81 (2024) 102613

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Reference architecture design for developing data management systems in smart farming

Ngakan Nyoman Kutha Krisnawijaya[a], Bedir Tekinerdogan[a,*], Cagatay Catal[b], Rik van der Tol[c]

[a] Information Technology Group, Wageningen University & Research, the Netherlands
[b] Department of Computer Science and Engineering, Qatar University, Qatar
[c] Farm Technology Engineering, Wageningen University & Research, the Netherlands

ARTICLE INFO

Keywords:
Domain analysis
Feature model
Reference architecture
Data management
Smart farming

ABSTRACT

The traditional data management systems prove inadequate to handle the volume, velocity, and variety of the data within farm business processes. Smart farming technologies offer advanced data management systems as a practical solution to these challenges. However, data is complex and originates from many sources; hence many aspects of data must be considered during the data management design of smart farming systems. This study proposes a reference architecture for data management in smart farming, developed through domain analysis and architecture modeling approaches. The domain analysis provides insights into the common and variant features and modules of the smart farming system, resulting in a blueprint representing family features across various smart farming domains. The effectiveness of the proposed reference architecture has been evaluated through two case studies, demonstrating its efficacy in designing data management systems for smart farming. The study found that the percentage of reused modules in the case studies, compared to the provided reference architecture, was 82.6%. The outcomes of this research will pave the way for further exploration in smart farming, particularly addressing data management issues within smart farming systems.

1. Introduction

The global food demand in 2050 is expected to increase significantly due to population growth and socio-economic factors, such as rising income, demographic structures, and urbanization (Food and Agriculture Organization of the United Nations (FAO), 2018). In order to feed the expanding global population, it is estimated that agriculture will need to produce around 50% more food by 2050 (Food and Agriculture Organization of the United Nations (FAO), 2020; Food and Agriculture Organization of the United Nations (FAO), 2017). However, due to several factors such as pressure on natural resources, lack of investment in agriculture, and technological gap, maintaining the pace of production increases may be challenging (Food and Agriculture Organization of the United Nations (FAO), 2017). Therefore, a profound transformation of the agriculture system is required to enhance agricultural productivity and sustainable food production (United Nations (UN), 2024).

One of the most promising approaches to maintaining or even increasing productivity is applying smart farming technologies. The smart farming system has been applied in different farming domains, such as crop (Bhat et al., 2023; Dong et al., 2016; López-Riquelme et al., 2017; Saranya and Nagarajan, 2020; Triantafyllou et al., 2019), livestock (Alonso et al., 2020; Kamilaris et al., 2018; Sant'Ana et al., 2022; Silva et al., 2014; Taneja et al., 2020; Wang et al., 2022), greenhouse (Subahi and Bouazza, 2020; Yang et al., 2017; Zamora-Izquierdo et al., 2019), and fish farming (Lee and Wang, 2020; Liu et al., 2023; Zhang et al., 2024). The system enhanced decision-making processes by assisting the farmers with invaluable information regarding their field status (Wolfert et al., 2014). Several data processes and advanced technologies are needed in a smart farming system in order to provide insightful information (Rutten et al., 2013). For instance, the Internet of Thing (IoT) devices are utilized in digital data acquisition to capture the actual field condition (Alonso et al., 2020; Köksal and Tekinerdogan, 2019; Zamora-Izquierdo et al., 2019) and then machine learning and, more recently, deep learning techniques are applied to analyze the generated data (Perakis et al., 2020; Swain et al., 2020). However, the traditional data management systems are inadequate to handle the volume, velocity, and variety of the generated data by the sensors and IoT devices (Wolfert et al., 2017). Hence the advanced data management

* Corresponding author.
E-mail addresses: ngakan.krisnawijaya@wur.nl (N.N.K. Krisnawijaya), bedir.tekinerdogan@wur.nl (B. Tekinerdogan), ccatal@qu.edu.qa (C. Catal), rik.vandertol@wur.nl (R. van der Tol).

https://doi.org/10.1016/j.ecoinf.2024.102613
Received 11 November 2023; Received in revised form 21 April 2024; Accepted 24 April 2024
Available online 27 April 2024

system is a vital element of a smart farming concept (Saiz-Rubio and Rovira-Más, 2020).

In smart farming, generally, data management can be divided into four elements: (1) data acquisition and storage, (2) data preprocessing, (3) data analysis, and (4) data visualization (Ouafiq et al., 2019). Data acquisition is the process of acquiring data from the field and then storing them in digital storage for future analysis purposes (Rawat and Yadav, 2021). The data are collected from various sources, such as sensors, electronic resources (e.g., text, video, image), and public resources, and as such, the data are generated in different forms (i.e., structured, semi-structured, and unstructured). The next element is data preprocessing, which can be defined as enhancing the quality of data that impacts the analytical processing (Meena and Sujatha, 2019). In this stage, unnecessary data (e.g., noise, missing value, inconsistent data, duplicate data) are removed from the dataset before applying data analysis techniques for knowledge discovery. After preprocessing, the data are then analyzed in order to produce valuable information for the users to make better decisions. Data analysis is the process of discovering hidden information from data. Data visualization helps visually present the data by using various graphs or charts for decision-making (Hashem et al., 2015). Rawat and Yadav (Rawat and Yadav, 2021) stated that visual report is more efficient than text documents for information seekers to gain informative knowledge. Therefore, these data management advancements enable farmers to manage their farms efficiently and effectively (Meeradevi et al., 2019).

Due to the complexity of the data, for instance, the data originated from various sources. Consequently, many aspects of data (e.g., type, structure, source, etc.) must be considered during the design of the data management of smart farming systems (Giray and Catal, 2021). Therefore, a proper reference architecture is needed to build a system that can utilize these technologies and satisfy both functional and non-functional requirements. Functional requirements refer to the specific functionalities or features that the system must have to meet the needs of its users. These requirements are derived from user interactions and describe the system's intended behaviour. Non-functional requirements, on the other hand, are not directly related to the specific functionalities of the system, but rather define the qualities and characteristics that the system should possess. These requirements address aspects such as performance, reliability, usability, security, scalability, maintainability, and availability (Robertson and Robertson, 2012). In addition, an appropriate reference architecture is a system design that follows the proper software architecture design guidelines that help define the gross level structures of the system. Therefore, the proper architecture is key to understanding the whole system, analyzing the flow of the data in the system, and helping further the system's development (Tummers et al., 2021).

A reference architecture is a standardized architectural blueprint or model that provides guidelines and best practices for designing and implementing a specific type of system or application. It serves as a template or guide for developers, architects, and stakeholders involved in building similar systems. A reference architecture is a reference model that can be represented by one or more architectural views (Cervantes and Kazman, 2016). The valuable reference architecture has several criteria (Tummers et al., 2021): (1) The design should be acceptable, understandable, and accessible to all stakeholders of the organization, (2) The critical aspects of the domain should be addressed, and (3) The design should be up-to-date, maintainable, and valuable for the organization. From the reference architecture, an application architecture can be derived. Application architecture is the software architecture of a particular application that is presented using several architectural views (Tummers et al., 2021). The application architecture differs from one case to another since the design depends on the needs of the stakeholders (Köksal and Tekinerdogan, 2019).

Previous studies have discussed related reference architecture studies in the literature. For instance, Taneja et al. (Taneja et al., 2019) and Righi et al. (Righi et al., 2020) presented a client-server architecture for a smart farming system and did not describe a reference architecture in their study. Giray and Catal (Giray and Catal, 2021) presented a reference architecture for smart farming and sustainable agriculture. They focused on designing data management due to the complexity of data management operations in the smart farming system. Their study is based on design science research (DSR), like domain scoping, domain modeling, and reference architecture stages, to establish the data management reference architecture. To evaluate the reference architecture, they used three different case studies using the recent literature on sustainable agriculture. In contrast to this article by Giray and Catal (Giray and Catal, 2021), in this study, the reference architecture is illustrated using two case studies based on the actual project to develop smart farming in Indonesia, one for smart dairy farming and the other for smart fish farming. In addition, the main focus was the sustainability aspects of agriculture in their study.

Köksal and Tekinerdogan (Köksal and Tekinerdogan, 2019) designed a reference architecture for IoT-based Farm Management Information Systems (FMISs). They proposed the architecture that includes main features such as data acquisition, processing, monitoring, planning, decision-making, and documenting. They followed an architecture design method using decomposition, layered, and deployment views to present the reference architecture. The design showed that the approach is effective, and their reference architecture can be used to derive a concrete FMIS. Their study, however, did not focus on data management design. Tummers et al. (Tummers et al., 2021) and Kruize et al. (Kruize et al., 2016) described a reference architecture for FMISs that is part of the smart farming system. Tummers et al. (Tummers et al., 2021) used context and decomposition views to represent the proposed reference architecture. Their study also presented a feature model based on the identified features in the existing FMISs in the literature. They evaluated the reference architecture using three case studies. In comparison, Kruize et al. (Kruize et al., 2016) proposed the reference architecture for farm software ecosystems, their scope is slightly wider than FMISs. However, both studies did not focus on data management design for smart farming systems. Santana et al (Santana et al., 2014) developed an automated system to identify bee species since they have a significant role in agriculture. They also established a reference process for bee classification based on wing images. The reference process is beneficial to help other researchers or stakeholders to understand this complex process by following the provided steps and experiments.

As such, the contribution of this study is to enhance the current smart farming reference architecture, particularly the data management infrastructure, by following the software architecture guidelines and using a multi-case study protocol to evaluate the proposed reference architecture. To the best of our knowledge, a few studies exist (Köksal and Tekinerdogan, 2019; Tummers et al., 2021) that discuss the proper architecture of the smart farming system. Specifically, a complete architectural view of a data management system in smart farming is still a relatively new endeavor (Giray and Catal, 2021). The main objective of this research is thus to develop a reference architecture for data management in the smart farming system. Based on this objective, the following research questions were formulated:

- RQ1. What are the common and variant features of data management in smart farming?
- RQ2. How to design a proper data management reference architecture for smart farming?
- RQ3. How to develop application architecture based on the designed reference architecture?
- RQ4. How effective is the designed reference architecture?

As a result of this study, we propose a data management reference architecture for smart farming. We will provide the approach for designing the general reference architecture for data management in smart farming. This study follows the software architecture guidelines and main steps (Giray and Catal, 2021; Köksal and Tekinerdogan, 2019; Kruize et al., 2016; Tummers et al., 2021), such as domain scoping, domain and features modeling, and using architecture viewpoints to present the proposed architectural design as shown in Fig. 1.

The design shows the generic reference architecture for smart farming, which is evaluated by the following two case studies: dairy and fish farming in Indonesian context within the project called Smart Indonesian Agriculture (Smart-In-Ag) (Wageningen University and Research (WUR), 2024). For two specific domains, Dairy Farming and Fish Farming, we will show how to derive the domain architecture. The domain architectures are more specific than the generic architecture but are generic for the specific concrete applications of the corresponding domain.

The reference architecture presented in this study reflects the overall software architecture of the Smart-In-Ag project. The contributions of this study are as follows:

(1) A systematic domain-driven architecture design method is applied to provide insights into the common and variant features of data management in a smart farming system.
(2) A novel reference architecture has been proposed for developing data management in the smart farming system.
(3) The proposed reference architecture is validated using two case studies in a project.
(4) The generic data management reference architecture is presented, which can then be used for different application domains.

2. Research methodology

As stated in the Introduction section, the main objective of this research is to develop a reference architecture for data management in the smart farming system. The following research questions were defined:

- RQ1. What are the common and variant features of data management in smart farming?
- RQ2. How to design a proper data management reference architecture for smart farming?
- RQ3. How to develop application architecture based on the designed reference architecture?
- RQ4. How effective is the designed reference architecture?

The first research question aims to perform domain analysis to support the architectural design process. Domain analysis is a systematic approach to deriving required knowledge in a particular domain. The results of this question are provided in Section 3. The second question aims to design a feasible reference architecture for data management in a smart farming context. This study follows a formal approach and guidelines in order to derive a well-established architecture design. The results of this question are discussed in Section 4. Section 5 discusses the application model to answer the third research question, and then the proposed design is assessed using multi-case study approaches to answer research question 4.

Fig. 2 presents the steps that we have followed in conducting this study. In the first step, systematic literature research (SLR) was done in our previous study (Ardagna et al., 2016; Ayaz et al., 2019; Krisnawijaya et al., 2022; Villa-Henriksen et al., 2020). In the SLR study, the state-of-the-art of data analytics platforms for the agricultural system was identified. The common features, stakeholders, adopted technologies, and architectural design were the results of the SLR. The obtained information from the SLR study was used as input to support the domain modeling.

In the second step, domain scoping, we define the overall scope of the systems that we focus on. This is followed by a domain modeling activity in which the key concepts of the selected domain are identified and modeled. Here we use feature modeling for modeling the common and variants features of data management in smart farming systems. The feature model can support the design processes and give system developers insights into selecting the features they wish in the system. In the next step, the reference architecture is designed using a set of viewpoints. Finally, we will perform a multi-case study to evaluate the overall approach. These steps will be described in detail in the following sections.

3. Domain analysis

In this research, the domain analysis process was applied to understand the data management in the smart farming of interest. Domain analysis is a systematic activity applied to derive and store required knowledge to support the architectural design process (Köksal and Tekinerdogan, 2019; Salma et al., 2017; van Geest et al., 2021). The

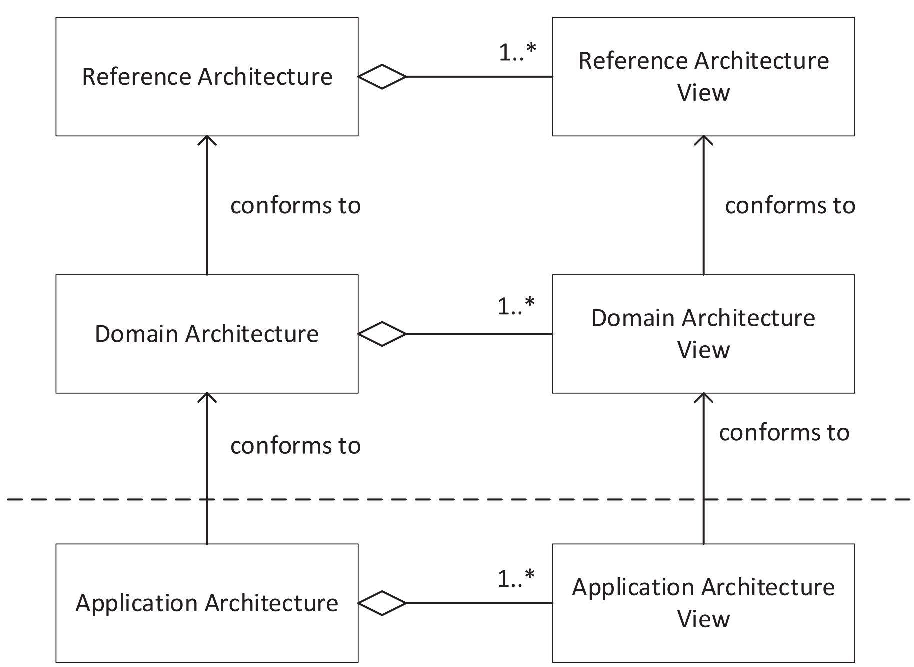

Fig. 1. Methods used for deriving the application architecture. Adapted from Tummers et al. (Tummers et al., *2021*).

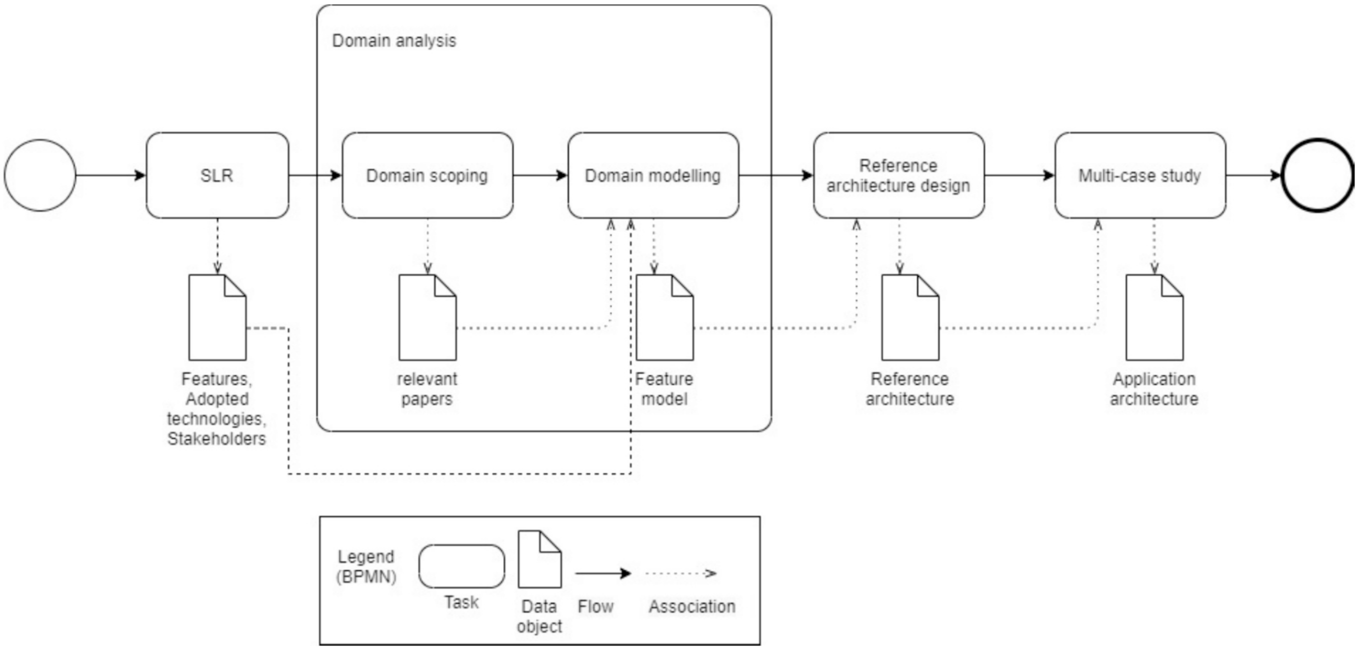

Fig. 2. The adopted research method for deriving the reference architecture.

definition of the term “domain” was adopted from Salma et al. (Salma et al., 2017), which is defined as follows: “*Domain is an area of knowledge or activity characterized by a set of concepts and terminology understood by practitioners in that area.*” In general, domain analysis involves two main activities (see Fig. 3): domain scoping and domain modeling, which is explained in the next subsections.

3.1. Domain scoping

In the domain scoping process, the domains of interest, stakeholders, and their goals were identified. This study followed the main steps presented in Fig. 3 to apply the domain scope. First, the domain is defined, and for this research, the domain is data management in smart farming. Understanding the recent factors for deriving knowledge to establish the reference architecture is crucial (Giray and Catal, 2021). The agricultural domain, in our case, involves both traditional and

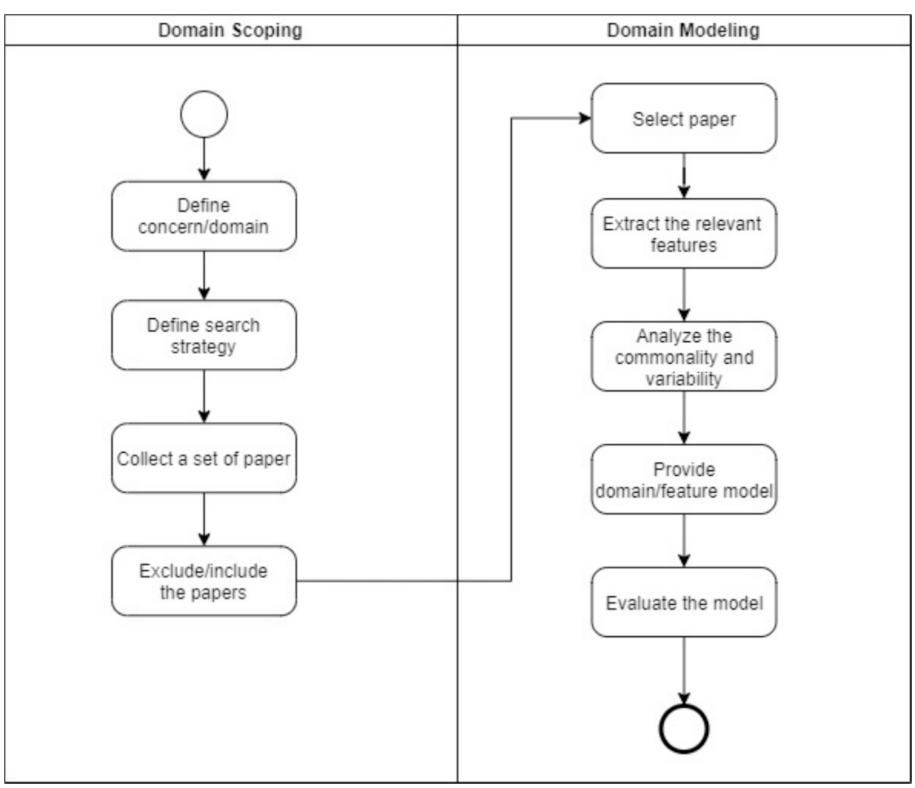

Fig. 3. Domain analysis process.

modern farming to get insight into the daily practices of its processes. Therefore, the phrase “farming”, “agriculture”,“ smart farming”, and “precision agriculture” are our main phrases to combine with the “reference architecture”. The next step is identifying the recent trends in data management. According to Giray and Catal (Giray and Catal, 2021), big data, machine learning, and data lake are emerging technologies that evolved from traditional data management in recent years. Based on their research, this study adopts the phrases used to represent data management and combines them with the other phrases explained earlier. This search strategy has been applied in the Scopus database to collect the papers. To assure sufficient coverage, the manual search method was also conducted.

The exclusion criteria were applied to the obtained papers to get only the relevant papers for designing the feature model. The papers that are not written in English, do not provide a complete text, and are duplicate publications were excluded from the analysis. The relevant papers were used as input for deriving a domain model and the reference architecture. The list of the selected papers is shown in Table 1.

3.2. Feature model

One of the common approaches to representing domain knowledge is a feature model. Feature modeling is one of the approaches in a domain model to show the familiar concepts of the identified domain knowledge by using a commonality and variability analysis approach (Tekinerdogan et al., 2005; van Geest et al., 2021). Thus, a feature model can be used to show common and variant features of a system or concept. This model is constructed as a tree, where the root element represents the system or concept, and its nodes represent the particular system’s features (Salma et al., 2017). Each feature has sub-features and has the following specific types to show the dependencies among features: mandatory, optional, or alternative. The analysis is explained in the following subsections.

3.2.1. Feature model for smart farming

Firstly, this study identified the implementation aspects of smart farming in agriculture. Smart farming is the concept that applies digital data to provide precise information to support decision-making for the primary farming process. In general, the feature model for the smart farming system is presented in Fig. 4. Various common features of a smart farming system are presented in this model. Several top-level features of the smart farming system have been identified, such as categories, adopted technologies, farm activities, goals, domains, field data, and a data management system (Saiz-Rubio and Rovira-Más, 2020).

3.2.1.1. Categories. According to Balafoutis et al., (Balafoutis et al., 2020) smart farming can be categorized into the following three interrelated main categories: farm management information system (FMIS), precision agriculture (PA), and agricultural automation and robotics. FMISs are a set of software systems to assist farmers in performing various agricultural tasks. The FMISs are used to collect, process, store and disseminate data to carry out operations and functions of the farm. Meanwhile, PA refers to optimizing the farm management input by using several emerging technologies, such as Unmanned Aerial Vehicles (UAVs) for aerial data, sensors to get ground data, and a decision support system (DSS) to optimize farming decision-making. These data are used to observe and measure the various field parameters to obtain insights regarding the precise time and moments to do specific tasks. The third

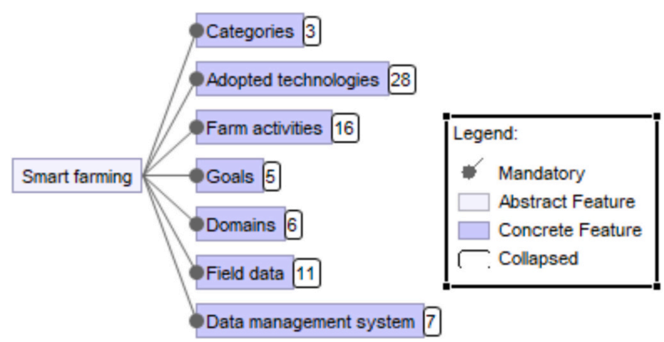

Fig. 4. Feature diagram of data management in agricultural systems.

Table 1
Papers to derive the data management reference architecture in smart farming.

J. Tummers, A. Kassahun, and B. Tekinerdogan, “Reference architecture design for farm management information systems: A multi-case study approach,” *Precision Agriculture,* vol. 22, pp. 22–50, 2021, doi: https://doi.org/10.1007/s11119-020-09728-0. (Tummers et al., 2021)
Ö. Köksal and B. Tekinerdogan, “Architecture design approach for IoT-based farm management information systems,” *Precision Agriculture,* vol. 20, pp. 926–958, 2019, doi: https://doi.org/10.1007/s11119-018-09624-8. (Köksal and Tekinerdogan, 2019)
G. Giray and C. Catal, “Design of a Data Management Reference Architecture for Sustainable Agriculture,” *Sustainability,* vol. 13, pp. 1–17, 2021, doi: https://doi.org/10.3390/su13137309. (Giray and Catal, 2021)
J. W. Kruize, J. Wolfert, H. Scholten, C. N. Verdouw, A. Kassahun, and A. J. M. Beulens, “A reference architecture for Farm Software Ecosystems,” *Computers and Electronics in Agriculture,* vol. 125, pp. 12–28, 2016, doi: https://doi.org/10.1016/j.compag.2016.04.011. (Kruize et al., 2016)
A. Triantafyllou, P. Sarigiannidis, and S. Bibi, “Precision Agriculture: A Remote Sensing Monitoring System Architecture,” *Information,* vol. 10, no. 11, 2019, doi: https://doi.org/10.3390/info10110348. (Triantafyllou et al., 2019)
J. A. 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: https://doi.org/10.1016/j.agwat.2016.10.020. (López-Riquelme et al., 2017)
V. Saiz-Rubio and F. Rovira-Más, “From Smart Farming towards Agriculture 5.0: A Review on Crop Data Management,” *Agronomy,* vol. 10, pp. 1–21, 2020, doi: https://doi.org/10.3390/agronomy10020207. (Saiz-Rubio and Rovira-Más, 2020)
D. J. McConnell and J. L. Dillon, “Farm Management for Asia: A Systems Approach,” in *FAO Farm Systems Management Series*. Rome: Food and Agriculture Organization of the United Nations, 1997. (McConnell and Dillon, 1997)
K. Demestichas and E. Daskalakis, “Data Lifecycle Management in Precision Agriculture Supported by Information and Communication Technology,” *Agronomy,* vol. 10, pp. 1–21, 2020, doi: https://doi.org/10.3390/agronomy10111648. (Demestichas and Daskalakis, 2020)
S. Wolfert, L. Ge, C. Verdouw, and M.-J. Bogaardt, “Big Data in Smart Farming – A review,” *Agricultural Systems,* vol. 153, pp. 69–80, 2017, doi: https://doi.org/10.1016/j.agsy.2017.01.023. (Wolfert et al., 2017)
M. A. Zamora-Izquierdo, J. Santa, J. A. Martínez, V. Martínez, and A. F. Skarmeta, “Smart farming IoT platform based on edge and cloud computing,” *Biosystems Engineering,* vol. 177, pp. 4–17, 2019, doi: https://doi.org/10.1016/j.biosystemseng.2018.10.014. (Zamora-Izquierdo et al., 2019)
R. S. Alonso, I. Sittón-Candanedo, Ó. García, J. Prieto, and S. Rodríguez-González, “An intelligent Edge-IoT platform for monitoring livestock and crops in a dairy farming scenario,” Ad Hoc *Networks,* vol. 98, 2020, doi: https://doi.org/10.1016/j.adhoc.2019.102047. (Alonso et al., 2020)
X. Yang, S. Zhang, J. Liu, Q. Gao, S. Dong, and C. Zhou, “Deep learning for smart fish farming: applications, opportunities and challenges,” *Reviews in Aquaculture,* vol. 13, pp. 66–90, 2021, doi: https://doi.org/10.1111/raq.12464. (Yang et al., 2021)
M. Taneja, N. Jalodia, J. Byabazaire, A. Davy, and C. Olariu, “SmartHerd management: A microservices-based fog computing–assisted IoT platform towards data-driven smart dairy farming,” *Software: Practice and Experience,* vol. 49, pp. 1055–1078, 2019, doi: https://doi.org/10.1002/spe.2704. (Taneja et al., 2019)
R. d. R. Righi, G. Goldschmidt, R. Kunst, C. Deon, and C. A. d. Costa, “Towards combining data prediction and internet of things to manage milk production on dairy cows,” *Computers and Electronics in Agriculture,* vol. 169, pp. 1–13, 2020, doi: https://doi.org/10.1016/j.compag.2019.105156. (Righi et al., 2020)

category is autonomous machines or robots applied in agriculture. These reacting technologies are interconnected to cover the automatic control of all agricultural production levels by using machine learning, computer vision or artificial intelligence algorithms (Balafoutis et al., 2020). The system designers are able to choose one of these categories or combine these categories when implementing smart farming. This feature diagram is shown in Fig. 5.

3.2.1.2. Adopted technologies. In this research, the adopted technologies were divided into the following five sub-features: data analytics, sensing technologies, computing infrastructures, communication technologies, user interface, and hardware systems. All these features are mandatory features in a smart farming system, and they are integrated into one another (Krisnawijaya et al., 2022). In data analytics features, machine learning, deep learning, statistical-based model, mathematical model and geospatial analysis are commonly used to analyze the data. The system designers can choose one or all of these approaches in their system. Sensors are mandatory tools in sensing technologies since they can be used almost in all agricultural domains as data sources. Meanwhile, UAVs, Unmanned Ground Vehicles (UGVs), and Geographic Information Systems (GISs) are optionally utilized in a particular agricultural domain. For instance, when the smart system needs spatial, location or area data, the designer can simply use UAV and GIS as data sources.

In a smart farming system, the architects should also consider the system infrastructures to develop and apply them. The cloud-based and standalone systems are the common infrastructure found in the current studies. However, some studies also implemented the hybrid system by using edge computing or blockchain to improve the system's computing performance. Furthermore, proper communication in computing infrastructure should be determined when developing a smart system since it also affects the system's performance. In this study, it is found that several communication technologies are used in smart farming systems, such as Wireless Sensor Networks (WSNs), Radio Frequency Identification (RFID), Near Field Communication (NFC), Bluetooth, and cellular network. They are mandatory in the system; however, the architects can choose the suitable ones.

User interface applications can be used to disseminate and visualize the data for the stakeholders, and it has been found that there are two application types: Web and mobile applications. A web application and mobile application are required in the smart farming system. In order to send several commands from the central processing unit to the actuators, Internet of Things (IoT) technology is required. IoT can also be utilized to control several automation systems or robots to do specific tasks. In this study, both IoT and robotics are categorized as hardware features in the smart farming system. Fig. 6 shows adopted technologies in smart farming.

3.2.1.3. Farm activities. In the farm system, farming practices mean a collection of production methods that are applied to produce agricultural products (Corporation, 2020). Ploughing, planting, fertilization, irrigation, pest inspection, and growing the seed are examples of the crop's daily agricultural activities (McConnell and Dillon, 1997). In animal farming, such as dairy, fish, and poultry, it is common to see the producer raising their livestock, disease inspection, checking the feed, and managing the waste. According to Wolfert et al. (Wolfert et al., 2017), these activities can be divided into primary and supporting farm processes. The primary activities are those involved in product creation. Furthermore, two daily sub-activities for controlling activity are monitoring and recording the field situation. However, even though the farmers have already implemented smart systems, for some reason, they still prefer to do several practices in traditional approaches (Mourik et al., 2021). Fig. 7 presents the feature diagram of farming practices.

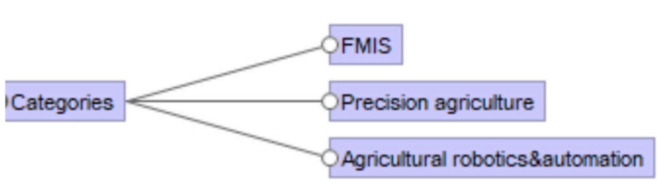

Fig. 5. Feature diagram of categories.

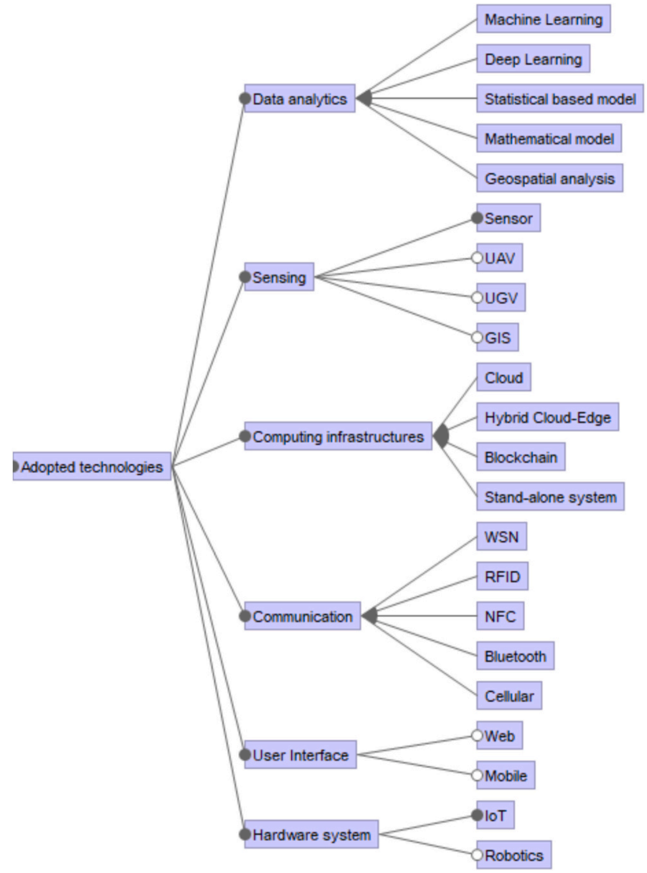

Fig. 6. Feature diagram of adopted technologies.

3.2.1.4. Goals. The results from our investigation of collected papers show five common goals of the smart farming system: productivity improvement, cost reduction, resource efficiency, prevention of diseases, and risk management. Fig. 8 shows the feature diagram of goals.

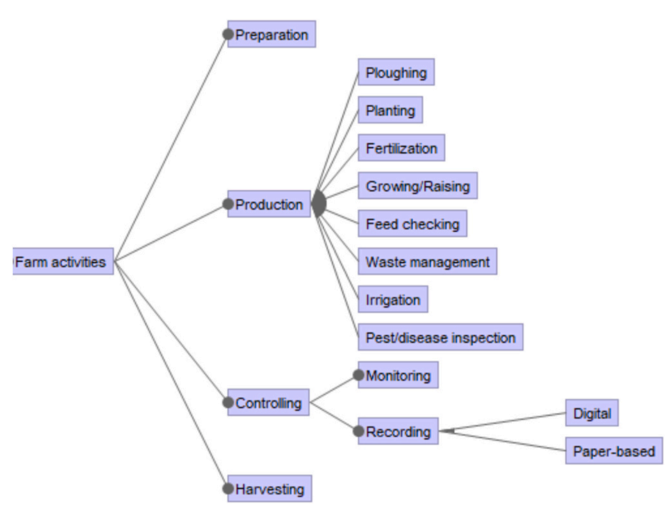

Fig. 7. Feature diagram of farm activities.

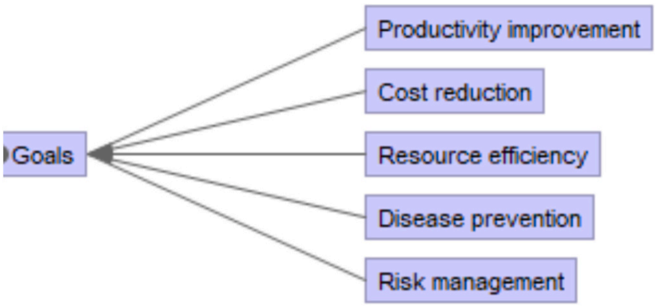

Fig. 8. Feature diagram of goals.

3.2.1.5. Domains. The domain feature in this research includes all of the mentioned agricultural domains in the literature that implemented smart farming as their management solution. In this feature, the domains are dairy, animal, fishery, arable, horticulture, and greenhouse. The feature is shown in Fig. 9.

3.2.1.6. Field data. The crucial feature in smart farming is field data since smart systems' decision-making relies on the actual data derived from the field (Collado et al., 2019; Krisnawijaya et al., 2022; Saiz-Rubio and Rovira-Más, 2020). The field data are retrieved from parameters measured from the crop, animal, soil, and environment. These measurements are based on the agricultural domains and the purposes of the system. The field data feature is presented in Fig. 10.

3.2.1.7. Data management system. The advanced data management lifecycle model usually consists of several activities, such as data acquisition, storage, processing, monitoring and reporting, and visualization (Pääkkönen and Pakkala, 2015). Another feature commonly found in the data management system in smart farming is data security. The data can be collected, processed, and analyzed to turn these features into valuable information for farmers. Therefore, data management can help farmers manage their farm operations better since they can make decisions tailored to their farms' specific needs. The features diagram of the data management system is presented in Fig. 11 and is discussed in the next section.

3.2.2. Feature model of data management

In this section, the common features and sub-features of the data management system are discussed. Both traditional and advanced data management systems, such as big data, were observed in this study. Fig. 12 presents the top-level features of the data management system in the context of smart farming. In addition, Fig. 13 shows the relationship

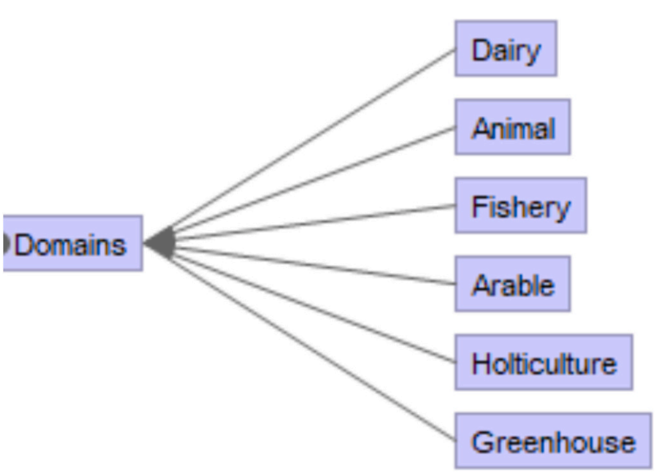

Fig. 9. Feature diagram of domains.

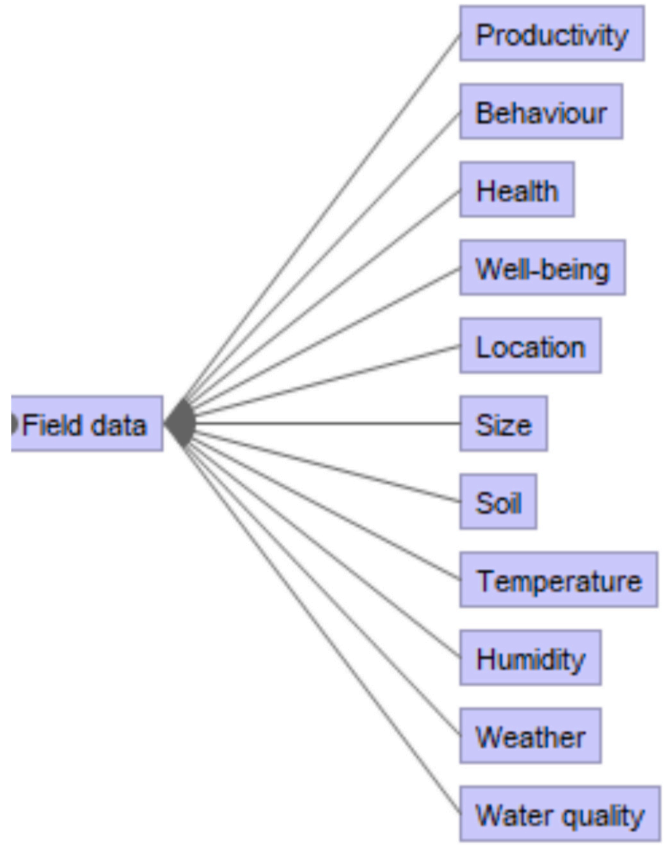

Fig. 10. Feature diagram of field data.

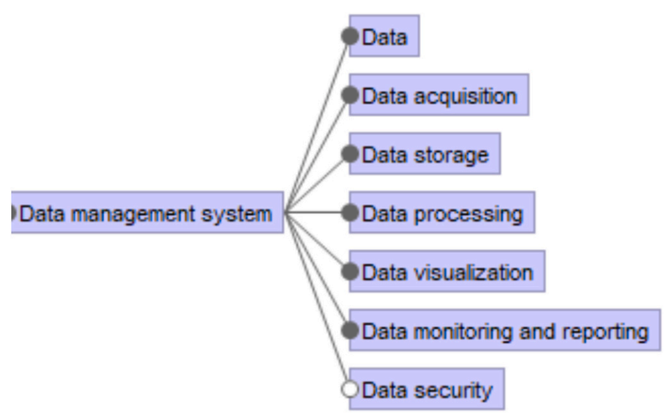

Fig. 11. Feature diagram of the data management system.

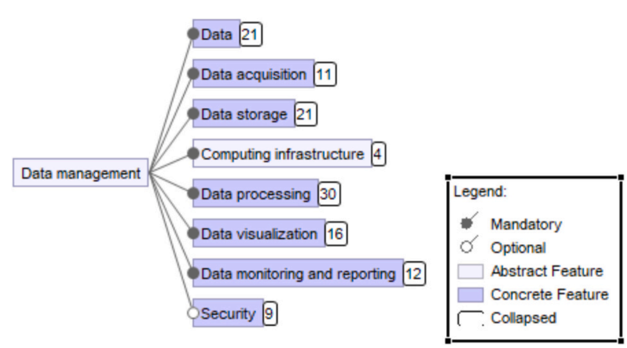

Fig. 12. The detailed features of a smart farming data management system.

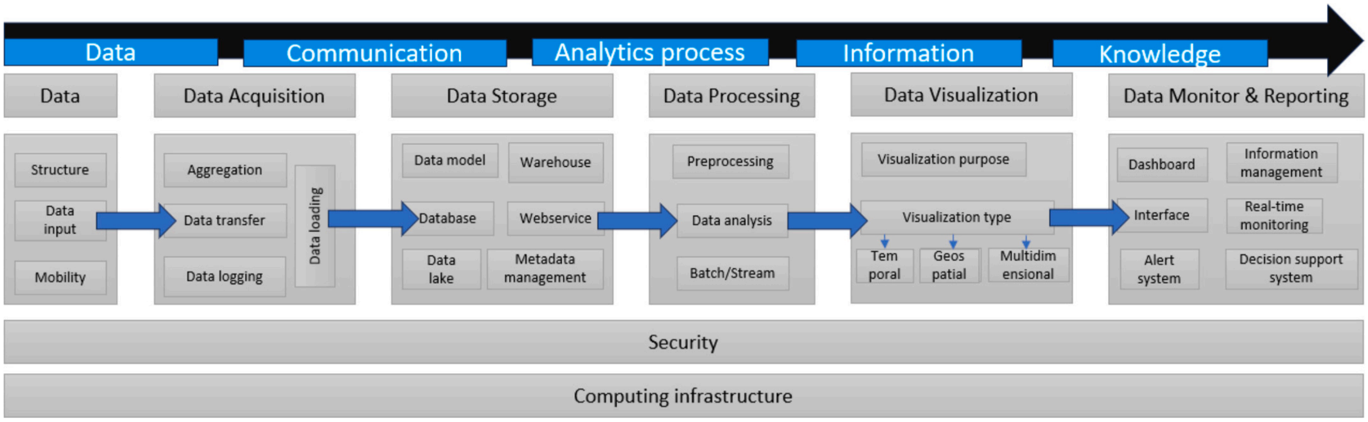

Fig. 13. The relationship between features of a smart farming data management system.

between features for data management systems. A data value chain starts from generating raw data and then transfers and loads into data storage. Web services can bridge data storage and processing features after the data is stored in a database system. The data can then be processed and analyzed in these features and transformed the data into valuable information. The obtained information is disseminated through data visualization features. Mostly, data analytics resulted in valuable information but also provided knowledge for decision-making processes. The features include an alert system, real-time monitoring, and information management.

3.2.2.1. Data. Traditionally, the farmers have applied handwritten notes to evaluate their farm operations. With the development of technology in modern agriculture, data can be generated from several sources and has become a mandatory feature of the agricultural system (Fulton and Port, 2018). Salma et al. (Salma et al., 2017) explained that this feature could describe data usage, state, and representation. In a smart farming system, there are several sub-features of the data feature, such as mobility, structure, and data input. Mobility of the data, which is batch and stream, can affect the system's data analysis and processing operations (Salma et al., 2017). Therefore, this sub-feature should be considered while designing a smart system. The next sub-feature is the data structure. There are three types of data structures based on the data sources: structured, semi-structured, and unstructured. This feature highly influences the development process of data storage, processing, and analysis (Salma et al., 2017).

Furthermore, the data input defines the process of how data can act as an input for the smart system. In general, a system can be accessed by other systems or people depending on the status of the data. For instance, the upon-requested data can be accessed only by asking permission from the data owner. On the other hand, everyone could have permission to get data from public sources. Therefore, the data access type is a part of the data input feature. Another sub-feature of data input is a modality to check data form, whether textual, visual, or audio. As mentioned, the data sources are also essential to explain the generated data and how to handle them. Data generation can be divided into agronomic, machine, production (Fulton and Port, 2018), and predictive data. The feature diagram of the data feature is shown in Fig. 14.

3.2.2.2. Data acquisition. This feature is responsible for generating new data or collecting the existing ones (Demestichas and Daskalakis, 2020). The data acquisition features consist of data loading, data logging, data aggregation, and data transfer. Data aggregation has the following sub-features: data fusion and data integration. Data fusion refers to concatenating two or more representations of two identical objects into single, clear, and consistent ones (Demestichas and Daskalakis, 2020). Data

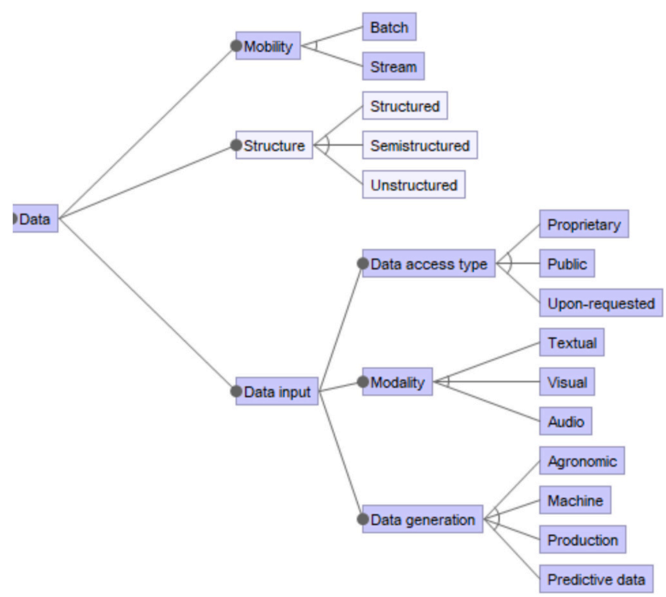

Fig. 14. Feature diagram of data.

transfer includes the following sub-features: Wireless Sensor Network (WSN), Bluetooth, Radio Frequency Identification (RFID) and Cellular network (see Fig. 15).

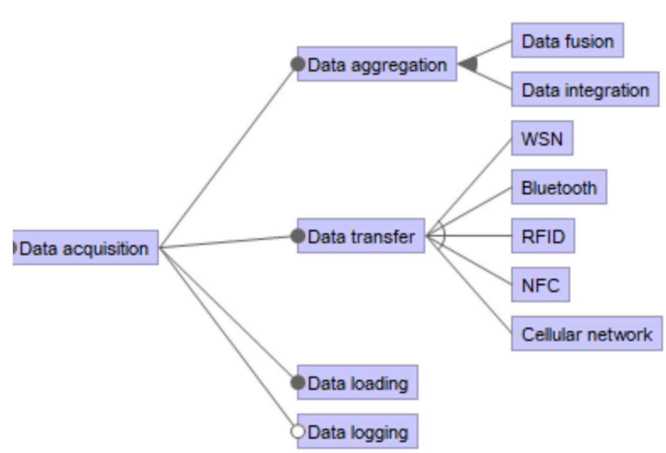

Fig. 15. Feature diagram of data acquisition.

3.2.2.3. Data storage. Fig. 16 presents a data storage feature that is vital in the smart farming system since it is a place to store all generated data needed by the actors involved in the system to share valuable knowledge and experiences (Demestichas and Daskalakis, 2020). It comprises seven sub-features: a data model, database, web services, data warehouse, data lake, and metadata management. The data model provides a framework for all stages of data to store and process according to business needs. This study considered both traditional data storage features (e.g., data warehouse, metadata management) and advanced big data system features, such as data lake, graph data model, and NoSQL database.

3.2.2.4. Computing infrastructure. Considering the system's infrastructure in the smart farming context is crucial since it affects scalability, cost, and performance. The computing infrastructure in this study has four widely used system infrastructure models for developing a smart agricultural system: cloud computing, hybrid cloud-edge computing, blockchain system, and standalone system (dedicated server). Fig. 17 shows the feature diagram of the computing infrastructure.

3.2.2.5. Data processing. In data processing features (see Fig. 18), data preprocessing aims to prepare and facilitate data processing operations (Salma et al., 2017; Yang et al., 2021). The data is cleaned, transformed, or compressed to be ready for the analysis stage. Data cleansing is a mandatory feature in data preprocessing, while other features are optional and depend on the needs or purposes of data analytics. Furthermore, in an agricultural system, it is common to see various types of analyzing the data, such as descriptive, diagnostic, predictive, or prescriptive analytics. Hence, the system uses advanced analytics techniques, namely geospatial analysis, mathematical model, statistical analysis, deep learning, and machine learning, depending on analytics types or purposes and users' needs. The machine learning method has its own types to analyze the data and specific tasks to generate the analytics results. Finally, as mentioned earlier in the data feature (see Fig. 14), there are two types of data processing: batch and stream processing.

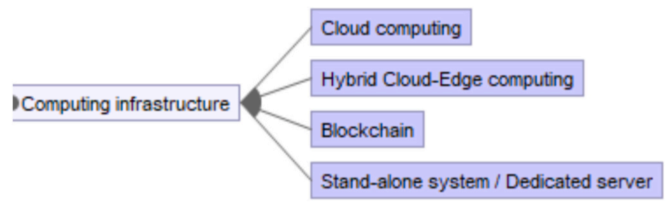

Fig. 17. Feature diagram of computing infrastructure.

3.2.2.6. Data visualization. This feature provides the way how the information is presented to the end-users. There are two common purposes of developing visualization systems in the smart system to help users identify the trend of data or the pattern of data. The data visualization types were also considered to deliver insightful information to users (Krisnawijaya et al., 2022). There are three types of visualization based on the generated information from the system, such as temporal information (represented by using a line graph or scatter plot), multi-dimensional information (using a pie chart, Venn diagram, bar graph, or histogram), and geospatial information (using heat or flow map). Fig. 19 shows the features diagram of the data visualization.

3.2.2.7. Data monitoring and reporting. The mandatory features in monitoring and reporting features are interfaces, dashboards, and information management. The interface provides the interaction of the smart system with the users and other systems (Salma et al., 2017). Furthermore, the information is delivered to the end-users through the dashboard and information management report. Real-time monitoring, decision support system, and alert system are optional in data monitoring and reporting features, as shown in Fig. 20.

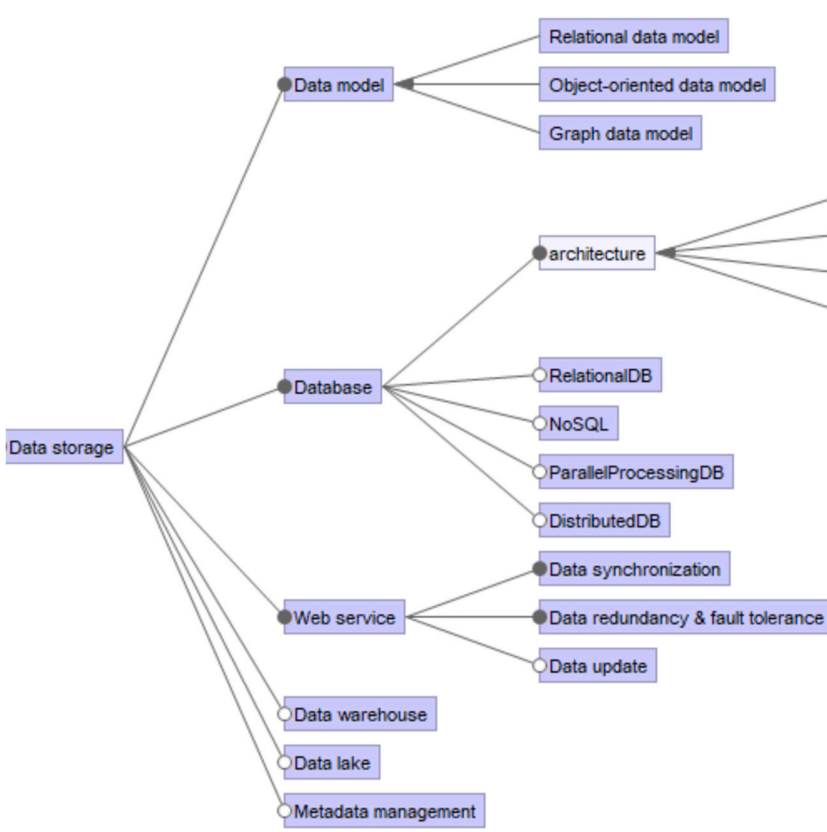

Fig. 16. Feature diagram of data storage.

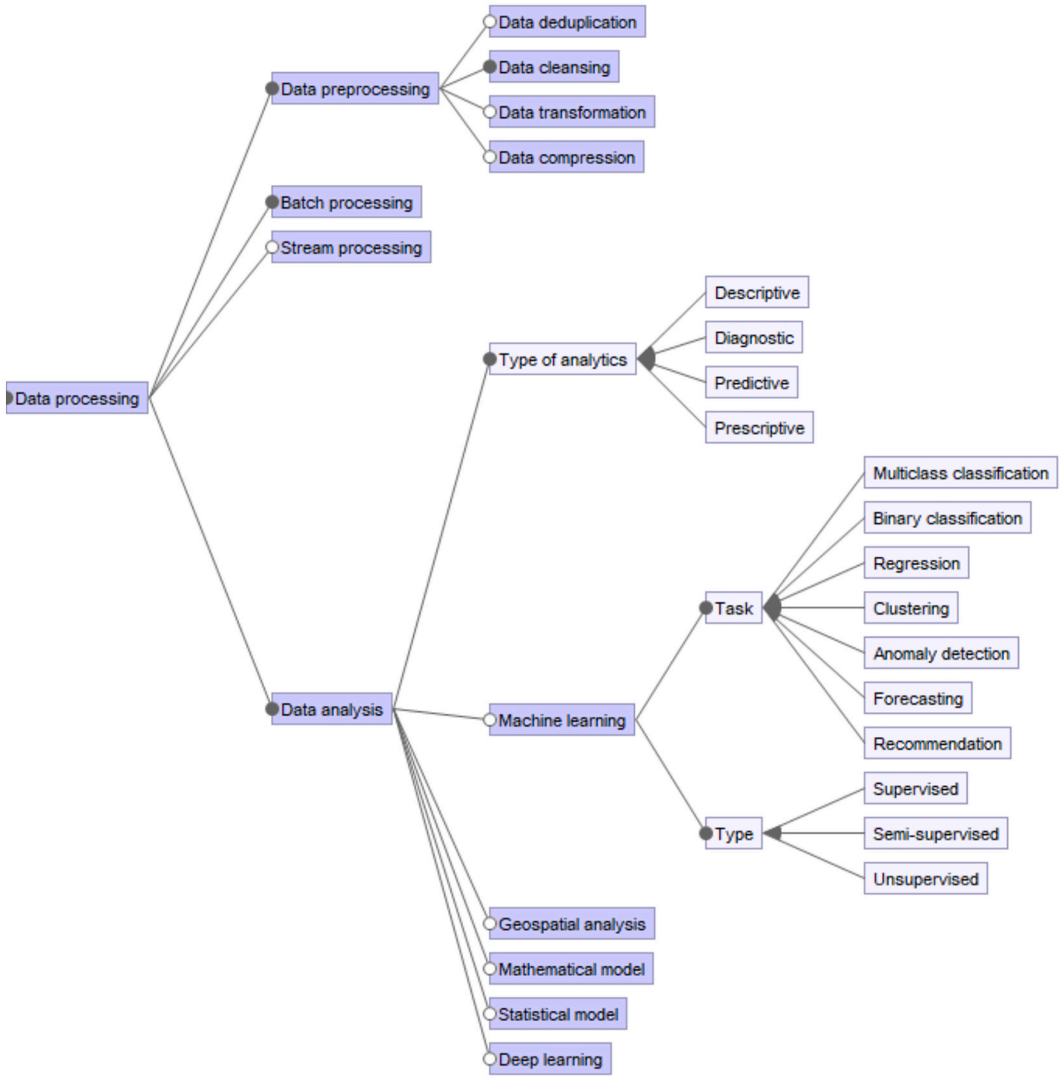

Fig. 18. Feature diagram of data processing.

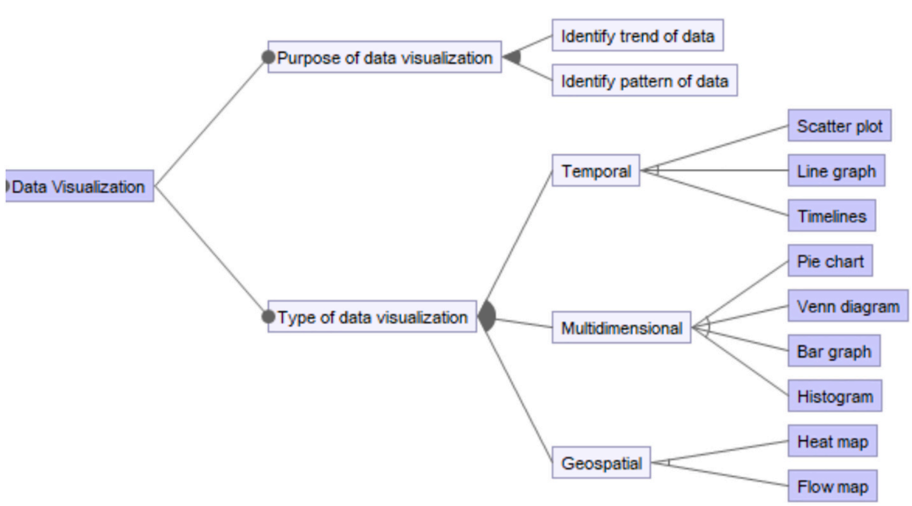

Fig. 19. Feature diagram of data visualization.

3.2.2.8. Data security. Data security is not explicitly mentioned in most existing data management studies. Hence, data security is optional in this study. However, data security is important and should be discussed and implemented to protect people's data. Fig. 21 shows several features of data security.

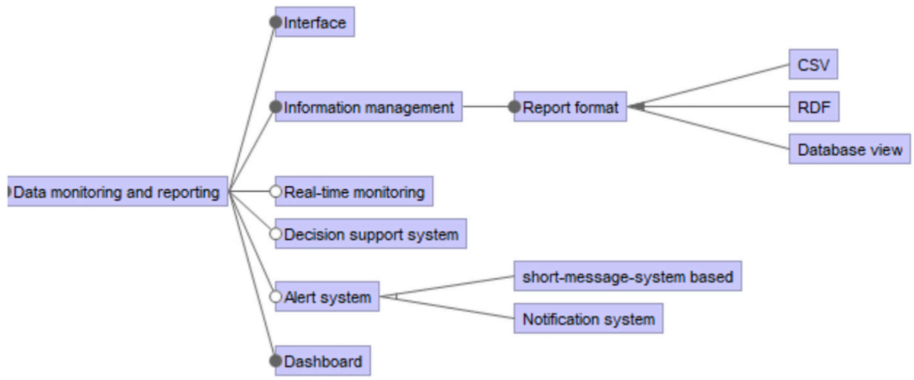

Fig. 20. Feature diagram of data monitoring and reporting.

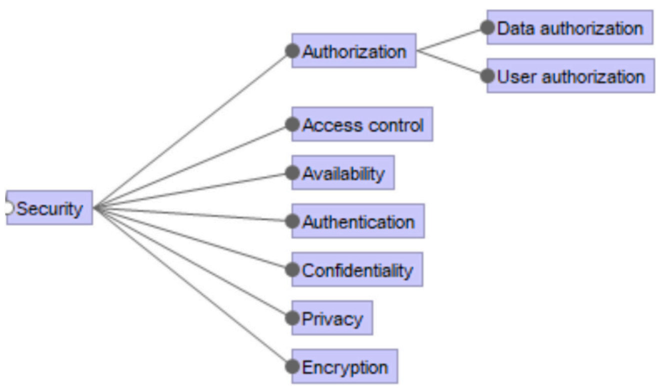

Fig. 21. Feature diagram of data security.

4. Reference architecture design

4.1. Selection of views

In this section, the three selected sets of viewpoints of the Views and Beyond (V&B) architecture framework (Clements et al., 2010) are used, including the context diagram, decomposition viewpoint, and deployment viewpoint. These views were used to represent the proposed reference architecture. In our review, the context diagram, decomposition view, and deployment view were used to represent the reference architecture. The context diagram depicts all interactions between the system and external elements. It is applied to show the project's scope and clarify various parts. Therefore, a context diagram illustrates what system is to be developed as well as which parts and components are needed (Clements et al., 2010; van Geest et al., 2021). The decomposition view depicts the system's modules and submodules. This view is used to show how the system's responsibilities are divided among them. The use of a decomposition view in understanding both similarities and dissimilarities across diverse modules enables the parallel implementation of responsibilities since separate modules can be allocated to different teams (Clements et al., 2010; van Geest et al., 2021). The decomposition view shows the software's structure by decomposing larger modules into smaller ones. It is perceived as a fundamental architecture view since it provides the input for the deployment view. Meanwhile, the deployment view is applied to analyze performance, reliability, security, and availability (Clements et al., 2010; van Geest et al., 2021). The reference architecture of the smart farming system is elaborated in the following subsections.

4.2. Context diagram

The context diagram of the data management system in the smart farming context is presented in Fig. 22. First, the data management system interacts with various data generated by IoT devices, sensors, UAVs, and UGS. UAVs and UGS are optional devices for data acquisition operations, depending on the agricultural domains and farming practices. The mandatory actors are farmers, while researchers and veterinarians are optionally involved in the system. However, these actors also provide some inputs to enhance the collected data for the system. The data management system also receives the data from other sources, such as weather services and intelligent machinery (e.g., milking robots). The former usually provides public data for people who want to access weather data for specific purposes. The latter is usually installed on the farm to assist farmer activities and generate their status to keep in the system. Furthermore, the generated information or action from the data analysis phase is sent to the IoT devices or web and mobile applications. The IoT devices do some tasks based on the commands sent by the system while showing generated information for the end-users through the web or mobile applications. The standalone system is required if no servers or network communications are installed on the farm. Lastly, the other mandatory stakeholders are the decision support and alert systems. These systems help farmers to monitor and manage their farms.

4.3. Decomposition view

In this section, all possible modules required for the data management system in smart farming are presented in a decomposition view (see Fig. 23). This view shows the decomposition of main modules into sub-modules in the overall system module, and all the modules are optional. The main purpose of presenting this view is to list entities that are supposed to be considered when designing the software architecture (Tummers et al., 2021). The top-level modules for data management systems are data acquisition, storage, processing, visualization, monitoring, and security. Besides, this view also presents the sub-modules from each main module mentioned before. For instance, the following sub-packages in the data processing package are shown in the decomposition view: data preprocessing, batch processing, stream processing, and data analysis. The system designer should take into account the data analysis techniques such as geospatial analysis, statistical-based models, mathematical models, deep learning, and machine learning when designing the smart farming system.

4.4. Deployment view

Fig. 24 shows the identified modules to the relevant hardware. The data acquisition, storage, visualization, and processing packages are placed on the data management server. On the user side, the application systems are deployed and installed. Other nodes are the devices that act as data generators for the smart farming system. The deployment view presents zero or more servers and one or more clients. The standalone system with all modules on the client side is deployed if there is no

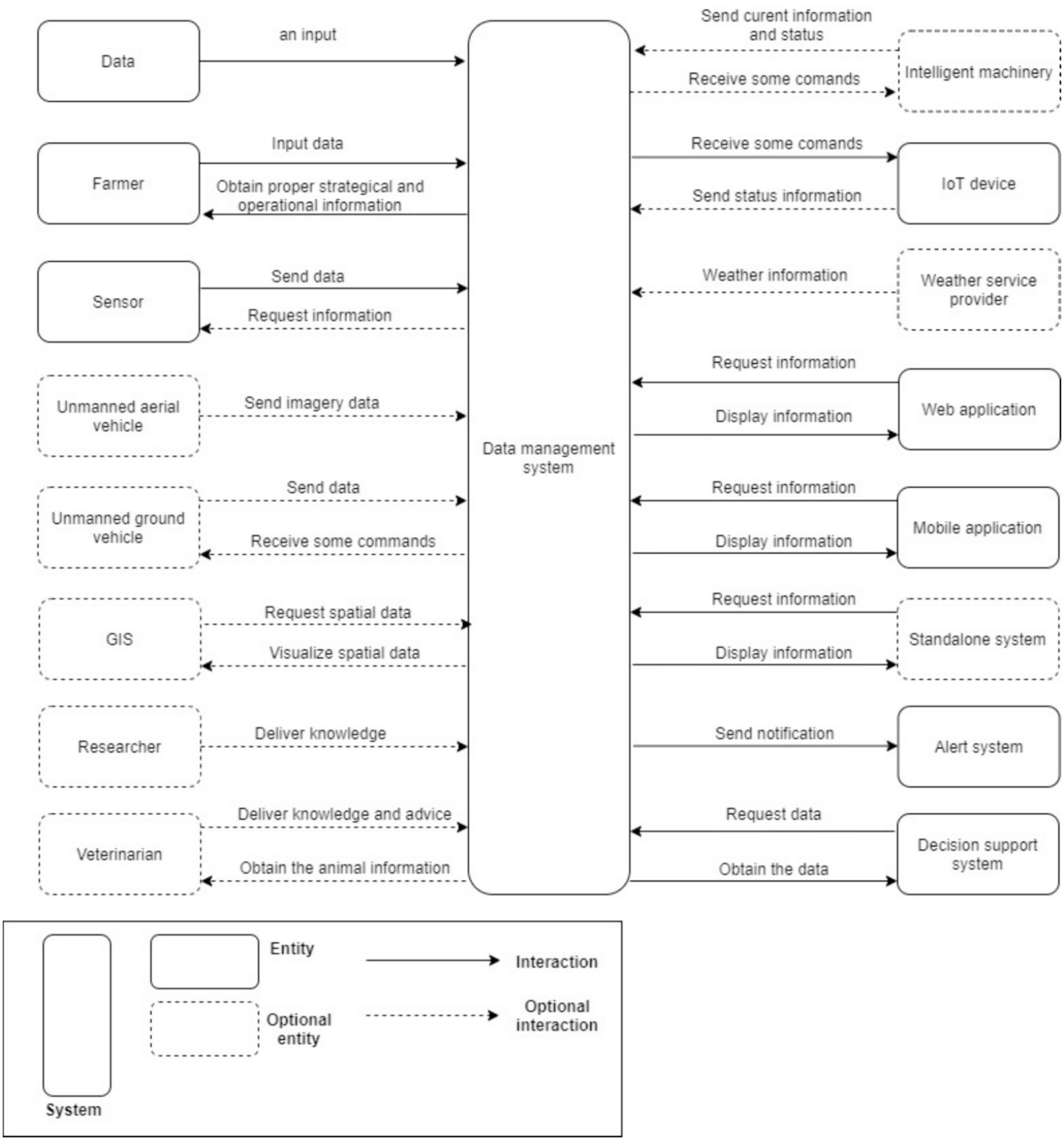

Fig. 22. Context diagram.

server. Meanwhile, a system with at least one server and multiple clients is a client-server system, i.e., mobile or web applications. Finally, the cloud-based system is a system that has multiple servers and multiple clients with advanced communication technology (Tummers et al., 2021).

5. Case study evaluation

5.1. Problem statement

In this section, the case study that describes the underlying problems is discussed. Two case studies of smart farming in Indonesia have been selected due to two main reasons. First, to the best of our knowledge, no literature discusses a proper reference architecture for the Indonesian smart farming system. Second, this study is funded by INREF to develop a smart farming system in Indonesia. Therefore, the application architectures resulted in this study are used as a blueprint for further system development. In the following, the details of two case studies are presented.

5.1.1. Case study: dairy farming

According to Statistics Indonesia (Statistics Indonesia - Badan Pusat Statistik (BPS), 2024a), milk production in Indonesia has significantly increased in the last three years, from 135 million litres in 2018 to 221 million litres in 2020. Indonesia's dairy industry is dominated by smallholder farmers, about 87% of the total production (Kementerian Perdagangan Republik Indonesia - Ministry of Trade Republic of Indonesia, 2010). Their characteristics are typically small, still, on average, maintaining traditional approaches to managing their farms and owning less than ten milking cows (Jahroh et al., 2020). The

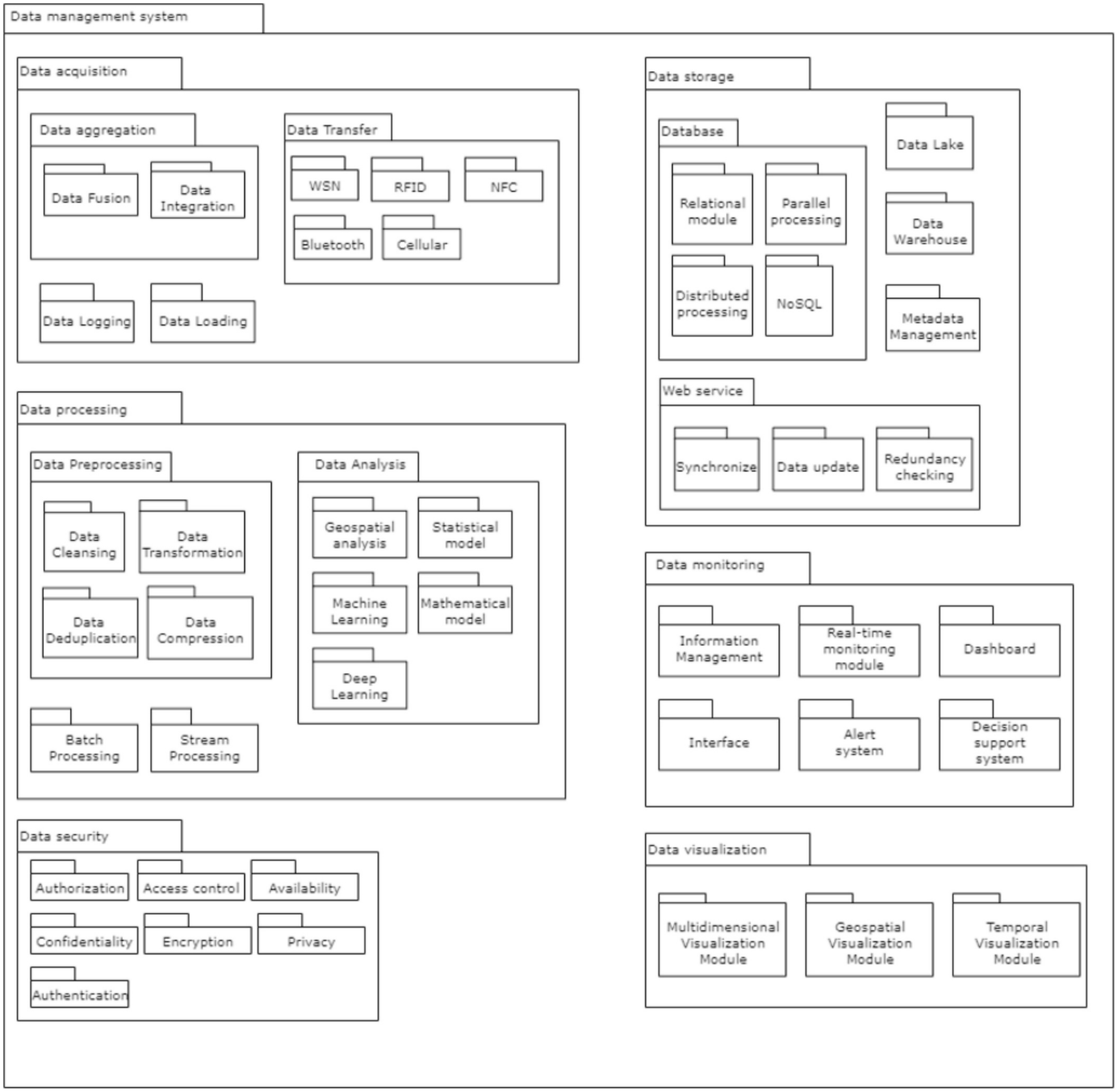

Fig. 23. The reference decomposition view of smart farming data management.

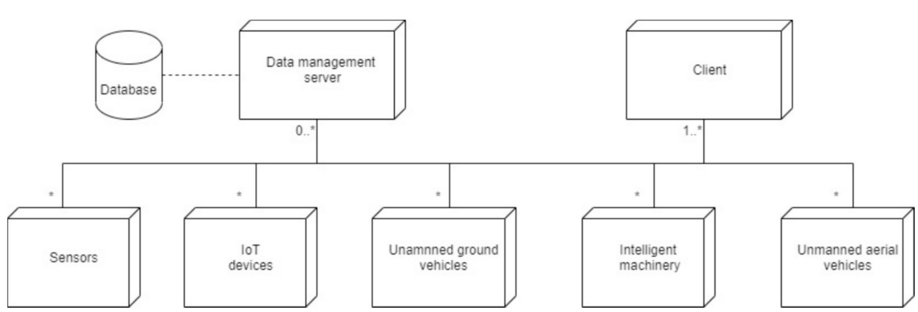

Fig. 24. Deployment view.

Indonesian government aims to increase domestic milk production by up to 50% by implementing an advanced dairy farming system. For instance, applying new technologies in feed management is considered to increase production (Jahroh et al., 2020). One of the central regions of dairy production in Indonesia is West Java, which has potential resources and a suitable climate for raising dairy cattle and became one of the regions with the largest dairy cattle population in Indonesia (Statistics Indonesia - Badan Pusat Statistik (BPS), 2024b). Furthermore, West Java is one of Indonesia's top three provinces, contributing to Indonesian milk production (Jahroh et al., 2020).

5.1.2. Case study: pond-based aquaculture

Indonesia plays an important role in global fishery production, contributing 5.8% of global production (Senff et al., 2018). It has a wide area suitable for aquaculture development (Food and Agriculture Organization of the United Nations (FAO), 2024). Besides dairy farming, West Java also has potential fishery production, with 321,000 tons of production in 2020 (Kementerian Kelautan dan Perikanan - Ministry of Marine Affairs and Fisheries Republic of Indonesia, 2024). Most fish farmers in West Java still use traditional approaches for their farms or ponds.

5.1.3. Problem statement

Mainly, the farming practices in Indonesia, including dairy and fish farming, are still using traditional approaches. The farmers usually faced challenges in making decisions regarding the problems in the field since they did not have accurate data and information to support their decisions. A proper smart system is essential in Indonesian agricultural practices nowadays. Therefore, in order to help the Indonesian government, a smart Indonesian agricultural project was established by the lead of Wageningen University and Research (WUR) and Institut Pertanian Bogor (IPB University Indonesia). This project aims to develop data infrastructure for Indonesian agriculture, specifically fish and dairy production. The whole project is funded by INREF (Wageningen University and Research, 2024a; Wageningen University and Research, 2024b; The Interdisciplinary Research and Education Fund (INREF), 2024).

5.2. Case study protocol

The case study is used to assess and evaluate the proposed architecture designs and feature models. The main goal of using these case studies is to evaluate the developed architecture designs and feature models. The case study research protocol in this research follows the protocol defined by (Köksal and Tekinerdogan, 2019; Runeson and Höst, 2009; Tummers et al., 2021; van Geest et al., 2021). Both dairy and fish studies are prospective cases, which include the system that are planned to be developed (Köksal and Tekinerdogan, 2019). Five process steps that have been followed when implementing the case studies are as follows: (1) Designing the case study, (2) Preparing data collection, (3) Collecting evidence, (4) Analyzing collected data, and (5) Reporting. Table 2 presents the case study design. The data collection is conducted by distributing questionnaires to the project manager, project members, and the smart farming experts outside the project. The questions for the survey are presented in Appendix A.

Table 2
Case study design.

Case study activities	Cases: Indonesian dairy and fish farming
Goal	To evaluate feature models and architecture designs
Research questions	RQ4. How effective is the designed reference architecture?
Sources	Project manager, project members, smart farming experts outside the project
Data collection	Survey through questionnaire
Data analysis	Qualitative data analysis

The survey is organized as follows: (1) The survey instrument is created in the first step. This research used a questionnaire as the survey instrument to observe and get information regarding dairy and fish farming practices and future plans. (2) Second, the questionnaire is pilot tested with several field experts to ensure that the questions are being interpreted as intended. The information gained from the pilot test process is essential to identify the revisions needed to increase the questionnaire's content validity. (3) The final questionnaire is distributed to the experts and researchers in the case studies domain, dairy and fish farming. Later, the survey results are used as a basis to develop application architecture. (4) In the final step, the researchers analyze and review the application architectures from the cases. The following sections discuss the results of the processes mentioned above.

5.3. Feature model for the case studies

5.3.1. Case-1: dairy farming

Fig. 25 shows the feature model for Indonesian smart dairy farming based on the conducted survey. This application feature model is derived from the family feature model for the smart farming system, as shown in Fig. 4. The family features are chosen to fit the case study's needs. As shown in Fig. 24, FMIS is the common smart farming type in Indonesia, while PA and automation systems are optional. FMIS in Indonesian smart farming systems focuses on managing and transforming raw data into valuable information for stakeholders.

Regarding data analytics features, machine learning, statistical, and mathematical models were commonly used to analyze the raw data. Sensors are the primary sensing technology to help generate the data to support data analytics performance. The UAVs, UGVs, and GIS can be used as additional data sources to get more precise data regarding farm conditions. For example, the UAV and GIS can be used as data sources when the smart system needs spatial, location or area data. A cloud computing platform and standalone system (i.e., dedicated server) are commonly found as the computing infrastructure to implement data processing and analytics. Furthermore, WSN and cellular networks are used to communicate among the devices in the field, while RFID is an optional approach. In order to disseminate and visualize the actual data, mobile applications are required in smart dairy farming systems. In this feature diagram, mostly the stakeholders of dairy farming choose a mobile application as a mandatory application. It does not mean that web applications are not required, but they prefer to have mobile applications to see the actual data from the field. It might drive due to users have a better experience by using mobile applications since they are simple, accessible, and user-friendly.

The generated data in the dairy domain are based on several measured parameters, such as cow's health, behaviour, well-being, productivity, and location. These data must be analyzed to achieve the system's goals, such as productivity improvement, resource efficiency, or disease prevention. Other measurements are optionally used depending on the research purpose.

5.3.2. Case-2: fish farming

In the fish farming case, automation systems or robotics are the most studied systems. This result is strengthened by Dzulqornain et al. (Dzulqornain et al., 2018) and Periyadi et al. (Periyadi et al., 2020), who also developed an automation system to help farmers control pond water quality by monitoring water's pH and salinity, water temperature, and water level using sensors. Therefore, sensors are needed to collect the field data and become the only source of data in smart fish farming. Contrary to smart dairy farming, developers tend to use mathematical models rather than other techniques as data analytics algorithms in smart fish farming systems. IoT is a mandatory feature in the domain to control hardware systems and robotics devices. For data communication, WSN is a preferred approach in smart fish farming. Mobile applications are still critical devices as user interfaces in dairy and fish

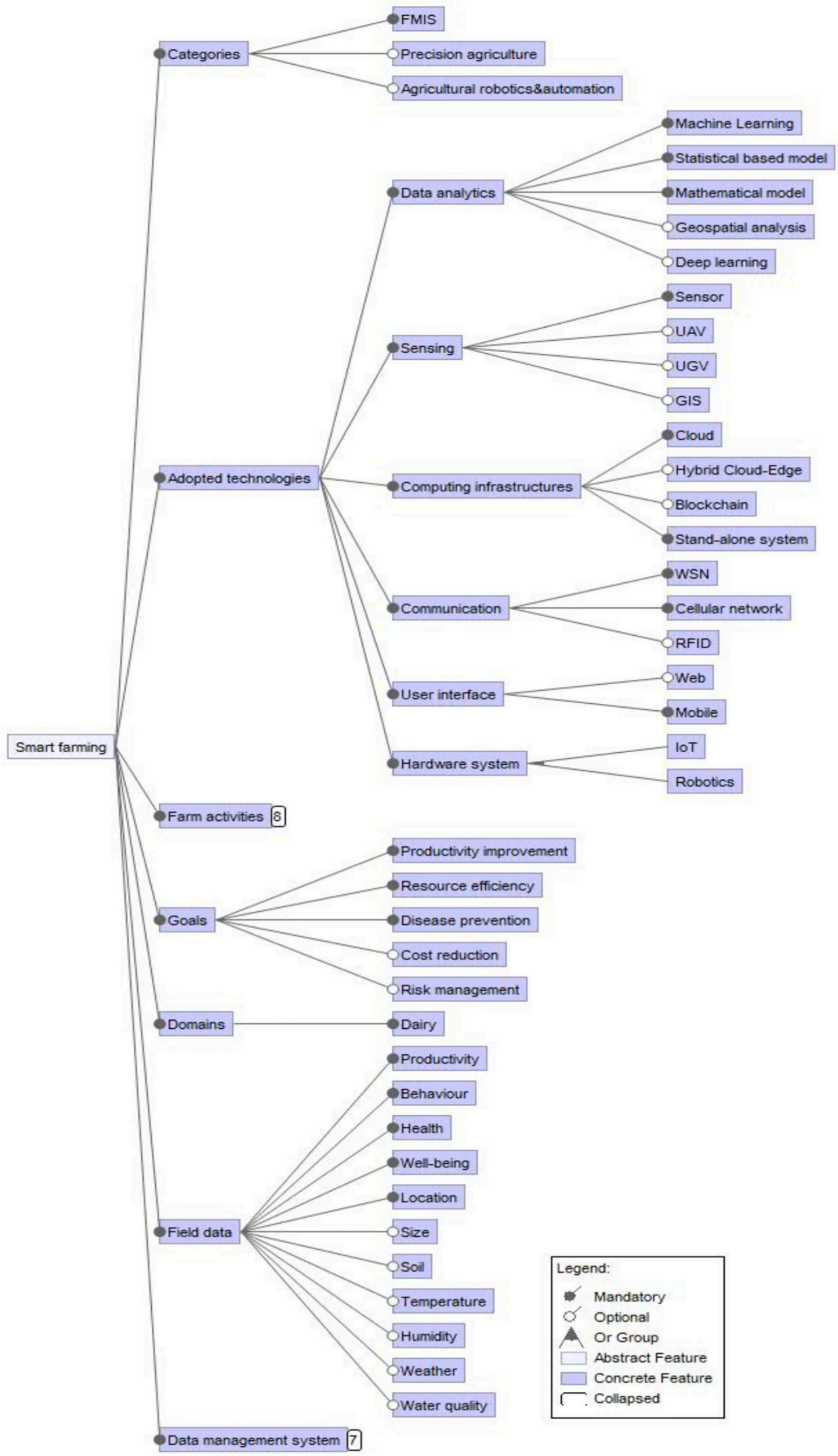

Fig. 25. Reference feature diagram for smart dairy farming of the INREF project.

farming to present the information.

The common goals of the smart system in aquaculture are productivity improvement, resource efficiency, and cost reduction. The system needs specific field data to achieve these goals, such as fish productivity, location, water temperature, and water quality. Fig. 26 presents the application features of smart fish farming.

5.4. Application architecture design

5.4.1. Case-1: dairy farming

5.4.1.1. Context diagram. Fig. 27 shows the context diagram for the smart dairy farming case. Firstly, the weather information and spectral

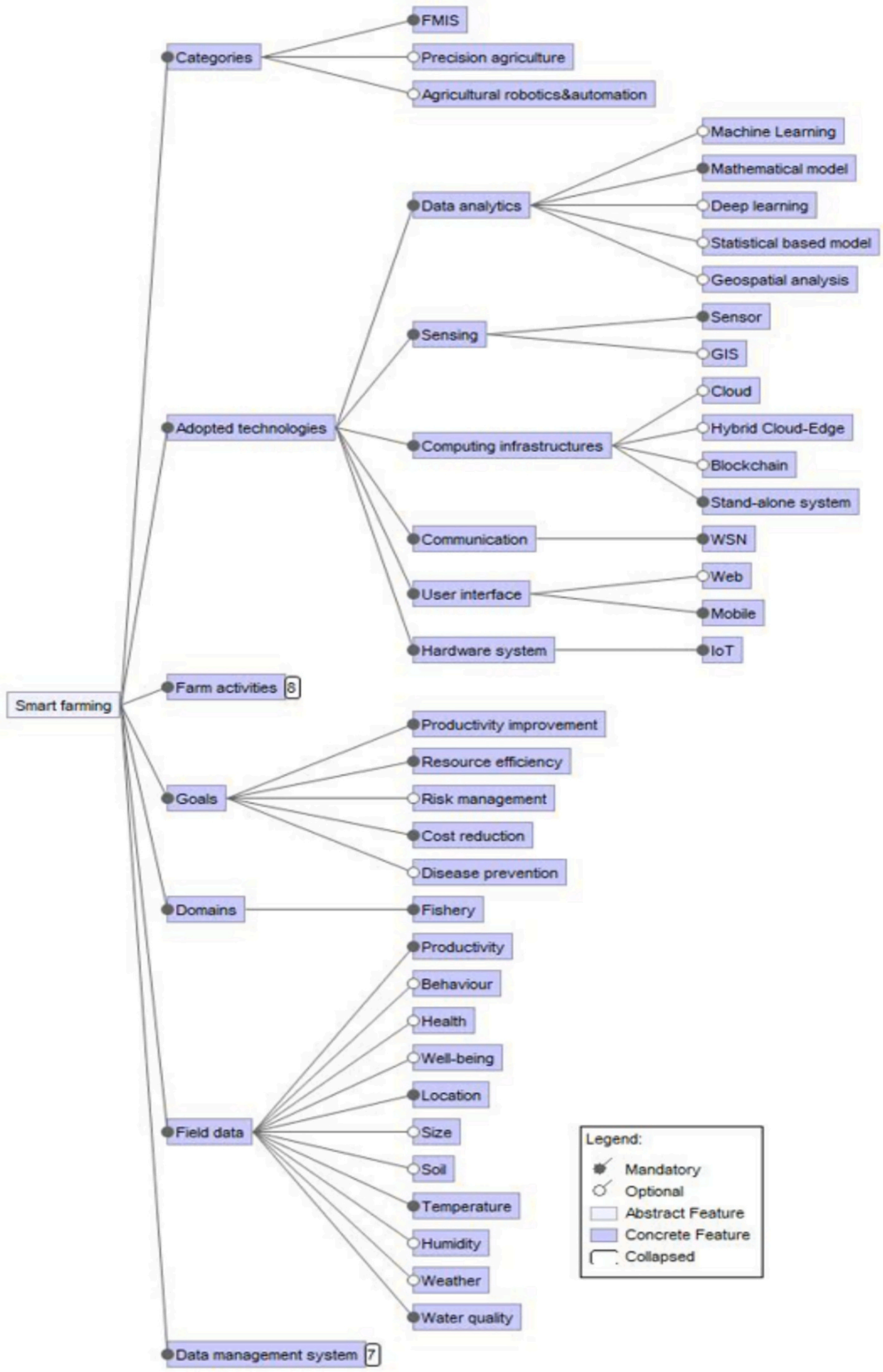

Fig. 26. Reference feature diagram for smart fish farming of the INREF project.

data from the weather service provider and the satellite data provider are not necessarily needed in this case. The veterinarian and researcher added information on dairy farming is essential for the system. Thus, the veterinarian is transformed from an optional to a mandatory stakeholder in the dairy domain. Another change is that intelligent machinery has become essential in smart dairy farming, while alert and decision support systems are optional. Generally, to access information, the survey participants tend to choose standalone and mobile rather than web applications.

5.4.1.2. Decomposition view. Fig. 28 shows the decomposition view of the smart dairy farming case. It can be seen that 37 modules are reused in this case study, which becomes the functional requirements from the stakeholders. WSN, cellular networks, and RFID can be implemented as communication protocols to transfer data among entities involved in the system. All the sub-features of the data acquisition and monitoring feature can be used. Furthermore, data lake, warehouse, metadata management, and all web services' functions are essential for the smart system to store and manage the data. Regarding the data analysis module, the experts in this survey tend to choose a mathematical, statistical-based model and machine learning to implement data analytics. Finally, visualization modules are selected to disseminate the information by using multi-dimensional (e.g., pie chart, Venn diagram, bar graph, histogram) and temporal (e.g., scatter plot, line graph,

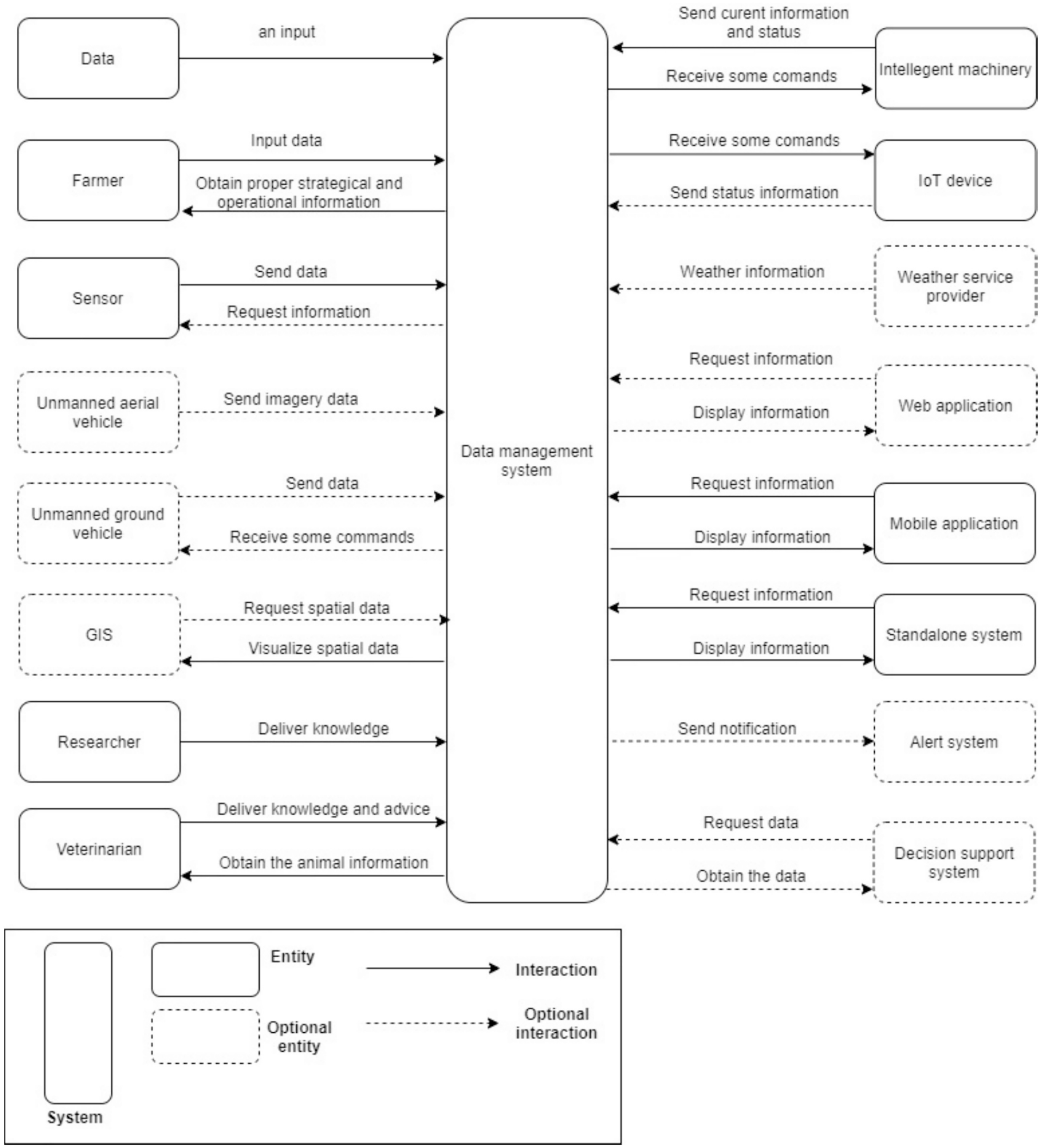

Fig. 27. Context diagram of smart dairy farming.

timelines).

5.4.1.3. Deployment view. The architecture deployment view of smart dairy farming is presented in Fig. 29. The required modules discussed in the decomposition view are deployed to a central server. In addition, the sub-modules of the system are not shown in the deployment view. In this case study, a thin client (fewer modules installed) is chosen for the client side since all modules and data are located on the central server. The deployment view presents zero and more servers, meaning standalone, mobile, or web applications are covered in this design. Finally, sensors, intelligent machinery, and IoT devices are also deployed in the system.

5.4.2. Case-2: fish farming

5.4.2.1. Context diagram. Fig. 30 shows the entities involved in the smart fish farming system. Overall, farmers, researchers, and

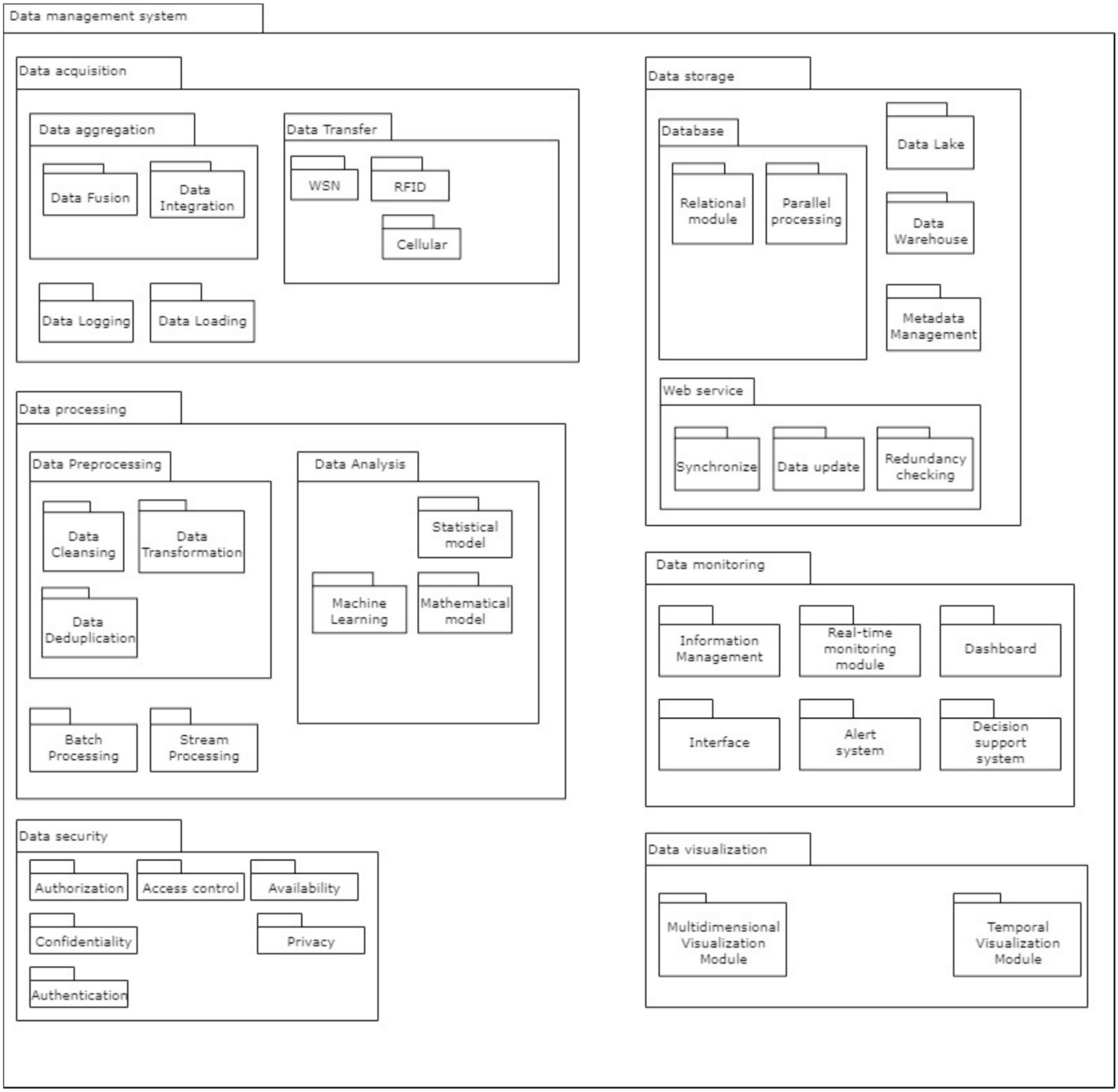

Fig. 28. The application decomposition view of smart dairy farming data management.

veterinarians are key actors in this domain. Sensors and IoT devices are essential to generate data both in dairy and fish farming. In contrast to dairy farming, weather service providers are optional for this case study, depending on the need for weather data. As in dairy farming, standalone and mobile applications are the preferred device to present information regarding the field's condition rather than the website. A decision support system and alert system are still needed and depend on the system's purpose.

5.4.2.2. Decomposition view. Fig. 31 provides the decomposition view of smart fish farming obtained using the reference decomposition view presented in Fig. 23. Out of 46 modules provided in the reference architecture, 28 are the functional requirements of the stakeholders in the fish farming case study. It is shown that WSN is the only protocol chosen by experts in the survey to transfer the data among entities. The data is then stored in a relational database module. The implementation of a data lake and warehouse is also needed in this domain. For analytics purposes, it seems that the mathematical model is the preferred approach for the experts. Finally, the temporal visualization module is needed to present the valuable information.

5.4.2.3. Deployment view. The modules shown in the decomposition view are deployed in a central server system. The central server employs sensors and IoT devices to help farmers manage their ponds. As in dairy farming, a thin client is selected, and zero or more servers are required for this case study. The deployment view is presented in Fig. 32.

6. Discussion

This research has developed a reference architecture for smart

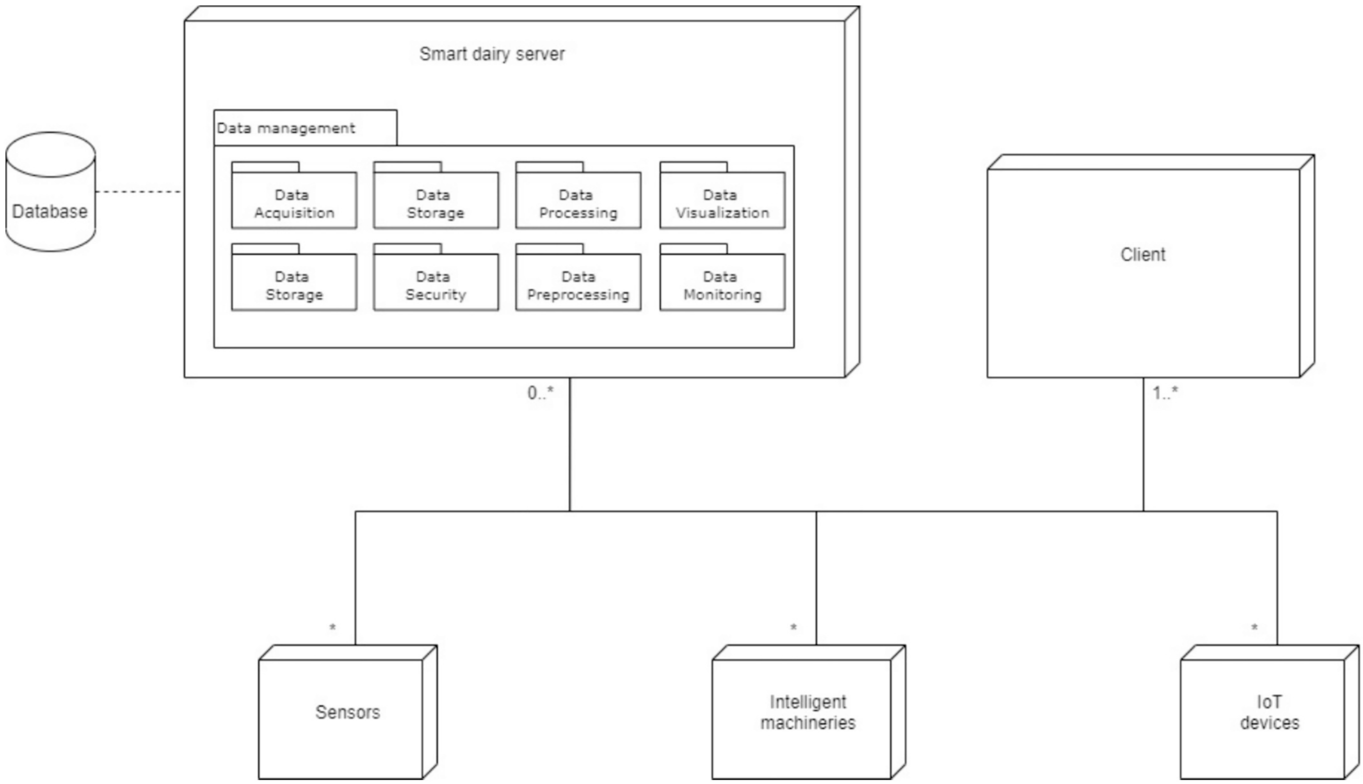

Fig. 29. Deployment view of the smart dairy farming.

farming data management through domain analysis and architecture modeling. Domain analysis provides insights into common and variant features and modules of smart farming systems, forming the basis for family features. These family features contribute to the development of the reference architecture. As a result, other smart farming application architectures can be designed using the reference architecture and family features outlined in this study.

The domain analysis in this study yields two family features: one encompassing general smart farming features and another focusing on data management. While the primary goal is to develop a data management architecture, features of smart farming systems from a broader perspective are also identified. These family features are derived from academic literature in the Scopus database and also from our previous SLR work (Krisnawijaya et al., 2022).

The presented reference architectures are designed to adapt to the dynamic nature of smart farming. Given the evolving nature of smart farming technologies, flexibility is essential to integrate new features, handle increasing data complexity, and incorporate modules from various sources. Supplementary literature, such as software requirement specifications (SRS) documents, vendor websites, and other grey literature, may introduce new functionalities and modules to enhance the proposed architecture.

To address data complexity, our reference architecture incorporates several big data features, including data lakes, NoSQL databases, and distributed databases. Additionally, data is highlighted as a mandatory feature in the feature diagram to emphasize its significance in data management.

The reference architectures are developed using an architecture framework with well-defined viewpoints, including context, decomposition, and deployment. Application architecture can be derived from the proposed reference architecture through a multi-case study method, offering flexibility for different concrete architectures based on specific needs.

It is crucial to note that presented modules and features are not absolute for certain domains, as different application domains may necessitate new features and functionalities. Therefore, further research is essential to enhance and adapt the architecture.

Two case studies were conducted to demonstrate the applicability of the proposed architecture in the dairy and fish farming domains. The results showed that the reference architecture facilitated the rapid design of application architectures, as validated by stakeholder feedback through a questionnaire. To mitigate the threat of misinterpretation, survey questions were thoroughly discussed among researchers and pilot-tested with field experts.

Based on the case study results, the application architecture was easily derived from the reference architecture. Stakeholders' responses through the questionnaire form indicated that the design method expedited the application architecture design process. All survey questions presented in Appendix A were accompanied by a description of the question's purpose to facilitate stakeholders in answering. Furthermore, the stakeholders were able to express their expectations and needs of the design using the space provided in the questionnaire. The application architecture view models were designed based on stakeholder requirements. Initially, the context diagram was crafted to illustrate the interaction between entities and the system. The input for this diagram came from stakeholders and their concerns, aiming to capture as many essential interactions as possible. However, it is acknowledged that there might be more possible interactions than those depicted in the diagram.

Another view model in this study is the decomposition view, presenting all possible modules for a specific domain. The view was designed as generically as possible by including all possible modules in the design. Despite careful and iterative discussions among the authors, there remains a possibility that some modules for a particular application are missing. Finally, the third view is the deployment view, mapping all features and modules into specific servers, applications, and devices.

In addition to qualitative evaluation, conclusions can be derived

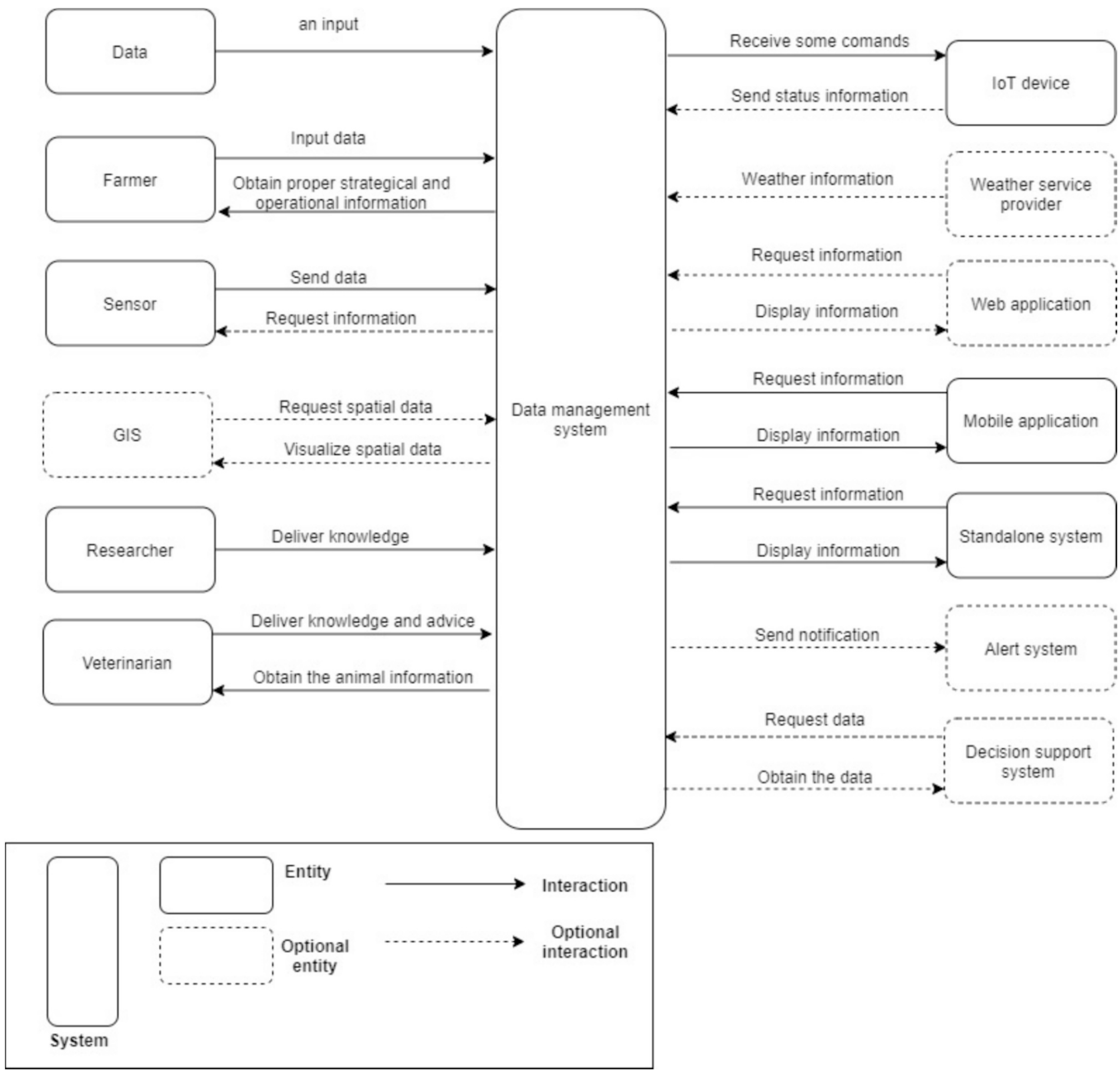

Fig. 30. Context diagram of smart fish farming.

from quantitative metrics, including the percentage of reuse given the provided reference architecture. For both case studies, we observed that the total reuse based on the reference architecture was quite high. In the context diagrams, the reference architecture could be reused without the need for adding new architectural entities. A similar result was found in the decomposition and deployment view, where the modules provided in the reference architecture can significantly be applied in both case studies, and additional modules are not required. Out of the 46 modules provided in the reference architecture, 10 modules were reused only in the dairy farming case study, one module was reused only in the fish farming case study, while 27 modules were reused in both case studies. The calculation of the percentage reuse modules can be seen below.

$$\%reuse\ modules = \frac{total\ reused\ modules\ in\ case\ studies}{total\ modules\ in\ reference\ architecture}\ x\ 100\%$$

$$= \frac{10 + 1 + 27}{46}\ x\ 100\% = \frac{38}{46}\ x\ 100\% = 82.6\%$$

Based on the calculation above, the total reuse of the modules was 82.6%. This is an essential and substantial improvement in cost savings and time reduction for the development of the systems.

Finally, this study proposes a novel method to derive the reference architecture by involving domain architecture in the process, which can be seen in Fig. 1. Although several previous studies have developed reference architecture in the agricultural field, Giray and Catal (Giray and Catal, 2021), Tummers et al. (Tummers et al., 2021), and Santana et al. (Santana et al., 2014), none of them applied domain architecture in deriving their architecture despite its importance in acquiring and storing required knowledge and information to assist and support the application design process in the domain of interest. Therefore, this research provides not only generic architecture for smart farming but also for the domains of dairy and fish farming.

7. Conclusion

This paper has presented and evaluated the proposed data management reference architecture for smart farming using two case studies. To the best of our knowledge, this is the first study that explicitly focuses on

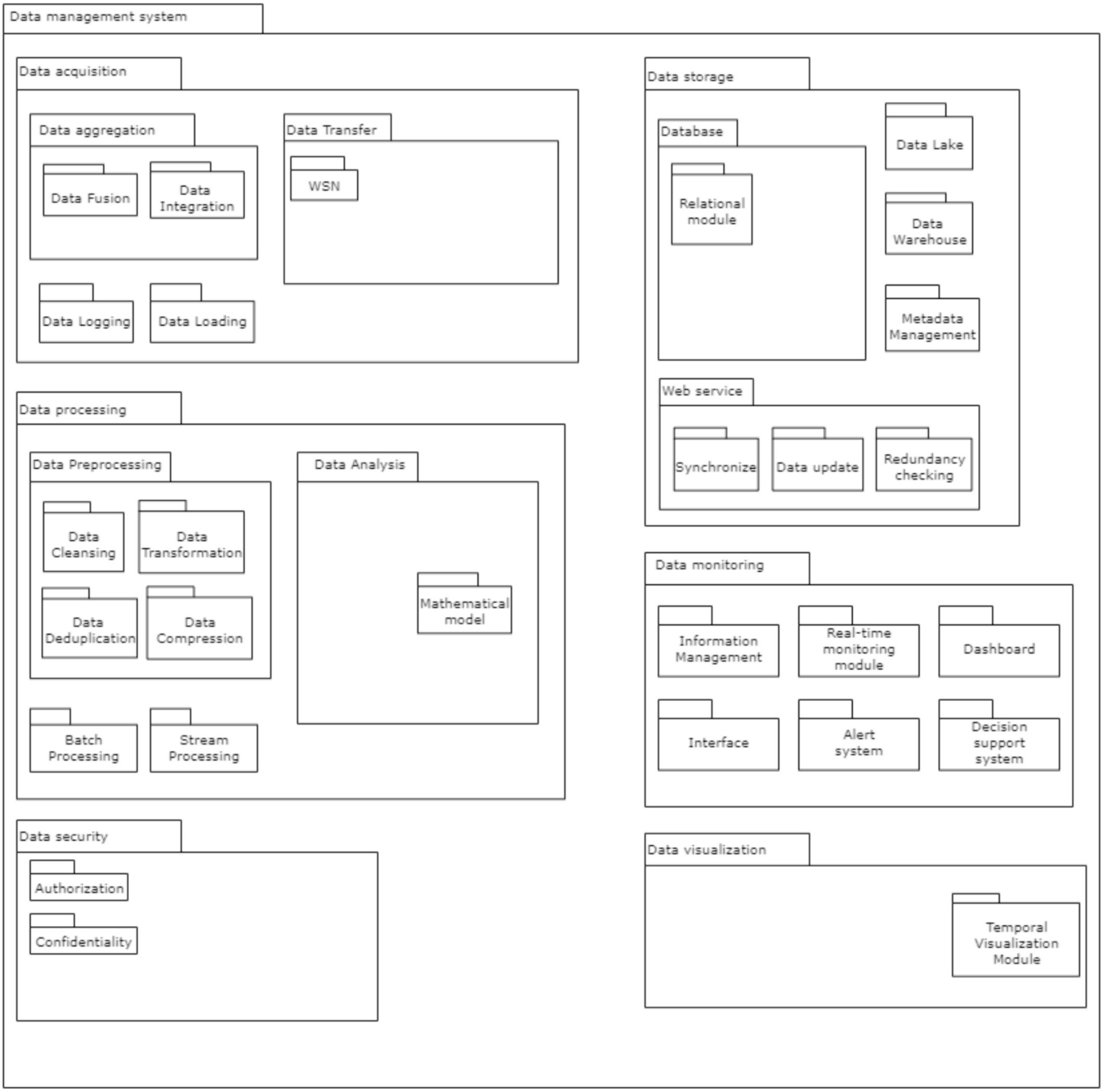

Fig. 31. The application decomposition view of smart fish farming data management.

the architecture design approach for the data management system in a smart farming context by incorporating domain architecture in the process. Overall, it was demonstrated that the proposed reference architecture is useful and practical for designing the smart farming system. The presented features and architecture views are beneficial when developing new systems in different agricultural domains, with a specific emphasis on data management and the smart farming system in general.

Formal architecture design viewpoints were chosen in order to develop the reference architecture in this research. The reference architecture can serve as a blueprint for designing new smart farming systems, as it has been validated using a multi-case study approach. The proposed reference architecture appeared to be successful and effective for designing smart farming systems. However, further research is needed to evaluate the proposed architectures for different domains. The authors expect this study to encourage more researchers and practitioners to develop and propose a novel design for smart farming systems.

Funding

This work is part of a larger set of projects that have been funded by the Interdisciplinary Research and Education Fund (INREF). This research is funded by the INREF of Wageningen University & Research (grant for the Smart-In-Ag program, https://www.wur.nl/en/Research-Results/Research-programmes/Cross-WUR-programmes/INREF.htm). The funder had no role in the study design, data collection, publishing decision, or preparation of the article.

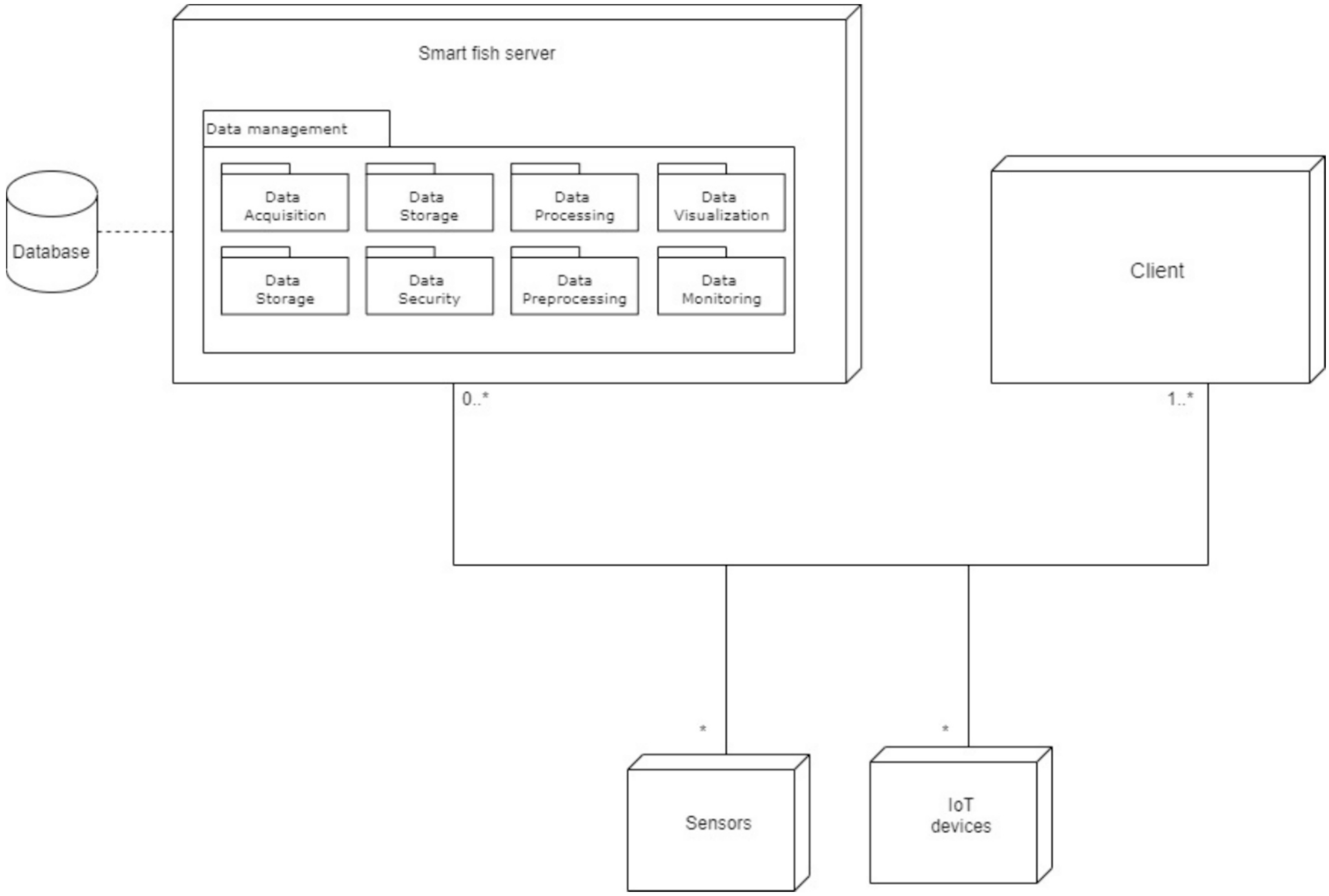

Fig. 32. Deployment view of the smart fish farming.

CRediT authorship contribution statement

Ngakan Nyoman Kutha Krisnawijaya: Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Bedir Tekinerdogan:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Conceptualization. **Cagatay Catal:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Conceptualization. **Rik van der Tol:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Conceptualization.

Data availability

No data was used for the research described in the article.

Acknowledgements

Not applicable.

Appendix A. Appendix

List of questions for the respondents (stakeholders):

1. What is your position in the domain(s)?
   - Farm manager
   - Data manager
   - Researcher
   - Data analyst
   - Other
2. What are the goal(s) of adopting smart farming technologies in your domain?
   - Productivity improvement
   - Cost reduction
   - Resource efficiency
   - Disease prevention
   - Risk management
   - Other
3. Which smart farming category do you find the most in your domain?
   - Farm management information system (FMIS)
   - Precision agriculture (PA)
   - Agricultural automation and robotics
   - I do not know
4. Which of the following data analytics techniques will you need for your domain/system?
   - Machine learning
   - Deep learning
   - Statistical based model
   - Mathematical model
   - Geospatial analysis
   - I do not know
5. In your opinion, what application(s) are needed in smart farming?
   - Web application
   - Mobile application
6. Do you need Internet of Things technology for your domain?
   - Yes
   - No
   - Not applicable

7. Do you need robotics for your domain?
   - Yes
   - No
   - Not applicable
8. Which of the following metrics do you think are important for your domain?
   - Temperature
   - Humidity
   - Weather
   - Water quality
   - Soil
   - Size
   - Productivity
   - Location
   - Behaviour
   - Health
   - Well-being
   - Other
9. Which of the following phases of the data science lifecycle are important for your domain?
   - Data acquisition (How the data are obtained)
   - Data storage (How the data are stored, e.g., cloud or local server)
   - Data processing (How the data will be transformed into valuable information)
   - Data visualization (How the data will be delivered to the stakeholders, mostly using graph, chart, etc.)
   - Data monitoring and reporting (How the data will be disseminated, e.g., reporting, notification, etc.)
   - Data security (How the data will be kept confidential)
   - I do not know
   - Not applicable
10. What data do you need for your domain?
    - Agronomic data (e.g., crop conditions, animal conditions, etc.)
    - Machine data (data from agricultural machinery)
    - Production data (e.g., planting dates, spraying records, etc.)
    - Predictive data (forecast future production under current conditions)
11. What type of data do you need?
    - Textual
    - Visual
    - Audio
    - Other
12. What type of data structure do you think is commonly processed in your domain?
    - Structured data (e.g., quantitative data, statistical results)
    - Semi-structured data (e.g., CSV, XML, JSON)
    - Unstructured data (e.g., qualitative data, text files, images, videos)
    - I do not know
13. Which features of data collection do you think are needed for your domain?
    - Data aggregation (Gathering data from multiple sources and presenting it in summarized format)
    - Data transfer (Securing exchange of data between systems or organization)
    - Data loading (Copying or loading data from a source to a database)
    - Data logging (Storing actions/events of a system or network over a period of time)
    - I do not know
    - Not applicable
    - Other
14. Which of the following techniques will you need for collecting data?
    - Sensor
    - Unmanned aerial vehicles (UAVs)
    - Unmanned ground vehicles (UGVs)
    - GIS
    - I do not know
    - Not applicable
    - Other
15. Which of the following processes do you need in data processing?
    - Collect data over a period of time, then process it
    - Collect data in a small time period, then process it (near real time)
    - Both processes are needed
    - I do not know
    - Not applicable
16. Which of the following computing infrastructures do you need?
    - Cloud computing
    - Edge computing
    - Blockchain
    - Standalone computer/Dedicated server
    - I do not know
    - Not applicable
    - Other
17. Which of the following communication technologies will you need for data transfer?
    - Wireless sensor network (WSN)
    - Radio Frequency Identification (RFID)
    - Near Field Communication (NFC)
    - Bluetooth
    - Cellular network
    - I do not know
    - Not applicable
    - Other
18. How do you access the dataset for research?
    - By public sources
    - Get permission from the data owner
    - Using proprietary data
    - Not applicable
    - Other
19. Which type of data model do you need for storing your data?
    - Relational data model (traditional model using tables with rows and columns)
    - Graph data model (NoSQL model using graph or tree to illustrate the relation between data)
    - Object-oriented data model (to work with complex data objects)
    - I do not know
    - Not applicable
    - Other
20. Which database type will you need?
    - Relational database (e.g., MySQL, Microsoft access)
    - NoSQL (e.g., MongoDB, HBase, Cassandra)
    - Parallel processing database (e.g., Oracle)
    - Distributed database (e.g., Hadoop, Apache Spark)
    - I do not know
    - Not applicable
    - Other
21. Which database architecture will you need for your domain?
    - Fileserver
    - Client-Server
    - 3-tier Architecture
    - Distributed system
    - I do not know
    - Not applicable
    - Other
22. Please select the function of the data storage that you think is necessary for a smart farming system.
    - Storing structured and filtered data (Data warehouse)

- Storing structured, semi-structured, and unstructured raw data (Data Lake)
- Storing information about the structure of the stored data (Metadata repository)
- I do not know
- Not applicable
- Other

23. Which web service functions do you need for your domain?
    - Data synchronization (Maintaining the consistency of among data)
    - Data redundancy & fault tolerance (Preventing data loss)
    - Data update (Updating changes automatically)
    - I do not know
    - Not applicable
    - Other
24. Which features will you need for data preprocessing?
    - Data deduplication (Eliminating excessive copies of data)
    - Data cleansing (Removing incorrect, corrupted, incomplete data)
    - Data transformation (Changing the structure or format of data)
    - Data compression (Modifying or converting the bits structure of data)
    - I do not know
    - Not applicable
    - Other
25. In data processing, which one do you choose?
    - Using well-labeled training data (Supervised)
    - Combining labeled with unlabeled training data (Semi-supervised)
    - Using unlabeled training data (Unsupervised)
    - I do not know
    - Not applicable
26. What kind of data analytics do you need?
    - To know what happen (Descriptive analytics)
    - To know how something happen (Diagnostic analytics)
    - To know what will happen (Predictive analytics)
    - To know what should be done if something happens (Prescriptive analytics)
27. Please select one or more machine learning tasks below that you need for data processing?
    - Multiclass classification
    - Binary classification
    - Regression
    - Clustering
    - Anomaly detection
    - Forecasting
    - Recommendation
    - I do not know
    - Not applicable
    - Other
28. For which purpose do you need data visualization?
    - Identify trend of data
    - Identify the pattern of data
    - I do not know
    - Not applicable
    - Other
29. What is the most common type for data visualization in your domain?
    - Temporal type (e.g., Scatter plot, line graph, timelines)
    - Multi-dimensional type (e.g., pie chart, Venn diagram, bar graph, histogram)
    - Geospatial (e.g., heat map, flow map)
    - I do not know
    - Not applicable
    - Other
30. Please select the function that you think is necessary for monitoring and reporting data
    - Interface
    - Dashboard
    - Information management (e.g., reporting)
    - Real-time monitoring
    - Decision support system
    - Alert system (e.g., short message system (SMS), phone notification)
    - Not applicable
    - Other
31. Which of the following data security approaches do you need?
    - Authorization
    - Access control
    - Availability
    - Authentication
    - Confidentiality
    - Privacy
    - Encryption
    - Not applicable
    - Other

References

Alonso, R.S., Sittón-Candanedo, I., García, Ó., Prieto, J., Rodríguez-González, S., 2020. An intelligent edge-IoT platform for monitoring livestock and crops in a dairy farming scenario. Ad Hoc Netw. 98 https://doi.org/10.1016/j.adhoc.2019.102047.

Ardagna, Claudio A., Ceravolo, Paolo, Damiani, E., 2016. Big data analytics as-a service: issues and challenges. In: Presented at the IEEE International Conference on Big Data (Big Data).

Ayaz, M., Ammad-Uddin, M., Sharif, Z., Mansour, A., Aggoune, E.-H.M., 2019. Internet-of-things (IoT)-based smart agriculture: toward making the fields talk. IEEE Access 7, 129551–129583. https://doi.org/10.1109/access.2019.2932609.

Balafoutis, A.T., Evert, F.K.V., Fountas, S., 2020. Smart farming technology trends: economic and environmental effects, labor impact, and adoption readiness. Agronomy 10, 1–26. https://doi.org/10.3390/agronomy10050743.

Bhat, S.A., Hussain, I., Huang, N.-F., 2023. Soil suitability classification for crop selection in precision agriculture using GBRT-based hybrid DNN surrogate models. Eco. Inform. 75 https://doi.org/10.1016/j.ecoinf.2023.102109.

Cervantes, H., Kazman, R., 2016. Designing Software Architectures: A Practical Approach. Addison-Wesley, Boston, p. 277.

Clements, P., et al., 2010. Documenting Software Architectures: Views and beyond, 2 ed. Addison-Wesley Professional, pp. 1–560.

Collado, E., Fossatti, A., Saez, Y., 2019. Smart farming: a potential solution towards a modern and sustainable agriculture in Panama. AIMS Agric. Food 4, 266–284. https://doi.org/10.3934/agrfood.2019.2.266.

Corporation, F.C.I., 2020. Good Farming Determination Standards. United States Department of Agriculture, Washington, p. 65.

Demestichas, K., Daskalakis, E., 2020. Data lifecycle management in precision agriculture supported by information and communication technology. Agronomy 10, 1–21. https://doi.org/10.3390/agronomy10111648.

Dong, J., et al., 2016. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. Remote Sens. Environ. 185, 142–154. https://doi.org/10.1016/j.rse.2016.02.016.

Dzulqornain, M.I., Udin, M., Al, H., Sukaridhoto, S., 2018. Design and development of smart aquaculture system based on IFTTT model and cloud integration. In: The 3rd International Conference on Electrical Systems, Technology and Information, 164, pp. 1–11. https://doi.org/10.1051/matecconf/201816401030.

Food and Agriculture Organization of the United Nations (FAO), 2017. The future of food and agriculture: Trends and challenges.

Food and Agriculture Organization of the United Nations (FAO), 2018. The future of food and agriculture: Alternative pathways to 2050.

Food and Agriculture Organization of the United Nations (FAO), 2020. The state of food security and nutrition in the world: Transforming food systems for affordable healthy diets.

Food and Agriculture Organization of the United Nations (FAO), 2024. Fishery and Aquaculture Country Profiles: Indonesia. https://www.fao.org/fishery/en/facp/idn?lang=en (accessed December 29, 2021).

Fulton, J.P., Port, K., 2018. Precision agriculture data management. In: Precision Agriculture. Wiley, Basics, pp. 169–187.

Giray, G., Catal, C., 2021. Design of a Data Management Reference Architecture for sustainable agriculture. Sustainability 13, 1–17. https://doi.org/10.3390/su13137309.

Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A., Ullah Khan, S., 2015. The rise of "big data" on cloud computing: review and open research issues. Inf. Syst. 47, 98–115. https://doi.org/10.1016/j.is.2014.07.006.

Jahroh, S., Atmakusuma, J., Harmini, Fadillah, A., 2020. Comparative analysis of dairy farming management and business model between East Java and West Java, Indonesia. In: Jurnal Manajemen & Agribisnis, Vol. 17, pp. 96–107.

Kamilaris, A., Anton, A., Blasi, A.B., Boldú, F.X.P., 2018. Assessing and mitigating the impact of livestock agriculture on the environment through geospatial and big data analysis. Int. J. Sustain. Agric. Manag. Inform. 4 (2) https://doi.org/10.1504/ijsami.2018.094809.

Kementerian Kelautan dan Perikanan - Ministry of Marine Affairs and Fisheries Republic of Indonesia, 2024. Produksi Perikanan. https://statistik.kkp.go.id/home.php?m=total&i=2#panel-footer (accessed December 29, 2021).

Kementerian Perdagangan Republik Indonesia - Ministry of Trade Republic of Indonesia, 2010. Profil Komoditas: Susu Sapi.

Köksal, Ö., Tekinerdogan, B., 2019. Architecture design approach for IoT-based farm management information systems. Precis. Agric. 20, 926–958. https://doi.org/10.1007/s11119-018-09624-8.

Krisnawijaya, N. Nyoman Kutha, Tekinerdogan, B., Catal, C., Tol, R.V.D., 2022. Data analytics platforms for agricultural systems: a systematic literature review. Comput. Electron. Agric. 195 https://doi.org/10.1016/j.compag.2022.106813.

Kruize, J.W., Wolfert, J., Scholten, H., Verdouw, C.N., Kassahun, A., Beulens, A.J.M., 2016. A reference architecture for farm software ecosystems. Comput. Electron. Agric. 125, 12–28. https://doi.org/10.1016/j.compag.2016.04.011.

Lee, C., Wang, Y.-J., 2020. Development of a cloud-based IoT monitoring system for fish metabolism and activity in aquaponics. Aquac. Eng. 90 https://doi.org/10.1016/j.aquaeng.2020.102067.

Liu, J., Zhang, L., Li, Y., Liu, H., 2023. Deep residual convolutional neural network based on hybrid attention mechanism for ecological monitoring of marine fishery. Eco. Inform. 77 https://doi.org/10.1016/j.ecoinf.2023.102204.

López-Riquelme, J.A., Pavón-Pulido, N., Navarro-Hellín, H., Soto-Valles, F., Torres-Sánchez, R., 2017. A software architecture based on FIWARE cloud for precision agriculture. Agric. Water Manag. 183, 123–135. https://doi.org/10.1016/j.agwat.2016.10.020.

McConnell, D.J., Dillon, J.L., 1997. Farm management for asia: a systems approach. In: FAO Farm Systems Management Series. Food and Agriculture Organization of the United Nations, Rome.

Meena, K., Sujatha, J., 2019. Reduced time compression in big data using MapReduce approach and Hadoop. J. Med. Syst. 43, 1–12. https://doi.org/10.1007/s10916-019-1369-3.

Meeradevi, M.A., Supreetha, M.R., Mundada, Pooja, J.N., 2019. Design of a Smart Water-Saving Irrigation System for Agriculture Based on a Wireless Sensor Network for Better Crop Yield. In: Lecture Notes in Electrical Engineering. Springer, Singapore, pp. 93–104.

Mourik, S.V., et al., 2021. Introductory overview: systems and control methods for operational management support in agricultural production systems. Environ. Model. Softw. 139, 15. https://doi.org/10.1016/j.envsoft.2021.105031.

Ouafiq, E.M., Saadane, R., Chehri, A., 2019. Data management and integration of low power consumption embedded devices IoT for transforming smart agriculture into actionable knowledge. Agriculture 12 (3), 329.

Pääkkönen, P., Pakkala, D., 2015. Reference architecture and classification of technologies, products and Services for big Data Systems. Big Data Res. 2, 166–186. https://doi.org/10.1016/j.bdr.2015.01.001.

Perakis, K., Lampathaki, F., Nikas, K., Georgiou, Y., Marko, O., Maselyne, J., 2020. CYBELE – fostering precision agriculture & livestock farming through secure access to large-scale HPC enabled virtual industrial experimentation environments fostering scalable big data analytics. Comput. Netw. 168 https://doi.org/10.1016/j.comnet.2019.107035.

Periyadi, G.I., Hapsari, Z. Wakid, Mudopar, S., 2020. IoT-based guppy fish farming monitoring and controlling system. Telecommun. Comput. Electron. Contr. 18, 1538–1545. https://doi.org/10.12928/TELKOMNIKA.v18i3.14850.

Rawat, R., Yadav, R., 2021. Big data: big data analysis, issues and challenges and technologies. In: IOP Conference Series: Materials Science and Engineering, 1022, pp. 1–9. https://doi.org/10.1088/1757-899X/1022/1/012014.

Righi, R.D.R., Goldschmidt, G., Kunst, R., Deon, C., Costa, C.A.D., 2020. Towards combining data prediction and internet of things to manage milk production on dairy cows. Comput. Electron. Agric. 169, 1–13. https://doi.org/10.1016/j.compag.2019.105156.

Robertson, S., Robertson, J., 2012. Mastering the Requirements Process: Getting Requirements Right. Addison-wesley.

Runeson, P., Höst, M., 2009. Guidelines for conducting and reporting case study research in software engineering. Empir. Softw. Eng. 14, 131–164. https://doi.org/10.1007/s10664-008-9102-8.

Rutten, C.J., Velthuis, A.G.J., Steeneveld, W., Hogeveen, H., 2013. Invited review: sensors to support health management on dairy farms. J. Dairy Sci. 96, 1928–1952. https://doi.org/10.3168/jds.2012-6107.

Saiz-Rubio, V., Rovira-Más, F., 2020. From smart farming towards agriculture 5.0: a review on crop data management. Agronomy 10, 1–21. https://doi.org/10.3390/agronomy10020207.

Salma, C.A., Tekinerdogan, B., Athanasiadis, I.N., 2017. Domain-driven design on big data systems based on a reference architecture. In: Software Architecture for Big Data and the Cloud, 1st ed. Elsevier Inc., pp. 49–68

Santana, F.S., et al., 2014. A reference process for automating bee species identification based on wing images and digital image processing. Eco. Inform. 24, 248–260. https://doi.org/10.1016/j.ecoinf.2013.12.001.

Sant'Ana, D.A., et al., 2022. Computer vision system for superpixel classification and segmentation of sheep. Eco. Inform. 68 https://doi.org/10.1016/j.ecoinf.2021.101551.

Saranya, C.P., Nagarajan, N., 2020. Efficient agricultural yield prediction using metaheuristic optimized artificial neural network using Hadoop framework. Soft. Comput. 24 (16), 12659–12669. https://doi.org/10.1007/s00500-020-04707-z.

Senff, P., Partelow, S., Indriana, L.F., Buhari, N., Kunzmann, A., 2018. Improving pond aquaculture production on Lombok, Indonesia. Aquaculture 497, 64–73. https://doi.org/10.1016/j.aquaculture.2018.07.027.

Silva, S., Alçada-Almeida, L., Dias, L.C., 2014. Development of a web-based multi-criteria spatial decision support system for the assessment of environmental sustainability of dairy farms. Comput. Electron. Agric. 108, 46–57. https://doi.org/10.1016/j.compag.2014.06.009.

Statistics Indonesia - Badan Pusat Statistik (BPS), 2024a. Produksi Susu Perusahaan Sapi Perah 2018–2020. https://www.bps.go.id/indicator/24/376/1/produksi-susu-perusahaan-sapi-perah.html.

Statistics Indonesia - Badan Pusat Statistik (BPS), 2024b. Populasi Sapi Perah menurut Provinsi (Ekor), 2018–2020. https://www.bps.go.id/indicator/24/470/1/populasi-sapi-perah-menurut-provinsi.html.

Subahi, A.F., Bouazza, K.E., 2020. An intelligent IoT-based system design for controlling and monitoring greenhouse temperature. IEEE Access 8, 125488–125500. https://doi.org/10.1109/access.2020.3007955.

Swain, M., Singh, R., Thakur, A.K., Gehlot, A., 2020. A machine learning approach of data Mining in Agriculture 4.0. Int. J. Emerg. Technol. 11 (1), 257–262.

Taneja, M., Jalodia, N., Byabazaire, J., Davy, A., Olariu, C., 2019. SmartHerd management: a microservices-based fog computing–assisted IoT platform towards data-driven smart dairy farming. Softw. Pract. Experience 49, 1055–1078. https://doi.org/10.1002/spe.2704.

Taneja, M., Byabazaire, J., Jalodia, N., Davy, A., Olariu, C., Malone, P., 2020. Machine learning based fog computing assisted data-driven approach for early lameness detection in dairy cattle. Comput. Electron. Agric. 171 https://doi.org/10.1016/j.compag.2020.105286.

Tekinerdogan, B., Bilir, S., Abatlevi, C., 2005. Integrating platform selection rules in the model driven architecture approach. In: Aßmann, U., Aksit, M., Rensink, A. (Eds.), Model Driven Architecture. MDAFA 2004, MDAFA 2003, Lecture Notes in Computer Science, vol. 3599. Springer, Berlin, pp. 159–173.

The Interdisciplinary Research and Education Fund (INREF), 2024. Smart-In-Ag Project Description. https://smart-in-ag.com/project-description/ (accessed December 29, 2021).

Triantafyllou, A., Sarigiannidis, P., Bibi, S., 2019. Precision agriculture: a remote sensing monitoring system architecture. Information 10 (11). https://doi.org/10.3390/info10110348.

Tummers, J., Kassahun, A., Tekinerdogan, B., 2021. Reference architecture design for farm management information systems: a multi-case study approach. Precis. Agric. 22, 22–50. https://doi.org/10.1007/s11119-020-09728-0.

United Nations (UN), 2024. Goal 2: Zero Hunger. https://www.un.org/sustainabledevelopment/hunger/ (accessed December 29, 2021).

van Geest, M., Tekinerdogan, B., Catal, C., 2021. Design of a reference architecture for developing smart warehouses in industry 4.0. Comput. Ind. 124 https://doi.org/10.1016/j.compind.2020.103343.

Villa-Henriksen, A., Edwards, G.T.C., Pesonen, L.A., Green, O., Sørensen, C.A.G., 2020. Internet of things in arable farming: implementation, applications, challenges and potential. Biosyst. Eng. 191, 60–84. https://doi.org/10.1016/j.biosystemseng.2019.12.013.

Wageningen University & Research, 2024a. INREF Projects (accessed December 29, 2021).

Wageningen University & Research, 2024b. Project: Smart Indonesian Agriculture (Smart-In-Ag). https://smart-in-ag.com/ (accessed December 29, 2021).

Wageningen University & Research (WUR), 2024. Interdisciplinary Research and Education Fund (INREF). https://www.wur.nl/en/research-results/research-programmes/cross-wur-programmes/inref.htm (accessed December 29, 2021).

Wang, Y., Li, S., Zhang, H., Liu, T., 2022. A lightweight CNN-based model for early warning in sow oestrus sound monitoring. Eco. Inform. 72 https://doi.org/10.1016/j.ecoinf.2022.101863.

Wolfert, S., Goense, D., Sørensen, C.A.G., 2014. A future internet collaboration platform for safe and healthy food from farm to fork. In: Proceedings of the Annual SRII Global Conference, IEEE, pp. 266–273. https://doi.org/10.1109/SRII.2014.47.

Wolfert, S., Ge, L., Verdouw, C., Bogaardt, M.-J., 2017. Big data in smart farming – a review. Agric. Syst. 153, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023.

Yang, J., Liu, M., Lu, J., Miao, Y., Hossain, M.A., Alhamid, M.F., 2017. Botanical internet of things: toward smart indoor farming by connecting people, plant, data and clouds. Mobile Netw. Appl. 23 (2), 188–202. https://doi.org/10.1007/s11036-017-0930-x.

Yang, X., Zhang, S., Liu, J., Gao, Q., Dong, S., Zhou, C., 2021. Deep learning for smart fish farming: applications, opportunities and challenges. Rev. Aquac. 13, 66–90. https://doi.org/10.1111/raq.12464.

Zamora-Izquierdo, M.A., Santa, J., Martínez, J.A., Martínez, V., Skarmeta, A.F., 2019. Smart farming IoT platform based on edge and cloud computing. Biosyst. Eng. 177, 4–17. https://doi.org/10.1016/j.biosystemseng.2018.10.014.

Zhang, T., et al., 2024. Fully automatic system for fish biomass estimation based on deep neural network. Eco. Inform. 79 https://doi.org/10.1016/j.ecoinf.2023.102399.