Springer Optimization and Its Applications 184

Dionysis D. Bochtis · Claus Grøn Sørensen · Spyros Fountas · Vasileios Moysiadis · Panos M. Pardalos *Editors*

Information and Communication Technologies for Agriculture—Theme III: Decision

Springer

Springer Optimization and Its Applications

Volume 184

Aims and Scope

Optimization has continued to expand in all directions at an astonishing rate. New algorithmic and theoretical techniques are continually developing and the diffusion into other disciplines is proceeding at a rapid pace, with a spot light on machine learning, artificial intelligence, and quantum computing. Our knowledge of all aspects of the field has grown even more profound. At the same time, one of the most striking trends in optimization is the constantly increasing emphasis on the interdisciplinary nature of the field. Optimization has been a basic tool in areas not limited to applied mathematics, engineering, medicine, economics, computer science, operations research, and other sciences.

The series **Springer Optimization and Its Applications (SOIA)** aims to publish state-of-the-art expository works (monographs, contributed volumes, textbooks, handbooks) that focus on theory, methods, and applications of optimization. Topics covered include, but are not limited to, nonlinear optimization, combinatorial optimization, continuous optimization, stochastic optimization, Bayesian optimization, optimal control, discrete optimization, multi-objective optimization, and more. New to the series portfolio include Works at the intersection of optimization and machine learning, artificial intelligence, and quantum computing.

Volumes from this series are indexed by Web of Science, zbMATH, Mathematical Reviews, and SCOPUS.

More information about this series at http://www.springer.com/series/7393

Dionysis D. Bochtis • Claus Grøn Sørensen
Spyros Fountas • Vasileios Moysiadis
Panos M. Pardalos
Editors

Information and Communication Technologies for Agriculture—Theme III: Decision

Springer

Editors
Dionysis D. Bochtis
Institute for Bio-Economy and
Agri-Technology (iBO)
Centre for Research and Technology
Hellas (CERTH)
Thessaloniki, Greece

Claus Grøn Sørensen
Department of Electrical
and Computer Engineering
University of Aarhus
Aarhus N, Denmark

Spyros Fountas
Department of Resources Management
and Agricultural Engineering
Agricultural University of Athens
Athens, Greece

Vasileios Moysiadis
Institute for Bio-Economy and
Agri-Technology (iBO)
Centre for Research and Technology
Hellas (CERTH)
Thessaloniki, Greece

Panos M. Pardalos
Department of Industrial
and Systems Engineering
University of Florida
Gainesville, FL, USA

ISSN 1931-6828 ISSN 1931-6836 (electronic)
Springer Optimization and Its Applications
ISBN 978-3-030-84151-5 ISBN 978-3-030-84152-2 (eBook)
https://doi.org/10.1007/978-3-030-84152-2

Mathematics Subject Classification: 68U35; 90B50

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Every action is based on a decision, while every decision requires acquisition and processing of the information available. The first book of the series on ICT for Agriculture (Theme I: Sensors) focuses on the data acquisition with the use of sensors, while the second one (Theme II: Data) focuses on data processing and utilization. This book here, the third of the series (Theme III: Decision) focuses on the transformation of the collected information into valuable decisions. The successful transition to the new digitized era of agriculture requires the implementation of elaborated decision-making for farmers, processors, and distribution channels of agricultural products. The focus is how to better use digital technologies to reduce cost, inputs, and time, and be more efficient and transparent. The book consists of 14 chapters relevant, and complementary to each other, to agricultural production and related products' distribution. Contributions are grouped in three distinct sections, namely: Value Chain, Primary Production, and Environment, based on the thematic area of each individual chapter, as well as the subject of the decision at hand.

The first section of the book is dedicated to decisions in the value chain of agricultural products. Value chain is a sequence of processes that add value to a product or commodity. The processes in the value chain are interconnected and cover the entire production cycle, including transportation, storage, processing, as well as promotion and marketing of products. Each process involves several stakeholders and products rendering a value chain a complex system that requires specialized organizing to make proper decisions. The application of ICT finds suitable grounds in agricultural value chain, improving the effectiveness and sustainability of production by the utilization of innovative techniques that can be integrated in traditional production processes. The key to the integrated processing of the value chain lays to the abundance of the available data. The proper manipulation of different types of data offers a promising potential for the improvement of production processes through custom in-field applications. Farm Information Management Systems are the means through which the data collected across the value chain can be processed for the extraction of useful and applicable information.

The success in the development of smart farming applications is a challenging task, and it is usually impeded by the lack of interoperability between the involved systems due to different technological architectures, standards, and communication protocols. The data collected in the agricultural sector derive from a variety of heterogenous sources. Additionally, there are not widely accepted standards for data collection and processing while there is not a wide application of interoperability mechanisms that allow for the connection of existing models. This fact highlights the lack of connectivity for the exchange and integration of the collected data, which is a matter of utmost importance in the development of smart farming applications. Towards that direction the chapter entitled "**Agricultural Information Model**" addresses these issues through the presentation of the respective model that was developed by the H2020 DEMETER project. The chapter elaborates on the design of an information model adhering to a layered and modular approach to develop a suite of ontologies that are in accordance with best practices by exploiting existing standards and well-scoped models. Aim is to establish alignments between these models to achieve interoperability and integration of existing data. The model is designed in an expandable manner, meaning that new concepts can be incorporated to the models, depending on the emerging needs. Therefore, the model is adjustable to farmers and stakeholders' preferences, through the realization of smart farming solutions that connect different systems and platforms of the agri-food sector. More specifically, the model acts as a mediator that supports the decision-making process by facilitating different systems in data exchange and providing access to different sources of data. The model is designed in three layers (meta-model layer, cross-domain layer, and domain-specific layer) while it is easy to expand and maintain. The chapter presents in detail the information model development process including the collected requirements and a description of the implementation details and the aspects of semantic interoperability that were investigated. Examples of the use of the model are also presented for the support of developers and users.

The increase in the consumption of protein products, along with the adverse environmental impacts that their production is connected to, calls for immediate actions with respect to the respective value chains. With respect to the value chain of beef cattle, Precision Livestock Farming (PLF) as well as Management Information Systems (MIF) are used to increase the efficiency of operations through improved decision-making. Moreover, the use of such systems enables the successful compliance to food quality and sustainability standards adding value to the final product, increasing at the same time consumers' trust. The integration of Internet of Things (IoT) technologies ensures the interoperability between the different points in the agri-food value chains. In the case of beef production, these points include all the stages of production (breeding, growing, finishing, slaughtering, meat processing, transportation, logistics, etc.) creating a value chain that covers the entire agroecosystem. Within this value chain, a vast number and forms of data can be gathered, analyzed, and stored, creating the need for systems for the assessment of the collected information.

Additionally, an increasing demand for improved traceability within the beef value chain is observed, leading to the eventual adoption of innovative technologies, towards a direction that creates a new data flow that needs to be handled properly. The proper processing of the collected data can offer a variety of benefits including the improvement of the livestock welfare, the increase in quality of the products, the reduction of waste, and the satisfaction of consumers' demands. Addressing the above issues, the chapter "**Development of a Framework for Implementing IOT-A on the Beef Cattle Value Chain**" presents the implementation of the IoT paradigm on the beef cattle value chain. The corresponding value chain is thoroughly described considering its main stages, stakeholders, processes, and the informational flow. In sequence, the requirements and services needed for implementing IoT were presented. A model in the form of IoT-A is presented. The presented model and methodology could be utilized to other agri-food sectors as it fulfills the identified requirements, considering also basic interoperability and security aspects.

Closing this section, the chapter "**Food Business Information Systems and Software in Western Greece**" attempts to examine the adoption of new information technologies by traditional food enterprises, having as a case study of the region of Wester Greece. The chapter, after providing an extensive review of the last 30 years' literature, presents the mapping and analysis of the software and food business information systems used by food businesses in Western Greece with the purpose to improve the management of their functioning. The outcome was that such systems are referred to human resources, in accounting and finance, in marketing and sales, in production and in operational functions. Most enterprises use Enterprise Resource Management (ERP) systems. Moreover, a few businesses use tailor-made packages mainly to manage production, human resources, and operational functions, while there is a growth prospective in operational, human recourses, and production functions to apply sophisticated software packages in the short-term future.

The Primary Production section elaborates on the decision-making for the improvement of the processes taking place within the farm, under the implementation of Information and Communication Technologies. The first chapter, entitled "**From Precision Agriculture to Agriculture 4.0: Integrating ICT in Farming**" focuses on the integration of Information and Communication Technologies in farming and the transformation of conventional agriculture to meet the need of increased effectiveness of agricultural practices. The chapter offers an overview of the basic technologies and provides examples of application in the agricultural sector alongside to the challenges of the Agriculture 4.0 era. Finally, the issue of the adoption of innovative technologies by the farmers is elaborated since it is highly linked to their successful penetration in the agricultural production.

The monitoring of the cultivation is an integral part in any agricultural Decision Support System (DSS) for the collection of the required data. The evolution of aerial monitoring technology has made unmanned aerial vehicles (UAVs or drones) a suitable solution for crop observation compared to satellite imagery. UAVs offer full coverage of fields upon request while they deliver on-the-spot images for further processing. UAVs are more flexible to user preferences; however, their limitations call for effective planning during their operation. Such limitations include their low autonomy as they are mostly battery-operated devices along with the need for long distance or heavy duty traveling for the execution of the required task (as for example mapping or spraying). The above constitute the routing process very important for a successful and efficient completion of the required missions. Providing further insight on the subject the chapter "**On the Routing of Unmanned Aerial Vehicles (UAVs) in Precision Farming Sampling Missions**" focuses on the use of UAVs within smart farming applications. The chapter presents the main applications of UAVs in the sector. Moreover, a profound examination of the methods of safe and optimal drones routing are presented along with the most widespread algorithms developed for that purpose. Lastly, representative applications of these algorithms are elaborated, drawing conclusions for further interventions and improvement.

Contributing to the monitoring and documentation in precision agriculture management systems, depth cameras have gained popularity during the last years. They are mostly utilized for the three-dimensional (3D) reconstruction of objects both in indoor and outdoor settings. As agricultural environments comprise of complex elements (e.g., trees or plants), the use of depth cameras, for example, to create models required for simulation purposes, is a significant challenge. Especially in outdoor environments that involve different object structures and uncertain conditions, the depth information collected may vary significantly. The chapter "**3D Scenery Construction of Agricultural Environments for Robotics Awareness**" examines the various technologies used by depth cameras and demonstrates indicative applications both in indoor and outdoor agricultural environments. A 3D orchard reconstruction is presented deriving from the processing of point clouds from Red Green Blue Depth (RGB-D) images collected in real fields. The orchard environment is simulated in Gazebo using the point cloud samples of trees collected by an UGV. This approach is considered significantly useful for the simulation and evaluation of the navigation of robotic systems even though tree characteristics' information (such as volume or height of tree canopies) can be extracted as well. The implementation of the approach shows a promising potential on environment simulation which can be utilized in several robotic applications despite any limitations such as the camera's limitations.

Aiming at controlling specific varieties of weeds, the chapter "**A Weed Control Unmanned Ground Vehicle Prototype for Precision Farming Activities: The Case of Red Rice**" attempts to address the damage of the spread of the red rice (*Oryza sativa f. spontanea*), which is a wild variety of rice, on the production of commercial rice. Red rice, with faster growing rate and better resilience to weather conditions compared to the ordinary variety, is characterized as a weed. Its population shows an increase on an annual basis, resulting in significant yield losses for farmers. What constitutes this species harmful for production, even though it is edible, is the morphology of its seeds. Being thin, covering entirely the plant, their collection is impossible while they tend to scatter on the field during harvesting getting mixed with the regular crops.

Rice is characterized as one of the most widely produced crops and constitutes a basic food source worldwide. In this manner, its effective production is of utmost importance. Its infestation with red rice is a considerable challenge since the commonly used conventional and chemical-based solutions cannot tackle the weed as it is genetically similar with commercial rice. Considering the above, the implementation of a mechanical solution is only accepted for this case. Towards that direction, the chapter presents the development of a prototype robot that handles the weed in field by applying herbicides "from above" at the stage of production when red rice is much higher than the conventional rice. The robot (an unmanned ground vehicle—UGV) carries a specially designed rod mechanism for the application of the herbicide only to the upper part of the weed. For that purpose, the rod carries a sponge that is drenched with herbicides while a sensor-based mechanism ensures the application of the agrochemical to the right plant elements. The study highlights the importance of innovative technologies in precision farming, that increase productivity, adding value to conventional farming processes while reducing the adverse impacts.

Monitoring of agricultural practices involves the collection of a variety of input data in multiple forms. The introduction of agricultural Decision Support Systems (DSSs) focuses on the organization and utilization of the collected information aiming at the optimization of the farmer's decision-making process. Additionally, DSSs can serve towards the sustainable transformation of modern agriculture by setting standards for environmentally improved processes and effective application of materials. One of the most crucial threats of agricultural production are weeds. Weeds compete with crops leading to reduced yield and poor quality, thus increasing production cost since further processing is required. Attempting to address that challenge, the chapter entitled "**Decision-Making and Decision Support System for a Successful Weed Management**," elaborates on the elements that affect decision-making and should be considered in the development of a DSS for weed management. The first step towards an effective weed control is to examine the environmental factors and the agricultural processes that favor the emergence of weed. Aim is to discover development patterns and predict weed growth for timely intervention suggestions from a DSS. Moreover, for the optimization of decision-making, the biological characteristics of weeds variates must be investigated. A thorough knowledge of the weed biology is considered as a prerequisite for the implementation of any weed control strategy. The chapter highlights the importance of the dissemination of such systems to the end-users/farmers. The trust of farmers in these systems is necessary to adopt them.

The chapter entitled "**Zephyrus: Grain Aeration Strategy Based on the Prediction of Temperature and Moisture Fronts**" presents a new aeration control strategy (called Zephyrus) based on the prediction of air velocity changes and changes in temperature and moisture contours, while air is passed through grain bulks. The presented control approach can be used with different aeration system designs and automatically adjusting its set points according to the geographic region and particular season.

The use of smart livestock technologies for decision-making has proved to benefit farmers by increasing the productivity and profitability of farming processes. The real-time farm scenarios that are examined allow for accurate and timely interventions providing several benefits for the animals and the farmer. Nevertheless, the monitored parameters are usually inaccurate and incomplete, while conflicted data may also be generated, constituting classical logic inadequate for their assessment. In that case, the use of non-classical logic in data processing may prove beneficial as it facilitates non-intrusive assessment, without the disturbance of the animals. The evolution of ICT tools has introduced the utilization of online mathematical identification techniques in smart farming. Such methods, can provide online estimations for the unknown parameters, offering models that can adapt to most of the biological processes, though the biological understanding of the causal mechanisms is not feasible due to their complexity and non-linearity. To deal with the emerging uncertainties, the employment of non-classic logic may allow for solutions for the managerial processes and the biological response models in smart farming. To obtain a complete picture of the state of the animals, a real-time continuous monitoring scheme is preferable to the view-in-time assessment that is used traditionally. In this case, farmers and animals are benefited by alerts that facilitate timely and targeted interventions.

The chapter entitled "**Decision-Making Applications on Smart Livestock Farming**" presents the fundamentals of smart livestock farming, introducing the managerial processes that apply non-classic logic and data mining through the demonstration of models implemented on farm actions. Additionally, tools for creating decision-making applications in smart livestock farming are presenting as, for example, paraconsistent logic applications and applications that use machine learning. The introduction of smart livestock farming leads to the automation of many procedures on the spot. The collection of information can be achieved from a variety of sources and different sensors resulting in powerful decision-making tools since farmers can use the information in conjunction with their observations and expertise. Such tools will continue to evolve along with the improvement of the respective technological elements.

The need for increase in food production and the use of resources and energy, in a sustainable manner, has driven the utilization of the recent innovations in ICT for the efficient assessment of the locally available resources (e.g., water, land, solar, and energy). Also considering the need to manage the environmental impact of agricultural production, the development of "climate-smart" solutions is becoming more important. Climate-smart systems that utilize ICT technologies bridge the gap between the fragile agricultural ecosystems and the non-agricultural processes that are used for their assessment. Nevertheless, such agri-environmental systems are characterized by high complexity and uncertainty, thus the establishment of new methodologies is challenging. The aim is to accomplish the sustainable cooperation of human-operated technologies and environment through the modeling of the diversified set of multidisciplinary processes that are involved. The last section of this book is dedicated to the development of innovative decision applications that also consider the protection of environment, recognizing its importance in the preservation and considerate use of resources, along with the mitigation of adverse impact that are related to agricultural production.

The emergence of the new generation of ICT technologies has promoted the development of tools for the support of decisions with respect to the planning and operation of complicated agricultural systems. However, the holistic management of these multifactorial systems requires the evaluation of the complex interactions that arise, as their statistical or structural analysis is not sufficient. To address the relationships that are created between the sub-processes that coexist, the development of predictive models is becoming very important. The amount of data collected increases exponentially, widening the gap with the potential of large-scale and long-term decisions. To bridge that gap, process model engineering of agri-environmental systems works towards the integration of the well-established and innovative frameworks. The complexity of the agricultural processes, that involves a variety of different disciplines, calls for the development of integrated models and methodologies. For the study of a variety of complex agri-environmental processes, the method of Programmable Process Structures (PPS) has been used in the past. The introduction of PPS has been driven by the need for the dynamic simulation of agricultural environmental processes that require the use of easily modifiable, extensible, and connectable models that represent the functional and structural characteristics in a unified manner. In the chapter entitled "**Programmable Process Structures of Unified Elements for Model-Based Planning and Operation of Complex Agri-environmental Processes**", the application of the PPS methodology is presented with the demonstration of three examples of increasing complexity, showing how PPS handles the structure and functionalities of agri-environmental processes. First, an overview of the existing functional and structural approaches is presented, while the PPS method is described along with its innovative characteristics. Moreover, through the demonstrated examples, the most important features of PPS are presented and more specifically its wide applicability in different process systems that facilitates the integrated assessment of dynamics of the interacting systems and the environment.

Planning and modeling the assessment of agricultural practices can provide farmers with valuable information prior to the execution of a task, thus the prediction potential is one of the most important domains to focus on the penetration of ICT technologies in agriculture. However, the assessment of already applied practices is equally important before the implementation of state-of-the-art ICT applications, especially when their evaluation is imminent. For example, sugarcane burning is a usual practice in South American countries. The use of fire in agriculture has been connected to several adverse impacts related to the increase in temperature and the decrease in natural soil moisture which leads to reduced soil fertility due to increased soil compaction, loss of porosity, and erosion. Moreover, the consequences of sugarcane burning are more severe on human health as indicated by many studies. The objective of the chapter titled "**Monitoring and Estimation of Sugarcane Burning in Brazil, Using Linear Mixed Models**" is to evaluate the spatial and temporal distribution of fire incidences in Paranapanema region in Brazil. For this purpose, images from the Landsat satellites, numerical data (regarding area and fire incidences), and categorical data (terrain slope) was also used. A statistical model was used to evaluate data, making it possible to identify a decrease in fires in smooth undulating terrains, corresponding to 99.9% per year, being characterized by the increase in agricultural machinery in these areas. The results were quite positive as the proposed model could forecast for the next 6 years, in which timeframe, considering causes/effects, there would be a high decrease. It is interesting to note that the direct change in land use can be assessed with the use of ICT technologies (such as remote sensing images). In this manner, the areas that were converted to sugarcane plantations, even though were destined to other uses, can be easily monitored demonstrating the applicability of such systems.

Considering the assessment of agricultural practices, the evaluation of on-farm efficiency is gaining the attention with the increasing concern caused by the rising energy cost and environmental impacts due to machinery and input use. The intensification of agriculture has led to an increase in the consumption of non-renewable resources with the well-known environmental, economic, and societal consequences that are related to conventional agricultural practices and threaten energy security and autonomy. The only way to address the imminent impacts is to move to more sustainable agricultural practices. The detailed supervision of agricultural operations by applying innovative ICT technologies is an essential step towards the realization of sustainable agriculture.

To achieve optimum management of agricultural operations, several tools and simulation methodologies have been developed. In these tools, system optimization is attempted at various levels and with different assessment targets such as the minimization of financial or environmental cost, the reduction of greenhouse gas emissions, or the increase of productivity. The common ground in all the tools developed is the large number of input parameters and processes that characterizes each agricultural system. Therefore, performance assessment tools mostly aim at identifying the weak spots of agricultural supply chain. Among the various agricultural in-field operations, fertilization (in its various forms such as organic or chemical fertilization) has proved to be one of the highest energy consuming. The chapter "**A Decision Support System for Green Crop Fertilization Planning**" presents an evaluation of chemical fertilization, using two distinct case studies that include an annual food crop (industrial tomato) and a perennial energy crop (Arundo donax) to highlight the differences in the energy cost on an annual basis. In this study, the energy performance of the two crops is compared with the use of a decision support system providing decisions with respect to the optimal crop allocation considering a set of available fields. The results demonstrate a high variance in energy consumption among various crop, highlighting the need for custom assessment with the employment of operation management tools. Such tools can be developed in the context of integrated Farming Management Information Systems (FMISs), creating comprehensive tools and services for stakeholders. Within an FMIS, innovative ICT technologies can be utilized towards facilitating decision-making for increasing operation efficiency, improving, as an aftermath, the sustainability of agricultural processes through the minimization of input materials.

It is worth noting that increasing the innovation penetration is a common policy in complex and progressive environments such as agricultural systems. Nonetheless, the tendency for development is also driven by the emergence of environmental, economic, and social problems. Agricultural stakeholders must adapt to the new challenges to safeguard a sustainable and viable future. Towards that direction the agroecological movement aims to a more sustainable agriculture by attempting to increase the efficiency of agricultural processes in a sustainable and ethical manner. However, the transition from conventional to sustainable agricultural processes is not always easy, thus appropriate tools, that facilitate the design and analysis of biodiversity friendly production systems and practices, are required. Along with the relevant tools, a change in the farmers' perception, with respect to what constitutes a process as sustainable, is required. More specifically, farmers should understand the functions of an agroecosystem, while also improving their managerial skills. In this way, they will be able to make beneficial decisions, reaching their production goals while abiding by the standards and principles of agroecology. The above highlight the benefits of the farmers' education and training, with respect to the consequences of an action, its suitability with short- or long-term goals, its compatibility with other actions, etc. The use of ICT technologies can facilitate the learning process through the involvement of the farmer to the assessment of the various components and processes of an agroecosystem.

The chapter "**Knowledge Elicitation and Modeling of Agroecological Management Strategies**" attempts to examine, the management strategies used by farmers, at a level that allows for the simulation of the various operational management processes. The concept developed borrows from the Belief-Desire-Intention (BDI) theory that conceptualizes the sequential decision-making behavior of rational agents. Even though a variety of simulation approaches have been proposed for the assessment of agricultural systems, only a few succeed on working up to farm scale. The aim of the chapter is to facilitate the holistic examination of farm management aspects such as the planning of tasks, the goal-based adaptation to circumstances, and the proper allocation of resources in the context of the implementation of a production strategy. The authors' objective is to establish the guidelines for the management of agroecological systems to test and disseminate the knowledge.

Concluding, appropriate decisions are necessary in the entire value chain of agricultural products and concern the planning as well as the assessment of the consequences of the processes, as highlighted through the chapters of this book. Thus, decisions are required for the proper and effective planning, monitoring and execution of agricultural tasks, for the efficient minimization of inputs as well as the evaluation of the processes. The new era of digital agriculture calls for elaborate decisions that optimize the consequent actions. This book leads the way towards the final volume of the series, under the theme **Action** that focuses on the implementation of cutting-edge technologies on real-world applications.

Thessaloniki, Greece — Dionysis D. Bochtis
Aarhus N, Denmark — Claus Grøn Sørensen
Athens, Greece — Spyros Fountas
Thessaloniki, Greece — Vasileios Moysiadis
Gainesville, FL, USA — Panos M. Pardalos

Contents

Part I
Value Chain

Agricultural Information Model

Raul Palma, Ioanna Roussaki, Till Döhmen, Rob Atkinson, Soumya Brahma, Christoph Lange, George Routis, Marcin Plociennik, and Szymon Mueller

1 Introduction

Smart farming or precision agriculture refers to the adoption of various digital technologies in the agriculture domain aiming at automation and improved efficiency of various farming operations and processes considering several aspects (e.g., increased yield quality and quantity, reduced cost and necessary resources, reduced environmental footprint) [1, 2].

There is a worldwide trend towards the smart farming paradigm and particularly in Europe, where EU Member States have joined forces and signed a declaration of cooperation on "*A smart and sustainable digital future for European agriculture and rural areas*"[1] in April 2019, which acknowledges the potential of employing digital

[1] The declaration has been signed by 25 Member States. For details please refer to https://ec.europa.eu/digital-single-market/en/news/eu-member-states-join-forces-digitalisationeuropean-agriculture-and-rural-areas

R. Palma (✉) · S. Brahma · M. Plociennik · S. Mueller
Poznan Supercomputing and Networking Center, Poznań, Poland
e-mail: rpalma@man.poznan.pl; sbrahma@man.poznan.pl; marcinp@man.poznan.pl; mael@man.poznan.pl

I. Roussaki · G. Routis
Institute of Communication and Computer Systems, Athens, Greece
e-mail: ioanna.rousski@cn.ntua.gr; georgios.routis@cn.ntua.gr

T. Döhmen · C. Lange
Fraunhofer Institute for Applied Information Technology FIT, Sankt Augustin, Germany
e-mail: till.doehmen@fit.fraunhofer.de; christoph.lange-bever@fit.fraunhofer.de

R. Atkinson
Open Geospatial Consortium Europe Technologielaan, Leuven, Belgium
e-mail: ratkinson@ogc.org

D. D. Bochtis et al. (eds.), *Information and Communication Technologies for Agriculture—Theme III: Decision*, Springer Optimization and Its Applications 184, https://doi.org/10.1007/978-3-030-84152-2_1

technologies in the agricultural sector and rural areas and supports the setting up of common data spaces. In this framework, the EC has issued the European Strategy for Data[2] in February 2020, which foresees the roll-out of common European data spaces in nine strategic sectors, including agriculture. This strategy considers data to be "*... one key element to enhance the sustainability performance and competitiveness of the agricultural sector. Processing and analyzing production data, especially in combination with other data on the supply chain and other types of data, such as earth observation or meteorological data, allows for precise and tailored application of production approaches at farm level.*" The production data are collected by Farm Management Systems (FMS) or similar applications during farm operations that will be coupled with open data (such as satellite images, weather data, soil maps that are for public use) to maximize the potential value of knowledge generated and introduce new opportunities for monitoring and optimizing the consumption of natural resources.

As mentioned in the concept note[3] issued by the EC in the framework of the Expert workshop on a Common European Agricultural Data Space that took place in September 2020, EC aims to "*set up, populate and operate a secure and trusted dataspace in order to enable the agriculture sector to access data relevant for agricultural production.*" However, before this common data space can be established, there are several technical challenges that need to be addressed jointly. Two major ones among these include the establishment of a common agriculture data model able to represent the wealth of information generated and used in the agri-food sector and the delivery of the necessary interoperability mechanisms. These are very difficult to address, given the wide heterogeneity of data models and semantics used to represent data in the agri-food domain, as well as the lack of related standards to dominate this space, that make full-scale interoperability almost impossible to achieve. The situation is further complicated by the growing number of stakeholder groups involved in the agricultural activities, and the increasing volume of data being collected by different devices and produced by a wealth of different systems and platforms. This chapter aims to present in detail the DEMETER Agricultural Information Model (AIM) that has been delivered by the H2020 DEMETER project (857202) supported by the European Union. AIM builds upon a thorough analysis of both the related state of the art and the state of the practice. It is driven by the elicited stakeholder requirements and, of course, by the DEMETER vision for the creation of a common agricultural data space, where semantic interoperability is ensured. AIM aims to establish the basis of a common agricultural data space and enable the interoperation of different systems, potentially from different vendors, and the

[2] "A European strategy for data," Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM/2020/66 final, URL: https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=COM:2020:66:FIN

[3] http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68838

analysis of data produced by those systems in an integrated manner in order to make economically and environmentally sound decisions.

More specifically, the rest of the chapter is structured as follows: Section 2 provides an analysis of the state of the art (and state of the practice) on related data models and semantic interoperability mechanisms that exist in the domain of Smart Agrifood. Initially, general data models that are widely used are presented, then specific ontologies are elaborated upon and, finally, mechanisms that allow for semantic interoperability are discussed. This analysis, together with the technical requirements (presented in the following section), drives the design and development of AIM.

Section 3 gives an overview of the technical requirements extracted by the project, based on input received from the stakeholders involved and driven by the DEMETER vision. This is a list of specific technical requirements that AIM needs to address, as well as requirements regarding the interoperability with existing systems and ontologies, including the mapping of these data models to AIM. Section 4 presents the design of AIM:

- Section 4.1 presents the core meta-model, which is based on the NGSI-LD information model [3].
- Section 4.2 presents the cross-domain ontology, i.e., the set of generic models that aim at providing common definitions that are not agri-food-specific and that need to be handled by AIM, avoiding conflicting or redundant definitions of the same classes at the domain-specific layer.
- Section 4.3 presents the domain-specific ontologies developed for AIM, which are mainly applicable for the agri-food domain and model information such as crops, animals, agricultural products as well as farms and farmers.

Section 5 elaborates on the approach enabling semantic interoperability between AIM and several existing ontologies and dominant agri-food data models detailing the semantic mapping of these to AIM.

Section 6 presents the implementation of the AIM. More specifically, it presents the implementations for the different AIM layers, highlighting the critical design and implementation choices made, the mappings implemented, and the various tools used during the implementation process.

Section 7 presents the methodology of profiling and demonstrates its usefulness on building AIM. Section 8 completes the circle elaborating on use cases where the AIM has been adopted in operational environments engaged in the pilots of the DEMETER project. Finally, Section 9 concludes the document presenting also future work towards the final version of AIM.

2 Related Work

The DEMETER approach to creating Agriculture Integrated Model can be related to advancements in several areas of scientific work: general data models, specific ontologies, and vocabularies as well as established mechanisms for semantic interoperability. On a broader scale, general data management solutions, such as FIWARE NGSI/NGSI-LD (Next Generation Sensors Initiative—Linked Data), are proposed for use in cross-domain use cases. NGSI is a protocol developed by OMA (Open Model Alliance Ltd.) to manage Context Information about entities: their characteristics and relationships with other entities. Context Information Manager (CIM) is responsible for providing services for exchanging context information (e.g., registration, discovery, notification) between interested parties. FIWARE NGSI[4] is a binding of OMA NGSI-9 and NGSI-10 abstract interfaces for CIM. Updates to the protocol provide further benefits: NGSIv2[5] was designed with RESTful (Representational State Transfer) principles in mind and provide support for JSON (Java Script Object Notation) data format. The extension of the NGSIv2 format to incorporate advances of Linked Data approach is ETSI's (European Telecommunications Standards Institute) NGSI-LD format [4].

NGSI is a protocol that was developed in order to manage Context Information. A Context Information Model (CIM) [5] is a platform or a system—also known as Context Broker that provides the following services: context information registry, discovery, publication, mediation, modification, or notification. An entity consists of a set of characteristics that describe it, including its dynamic state. It also consists of other entities with which it has defined relationships, and the nature of those relationships.

NGSI-LD [6, 7], is an OMA NGSI information model, known as the evolution of NGSI, in order to better support linked data (entity's relationship), property graphs and semantics, so that it can benefit the capabilities offered by JSON-LD (Java Script Object Notation—Linked Data). The NGSI-LD consists of two levels; the first one is the foundation classes, which correspond to the core meta-model, and cross-domain ontology and the second is the domain-specific ontologies or vocabularies.

FIWARE [8, 9] Agrifood data models enable data portability for different applications in the domains of Smart Cities and Smart Agrifood. They are used with FIWARE NGSI 2 and NGSI-LD. FIWARE Agrifood data model represents a standard format which provides support to develop FIWARE solutions in the Smart Agrifood domain. By adopting a standard model, there is improvement in the standardization of information coming from IoT Networks that includes IoT (Internet of Things) sensors, wearables, GPS (Global Positioning System) services, UAVs (Unmanned Aerial Vehicles), robots, and farm machines. It increases the uniformity and interoperability of the data in the FIWARE technologies and ecosystem's applications. FIWARE Data Models that exist in the Agrifood field are the

[4] https://fiware-orion.readthedocs.io/en/1.8.0/user/walkthrough_apiv1/index.html

[5] http://fiware.github.io/specifications/ngsiv2/stable/

following: AgriApp, AgriCrop, AgriFarm, AgriGreenhouse, AgriParcel, AgriParcelOperation, AgriParcelRecord, AgriPest, AgriProductType, AgriSoil, Animal, WeatherObserved, WeatherForecast.

Saref4agri [10] is an OWL-DL (Web Ontology Language Description Logic) ontology, which extends SAREF (Smart Appliances REFerence) for the Smart Agriculture and Food Chain. ETSI's specialist task force STF 534[6] was established in order to extend the SAREF ontology for the domains of Smart Cities Industry and Manufacturing and Smart Agrifood. As it is mentioned in the associated Saref4agri requirements document ETSI TR 103 511, the intention of Saref4agri is to connect SAREF with ontologies that exist, such as W3C (World Wide Web Consortium), SSN (Semantic Sensor Network), W3C SOSA (Sensor, Observation, Sample, and Actuator), and GeoSPARQL (Geographic Simple Protocol and RDF Query Language).

The existence of different brands of farm equipment and software currently collects and uses data in range of proprietary file formats. This is the normal consequence of the industry's growth; however, it makes it hard for the end-users to "construct the full image" and output value from the data. It is clear that there exists the absence of interoperability in agricultural field operations, something that is not only a problem of a deficiency of common data formats or syntax. There has also been a lack of a shared understanding, or semantics in the field operations among the different industry actors, that are involved in field operations. So, there is use of multiple terms or codes to refer to the same concept or use of the same term to refer to multiple concepts. An outcome of the situation described previously, meaning the non-existence of common semantics, is the lack of common Reference Data. For instance, a usual set of code lists, unique identifiers, and controlled vocabularies which could be used in order to identify crop inputs, farm machine, implement and sensor models, so that every part involved can understand. AgGateway,[7] which represents a non-profit consortium of over 200 companies occupied with the implementation of standards aiming in advancing digital agriculture, realized its Precision Agriculture Council in 2011 so that it can tackle the previously described problems. As a result, the Agricultural Data Application Programming Toolkit (ADAPT) [11] was created targeting to realize common object model operations as well as a set of format conversion tools. ADAPT is an object model that represents some part of the world that is of interest. Object models are consisted of classes that represent groups of related data; an object-oriented version of the standard concept of a data type, and their relationships.

Another model is the IDS Information model [12–14], which is an RDF/OWL (Resource Description Framework/Web Ontology Language)—ontology and covers fundamental concepts of the International Data Spaces (IDS) software architecture for data exchange, such as types of digital contents that are exchanged by IDS participants via IDS base components. IDS addresses challenges such as

[6] https://portal.etsi.org/STF/stfs/STFHomePages/STF534

[7] http://www.aggateway.org

interoperability and security, which are related to DEMETER. The IDS Information Model is an extension of DCAT (Data Catalog Vocabulary) [15], DQV (Data Quality Vocabulary) [16], ODRL (Open Digital Rights Language), [17] and related standards for the description of data resources. The rationale behind this model is based on an exchange of ideas with the working group that specified DCAT. From an NGSI-LD point of view, metadata transfer high-level information about content of datasets, while the IDS Information Model takes contextual properties of datasets, and more generally, digital resources. The IDS Information Model supports extension points in order to transfer information about the domain-specific structure and semantics of the content of datasets. More specifically, it reuses the W3C standards, VoID, for referring to domain ontologies and their terms that are used in a dataset, Data Cube in order to describe the structure of tabular or matrix-like datasets, and SHACL (Shapes Constraint Language) to describe general graph-shaped datasets' structures [18].

The European data portals[8] provide a way to expose provided datasets, both metadata and data to reusers. Generally speaking, the federated architectures of International Data Spaces, that was described previously [13], and the more recent GAIA-X, which explicitly aims at delivering a technical solution for the data spaces envisaged by the European Data Strategy, have been created with the rationale to accommodate a wide range of legacy systems and platforms. It has been proven that a federation of existing platforms is possible when following these approaches. As far as DEMETER is concerned, it is committed to building on previous work like IDS. Concerning data exchange, DEMETER will require both service and API brokerage, which is what GAIA-X targets at [19].

GS1 [20], which is a global model, is an identification standard that provides the means to identify entities of real world, so that they may be the subject of electronic information. This information is stored and/or communicated by end-users. GS1 includes unique identifiers that can be used by an information system to refer unambiguously to a real-world entity such as a trade item, logistics unit, physical location, document, service relationship, or any other entity.

FOODON [21] project targets on building an easily accessible and comprehensive global farm-to-fork ontology concerning food, in a way that it accurately and consistently describes foods commonly known in cultures from around the world. It covers terms for the origin of food sources, the processing and cooking that took place in a product (such as heat processing), covering how it is packaged (meaning the container or the wrapping of this product).

FOODIE ontology [22] represents an application vocabulary including different categories of information dealt by typical farm management tools/apps for their representation in semantic format, and in accordance with existing standards and best practices such as INSPIRE, ISO/OGC (International Organization for Standardization/Open Geospatial Consortium) standards. This ontology aims to enable the representation of farm-related data in a semantic format, and also to enable the

[8] https://data.europa.eu/

advantage of semantic technologies for different tasks, such as transformation of semi-structured data to semantic format, ontology-based data access, data integration using linked data as a federation layer, knowledge discovery through inference, interlinking data with established vocabularies and relative datasets in Linked Open Data cloud.

3 Technical Requirements

This section presents the technical requirements extracted, which drive the design of a common Agricultural Information Model. These have been divided into two main classes, the first focuses on the core data modeling aspects and the second on aspects related to the support for semantic interoperability with existing systems and ontologies. These two requirement classes comprise functional requirements and are presented in the next two subsections.

3.1 Core Data Modeling Requirements

DM1. Common data view for heterogeneous models (Mandatory): AIM needs to define a common data model that specifies a common view on all heterogeneous entities connected and all the data involved in the pilots. This common data model shall be used for all data exchanged between software components. Therefore, it needs to support the translation of the obtained data streams to a common data model.

DM2. Representation of crop farms data (Mandatory): AIM needs to enable the common representation of agronomic data (e.g., crops, sensor data from the field, thermal/multispectral imagery from UAVs, production data, geolocation data, planting data, irrigation logs, fertilization logs, spraying logs) including: (1) Farm and economics modeling: agricultural type and economic size, production volumes and types, calculations according to results, etc.; (2) Field data modeling: location and geometry of the field, planting date, planting distance, detailed yield information; (3) Field status modeling, for example, water- or nitrogen-stressed fields, appropriate evaluation indices (e.g., Normalized Difference Moisture Index (NDMI)), need for fertilizing; (4) Soil data modeling: soil temperature and moisture, soil physical and chemical analysis; (5) Crops, treatment and fertilization modeling: crop type, crop developing stages, crop cultivar or variety, crop health status and pests, pesticides, nitrogen levels, information from counting devices used for the control of insects or plagues; (6) Traceability information of crops (production, transport, retail) to be used in the product passport information, (7) Water management modeling: water and energy consumption, water quality (e.g., salinity levels).

DM3. Earth Observation Data Representation (Mandatory): AIM needs to enable the representation of current earth observation (EO) data as well as historical

EO data, including for example satellite data, remote sensing imagery, soil maps, vegetation indices, such as NDVI, EVI, NDRE, and NDMI. It needs to also get EO metadata, for example, through interfaces compliant with the OGC 13-026r8 specification.[9]

DM4. Representation of livestock data (Mandatory): AIM needs to enable common representation of livestock data and traceability of products including: (1) Modeling of dairy and beef farms and data from farm robots: milk and meat production and quality, milk properties and quality (fats, proteins, somatic cells, and bacterial content), economic data; (2) Modeling of data from cows' wearables: animal ID, location, temperature, pedometer data, movement; (3) Modeling of animals' welfare, behavior, and habits: eating habits, respiration monitoring data, rumination, activity, rectal temperature control data, feed and water consumption data, biomarkers related with animal well-being and welfare (e.g., cytokine markers); (4) Food traceability information of dairy products and pastries (tracking of ingredients and supply chain); (5) Modeling of poultry farms: animal welfare, habits, living conditions, stress levels, medical treatment, feeding patterns, feed origin; (6) Traceability information of poultry products (production, transport, retail) to be used in the product passport information; (7) Modeling of apiary and hives: location of hives, apiary weight.

DM5. Representation of meteorological and open spatial data (Mandatory): AIM needs to enable representation of weather data (e.g., temperature, humidity, wind speed/direction, solar radiation, pressure) and open spatial data modeling. Meteorological data will be collected by interfacing with existing sensors, or new sensors that will be provided for this purpose.

DM6. Representation of agricultural machinery data (Mandatory): AIM needs to enable common representation of agricultural machinery data such as: engine data, fuel consumption, emissions, exhaust gas, NOx-conversion, and exhaust temperatures. The data are defined by Controller Area Networks Bus (CAN) protocol specifications; consequently, it will be necessary to consider that the translation of CAN-Bus Model into AIM involves understanding the specific CAN-Bus information (the message set for subsystem data exchange) of the supplier to the vehicle subsystem into new information according to the AIM communication specifications. Finally, AIM needs to represent entities types and formats, relationships among them, possible range between the values (if any).

DM7. Representation of farmer's preferences and DSS recommendations to them (Mandatory): AIM needs to enable common data model able to interpret farmers' needs and preferences including: (1) farmers' needs related to cost optimization (e.g., linking economical aspects of wholesale and retail prices), production issues (better quality of their products, crop variety per field, optimal date for planting and harvesting), cost/benefit analysis of field operations (irrigation/fertilization), optimization on irrigation/fertilization strategies, disease monitoring, yield analysis (e.g., the estimation of crop yield according to climate conditions), animal

[9] http://www.opengis.net/doc/is/opensearch-eo/1.0

welfare tracking; (2) production preferences (e.g., the use of non-chemical pesticides, attention to animal welfare, transparency to the consumers); (3) any other relevant data input collected during farm operations (related to animal welfare, crop production, product's characteristics). AIM should also enable common representation of recommendations and notifications to farmers, as well as the metadata used for providing recommendations to farmers through the DSS system and analytics tools. In this way, farmers' needs and preferences will be adequately analyzed (data integration and analysis) and decision support (visualization) will be provided.

DM8. Flexible and extensible model representation (Mandatory): AIM needs to support flexibility and extensibility in the representation of AIM through the use of a modular approach, the reuse or alignment with thesauri/classifications available as linked data, the use of property graphs and semantics, the use of appropriate data interchange models (e.g., RDF), knowledge representation languages (e.g., SKOS, RDFS, OWL), and rule languages (e.g., SWRL or OWL-RL), which would enable the semantic querying of data.

DM9. Representation of data quality metrics (Mandatory): AIM needs to include quality metrics in its data model. These data will be used for evaluating the accuracy, precision, granularity, completeness, consistency, timeliness, validity, uniqueness (where applicable) of the agri-food data and will be used by the system.

DM10. Provide a basis for data exchange across stakeholders (Mandatory): AIM needs to enable data exchange across authorized stakeholders. To facilitate this, it needs to include data regarding the supply and usage of agri-data and any other type of data that is stored in the AIM unified ontology including any economic transactions regarding the usage of such data.

DM11. Data Models enabling analysis of large volumes of heterogeneous data (Mandatory): AIM needs to enable the analysis and processing of large amount of heterogeneous data, including their storage and transfer. This is paramount as the agri-food domain involves numerous data sources, some of which collect very large data volumes (e.g., satellite data, remote sensing imagery, soil maps).

DM12. General Model for Interoperability (Mandatory): AIM needs to provide a general model for data interoperability, which should be flexible and extensible for all use cases. More specifically: (1) It will be composed of discrete modules addressing specific "competency questions," following best practices in ontology engineering—allowing these to be adopted standards or tightly managed development efforts with clear testability; (2) It needs to handle interoperability for different implementation aspects; (3) Meta-models, domain models, profiles, and vocabularies need to be handled individually using appropriate specialized modeling mechanisms.

DM13. Simplified Profiles of Data Model (Mandatory): AIM needs to provide simple profiles of the general model suitable for individual pilot cases; these profiles will define "schemas"—or views, while the general model will define semantics what objects can be identified and reused in different views.

DM14. Semantic model that supports scalability and support of legacy systems (Mandatory): AIM needs to implement semantic interoperability in a scalable and sustainable way, for example, by maintaining a dependency graph at the module level within each implementation rather than creating a temporary (project scoped) aggregated knowledge graph with no transparency of scope or provenance. It should support semantic interoperability for data originating from existing systems involved in the pilots (legacy systems). It should publish all domain-specific semantic interoperability resources in a canonical standards-based and interoperable fashion appropriate to the type of resource (e.g., vocabulary, schema, object model, profile, data type).

DM15. Governance Arrangements (Mandatory): AIM needs to specify governance arrangements for each component, determining who, when, and how updates to the included components should be handled. This includes pragmatic project scope governance of temporary resources, as well as requirements for governance of project resources that would ensure future interoperability.

DM16. Abstract model for integrating sensors, processing, and decision support systems (Mandatory): AIM needs to provide an abstract model for the general process of linking sensor data through processing chains into decision support systems, including how intermediate data products relate to sources and outputs. This can be based on an existing general model, or, if necessary, to create something new, to be pushed as an OGC (Open Geospatial Consortium) and/or W3C (World Wide Web Consortium) general model specification.

3.2 Semantic Interoperability Requirements

SI1. Service wrappers and translators (Mandatory): AIM needs to facilitate the development of service wrappers and translators, also known as data providers and consumers, which will enable the different tools/platforms in a (regional/national) AKIS to expose and consume data in interoperable forms.

SI2. Mapping AIM with standard models (Mandatory): AIM should implement (semantic) mappings from standard and/or widely used ontologies/vocabularies with the AIM, enabling the semantic integration of data represented using any of these models. As part of the semantic mapping, AIM needs to identify logical connections between classes, properties, and objects across ontologies. The mappings will deal with cases in which, e.g., a class in one ontology is the intersection (or union) of two classes in another, or the complement of another class, or a simple object needs to be mapped to a complex class in another ontology.

SI3. Semantic Interoperability support (Mandatory): Support semantic interoperability, encompassing semantic integration. This can be realized via the implementation of suitable data provider and consumer services, new ontologies, and the mappings with existing ontologies/vocabularies, as well as the other mechanisms developed to facilitate data integration.

SI4. Semantic mapping best practices (Mandatory): Follow best practices and approaches for generating the mapping between AIM and existing ontologies/vocabularies, including: (1) Transformation of existing ontologies into common format, for example, OWL, use of semantic rules or annotations/punning; (2) Reuse of AIM terms and only extend it if necessary. In the latter case, reuse existing terms whenever possible, and only otherwise create new terms/extensions; (3) Use of appropriate mapping constructs/axioms, such as owl:equivalentClass and owl: equivalentProperty with OWL classes/properties; skos:closeMatch, skos: exactMatch, skos:broadMatch, skos:narrowMatch, and skos:relatedMatch with SKOS concepts; owl:sameAs for individuals, etc.; (4) Treating of the mappings as "first class" components of a modular knowledge graph, making them available in line with FAIR principles, and governing them appropriately and transparently. Consider mappings across different levels of specification granularity as well of abstractions using the appropriate mechanisms in a standardized way, for example, mappings from meta-models to models (OWL subclassing); mappings between concepts at the same level of abstraction; mappings between controlled vocabulary terms; mappings between measurements and classifications (e.g., threshold values for "good,"); soft- vs. hard-typing mappings with classes with a subtype property vs. specific subclasses.

SI5. Tools for generating ontology mappings (semi-) automatically (Mandatory): Identify and select, if possible, suitable tools for the (semi-) automatic mapping of ontologies/vocabularies. Some example tools to be analyzed include the Alignment API, PARIS, Map-On, etc.

SI6. Identify tools to validate mappings (Mandatory): In order to facilitate the mapping between the AIM and existing ontologies, it is necessary to identify and select, if possible, suitable tools to validate the generated mappings. This is necessary because some of the mappings may be quite complex. For example, when a specific schema is mapped to a more general schema, then some schema elements may be replaced by use of a qualifying term in corresponding more abstract elements. In such cases, the coverage of the mappings as well as the result of exercising a mapping against the target model need to be validated. It would also be desirable to define a validation process and a simple reference implementation that can define test procedures to be integrated into traditional development tooling.

SI7. Select relevant existing ontologies to align with AIM (Mandatory): Identify and select relevant standards and/or widely used ontologies/vocabularies to align with the AIM and identify the key terms in each of them that would need to be aligned. Examples of dominant, often standardized, ontologies that need to be aligned to AIM to ensure sufficient semantic interoperability include: Saref4agri, FIWARE, AGROVOC, INSPIRE, SOSA/SSN, FOODIE, rmAgro, drmCrop, Soilphysics, AgriFarm, FOODON, etc.

4 AIM Design

In line with best practices and recommendations (e.g., [23–25]), the specification of DEMETER AIM follows a modular approach in a layered architecture, enabling among others:

- to facilitate the interoperability with existing models by reusing available (standard and/or well-scoped) vocabularies and ontologies in the modules, instead of defining new terms, whenever possible
- to facilitate the mapping/alignment with other models, by module instead of the whole model
- to facilitate the extension of the domain/areas covered in AIM with additional modules
- to facilitate its maintenance, by updating and/or modifying only specific modules
- to facilitate the mapping to top-level/cross-domain ontologies
- to foster the reuse of AIM, by enabling reuse at module level

More in detail, AIM has been designed in three layers, namely the core meta-model layer (Sect. 4.1), the cross-domain layer (Sect. 4.2), and the domain layer (Sect. 4.3). This three-layer approach is inspired by and similar to the NGSI-LD information model [3]. Moreover, AIM adopts the same meta-model as NGSI-LD, as discussed in the next section. However, as opposed to NGSI-LD, AIM exploits semantic referencing to foster interoperability and reuse. Accordingly, the cross-domain and domain models are built by reusing, as much as possible, existing standards and/or well-scoped vocabularies and ontologies with well-defined semantics, in addition to include mappings between these models to enable their interoperability and integration of derived data. Briefly, the cross-domain ontology comprises the set of general concepts that aim at providing common definitions for the whole agri-food domain handled by AIM and at avoiding conflicting or redundant definitions of the same classes at the domain-specific layer. The domain ontologies, on the one hand, model information such as crops, animals, agricultural products as well as farms and farmers just to mention some of the most important concepts. In addition to the three AIM-specific layers, a meta-data schema has been specified to describe various concerns of data exchange in AIM-related settings. For the latter, the AIM defines a profile of the IDS Information Model reviewed in Sect. 2, focusing on those aspects of the concerns specified by IDS (“content,” “concept,” “community of trust,” “commodity,” “communication,” “context”) that are relevant to DEMETER. Figure 1 provides an initial overview of the AIM model.

4.1 Meta-model Layer

A meta-model, as its name implies, is a model of a model. Meta-models are typically used for different purposes. For instance, they can be used for the specification of modeling language constructs in a standardized, platform independent manner [26], to specify and restrict a domain in a data model and systems specification [27], or to provide an explicit model of the constructs and rules needed to build specific models within a domain of interest [28]. In fact, as noted in [28], meta-models can be viewed from three different perspectives: (1) as a set of building blocks and rules used to build models; (2) as a model of a domain of interest; (3) as an instance of another

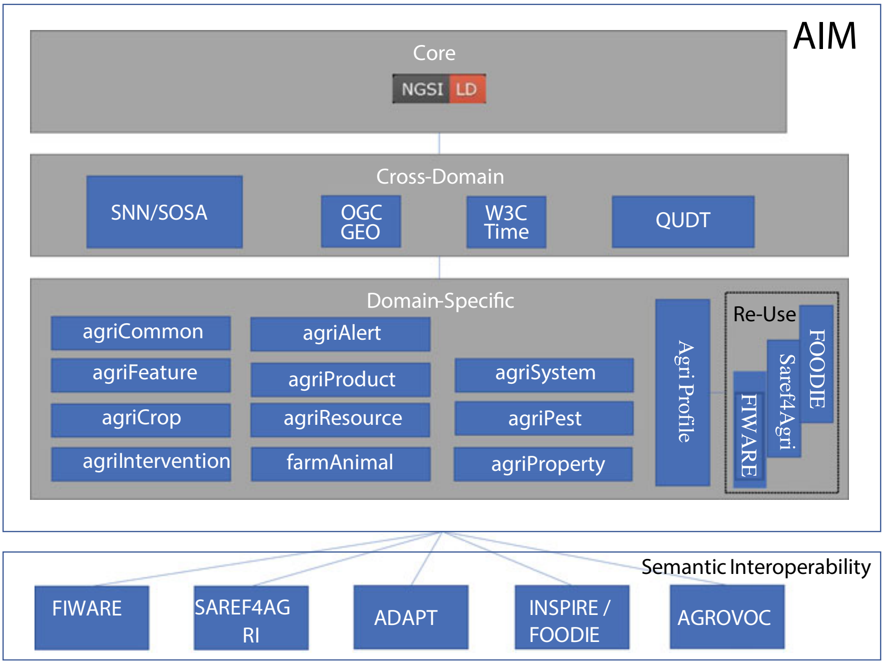

Fig. 1 AIM model overview

model. In the context of the DEMETER meta-model, the first perspective is being considered.

After analyzing different [meta-]models and approaches (e.g., [26, 29, 30–32]), it was decided to follow the meta-modeling approach of NGSI-LD [6], a standard of the European Telecommunications Standards Institute (ETSI) whose mission is to make it easier for end-users, city databases, IoT, and third party applications to exchange information. NGSI-LD is an evolution of the NGSI context interface family, particularly the FIWARE NGSIv2 information model,[10] which was evolved by ETSI ISG CIM initiative to support property graphs, semantics, and linked data by adopting the increasingly popular JSON-LD serialization format (also a priority in DEMETER project).

Property graphs are composed of nodes (vertices), relationships, and properties, where nodes may have properties in the form of key-value pairs, keys are strings, values are arbitrary data types, and relationships are arc (uni-directional, i.e., directed edges) that have an identifier, a start node, and an end node. Like nodes, relationships can have properties attached to them. Linked data, on the other hand, is based on the RDF model, which expresses data as triples of the form <subject, predicate, object>. RDF, however, does not support directly attaching properties to predicates

[10] https://fiware-tutorials.readthedocs.io/en/latest/