On Designing and Implementing Agro-ecology IoT Applications: Issues from Applied Research Projects

Sandro Bimonte, Amina Belhassena, Christophe Cariou, Jean Laneurit, Rim Moussa, Gerard Chalhoub, Robert Wrembel, Gauthier Picard, Ladjel Bellatreche, Alexandre Journaux, et al.

On Designing and Implementing Agro-ecology IoT Applications: Issues from Applied Research Projects

Sandro Bimonte*, Amina Belhassena*, Christophe Cariou*, Jean Laneurit*, Rim Moussa†, Gerard Chalhoub‡, Robert Wrembel§, Gauthier Picard¶, Ladjel Bellatreche‖, Alexandre Journaux**, Thierry Heirman**, Ali Hassan ††, Stefano Rizzi‡‡, Jean-Pierre Georgé[x]

* Université Clermont Auvergne, TSCF, INRAE, France
name.surname@inrae.fr

†University of Carthage, Tunisia
rim.moussa@enicarthage.rnu.tn

‡Université Clermont Auvergne, LIMOS-CNRS, Clermont-Ferrand, France
gerard.chalhoub@uca.fr

§Poznan University of Technology, Poland
robert.wrembel@cs.put.poznan.pl

¶ONERA/DTIS, Université de Toulouse, France
gauthier.picard@onera.fr

‖ ISAE-ENSMA, France
bellatreche@ensma.fr

**GenPhySE, INRAE, France
alexandre.journaux@inrae.fr, thierry.heirman@inrae.fr

††Umanis, France
ahassan@umanis.com

‡‡DISI, University of Bologna, Italy
stefano.rizzi@unibo.it

[x]IRIT, University of Toulouse III, France
jean-pierre.george@irit.fr

***Abstract*—The design and implementation of agro-ecology IoT applications is a non-trivial task since the data processed in such applications are typically complex and heterogeneous. Moreover, these applications are implemented using different systems and technologies, over complex IoT communication network layers (edge, fog, cloud). The existing system design methods fail to effectively represent data in such a scenario. In this position paper we report and discuss the open issues for a new, dedicated design method, based on our initial experience in implementing an agro-ecology IoT system.**

***Index Terms*—Internet of Things, Big data, smart farming, Agriculture robots**

I. Introduction

In the recent years, the Internet of Things (IoT) has been successfully applied in several different application domains, as for example healthcare, environment, mobility, and even agriculture [1]. IoT is the set of physically connected devices that support computation and communication by means of different communication networks (e.g., ZigBee, Wi-Fi, ADSL). As described in [2], IoT produces Big Data, which are data mainly characterized by (at least) the 3Vs, namely *Volume*, *Variety*, and *Velocity*. The usage of IoT in the agricultural business is needed and promising. Indeed, one recent report[1] estimates that in 2027 this sector will reach 34 billion USD. Agro-ecology aims to develop new cultural practices that respect the environment and at the same time save production and biodiversity [3]. Agro-ecology has been recognized by all governmental, economic, social and environmental institutions as one the main challenges of humanity for the next 30 years. Data used by agro-ecology models are very diverse, including environmental, agricultural, and socio-economic data, at different (micro and macro) spatial and temporal scales and also data coming from mobile autonomous robots and drones.

In the context of agriculture, IoT has been successfully employed for different applications, e.g. agronomic surveillance and livestock production. IoT leads to a revolutionary approach for agro-ecology since it provides the stakeholders with more precise, complete, and innovative data and their associated analysis. In particular, at the crossroads between

[1]https://www.marketsandmarkets.com/Market-Reports/iot-in-agriculture-market-199564903.html

breeding and agro-ecology, two main research topics emerge: image recognition via neural networks to detect and recognize the parasites on the legs of a grazing animal, and then the geo-localization of parasites on a plot or a territory. The monitoring of crops development and agricultural practices using autonomous robots is another hot research topic. Agro-ecological animal and plant breeding in the era of IoT and Artificial Intelligence implies the usage of wireless sensors, drones, satellite images, multimedia data, and classical data in an integrated, coherent, and effective way.

An example of a classical IoT architecture in the agricultural context is illustrated in Figure 1, which shows the data and the network connections involved.

Data is collected, and sometimes computed, by IoT devices (such as autonomous robots, tractors, meteorological sensors, drones, etc.) deployed in the field. These IoT devices produce real-time stream data which, when combined with other data (such as farm data, geospatial data, images, etc.), can be used for online analyses at the farm level. Moreover, historical IoT data and other external data can be used to provide more complex analyses (such as prediction models, OLAP, etc.). Therefore, an IoT agriculture application is usually fed with data coming from the field and historical external data in a real-time way. All these data are deployed in different data management systems (sensors devices, tractors' laptops, classical PCs, distributed servers, etc.). These data management systems are deployed at different levels of the network architecture (directly on the field, in the farm, in the cloud, etc), and they communicate by means of various network communication protocols (for example, ADSL, Wi-Fi, etc.).

Overall, agricultural IoT applications require:

- the use of complex spatio-temporal data (e.g., robot trajectories, meteorological data);
- the use of stream data (e.g., from sensors deployed in fields) and historical data (e.g., warehoused data on all the aspects of an IoT system).

Moreover, agro-ecology IoT applications seem to be more challenging than in Industry 4.0 in the following aspects:

- the use of autonomous robots and vehicles that operate in an uncontrolled environment;
- the limited computation and communication resources (ADSL networks, low-quality Wi-fi connections, small laptops) deployed in rural areas;
- the involvement of stakeholders (such as farmers, researchers, managers, etc.) who have heterogeneous profiles with different knowledge and experience in smart farming (from farmers not skilled in IT to researchers in robotics).

When dealing with IT applications that process complex and heterogeneous data, the adoption of a conceptual design step using formalisms such as UML or E/R has been widely proved to be necessary to grant the success of projects [4]. Indeed, these formalisms make the implementation and technical issues transparent, allowing database designers and IoT experts to focus exclusively on the functional requirements provided by end-users. However, to the best of our knowledge, data modeling methods for IoT agricultural applications have not been deeply investigated so far (see [5] for a complete survey). In this position paper, we motivate the need for a new methodology for agro-ecology applications' design and implementation (Sec. II), then we present the modelling and implementation requirements and some envisaged solutions. Some relevant works are presented in Sec. IV. This contribution is based on our findings while realizing some agro-ecology research projects.

II. Motivation

IoT in the agro-ecological context comes with new issues that we discuss in this section.

A. What

Agro-ecology IoT data have a spatio-temporal nature since all agronomic and bravery phenomena are geolocalized (e.g., plots, positions of animals and robots). These data are complex, ranging from images and videos to time series produced by sensors and autonomous robots. Therefore, they need *ad-hoc* conceptual representations and implementations. Indeed, IoT systems typically rely on relational or NoSQL database managment systems (DBMSs), data stream management systems (DSMSs), and other components implemented in different technologies and supporting different programming languages, which run on heterogeneous hardware (IoT devices, personal computers, cloud servers). Here, such complex, heterogeneous, and changing data will be called *polyglot data*.

Quality of Service (QoS) features (such as latency, data loss, etc.) play a major role in IoT data architectures. Data provided by a system and their QoS features are strictly related. For example, it is likely to send images from animal drinkers with different resolutions; robots can send one aggregated odometry data per minute instead of one data per second, according to the available network bandwidth. This means that, for each piece of data, different reliable representations must be considered by end-users. Thus, IoT data can be represented in different ways and at different abstraction levels (*multi-representation data*) according to the physical constraints imposed by the network architecture. Clearly, each representation can be implemented in different ways in its corresponding system (e.g. DBMS, DSMS, sensors).

These polyglot and multi-representation data must be correlated to provide a global data-centric representation of IoT data. These correlations raise several design and implementation issues since they can involve different data management systems (collection, storage, and computation). Noticeably, according to [6], no conceptual model allows representation of polyglot and multi-representation data.

At the conceptual design level, the main research questions to be faced are:

- "*How to define an integrated, polyglot meta-model that conceptually represents agro-ecological data together with data obtained from different kinds of computations independently of all implementation details?*",

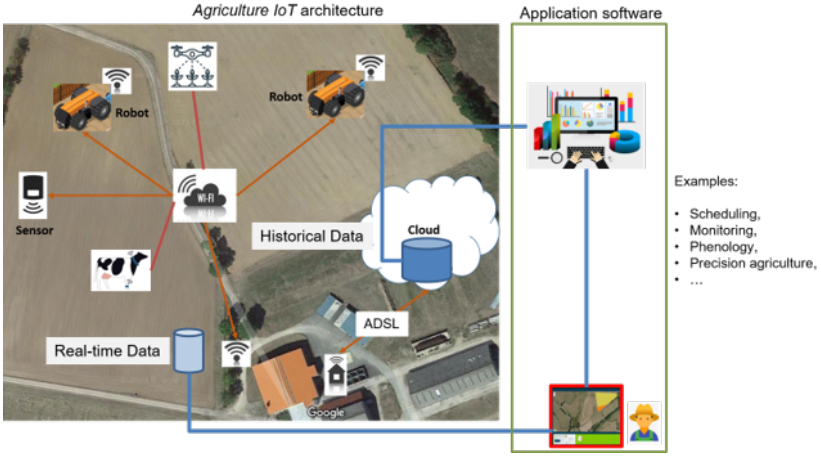

Fig. 1: An example of an agricultural IoT architecture.

- "*How to conceptually represent each agro-ecological data entity at multiple abstraction levels, and what policies should be defined to seamlessly switch from one level to another?*"
- "*Which QoS features can be specified by end-users during design, and how to integrate them with the meta-model and with the multi-representation policies?*"

At the implementation level, the main research questions are:

- "*How to generate (semi)automatic implementations of these polyglot and multi-representation IoT data over different data management (collection, storage, and computation) systems and programming languages?*"
- "*How to choose the most suitable technology for agro-ecological data management (collection, storage, and computation) and its deployment locations over the network?*"

B. How

Several different data management (collection, storage, and computation) systems have been proposed to take into account the particularities of the data stored and queries processed (data *workload*). For example, to handle very high volumes of robots data, NoSQL DBMSs seem better suited than classical relational ones. Therefore, in order to select the system that best fits the type of data, we can consider the workload as a "metadata" that must be also represented in IoT systems. To the best of our knowledge, only [7] introduces the workload at the conceptual level, but it addresses only NoSQL document DBMSs.

Thus, the research questions associated to the workload are:

- "*Which workload features are relevant at design time?*"
- "*How to integrate them in the conceptual meta-model?*"

C. Where

IoT applications are characterized by a geographically-distributed deployment of (potentially moving) devices, and a communication network continuum over different layers (from edge to cloud). The network layer where the data management system is deployed must take into account the QoS features. Therefore, QoS features on every layer play a major role in IoT data architectures. Indeed, they can respect some functional and non-functional requirements, such as bandwidth, which lead to a particular placement of data and computation over the different layers. For example, in the context of hard real-time applications, data and computation can be deployed at the edge level (for example on a robot) to improve performances.

Therefore, the research questions associated to QoS are:

- "*Which are the relevant QoS performance indicators to guide the deployment and the functioning of data management systems over the different network layers (edge, fog, or cloud): access delay, data rate, packet loss ratio?*"
- "*How to obtain these indicators in a reliable way for all layers of the network?*"
- "*How to exploit these QoS indicators with regard to user experience?*".

Moreover, agro-ecology IoT data can be implemented in different ways and locations in the IoT architecture, and only at run time the best data management system configuration of each data can be chosen to make the overall system resilient to network problems. Therefore, the research question associated to the run-time execution of the application is:

- "*How to define and implement an algorithm for dynamically choosing the most suitable configuration for the overall system at run-time, making it resilient according to the QoS indicators?*"
- "*Which configuration mode is most suited to enhance the user experience depending on the application use cases?*"

III. Issues for agro-ecology IoT applications' design and implementation

In this section we present two representative scenarios, as well as the requirements and some envisaged solutions for a method to design and implement agro-ecology IoT applications. Some relevant works are cited that could be extended to meet these requirements. Figure 2 shows an overview of

our proposal, and can be used as a reference for the whole section.

A. Scenarios

Here we briefly discuss the interest of IoT for agro-ecological practices in crops and breeding.

Crops: the use of autonomous robots and drones is more and more frequent in agro-ecology applications, for example for repetitive and long tasks such as plowing, picking, and harvesting needed by agro-ecology crops practices [8].

To support this transition, autonomous robots have an essential role to play, as they have low impact on the environment (they are light and can operate in fleets) and are able to perform repetitive and accurate farming operations over a long time. With special equipment and combined with data acquisition and data processing technologies, robots are able to autonomously perform efficient and targeted tasks in the fields, e.g., within inter-cropping systems, while optimizing the use of resources and main training at a high level of productivity. When robots cannot communicate over the network due to some of their own electronic or mechanical problems, drones could be used in place of the robots to send data to an information system.

Breeding: New environmental and agronomic resilient breeding practices apply to breeding animals outdoors. Rabbits and pigs can thus be bred outdoors, but their epidemiological monitoring must also be done on the field. This implies that animal health must be checked by instruments that are placed in the field using some classical measurements tools (like weight or movement behaviour), but also cameras. Moreover, disease vectors can also be present on the vegetation, which must also be monitored. Thus, outdoor breeding implies an advanced monitoring of animals and their environment to prevent contamination and make their natural environment safe.

B. Design

In the context of complex polyglot data, it is important to decouple functional requirements from the architectures and technologies used to implement them. In this direction, we propose to adopt the Model-Driven Architecture (MDA), an OMG-supported approach to software design, development, and implementation which encourages this decoupling. In our context, MDA would allow to design polyglot and multi-representation data-centric IoT applications, also considering QoS network features, which are relevant at the conceptual level for end-users. For example, the conceptual representation of the red data entity of Figure 2 should be augmented with multi-representation, workload, and QoS features. Moreover, its multi-representation-aware implementation can be done in two different data management systems, namely, a classical PC and a sensor (polyglot data) as shown by the two red rectangles of the *Design and (semi)automatic implementation* step in Figure 2.

MDA also aims at producing rapid and error-free implementations that comply with functional requirements; this is a mandatory feature for the development of complex systems. Noticeably, MDA promises to streamline the design iterations, which is very relevant in the agro-ecological context since functional requirements are usually not clearly defined from the beginning of the project.

Combining MDA with UML profiles, which provide a formal language to design data and computation, seems a natural choice. In particular, we propose to define a UML profile based on a data representation with UML Class element, which has already been successfully used to model IoT nodes associated to other complex data (stream, spatial data, etc.) in [5]. To this end, we plan to extend the Platform Indipendent Model (PIM) of [5] with data types other than the ones collected by sensors (polyglot data). A first attempt in this direction has been done in [9]. For multi-representation data, an approach similar to the one proposed for spatial databases could be adopted [10], since they are based on UML Class element too. Representations could be changed by means of ad-hoc Class methods or OCL constraints. QoS and workloads could be added as tagged values of these Class elements. A promising direction to obtain a (semi)automatic implementation of these data is to extend the Platform Specific Model (PSM) of [5], which provides a UML model of sensor devices used to define the PSM from the PIM representing IoT data. The PSM and the device model proposed in [5] could be extended by considering data management systems and networks. Finally, the workload and QoS features defined at the PIM level will also be included in the PSM and device model, in order to obtain coherent PSM models and implementations with PIM models.

C. Implementation

IoT data are distributed and exchanged over a communication network. Therefore, a set of advanced network performance indicators to help the deployment process of data management systems, and to choose the right configuration at run-time, must be proposed. For example, as shown in the *Run-time* step of Figure 2, the PC data management implementation for the red data entity, and the cloud for the blue data entity are chosen at run-time among all the implementation solutions defined at design time.

In particular, these performance indicators will give us an idea about the QoS that can be expected from the network. To meet this requirement, a promising direction is to extend the approach presented in [11]. Because of multi-representation of data and QoS, the corresponding data management systems can be deployed in different network layers. This leads to different possible implementation configurations of the same IoT data-centric conceptual model. Therefore, an *intelligent agent* methodology to choose the best implementation configuration at run-time is needed. The envisioned distributed and decentralized infrastructure requires locally implementing online decisional algorithms aimed at operating optimal allocation of resources and services to specific devices and computation nodes. This will encourage resilience and adaptation in case of services or devices failure or disconnection. In this context, the multi-agent systems paradigm is a perfect

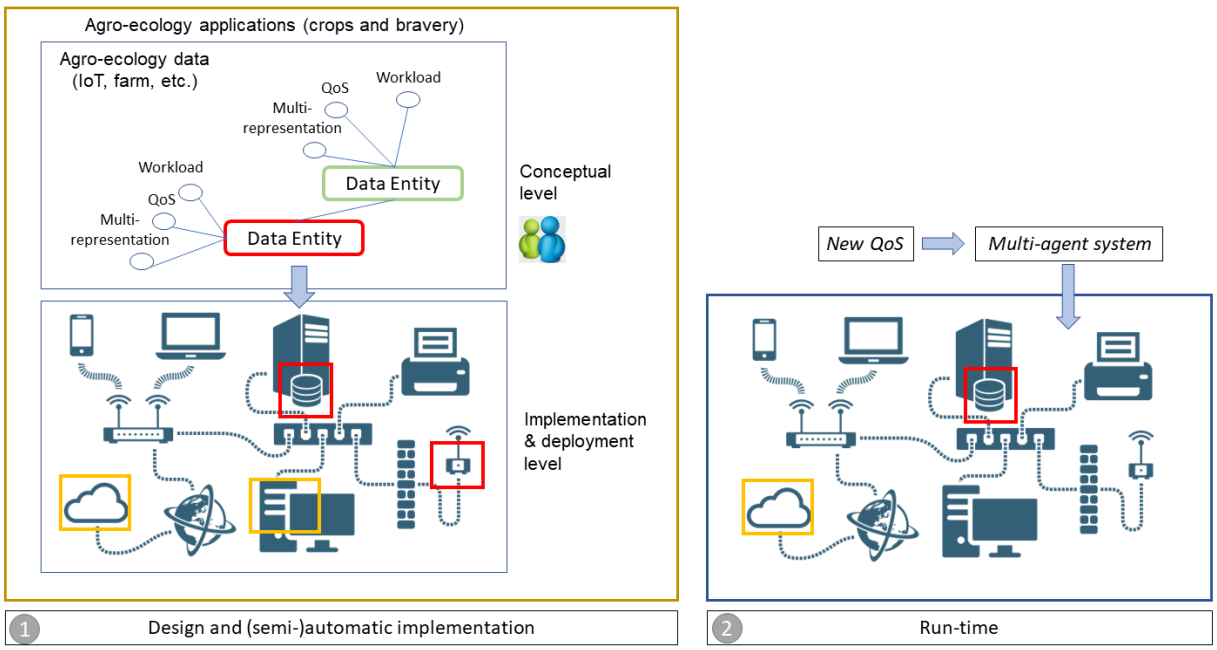

Fig. 2: An overview of the envisaged proposal.

fit, as it provides both theoretical and practical tools to design and develop complex systems composed of several decision nodes. A possibility is to rely on resilient deployment and self-organization of intelligent systems [12], where agents with local decision rules and decision redundancy allow the system to adapt to unexpected events. Finally, to coordinate the autonomous mobile sensors, multi-robot and multi-agent planning techniques could be adopted [13]. Noticeably, such a multi-agent implementation will challenge the state-of-the-art techniques, due the large scale of the envisioned system, the presence of unpredictable queries and communication failures, and the use of autonomous vehicles as mobile sensors, which goes far beyond classical IoT-based sensing systems.

IV. Related work

The seminal paper of [14] suggests IoT and Big Data as very promising approaches for the development of smart agro-ecology solutions. According to the authors, the usage of simple sensors devices must be coupled with more sophisticated devices such as smartphones, and equipment such as drones and autonomous robots. [15] provides a complete study about interdisciplinary of IoT and crop management, and it also highlights the importance of the usage of mobile equipment coupled with Big Data analysis tools. Several other works recognize the importance of this integrated usage of different data sources [16], [17].

Some works have been proposed for the modelling of IoT-based applications using UML. Recently, [18] and [19] surveyed studies on the development of IoT applications, and classified them according to the main steps used, such as identifying the actors, the requirements, the implementation of a proof of concept, until the study of technical implementation issues. [5] provides a complete survey about existing work using UML to represent data used by IoT applications. The authors conclude that all existing works do not provide a sufficient abstraction level allowing to take apart technical details of the IoT devices and networks from the data representation. According the authors, this makes difficult an agile and effective design of the data-centric application with end-users.

The usage of multi-agent systems in IoT applications, and recently in agriculture is more and more present, such as in [20], where agents are used to manage the communication network of green houses, or in [21], which uses deep reinforcement learning to improve decision-making irrigation process of crops, and [22] for pesticide use reduction. However, in the context of autonomous agricultural robot, to the best of our knowledge, no work applies multi-agent systems to coordinate them with other IoT devices deployed in the field.

When dealing with real-time applications, knowing the state of the network in terms of available throughput, delay, and packet loss helps to adapt the required QoS. Indeed, when the performance of the network is degraded, lowering the QoS requirements in terms of data rate would help reduce the packet loss. This is known as adaptive bitrate streaming and is mainly based today on the HTTP protocol under the name DASH (Dynamic Adaptive Streaming over HTTP) [23]. Supervising the performance of a wireless network is a challenging task [24]. For example, in Wi-Fi networks, many rate adaptation methods exist that try to maximize the data rate of transmissions depending on the quality of the link between nodes [11]. Some of these methods are based on implicit observations that do not require additional overhead, but they lack in reactivity to network changes. Other methods are explicit and react faster to network changes, but they

require sending feedback in order to inform the sender about the quality of the link from the receiver perspective. When the HTTP protocol is not used, new methods are needed to adapt QoS requirements in wireless networks based on observations of network performances. When audio and video are transmitted, methods based on RTP (Real-Time protocol) in conjunction with the RTCP (RTP Control Protocol) are generally used. Other types of traffic are also critical from the application point of view, such as control/command traffic for remotely guiding mobile robots. Hence, the need for a general-purpose QoS monitoring method for wireless networks.

V. Conclusion

Making all implementation and technological details related to IoT transparent to agro-ecological decision-makers is a crucial key factor to successfully create effective new agro-ecology applications. This because decision-makers are usually not skilled in IT, hence, adopting a simple but formal design formalism focused on the data they use will allow them to easily interact with IT and IoT experts to define their own agro-ecology application. In this position paper we have listed the related requirements and research challenges from two points of view: design and implementation. We have also introduced two real case studies that could be used as proof of concept for our future proposals.

Acknowledgements

This work is supported by the French National Research Agency project ANR-19-LCV2-0011 *Tiara*, and French government IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25).

References

[1] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E. M. Aggoune, "Internet-of-things (IoT)-based smart agriculture: Toward making the fields talk," *IEEE Access*, vol. 7, no. 1, pp. 129551–129583, 2019.

[2] M. Marjani, F. Nasaruddin, A. Gani, A. Karim, I. A. T. Hashem, A. Siddiqa, and I. Yaqoob, "Big IoT data analytics: Architecture, opportunities, and open research challenges," *IEEE Access*, vol. 5, pp. 5247–5261, 2017.

[3] T. Dalgaard, N. Hutchings, and J. Porter, "Agroecology, scaling and interdisciplinarity," *Agriculture, Ecosystems & Environment*, vol. 100, no. 1, pp. 39–51, 2003.

[4] S. Robinson, G. Arbez, L. G. Birta, A. Tolk, and G. Wagner, "Conceptual modeling: definition, purpose and benefits," in *Proce. Winter Simulation Conference*, (Huntington Beach, CA, USA), pp. 2812–2826, 2015.

[5] J. E. Plazas, S. Bimonte, C. de Vaulx, M. Schneider, Q. Nguyen, J. Chanet, H. Shi, K. M. Hou, and J. C. Corrales, "A conceptual data model and its automatic implementation for IoT-based business intelligence applications," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10719–10732, 2020.

[6] J. E. Plazas, S. Bimonte, M. Schneider, C. de Vaulx, and J. C. Corrales, "Self-service business intelligence over on-demand IoT data: A new design methodology based on rapid prototyping," in *Proc. ADBIS*, pp. 84–93, 2020.

[7] C. de Lima and R. dos Santos Mello, "A workload-driven logical design approach for NoSQL document databases," in *Proc. iiWAS*, pp. 73:1–73:10, 2015.

[8] R. D. Christian, J. Lyon, A. de Boon, M. Hanheide, and S. Pearson, "Responsible development of autonomous robotics in agriculture," *Nature Food*, vol. 2, p. 306–30, 2021.

[9] A. Belhassena, S. Bimonte, P. Battistoni, C. Cariou, G. Chalhoub, J. C. Corrales, J. Laneurit, R. Moussa, J. E. Plazas, R. Wrembel, and M. Sebillo, "On modeling data for IoT agroecology applications by meansof a UML profile," *Proc. International Conference on Management of Digital EcoSystems*, 2021.

[10] E. Zimanyi, C. Parent, S. Spaccapietra, and C. Vangenot, "Multiple representation modeling," in *Encyclopedia of Database Systems, 2nd Edition* (L. Liu and M. T. Ozsu, eds.), Springer, 2018.

[11] I. Sammour and G. Chalhoub, "Evaluation of rate adaptation algorithms in ieee 802.11 networks," *Electronics*, vol. 9, no. 9, 2020.

[12] P. Rust, G. Picard, and F. Ramparany, "Resilient distributed constraint optimization in physical multi-agent systems," in *Proc. ECAI*, pp. 195–202, 2020.

[13] P. Bechon, C. Lesire, and M. Barbier, "Hybrid planning and distributed iterative repair for multi-robot missions with communication losses," *Autonomous Robots*, vol. 44, pp. 505–531, 2020.

[14] V. B. Maurel and C. Huyghe, "Putting agricultural equipment and digital technologies at the cutting edge of agroecology," *Oléagineux, Corps Gras, Lipides*, vol. 24, no. 3, pp. 1–7, 2017.

[15] G. Vitali, M. Francia, M. Golfarelli, and M. Canavari, "Crop management with the iot: An interdisciplinary survey," *Agronomy*, vol. 11, no. 1, 2021.

[16] P. Gonzalez-de Santos, R. Fernández, D. Sepúlveda, E. Navas, L. Emmi, and M. Armada, "Field robots for intelligent farms—inhering features from industry," *Agronomy*, vol. 10, no. 11, 2020.

[17] A. Krishnan, S. Swarna, and B. H. S, "Robotics, IoT, and AI in the automation of agricultural industry: A review," in *Proc. B-HTC)*, pp. 1–6, 2020.

[18] B. Tekinerdoğan, E. Tüzün, and G. Giray, "Iot system development methods" in internet of things challenges advances and applications," in *Internet of Things Challenges Advances and Applications*, pp. 141–159, 2018.

[19] G. Fortino, C. Savaglio, G. Spezzano, and M. Zhou, "Internet of things as system of systems: A review of methodologies, frameworks, platforms, and tools," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 223–236, 2021.

[20] J. Wang, M. Chen, J. Zhou, and P. Li, "Data communication mechanism for greenhouse environment monitoring and control: An agent-based iot system," *Information Processing in Agriculture*, vol. 7, no. 3, pp. 444–455, 2020.

[21] F. Bu and X. Wang, "A smart agriculture iot system based on deep reinforcement learning," *Future Generation Computer Systems*, vol. 99, pp. 500–507, 2019.

[22] O. Bahri, A. Mourhir, and E. Papageorgiou, "Integrating fuzzy cognitive maps and multi-agent systems for sustainable agriculture," *Euro-Mediterranean Journal for Environmental Integration*, vol. 5, no. 7, 2020.

[23] I. Sodagar, "The MPEG-DASH standard for multimedia streaming over the internet," *IEEE Multim.*, vol. 18, no. 4, pp. 62–67, 2011.

[24] I. Sassi, A. Gouin, and J. Thiriet, "Wireless network performance evaluation for networked robots," in *Proc. ETFA*, pp. 1–5, 2017.