Journal of King Saud University – Computer and Information Sciences xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

Cloud and distributed architectures for data management in agriculture 4.0 : Review and future trends

Olivier Debauche [a,b,c,*], Saïd Mahmoudi [a], Pierre Manneback [a], Frédéric Lebeau [b,c]

[a] University of Mons, Computer Science Unit, 20 Place du Parc, Mons 7000, Hainaut, Belgium
[b] University of Liège, GxABT – BioDynE/DEAL, 2 Passage des Déportés, Gembloux 5030, Namur, Belgium
[c] University of Liège, GxABT – TERRA, 2 Passage des Déportés, Gembloux 5030, Namur, Belgium

ARTICLE INFO

Article history:
Received 8 April 2021
Revised 8 September 2021
Accepted 10 September 2021
Available online xxxx

Keywords:
Agriculture 4.0
Smart farming
Smart agriculture
Lambda architecture
Kappa architecture
Edge computing
Fog computing
Micro-service architecture
Data lake
Data house
Blockchain
Osmotic computing
Dew computing

ABSTRACT

The Agriculture 4.0, also called Smart Agriculture or Smart Farming, is at the origin of the production of a huge amount of data that must be collected, stored, and processed in a very short time. Processing this massive quantity of data needs to use specific infrastructure that use adapted IoT architectures. Our review offers a comparative panorama of Central Cloud, Distributed Cloud Architectures, Collaborative Computing Strategies, and new trends used in the context of Agriculture 4.0. In this review, we try to answer 4 research questions: (1) Which storage and processing architectures are best suited to Agriculture 4.0 applications and respond to its peculiarities? (2) Can generic architectures meet the needs of Agriculture 4.0 application cases? (3) What are the horizontal development possibilities that allow the transition from research to industrialization? (4) What are the vertical valuations possibilities to move from algorithms trained in the cloud to embedded or stand-alone products? For this, we compare architectures with 8 criteria (User Proximity, Latency & Jitter, Network stability, high throughput, Reliability, Scalability, Cost Effectiveness, Maintainability), and analyze the advantages and disadvantages of each of them.

Contents

* Corresponding author at: University of Mons, ILIA Lab, 20 Place du Parc, Mons 7000, Hainaut, Belgium.
E-mail address: olivier.debauche@umons.ac.be (O. Debauche).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2021.09.015

1. Introduction

Nowadays, the Internet of Things (IoT), also formerly named pervasive Internet, is present in all domains of our daily life and follows exponential growth. The number of connected devices is estimated at the horizon of 2022 at 42.5 billion and at the horizon of 2025 at 75.5 billion[1]. The global IP traffic is estimated to 333 ZB per month in 2022[2] with the need to store and treat this data (Carnevale et al., 2019). The European Commission has predicted that 18 billion of 29 billion connected devices will be related to the IoT in 2022 (Agency, 2020). Cisco in a white paper has announced that connected devices to the Internet will generate 850 ZB/year by 2021 (Cisco, 2018). It is difficult to precisely determine the number of connected devices on a world scale, but their number is about several billion. In addition, McKinsey Global Institute predicts a total economic impact of IoT and Edge Computing devices that will reach 11 trillion USD by 2025 (Manyika and Chui, 2015). In the sector of the agriculture, nearly 12 million agricultural sensors installed globally by 2023 with an increase of 20% annually, which is predicted by the Business Insider Intelligence Service (Meola, 2021). Moreover, the smart agriculture business was estimated at USD 13.8 billion in 2020 and is projected to reach USD 22 billion by 2025 at a Compound Annual Growth Rate (CAGR) of 9.8% (Meola, 2021).

Within the IoT era, the type of clients is becoming increasingly lightweight. IoT devices and the network environment is gradually changing from high-speed wired networks to unstable wireless communication. Meanwhile, users' demand IoT applications is also shifting to real-time and context-aware service provisioning, making the focus moving progressively from the cloud to the edge (Ren et al., 2017).

The cloud is located within the Internet and is geographically centralized, is constituted of a few resourceful server nodes, and is inserted in multi hops in terms of distance among the clients (Munir et al., 2017). Cloud Computing (CC) is a paradigm widely available that offers benefits like minimal management effort, convenience, rapid elasticity, pay per use, ubiquity (Ai et al., 2018), easy maintenance, centralized management, and high server utilization (Shi et al., 2016). Furthermore, resources centralization implies an increase of average network latency, heavy bandwidth utilization, and high processing delay. Indeed, the tremendous amount of data handled in a unique server point can create congestion in the cloud servers and backhaul links (El-Sayed et al., 2017).

Nevertheless, the rapid parallel development of the pervasive intelligent device, ubiquitous network, growth in popularity of virtual and augmented reality, self-driving vehicles, UAVs, social networks, networks applications, and services are not without consequences. As a matter of fact, the network bandwidth and speed limit performance and effectiveness of cloud computing especially for real-time and mission-critical applications cannot be guaranteed. Moreover, cloud computing can be hardly adapted or applied to various types of technologies and applications scenarios (Zhou et al., 2017). To address these issues, various extensions of central cloud computing have been proposed by industrial and academics to move computing and storage at the edge of the network close to users. Fog computing uses network elements between the central cloud and the edge of network and absolute edge elements such as microcontrollers close to sensors to process and store data with a distributed manner close to nodes.

Whereas, with the developments of mobile devices, some new paradigms close to mobiles users have been proposed. For example, cloudlets or micro data centers are geographically implanted and accessible by means of Wi-Fi protocols; but, this approach does not always guarantee enough network quality. Manufacturers of cellular network equipment have proposed the Mobile Edge Computing (MEC) paradigm that associates fog servers with base stations to provide services to mobile devices. MEC associated with 5G allows to combine an ultra-low latency network with high available bandwidth, and processing resources accessible in the vicinity. The MEC original concept has been extended then to wireless networks and consequently renamed in "Multi-access Edge Computing" (Wang et al., 2020).

Agriculture has previously undergone two waves of revolution. The first one was mechanization and the second was called the green revolution with genetic modifications (Saiz-Rubio and Rovira-Más, 2020). Since the late 1990s, the digital transformation of the agriculture in Agriculture 3.0 also called Precision Agriculture has begun with the integration of Geographical Information System (GIS), Global Positioning Systems (GPS), and the usage of sensors have invaded agriculture. They allowed the emergency of image processing, techniques using deep learning, and machine learning in the field of computer vision. This latter is implemented to discriminate weed, identify crops, detect diseases,...etc. The production of a large amount of data by agriculture 3.0 has required the development of big data technologies to process them, reflecting important changes in various fields of research. Collected data must be recorded in a specific format in order to discover patterns, curate errors, eliminate duplicated or inconsistent data, or solve noise problems (Triantafyllou et al., 2019).

Smart Farming also called Smart Agriculture or Agriculture 4.0 is a domain of IoT in full growth which bring innovative paths to improve the adaptability, the efficiency, and the resilience of the agriculture of production systems (Iaksch et al., 2021) boost competitiveness and profit (Triantafyllou et al., 2019), allocate resources reasonably, and avoid food waste (Zhai et al., 2020; Wolfert et al., 2017) thanks to the contribution of autonomous context awareness provided by sensors and the capability to execute autonomous or remote actions (Wolfert et al., 2017). Smart Farming displaces the strict application from the farm location to affect related fields such as decision making by farmers, biodiversity,

[1] https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/
[2] https://www.statista.com/statistics/499431/global-ip-data-traffic-forecast/

supply chains management, food availability and quality, insurance, and research in environment and earth sciences,...

Smart Farming is distinct from other domains of the Internet of Things (IoT) by the observation and action of biological objects (animals or plants). It differs from medical IoT by the fact that there are no issues related to privacy; but, the confidentiality of data is related to production processes. Like most areas of the IoT, Wireless Sensing and Actuating Network (WSAN) use Low-power and Lossy Network organized in hierarchical routing to collect data and actuate devices. Multi-path routing protocols can also be implemented to balance the data transfer load and conserve the energy of limited battery life, basic computational skills, unique communication identifier, and resources-constrained nodes. Due to the limited battery life, it is difficult and sometimes impossible to recharge or replace (Debauche et al., 2021). Moreover, energy-saving and ambient energy techniques must be applied to deal with the active and inactive operational time and schedule information transmission (Triantafyllou et al., 2019). To which objects can be added like connected agricultural vehicles, milking robots, Unmanned Aerial Vehicle (UAV) commonly known as drones, Unmanned Ground Vehicle (UGV) also called robots, mobile devices such as tablets used to encode punctual observations (Debauche et al., 2021), and external sources such as public geo-services (Triantafyllou et al., 2019).

The use of IoT in agriculture 4.0 ranges from family farming as for example in India on a very small scale with a few low-cost sensors and actuators to very large scales with thousands of expensive commercial sensors and many connected agricultural pieces of machineries as in the American mid-west. Smart Farming is characterized as aforementioned by a wide variety of objects that can produce the highly contrasted amounts of data from few bytes/s to Gb/s. In addition, the availability of network protocols in rural areas to transmit this data impact also the type of architecture to implement. Applications need treatments in real-time and/or at a different time. The "real-time" requirements are also very variable depending on the use case. For instance, remote control of drones requires reaction times of at most a few milliseconds while the Variable Rate Fertilizer (VRF) or Variable Spraying (VS) application aim to optimize nutrients and herbicides application respectively need reaction time in a range of few milliseconds to few seconds. The real-time processing for monitoring a herd of cattle is of the order of a few minutes to a few hours. The data retention time is very variable and is highly dependent on each use case. For example, UAVs produce tremendous quantities of images to transfer to the cloud in real-time where they must be quickly processed and stored. They can also be post-processed to extract additional data in batch processing. While UGVs images lose their value after processing and eventually actuating. However, if data is of a special, new or exceptional nature, it can be stored with a view, for example, to improving artificial intelligence algorithms. Other sensors transmit data only when anomalies are detected while others transmit at regular intervals a tiny amount of data.

However, the adoption of Smart Farming is hampered by the lack of models to guide stakeholders on how to implement and to deploy dense and heterogeneous IoT-based monitoring systems and manage their interoperability (Triantafyllou et al., 2019). Commercial sensors are very expensive making it impossible for small farms to implement them (Garcia et al., 2020). In addition, two trends are currently opposed. That coming from the manufacturers of agricultural machinery who have developed their ecosystems and who want to extend the services offered to farmers by attracting them into the captive ecosystems in which they are locked. Furthermore, another trend is the development of open ecosystems in which farmers can preserve the ownership of their data and keep control of the processing carried out on this data and of their use. On one hand, farmers are therefore faced with a dilemma where they are in any case forced to use agricultural equipment that collects their data against their will; on the other hand, they want to keep control of their data collected through IoT sensors. Currently, it is difficult to predict which of these two trends will take precedence over the other or whether one of the two will coexist (Wolfert et al., 2017). In this context, both private and public researchers can either use generic commercial platforms offered by cloud players on which they have limited possibilities of adaptation or develop their own architecture on the basis of commercial or free bricks, but with much greater possibilities of adaptation. In this case, the choice is also delicate, and a bad evaluation of the constraints can jeopardize the research project.

Due to the recent advances in big data, we present a survey that provides an overview of the state of the art regarding Smart Farming. It aims at summarizing parameters that condition the choice of architecture to collect, process, and store agricultural data. Since there is a wide variety of use cases, it is important to make an informed choice when it comes to architecture. In this way, we address the current gap in the literature with a review of cloud architecture used in Agriculture 4.0 to collect, process, and store data to enlighten the reader about the possible choices and the new trends that emerge. The rest of this paper is structured as follows: The second section is composed of two parts. In the first part, we summarize related previous review in the domain and their contributions, in order to contextualize our contribution to the literature. In the second part, we identify architectures implemented in Agriculture 4.0 use cases. In the third section, we describe the methodology used to identify papers, the conceptual framework used to analyze the literature, and the criteria used to compare the selected architectures. In the fourth section, we present architectures used to collect, process, and store data. We describe successively the cloud-centric architectures, the extension of cloud paradigm, the distributed architecture. In the fifth section, new trends and futures directions are presented. In the sixth section, we discuss the future evolution of Agriculture 4.0 to Agriculture 5.0. Finally, the last section concludes this paper with recommendations and perspectives.

2. Related works

We begin our review by identifying the previous review realized in the field of Internet of Things applied to Smart Agriculture to take stock of the state of art and highlighting aspects that have not been explored at the present time. In this section, we focus to achieve two objectives. The first aims to position our work in relation to the existing literature. The second aims to identify architectures commonly used in the case of applications in Agriculture 4.0.

2.1. Previous reviews

Reviewed papers presented in Table 1 were selected in the timeframe from January 2017 to July 2021. The major contribution of each paper was extracted and highlighted to show our contribution to the literature.

In the following paragraphs, we will draw a panoramic summary of the existing reviews during the past four years (2017–2021). In 2017, Ray (Ray, 2017) reviewed throughout his paper IoT applications and the challenges that have been faced while IoT deployment to improve farming. Talavera. et al. (Talavera et al., 2017) reviewed agro-industrial and environmental applications that are using the Internet of Things (IoT) for monitoring, control, logistics, and prediction. Tzounis et al. conducted a survey of IoT technologies in agriculture and the challenges that farmers face going forward (Tzounis et al., 2017). Elijah et al. identified the most encountered challenges in the field of IoT applications

Table 1
Summary of previous review achieved on big data management in a context of Smart Farming.

Major Contribution	Reference
Survey of agro-industrial and environmental solutions for monitoring, control, logistics, and prediction.	(Talavera et al., 2017)
Diagnosis and analysis of existing IoT deployments in regards to communication protocols.	(Ray, 2017)
Survey of IoT technologies in agriculture and highlighted the challenges going forward.	(Tzounis et al., 2017)
Identification of IoT challenges, its application in smart agriculture, and presentation of trends and technological innovation	(Elijah et al., 2018)
Review of IoT applications in Precision Agriculture, evaluation of previous contributions by researchers, and pathways to future innovation	(Khanna and Kaur, 2019)
Review of IoT deployment in protected agriculture, identification of its challenges, and prospection of the new research domain.	(Shi et al., 2019)
Review of existing IoT-based precision agriculture solutions for further achievement.	(Ruan et al., 2019)
Review, comparison, prospection, and challenges of wireless communication technologies applications in the field of Precision Agriculture.	(Feng et al., 2019)
Review, case study, and challenges of WSN in environmental behavior.	(Shafi et al., 2019)
Review, identification, challenges of current and future trends of IoT agriculture.	(Ayaz et al., 2019)
Survey of IoTbased agriculture, presentation of connection between IoT, big data, and cloud computing, regulation and policies of IoT, and its application in the field of agriculture.	(Farooq et al., 2019)
Survey of the use of UAVs, an overview of PA, and investigation of 20 UAV applications.	(Radoglou-Grammatikis et al., 2020)
Challenges of IoT-based agriculture architecture, a summary of existing surveys of smart agriculture. and classification of threats models, study, analysis of challenges and future works of security and privacy of green IoT-based agriculture.	(Ferrag et al., 2020)
Discuss the role of IoT and big data analysis in agriculture with an emphasis on the commercial status of applications and translational research outcomes.	(Misra et al., 2020)
resent different solutions to address IoT in arable farming challenges.	(Villa-Henriksen et al., 2020)
Systematic review presenting how IoT is used with smart farming	(Navarro et al., 2020)
Methodological review and analysis of IoT components and their applications in smart farming.	(Debauche et al., 2021)
Review of emerging technologies towards agriculture 4.0 and new pathways to agricultural practitioners.	(Liu et al., 2020)
Review, classification, presentation, comparison, and challenges of emerging technologies for IoT-based agriculture.	(Friha et al., 2021)

in smart agriculture and presented common trends for innovative ideas (Elijah et al., 2018). In 2019, Ayaz et al. provided a state-of-art about IoT-based architectures applied in agriculture and identified present and future trends in the same field of study (Ayaz et al., 2019). Farooq et al. presented the ingredients of IoT-based smart farming with used technologies that apply the utilization of network architecture and protocols; in addition to that, they provided an overview of the regulations and policies of the use of IoT in farming regarding security and privacy. They concluded their study by summarizing the main challenges encountered in this discipline (Farooq et al., 2019). Feng et al. provided an overview of the wireless communication technologies in the precision agriculture domain. They benchmarked the prospection and challenges of existing technologies with the regular communication time used (Feng et al., 2019). Shafi et al. conducted a literature review about IoT-based automation of agriculture along with Wireless Sensor Network (WSN). These authors presented a case study based upon two models: 1- a WSN to monitor real-time crop of health conditions, 2- system-base remote sensing imagery to classification between healthy and unhealthy yield (Shafi et al., 2019). In terms of agriculture protection, Shi et al. drew a panoramic review during the last decade to address the challenge and future works to further the research in the field of protected agriculture (Shi et al., 2019). Khanna et Kaur called into an evolutionary scenario to highlight the most significant impact of IoT in Precision Agriculture (PA). They evaluated the contribution of their predecessors and enhanced the challenges to open up a new direction of inspiration and innovation in IoT applied to PA (Khanna and Kaur, 2019). Ruan et al. reviewed literature works from 2009 to 2018 to suggest new ideas for folks interested to conduct research in the field of agriculture IoT, infrastructures, data security, and data sharing (Ruan et al., 2019). In 2020, two studies have been carried out about 20 UAV applications that are devoted to either aerial crop monitoring processes or spraying tasks (Radoglou-Grammatikis et al., 2020) and about the dilemmas that researchers must overcome while deploying IoT in the green agriculture domain (Ferrag et al., 2020). Villa-Henriksen et al. identified different challenges encountered during the implementation of IoT in various applications and proposed different solutions to address them (Villa-Henriksen et al., 2020). Misra et al. discuss the role of IoT and big data analysis in Smart Farming (Misra et al., 2020). In 2021, a recent study conducted by Friha et al. hypothesize the use, application, classification, and comparison of the most developed emerging technologies such as Internet of Things (IoT), Unmanned Aerial Vehicles (UAV), Wireless Technologies, open-source IoT platforms, Software Defined Networking (SDN), Network Function Virtualization (NFV) technologies, cloud/fog computing, and middleware platforms (Friha et al., 2021). In the same year, Debauche et al. conducted a literature review to describe the main components of IoT and its applications in the field of Smart Farming (Debauche et al., 2021).

2.2. Platforms implemented in use cases

We grouped applications into 4 categories: (1) **Water Management** in which we have aggregate all types of water use such as irrigation and watering animals. (2) **Plant Disease and Pest** groups all use cases in plant's pathologies detection and treatment of plant pathologies (spraying of fungicides, pesticides, etc). (3) **Crop Management** brings together all the use cases relating to cropping operations: soil management (plowing, fertilizer application), sowing, weeding, and harvesting. (4) **Livestock** includes everything related to the breeding of farm animals (nutrition, behavior, diseases, treatments). Table 2 summarizes platform used to implement use cases in Smart Farming classified following our four categories.

Garcia et al. give an overview on trends in Smart Irrigation in which they showed that data is stored in the database or in the cloud. On 151 reviewed papers, one uses Raspberry Pi, 18 databases, 53 clouds, and 79 are self-developed or not mentioned (Garcia et al., 2020). Navarro et al. identified 21 Platforms used in 50 various use cases classified into 5 categories: Artificial Intelligence, Big Data, Machine Learning, Computer Vision, and Other/Not Identified (Navarro et al., 2020). Jayaraman et al. present SmartFarmNet, an IoT platform offering effortless integration of sensors, supporting scalable data analytics, and proposing do-it-yourself tools to analyze and visualize data (Jayaraman et al., 2016). Codeluppi et al. describe LoRaFarM a general architecture modulated depending on the farm's characteristics and requirements (Codeluppi et al., 2020).

Table 2
Summary of cloud platforms, databases mentioned in Smart Farming reviews.

	Water Management	Plant Diseases & Pest	Crop Management	Livestock	Reference
IoT platform					
Thingspeak	x				(Maureira et al., 2011)
FIWARE	x				(Rodriguez et al., 2018)
NETPIE	x		x		(NECTEC, 2020)
Ubidots	x				(Ubidots, 2021)
SmartFarmNET	x		x		(Jayaraman et al., 2016)
Thinger.io	x		x		(Luis Bustamante et al., 2019)
Kaa IoT Platform	x			x	(KaaIoT, 2021)
IBM Watson IoT Platform	x	x	x	x	(IBM, 2015)
Microsoft Azure IoT Platform	x		x	x	(Microsoft, 2021b)
AT&T M2X Cloud			x		(AT&T, 2021)
Blynk			x		(Blynk, 2021)
MACQU			x		(Sigrimis et al., 2002)
ERMES			x		(Granell et al., 2017)
Agrocloud		x	x	x	(Kodati and Jeeva, 2019)
CropInfra			x		(Pesonen et al., 2014)
SensorCloud			x		(Corp, 2020)
LoRaFarM	x		x	x	(Codeluppi et al., 2020)
Cloud platform					
Amazon Web Service	x	x	x	x	(Amazon, 2021b)
IBM Cloud	x	x	x	x	(IBM, 2021)
Microsoft Azure	x	x	x	x	(Microsoft, 2021a)
Integra	x		x		(Souces and I., 2021)
Cloud Database					
DynamoDB	x		x	x	(Amazon, 2021)
MongoDB Atlas	x		x	x	(Mongo, 2021)
Firebase	x		x	x	(Google, 2021)
InfluxDB Cloud	x		x	x	(Influxdata, 2021)
Local Database					
MySQL	x				(Oracle, 2021)
SQLite	x				(SQLite, 2021)
PostgreSQL/PostGIS	x	x			(The PostgreSQL Global Development Group, 2021)
Apache Cassandara				x	(Apache Software Foundation, 2021a)
Apache Druid			x	x	(Apache Software Foundation, 2021b)

The monitoring of crops particularly more sensitive to them as saffron is crucial. The **DIAS Architecture** (Triantafyllou et al., 2019) uses different ground and leaf sensors to monitor the real-time 24/24 h cultivation process of saffron. This data is transmitted by LoRaWAN with IPv6 protocol and MQTT-SN protocol to FIWARE's context broker. The broker manages all networking devices by means of sixteen types of messages exchanged following publish-subscribe model. The FIWARE NGSI API of oversees the consumption, subscription, and processing of all the information collected and its publication. Afterward, the data is stored and analyzed with a random forest algorithm which allows extracting information about the crop growth and health. Vegetation indexes: Normalized Difference Index (NDI), Excess Greenness Index (ExG) are calculated with PiX4D[3] image processing tools. Object-based image analysis (OBIA) is used to recognize weeds or discriminate species. Finally, collected data are categorized and evaluated accordingly with vegetation index values, moisture level, and plant developing state by means of the Apache Spark framework for the Big Data analysis and Waikato Environment (WEKA) a framework specialized in data mining to produce reports and predictions.

Decision-making is a very important task in the farmers' activities but with the amount of data always increasing, they encounter difficulties on one hand to make proper decision about agricultural management and on the other hand translate this data into practical knowledge (Zhai et al., 2020). On the other hand, there is a need for platforms of the Agricultural Decision Support System (ADSS) to assist farmers to make precise decisions evidence-based. For example, **Watson Decision Platform for Agriculture** combines IBM Watson with IoT and Cloud Computing to detect crop disease from UAV images. It is also possible to optimize time for crop operations to obtain a better price on trading market. The second example is **Digital Farming System**[4] takes advantage of computer vision, cloud computing, and AI to propose a better timing for corp operations, notify when a crop is infected by any disease. **Smart Irrigation Decision Support System** (SIDSS) is composed on one hand of a set of sensors and a weather station and on the other hand a DSS based on two machine learning algorithms. Partial Least Squares Regression (PLSR) to deduct unnecessary variables and Adaptive Neuro -Fuzzy Inference Systems (ANFIS) used to minimize estimated errors under a target threshold (Navarro-Hellin et al., 2016). SIDSS generates planning of water amount and time for irrigation. **Multi-robot sense-act system** (Conesa-Munoz et al., 2016) is a planner of aerial and ground vehicles which assign tasks to the most appropriate work units. A Harmony Search Algorithm is used to optimize plans for UAVs while meta heuristic is running for ground vehicles.

2.3. Analysis of previous literature

The analysis of existing reviews about smart farming shows that applications use whether open source or commercial cloud architecture whether developing specific architecture responding to their aims or do not describe their storage and processing system. The latter represents more than half of the papers and means that some of the processing architectures remain unknown because they have never been specifically described and studied.

[3] https://www.pix4d.com/

[4] http://prospera.ag/

Moreover, the fact that further development is being made in architecture may be the fact that commercial platforms do not fully address the needs of Agriculture 4.0. This brings us to our research questions and their respective motivation:

1. *Which storage and processing architectures are best suited to Agriculture 4.0 applications and address its particularities?* Motivation: On one hand generic architectures dedicated or not to IoT are able to address a large number of use cases but not specifically the needs of Agriculture 4.0 exist. On the other hand, researchers develop architectures to address specific issues or requirements of use cases. The selection of an adapted architecture is crucial for the correct implementation of identified use cases.
2. *Can generic architectures meet the needs of Agriculture 4.0 application cases?* Motivation: Agriculture 4.0 has specific requirements described in the introduction section which cannot all be addressed by a single classical generic architecture. A comparison between the pros and the cons of major generic architecture in the context of agriculture 4.0 is important to highlight the choice during the conceptualization step.
3. *What are the horizontal valuation possibilities that allow the transition from research to industrialization?* Motivation: The use of architectural solutions which can be for example free of fees during the research phase but needs a reimplementation caused by license limitations, the cost of the license in the use cases budget, etc. The use of products in closed or semi-closed ecosystems is a barrier to the research valuation.
4. *What are the vertical valuation possibilities to move from algorithms trained in the cloud to embedded or autonomous products?* Motivation: The massive collection of data in the cloud allows to development of complex algorithms that need a large amount of computing resources to be elaborated. Afterward, they can be compressed, reduced, optimized in order to be deployed in embedded devices or divided and establish a collaboration between devices and computing resources such as cloud, fog, etc.

In order to answer these questions, a review of the literature will make it possible to synthesize the different approaches currently used, to identify new trends and to consider new lines of research to be explored.

3. Methodology

In order to address, our first and second research questions, we achieve a systematic review to identify generic architectures and combination of architectural elements used by researchers to implement concrete use cases. Moreover, we attempt also identify commercial products and existing services/ platforms used to implement projects in agriculture.

3.1. Systematic review methodology

The research questions outlined at the end of the related work section has been addressed by combining keywords of the first group that refers to architectures (i.e. cloud architecture, distributed architecture, big data, Internet of Things, IoT) and of the second group contained keywords related to agriculture (i.e. agriculture, smart farming, food, agri-food, precision agriculture).

Our methodology is based on 3 consecutive steps: literature identification, reading literature, and information extraction.

During the first step, we have read and have collected individual papers based on the achieved of previous papers. We have reviewed and completed by a systematic survey of white literature (full articles and conference papers) from January 2016 to December 2020. In addition, we targeted solely and exclusively papers written in English and focusing on architecture design have been considered. Our bibliographic review was limited to the last 5 years because the rapid development of IoT. The systematic review was retrieved from the following major bibliographic databases: Scopus (Elsevier), IEEE Xplore Digital Library, Wiley Online Library, ACM Digital Library, and Springer. These bibliographic databases have been chosen widely covering relevant bibliography and relevant advanced bibliometric features especially number of citation and relevant literature suggestion. From these databases 1058 peer-reviewed articles were retrieved. After their screening 55 papers were classified relevant while remaining articles were considerate not relevant and therefore excluded from further reading and analysis. The high number of excluded papers is due to numerous papers describe i.e. conceptual or theoretic architectures which were never implemented, experimental architectures that have been the subject of a single article or that have never been proven by other research teams. We discard also papers that were not a directly related Big data and the agricultural sector. Table 3.

In a second step, we included English grey literature (reports, blogs, magazines, and web-items) into our review using Web of Science and Google Scholar. Table 4. We discarded papers that were written in other languages than English, Master and doctoral dissertation, and duplicated articles gathered from Google Scholar. Afterward, we have selected literature that has carefully been read in detail to extract relevant information of research questions. The extracted information was analyzed and summarized in a conceptual framework illustrated in the Fig. 1.

Three ways of treatment of data are possible. The first process data in real-time (left branch identified by **(1)** on Fig. 1), this one is generally not stored except eventually particular or exceptional data in order to enrich the training database of artificial intelligence algorithms. This way of data treatment is used for example by robots that inspect a crop, discover a pest, and then eliminate it. After intervention the value data is near null. The second way is a mixed way in which data must be processed as quickly as possible. This one addresses use cases where latency required must be comprised between few milliseconds to few seconds with data, which conserves a value during a certain period of time. This latter justifies its storage according to the use case data management plan that predicts the time after which the data will be aggregated and then deleted. This way in identified by **(2)** on Fig. 1. It addresses use cases where all data must be processed and then stored for eventual post-processing for example to estimate trends of parameters such as the milk quality, volume of palatable species available in a pasture. The third way is stored data theirs native format without transformation (Identified by **(3)** on Fig. 1). This way is implemented on use cases that do not require real time processing or use cases where the amount data is so important, which makes treatment impossible. In this latter case, data are consumed by micro services that sample data to exact knowledge. This way is also employed for data which have a low value or lose their value so quickly that there is no point in transforming them for long-term storage. For instance, a UGV identifies and eliminates a pest. The image of the insect is no longer relevant after its elimination.

3.2. Architecture comparison criteria

In order to compare selected architectures, we chose to select 8 criteria:

(1) **User Proximity** expresses the necessity to be close to the user. This criterion is important for applications where privacy and response time to query are critical. Attribute a value of one * when privacy is not crucial; ** when the proximity with user is

Table 3
Keywords used for achieved the systematic review.

Area	Keywords	Related concepts
Agriculture	Agriculture, Agricultural Agri-Food Smart Farming Precision Agriculture, Precision Farming	e-Agriculture Agribusiness Farming
Internet of things	IoT, Internet of Things, internet-of-things	
Big data	Big Data Data Management Architecture	Big Data Data Management Cloud Architecture, Distributed Architecture

Table 4
Sources of collected literature.

Data source	URL
IEEE Xplore Digital Library	https://ieeexplore.ieee.org
Scopus	https://www.scopus.com
Springer	https://link.springer.com/search
Wiley Online Library	https://onlinelibrary.wiley.com
Google Scholar	https://scholar.google.com
Web of Science	https://publons.com/publon
ACM Digital Library	https://dl.acm.org

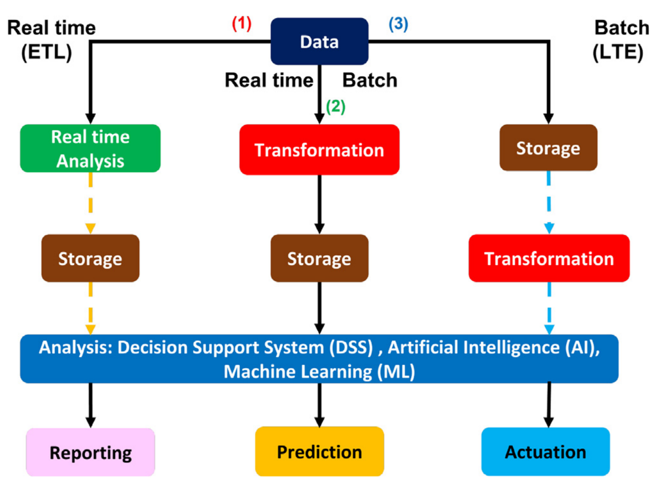

Fig. 1. Conceptual framework of data processing.

desirable but not crucial for the development of the use case; *** when the user proximity is the corner stone of the application.

(2) **Latency** & **Jitter** criterion describes the importance for the architecture to have a minimal latency and jitter. This criterion is particularly important for use cases where response time to query in quasi (real time) is required and/or time between data production and ingestion by the processing and storage architecture is essential.

(3) **Network stability** criterion translate the necessity to have a stable network or if is interruption can be tolerated. Use a value of * if the use case implemented can tolerate the absence of network during few hours; ** if few minutes of interruption are tolerable; *** is stability of the network is an essential element of the use case.

(4) The **high throughput** criterion expresses the capability of the architecture to process quickly a wide amount of data arriving at high frequency; Use a value of * if the data arrive mostly at regular intervals; a value of ** if the data arrive in bursts, and *** if the data arrive continuously at high frequency (>10 Hz).

(5) **Reliability** is a criterion that expresses if the infrastructure is critical in other terms whether an interruption in infrastructure could cause loss of life or not. Attribute a weight of * if the data is not critical and potential damages caused by an interruption of the architecture are minors or null; ** if potential damageable but tolerable if they occur more than once a year; *** if the application cannot tolerate any interruption which would cause irreversible damage or loss of human life.

(6) **Scalability** is a criterion that expresses the regularity of the evolution in terms of processing and storage during a period of one year. If the scalability must be achieved at most once a year use a weight of *; if the scalability is achieved at most twice a year use **; if the scalability must be achieved more than two times by year use a weight of ***.

(7) **Cost-Effectiveness** criterion reflects the need to control infrastructure costs. This criterion is more important as the infrastructure is brought to evolve both in terms of scale and complexity. Use the weight of * if the project will remain in a relatively constant size and do not need to be scaled or dramatically modified; Use **, if the project evolves reasonably, i.e. should not undergo significant modification more than once a year. Use a weight of *** if the size of the project and/ or its complexity need a fine study of cost.

(8) **Maintainability** criterion is directly linked to the sustainability of the project. If the sustainability of the project will not exceed two years to allocate a point of *; if the life of the project is between 2 and 5 years, assign a score of ** beyond 5 years, assign ***.

4. Architectures

The numerous publications dealing with cloud architectures relating to Agriculture 4.0, summarized in Table 2, show that a great deal of effort has been devoted to solving a whole range of problems related to many use cases. Indeed, a universal and a unique architecture do not exist for IoT applications in Smart Agriculture which ensure all needs of all use cases. This is the reason why several researchers have proposed various architectures which address specific issues of generic architectures.

The Fig. 2 gives a global overview on Agriculture 4.0 organization.

4.1. Central Cloud Architectures

Central Cloud Architectures are based on two basic architectures that are associated or combined in order to form modern architectures. These two architectures are:

Batch Architecture aims to process an entire dataset in an offline mode. For this type of architectures, as long as the processing of the dataset is not finished, it continues and produces results only when it has reached its end. Generally, the data is selected and distributed to different nodes in order to be processed more quickly. When all the treatments are achieved on all nodes, the results are sorted and aggregated to obtain a global output. This architecture is easily implemented, and the aggregation is done by a framework, but processing times can be long, and data extracted during the treatment cannot be processed before the end of the treatment in progress. Furthermore, it is possible to increment results of

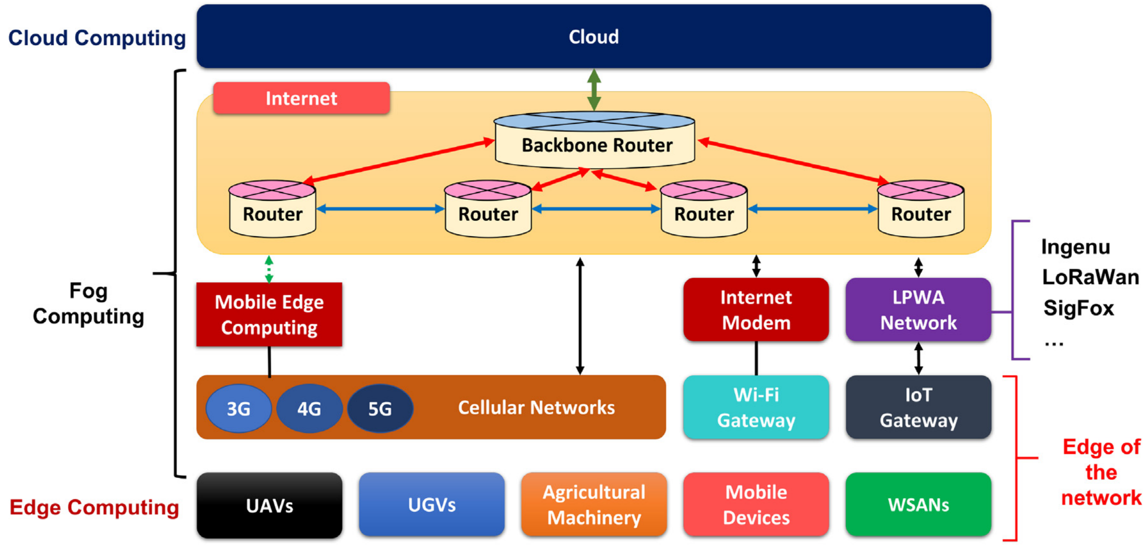

Fig. 2. Global structure of IoT in Agriculture 4.0.

Table 5
Pros and Cons of Batch Architecture.

Pros	Cons
- Easy to implement and maintain.	- Process only data previously stored in another form (file, database, etc).
- Able to achieve long term treatments (several hours or days).	- Processing cannot be modified before the end of the treatment.
- Reprocessing of old data that are easy to achieve.	- Results available only at the end of the treatment.

Table 6
Pros and Cons of Real-time Architecture.

Pros	Cons
- Allow a rapid treatment of newly arrived data.	- Not able to achieve processing on large size of the batch.
- Batch processing can be emulated using micro batches but not all algorithms can be implemented.	- Reprocessing of old data difficult to implement.
- Easy to implement and maintain.	- The need for real-time processing involves the use of an estimator rather than the precise values that would take too long to be calculated.

previous batch and produce a result that integrates treated data in progress.

Sallah et al. used a batch architecture to update data within the AquaCrop model (FAO) embedded in R-environment in order to facilitate model calibration and validation, run and evaluate all fields in a single run (Sallah et al., 2019). Nolack Fote et al. presented an architecture to extract knowledge on the long term from data in Precision Livestock Farming (PLF) (Fote et al., 2020). Table 5.

Real-time Architecture also named Streaming Architecture processes data as it arrives, and results are progressively available by opposition to the batch architecture where it is not necessary to wait for the end of ingestion of all input data to obtain a result. The notion of real-time is strongly dependent on the analysis context with a processing time from a few milliseconds to a few minutes. Real-time architecture can be implemented in two different ways. On one hand with micro-batch in which a tiny amount of data is processed each n seconds and a result is obtained at the end of the treatment or on the other hand with a streaming approach in which each new data is immediately processed and output is quickly produced. This architecture is limited to data flow processing (Miloslavskaya and Tolstoy, 2016). Table 6.

Various data are produced by different fields or animals sensors, vehicles, and robots of the Agriculture 4.0. Afterward, this data must be on one hand stored in a raw state and processed in an offline way where long and complex treatments can be achieved. On the other hand, data can be processed before its storing with offline processing, streaming processing, or a combination of these ones. The storage time is extremely variable following the nature of the data and their loss of value over time. **Offline processing** is classically used to process images from UAVs, UGVs, or satellites, for example, to determine photosynthesis activity, evaluate the canopy development or stocks of palatable species available in a pasture, etc. While **Streaming processing** allows detecting anomalies in animals' behaviors in real-time, or during agricultural operations such as the harvesting, disease and pest detection, weeds elimination. In these last cases, data is not stored because it quickly loses all value after its ingestion. Finally, a combination of the two previous ways i.e. **Offline and Streaming processing** is used to estimate real-time metrics and achieve complex treatments in an offline way at the same time. This approach is used by milking robots which detect anomalies in the production in real-time while the offline processing estimates the future production of each cow based on previous milking (Debauche et al., 2021).

Lambda architectures are used in systems that need to process and expose quickly massive amounts of streaming data. This cloud architecture was proposed by Nathan Marz and James Warren (Marz and Warren, 2013) to handle tremendous quantities of data and resolve complex problems combining processing large volumes of data (Batch) while incorporating the most recent data processed in real-time processes (Singh et al., 2019). This architecture is generic, scalable, and fault-tolerant against hardware failures and human mistakes. The architecture is composed of three layers: (1) batch layer process very large quantities of data by batch; (2) speed layer which processes data in real-time and provides views based on the most recent data and (3) serving layer responding to queries. Data comes from either a data source or a message queue.

This paradigm allows executing arbitrary queries over any real-time data and is particularly adapted for critical infrastructure and

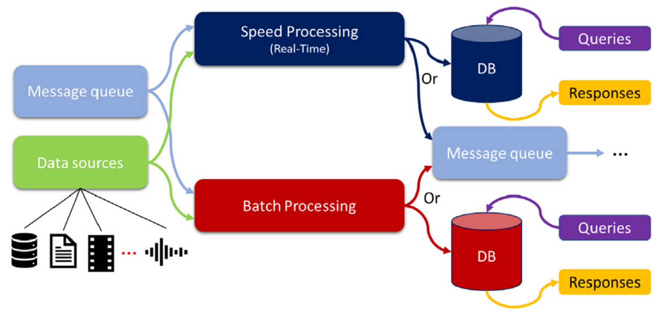

Fig. 3. Lambda Architecture General Scheme.

Table 7
Pros and Cons of Lambda Architecture.

Pros	Cons
- Process data in real-time or in batch processing in separate ways.	The reliability of two ways of treatment is most costly than other architectures if the two execute the same treatment.

health systems (Diaz et al., 2016). Several implementations of Lambda Architecture in smart Environment management, big data storage and analytics can be found in (Villari et al., 2014). Among the criticisms that have been made against lambda architecture is the need to make twice the developments for the real-time branch and the batch branch. It is possible to perform a batch processing and in real time with flow processing is what the Kappa architecture described below does (Kreps, 2014). Fig. 3 Table 7.

Among use cases in agriculture 4.0 using a lambda, we would like to highlight: Roukh et al. proposed WALLeSMART, a cloud platform based on lambda and specifically developed for Smart Farming. This platform implements Apache Kafka to store temporary data before their treatment. Apache Hadoop and the programming model Mapreduce is used for the batch processing while Apache Storm process data in realtime. The originality of this architecture is the coupling of a NoSQL database Apache Casandra and a SQL database, PostgreSQL where data is stored in the function of its nature. The GraphQL query language allows to querying databases. (Roukh et al., 2020; Roukh et al., 2020). Debauche et al. describe a lambda architecture for digital phenotyping (Debauche et al., 2020) and farm animals' behaviors coupled with an Application Hosting Architecture based on Apache Mesos and Docker containerization to facilitate the deployment of various applications. An API interconnects and controls accesses between the Lambda Architecture and the Hosting Application Architecture. The Lambda architecture is based on Apache Beam to easily change the runner in the function of the technology evolution and improve its sustainability. Apache Druid is used to store time series data (Debauche et al., 2019) and metadata of data stored in the Datalake based on Apache Hadoop (Debauche et al., 2018).

Table 8
Pros and Cons of Kappa Architecture.

Pros	Cons
- Very efficient for real-time processing thanks to in-memory processing. - Optimized cost because allows real-time and batch processing.	- Batch processing emulates thanks to micro-batch treated via the real-time way. - Not able to process large batch size. - Must be finely tuned from data to obtain the best performances (Nkamla Penka et al., 2021).

A variant of this architecture, named Unified Lambda architecture combines batch and stream pipelines which runs concurrently, and then the results are merged automatically (Siciliani, 2015). AllJoyn Lambda integrates AllJoyn a framework that offers: (1) proximal devices and applications discovering; (2) specific devices framework adapting; (3) transmission between devices with Bluetooth, Wi-Fi, etc.; (4) interoperability between operating systems; (5) efficient and secure data exchange through D-BUS (Villari et al., 2014).

The **Kappa architecture**, proposed by Jay Kreps from LinkedIn (Kreps, 2014), simplifies the Lambda architecture by combining real-time and batch layers. This cloud architecture differs from the Lambda architecture by using a non-permanent storage system of data in an unchangeable log file such as system as Apache Spark or Apache Kafka, and consequently allow only storage for a limited time in order to allow an eventual reprocessing of these data. Batch and Speed Layers are also replaced by a stream processing engine. So, the Kappa Architecture is composed of two layers: streaming and serving layers and can be implemented with a publish-subscribe messaging like Apache Kafka, which facilitates data ingestion. Fig. 4.

The main advantage of this architecture is its simplicity. It avoids having to maintain two separate code bases for the batch and speed layers. When processing on real-time and historical data are the same, a Kappa Architecture must be used. Fast Data Architecture is a variant of Kappa Architecture in which the data are no longer read from files but from an additional mechanism like Kafka that captures multiple streams combines them before being processed by the speed layer (Lakhe, 2016). Persico et al. achieved a benchmark of Lambda and Kappa architectures and show that Lambda outperforms Kappa for social networks data (YFCC100M) processing (Persico et al., 2018). Table 8.

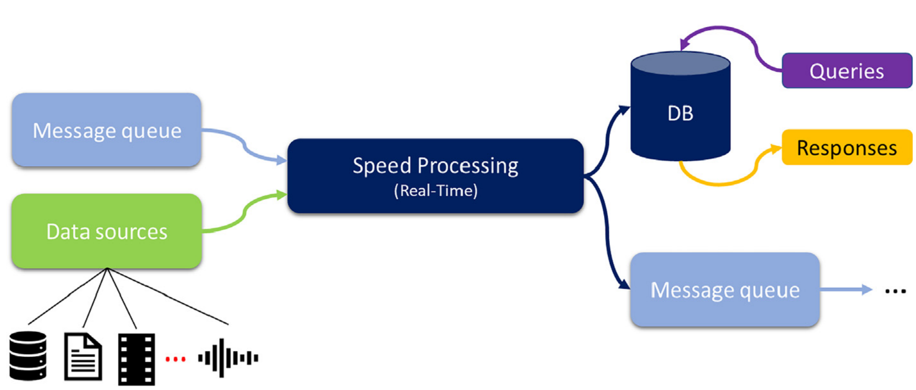

Fig. 4. Kappa Architecture General Scheme.

Table 9
Qualitative evaluation of cloud-centric architecture.

Criterion	Batch	Stream	Lambda	Kappa
User Proximity	*	*	*	*
Latency & Jitter	*	*	*	*
Network Stability	*	*	*	*
High throughput	***	***	***	***
Reliability	***	***	***	***
Scalability	***	***	***	***
Cost Effectiveness	***	***	*	**
Maintainability	***	**	*	**

Other Architectures derived or inspired of the previous architectures have been developed to address specific problems such as (1) **SMACK** (Estrada and Ruiz, 2016) which attempts to propose an optimal architecture with fixed components; (2) **Liquid** (Fernandez et al., 2015) is an architecture which provide low latency, incremental processing, high available with isolated resource, and able to store high throughput data at low operational cost architecture; (3) **Butterfly** (Lakhe, 2016) proposes to unify batch, speed and serving layers in a unique platform in which data are organized as a collection of three types of abstractions; (4) **Zeta** (Scott, 2015) which integrates a lambda architecture with business aspect of the enterprise; (5) **BRAID** (Giebler et al., 2018) is a hybrid processing architecture where all coming data and configuration file of processing, and eventually processing results written back are stored in a shared storage; (6) **IoT-a** (Hausenblas, 2014) is composed of three blocks: Ad-hoc queries, a Database, and a Distributed File System; (7) **Polystore** (Meehan et al., 2016) implements a multiple database system PostgreSQL, SciDB and Accumulo because a database alone cannot store all types of data efficiently. Table 9.

The analysis of the literature achieved shows that two major generic architectures: Kappa and Lambda allows to address of various use cases and are widely implemented and proven in other domains of the Internet of Things. The Lambda is more expensive to implement than the Kappa because of the need to maintain two separate parallel processing branches for stream processing and batch processing. It is interesting if different processing are carried out on the two processing branches. Otherwise, a Kappa architecture with a single processing branch that processes both the streams and the data in batches is more appropriate in most cases because it is cheaper and easier to maintain because a single code performs both types of processing (stream and batch). Looking at our first two research questions, we observe that Lambda and Kappa cloud architectures are efficient but these architectures alone operating in central cloud cannot address, for example, use cases where very low latencies are required. They will have to be hybridized and completed to address these particular cases. Two possibilities are available to us. The first way consists in associating several specialized cloud platforms to make it possible to obtain greater genericity or at least to better cover a domain. The second consists of supplementing the cloud-centric architectures that we have just mentioned with other architectural elements in order to better address the specific needs of Agriculture 4.0.

4.2. Extension of the cloud paradigm

With the increase of the amount of data produced by the myriad of connected things, the amount of data to process, to transfer by network, and to treat in the cloud computing have called into question the architecture of storage and data processing. To solve the problem, two ways have been proposed, the first is Multi-Cloud Computing, the objective of which is to ensure redundancy in order to improve latency. The second is the Federated Cloud with the aim of pooling resources for better use.

Multi-Cloud Computing (MCC) (Manyika and Chui, 2015) is an extension of Cloud Computing paradigm where services are distributed on multi-clouds. In this architecture the workflow is distributed entirely in the cloud, data redundancy is also verified. One advantage of the MCC is the high recovery rate but it has the same disadvantages as Cloud Computing, along with complexity and portability issues.

Kazim et al. proposed a framework to deliver IoT services and establish cooperation across multi-clouds. An authentication allows communicating cloud to authenticate each other cloud dynamically. While a service selects the best IoT service matching with user requirements among multiple clouds and taking into account the SLA parameters agreed between the user and the provider (Kazim et al., 2018).

Federated Cloud (FC) aggregates resources of multiple cloud providers to improve users' freedom and allows users to choose where they want to deploy their applications. A Federated cloud can be defined as a voluntary collaboration between heterogeneous cloud providers collaborating to share their own unused resources. Using a cloud federation helps to ensure service performance during load ups with resources borrowed from other clouds. In addition, the geographical dispersion of the installations makes it possible to migrate to another installation and to guarantee the service in case of breakdown. A unified interface allows to use it an easy consultation of the offered services. Finally, thanks to the dynamic distribution of the load, it is possible to bring the treatment closer to the user and consequently improve the Quality of Service (Assis and Bittencourt, 2016). Cloud federations include European Federated Cloud (Sipos et al., 2013), Massachusetts Open Cloud, Mosaic (Petcu et al., 2013), IEEE P2302, and Open stack Keystone.

Drakos et al. described agINFA, a common research data infrastructure for agriculture, food and the environment using EGI Federated Cloud. This infrastructure allows to partner to share research infrastructure components, APIs, a registry of web-based information service and dataset for agriculture (Drakos et al., 2015).

4.3. Distributed architectures

The post-cloud approaches allow to improve latency and jitter for immobile entities but do not provide an answer adapted for mobile devices and local awareness. The large amount of data generated at the edge has increased the speed of data transportation that is becoming the bottleneck for the cloud-based computing paradigms (Shi et al., 2016). Moreover, the treatment of data in the cloud does not offer any guarantees about privacy, on the response time and real-time actuation because the huge number of devices increases the latency and jitter. Moreover, the mobility of devices and power constraints makes the communicaion difficult with the cloud all the time (Botta et al., 2016; Zhou et al., 2017). The aim has been to bring data storage and processes data, filtering, and data analysis closer to data-producing objects to limit bandwidth consumption and relieve the cloud. Three major paradigms have been proposed to address these issues and bring cloud

Table 10
Pros and Cons of Fog Computing.

Pros	Cons
- Fast response time in avoiding transmission of data to the cloud (Sharofidinov et al., 2020). - The local storage and processing capabilities prevent data loss and outages when the Internet connectivity is limited (Sharofidinov et al., 2020). - Sensitive data can be filtered locally. In this case, only the data model is moved in the cloud (Sharofidinov et al., 2020), and data validation, compression, and encryption. - Gateway at fog level ensure the compatibility between old and modern devices (Sharofidinov et al., 2020) and various protocols for communication. - Improve the resilience thanks to the decentralization of the treatment on network devices (Sharofidinov et al., 2020).	- Failure or outage of the gateway can defeat thousands of devices. - The limited processing and memory capacities do not allow the deployment of algorithms requiring significant resources or the carrying out of long-term processing.

computing-like capabilities to the edge of the network. All these infrastructures manage mechanisms of Virtual Machine (VM) or containers migration and adjust if needed, the provisioning of capabilities where users are located. Moreover, the three paradigms allow the creation of federated infrastructures in which can coexist multiple edge infrastructures which can exchange information and services (Roman et al., 2018).

4.4. Elements of distributed architectures

In order to always bring closer, the processing capacities of intermediate processing have been set up between connected objects and the cloud at the network level (Fog Computing) and at the level of telephony providers (Mobile Edge Computing).

Fog Computing is a concept created by Cisco Systems and is an extension of the cloud computing paradigm (Munir et al., 2017) in which computation, storage and network services are provided between end devices and cloud/ classify and analyze the raw IoT data streams at near-edge and edge network level (Cisco, 2018). Fog nodes are either physical components such as gateways, switches, routers, servers etc. or virtual components such as virtualized switches, virtual machines, cloudlets, etc.; deployed following private, community, public or hybrid. Private nodes are reserved for a single organization, community nodes are used by a community, public nodes are dedicated to the general public, and hybrid mix the third previous modalities (Uehara, 2017). This paradigm allows to limit data transfer on cloud, reduce latency (Sethi and Sarangi, 2017), and jitter thanks to a three-tier architecture (Roman et al., 2018). In this hierarchical architecture, the analysis of local information is achieved at the low level and the coordination and global analysis are performed at the top level. The Fog Computing supports mobile devices (Sethi and Sarangi, 2017), response time in real-time or predictable latency (Lopez et al., 2015), bandwidth saving, an improving of security and resilience, scalability, multi-tenancy, advanced analytics, and automation (Byers, 2017), cost-effective services (Yang, 2017). Fog Computing allows also the federation of fog infrastructures in order to allow cooperation between multiple organizations (Roman et al., 2018). Furthermore, the architecture is optimized for a use case and applications which must run on them (Byers, 2017). Fog Computing differentiates from cloud computing mainly by the proximity with end-users at the edge of networks localized or distributed geographically consisting in many relatively less resourceful (Munir et al., 2017). In addition to network equipment, fog computing can also be carried out in cloudlets and micro data centers. Cloudlets were proposed to address the end-to-end responsiveness between mobile devices and associated clouds. Cloudlets (Mach and Becvar, 2017) are micro data center geographically deployed in vicinity of End Users. This mobility-enhanced small-scale cloud data center is composed of computers with high computation power which provide both computation resources and storage. Cloudlet is much more agile (highly dynamic provisioning) than cloud due to user mobility churning. The mobility of users implies the use of a virtual machine to rapidly instantiate compute-intensive and latency-intensive applications and migrate the offloaded services between different cloudlet in the function of the user mobility. Cloudlets must be firstly discovered, selected among several candidates before starting provisioning. At the end of the session, the instance is destroyed (Ai et al., 2018). Cloudlets are accessed by mobile user equipment via Wi-Fi imply a high latency caused by the network and switch between mobile network and Wi-Fi and by consequence Quality of Service (QoS) and Quality of Experience (QoE) are hard to fulfill (Mach and Becvar, 2017; Manyika and Chui, 2015). Moreover, Cloudlets cover usually a small region and do not offer any guarantee on ubiquitous computing and scalability in service (Manyika and Chui, 2015). MicroData Centers (MDCs) were proposed by Microsoft Research. It is designed to extend cloud data centers as cloudlets. MDCs are enclosures contemning all types of equipments (computing, storage, network) needed to provide a secure computing environment in order to run customs applications requiring low latency. MDCs are also well adapted to provide processing resources to end devices on battery or with limited computing capabilities. MDCs can be adapted in function network bandwidth and user needs thanks to certain flexibility in terms of latency and scalability of the capacity (Wang et al., 2020).

Guardo et al. proposed a framework composed of two fog layers respectively filtering and aggregating data, and clustering analysis, actuation management, and alert. The framework aims to improve computational load balancing between fog and cloud in order to reduce the amount of data to transmit to the cloud, reduce the waiting time for the user (Guardo et al., 2018). Taneja et al. proposed a SmartHerd an IoT platform dedicated to smart dairy farming based on microservices and Fog-assisted. The IoT gateway received data from transceivers, archived data aggregation, preprocessing, classification, feature selection, send critical alerts to farmers, and transmit data to IBM Watson IoT platform via MQTT protocol. In the IBM Watson IoT platform, a broker picks up data and store them in a Cloudant NoSQL JSON Database. Python Virtual Machine and Java Virtual Machine were used as containers equivalent for microservices deployment at fog level (Taneja et al., 2019). Sharofidinov et al. described a 4 layers architecture (Sensors Layers, Fog Layer, Network/Cloud Layer, and Application Layer) based on LoRa to monitor and predict the state of a greenhouse from a random forest algorithm. In the Sensor Layer, sensors acquire temperature, soil and air humidity, CO2 rate, and illumination connected to TTGO LoRa32 (ESP32 with LoRa Sx1276 chip) which are transmitted to the gateway by LoRa. At Fog Layer, preliminary analysis with Machine Learning algorithm, diagnosis of sensor status, and data compression are achieved. In the Network/Cloud Layer, compressed data are transmitted in order to be deeply analyzed and stored. Finally, in the Application Layer, analyzed data are converted in readable form to allows the monitoring and the control of the greenhouse (Sharofidinov et al., 2020). Table 10.

Mobile Edge Computing (MEC) was proposed by ETSI and is deployed by telecommunication companies on the edge of the network, which is characterized by ultra-low latency and high

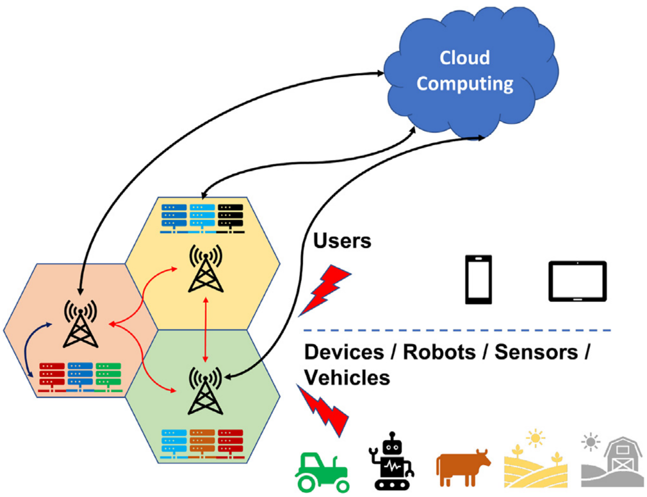

Fig. 5. Mobile Edge Computing General Scheme.

Table 11
Pros and Cons of MEC.

Pros	Cons
- Reduces needs in connection, response time delay, the congestion of other parts of the network (Valecce et al., 2019). - Use low level message from Wi-Fi to determine the location of each device (Location awareness) (Valecce et al., 2019). - MEC Server can be used as power open to applications and services (Valecce et al., 2019).	- Usable only for devices connected in Wi-Fi or 3GPP.

bandwidth. (Roman et al., 2018; Zhou et al., 2017). At the very beginning, Mobile Edge Computing (MEC) aims to bring real-time, high-bandwidth, and low-latency access to dependent applications known as cloud computing capabilities; in addition to, information technology (IT) features of cloud computing. MEC is distributed at the edge of the network. In fact, a new class of cloud-native applications are easily accessible, because of the close position of Edge Computing to the end user and apps. Also, it allows network operators to open their environment to a new ecosystem. As a result of this significant change, MEC application can be used in LTE macro base stations (eNBs), 3G radio network controllers (RNCs), Wi-Fi access points, edge network routers, and enterprise edge servers. MEC platform contains two main hosting infrastructures. The first is formed by hardware resources and a high-resolution screen. The second is composed of manageable applications with numerous capabilities such as the application of virtualization manager and platform services (Zhou et al., 2017). An important challenge for the MEC is the VM migration that must optimize the tradeoff between migration gain and migration cost and select optimal location (Ai et al., 2018).

Tran et al. investigated the collaborative Mobile Edge Computing in 5G Networks. MEC extends processing and storage resources at the edge of the Radio Access Network (RAN) while C-RAM is based on centralization of the base Station by means of the virtualization. Authors argue that both technologies are complementary in the 5G ecosystem (Tran et al., 2017). Fig. 5 Table 11.

Fan et al. combined MEC with data link management, combining with the industrial CAN bus characteristics to monitor water. Field Programmable Gate Arrays (FPGA) Altera implementing the AVALON bus was used to implement the system. Moreover, they

Table 12
Evaluation of distributed architecture with our criteria.

Criterion	Fog	MEC
User Proximity	**(*)	***
Latency & Jitter	*	*
Network Stability	***	**
High throughput	**	**(*)
Reliability	***	**
Scalability	*	*
Cost Effectiveness	**	**
Maintainability	**	**

propose a protocol to model random network disturbances and an online task offloading algorithm based on the monitoring of task execution (Fan and Gao, 2018). Valecce et al. proposed a 5G-robotics reference architecture for smart agriculture composed of UAV-Based Monitoring and connectivity, Machinery automation, and MEC Applications Server. UAVs/satellites capture high-resolution images during patrolling, which coupled with sensors data trigger a precise crop management. UAVs can also collect data or serve as a 5G mobile station. In field, image processing coupled with sensors data can be used for decision making. MEC allows to process gigabyte/s of data produced by autonomous vehicles and robots (Valecce et al., 2019). Table 12.

The development of fog computing and its counterpart for MEC wireless networks allow processing capabilities closer to users to improve response time but with lower computational capacities compared to the cloud. There are inherently two questions: Which association strategies to use between the cloud and the other levels of processing in the network? How to distribute the load between these different levels: local (Edge), network (Fog), and Cloud processing.

4.5. Collaborative computing strategies

In order to address, our fourth research question, we try to identify different possibilities to compose architectural elements. Indeed, different collaboration strategies between the different levels of data processing (cloud, fog, edge) can be considered depending on the particularities of the use cases. In the next paragraphs, we describe possibilities of collaboration between different treatment entries, and we illustrate each one with few examples.

Edge-Cloud aims to connect devices directly with the cloud that performs data processing. This strategy is often used by UAVs and UGVs which preprocess data before its transfer to the cloud because image treatment needs processing power and storage capabilities. The default of this approach is that the delay of the whole process from data transfer via high throughput wireless or cellular protocol to the transmission of processing results cannot be guaranteed because of the fluctuation of data rates linked to wireless networks (Wang et al., 2020). The processing of data can be achieved in an online mode with a real-time data transmission and processing by a stream, Lambda, Kappa or derived architecture of these one. An offline strategy with a data transfer by means of a computer and Internet connection on the cloud after the UAV fly and processing with a Batch, a Lambda, or a Kappa architecture or a derived architecture of these one is also possible. This latter costly avoid data transmission and is suitable for monitoring crops or livestock that do not require direct action.

Agriculture 4.0 uses in particular Unmanned Aerial Vehicles (UAVs) equipped with various sensors in order to improve the time of data collection, in reducing the cost of acquisition compared to traditional field phenotyping technologies. According to Tang et al., edge-cloud is majorly used in smart robots to reduce complexity (Tang et al., 2021). Indeed, the images of drones to be used must be orthorectified and assembled. These operations require significant resources in terms of computing power, and memory. All these collected data must be rapidly processed, analyzed, and visualized. **Agroview** (Ampatzidis et al., 2020) is a platform that developed a cloud and AI-based application to survey and assess the agriculture field, deployed on Amazon Web Services (AWS). A website allows the upload of images or existing orthomosaic, the consultation for each tree field e.g., number of trees, tree gaps count, area of the field, the average height of trees, canopy area, etc. The website also allows the stitching of an orthomosaic and the generation of a Digital Surface Model (DSM). A tree detection algorithm developed in C allows the detection of individual tree and tree gap, and estimate tree parameters such as height, canopy area, health/stress estimation. The pipeline of treatment uses a Faster R-CNN to detect the region of interest (ROI) and the ResNet101 network allows to detect trees and row orientation. Afterward, the Yolo classifier using Darknet19 was applied along each row of trees to obtain a more precise detection. Debauche et al. presented an Edge-Cloud architecture for the analysis of cattle behavior from 9-DOF IMU data sampled at 100 Hz and GPS location sampled at 0.5 Hz that is then processed with an algorithm proposed by (Andriamandroso et al., 2017) in batch processing (Debauche et al., 2019; Debauche et al., 2020). Popescu et al. proposed an integrated system UAV-WSN-IoT where WSN data is collected by UAVs before their transmission to the ground control station and afterward to the cloud (Popescu et al., 2020). Debauche et al. proposed an architecture for scientific research dedicated to honeybee Colony Collapse Disorder. In this architecture, data is compressed on LoPy at the edge level before its collection by the LoRaWan gateway and its transmission to the Lambda architecture in the cloud where it is processed (Debauche et al., 2018).

Edge-Fog aims to connect devices directly with network components such as gateways, routers that perform data processing. The major benefits of this approach are an optimization of the bandwidth, a reduction of traffic and latency, a better privacy, and an improved security level (Badidi, 2020). Fog nodes collect, aggregate, filter, encrypt, compress, and process IoT data (Gupta et al., 2020). This way is used for example by milking robots where data are processed by a computer close the robot and can be viewed remotely by the farmer. 5G also promotes mobile edge computing (MEC).

Debauche et al. presented an AI-IoT architecture for the deployment of Artificial intelligence algorithms and Internet of things services at fog level using docker containerization and Kubernetes orchestration. This architecture has been developed to automatically deploy AI algorithms after retraining when performances (accuracy, recall, precision) are improved (Debauche et al., 2020). Debauche et al. proposed a Multi-Agent System (MAS) deployed at edge level allowing to control abnormal data present in sensed data and eventually cure this data when it is possible. The MAS simultaneously manages pivot irrigation, plant diseases and pests' detection, and their curation. The data is partially transmitted to the cloud to improve the detection of diseases and pests and retrain AI algorithms before their redeployment at the edge level (Debauche et al., 2020). Debauche et al. described a fog architecture in which a Gated Recursive Unit (GRU) algorithm is deployed on NVIDIA Jetson Nano for real-time poultry monitoring. GRU is simpler than LSTM algorithm. GRU is built to avoid varnish gradient problems. Periodically data is transmitted to the user interface implemented in NodeJS in the cloud (Debauche et al., 2020).

Edge-Fog-Cloud is a paradigm in which data are partially processed in the fog and more complex treatments are achieved in the cloud. This way is used by wireless Sensor and Actuator Network (WSAN), which passes through a gateway that provides interconnection between the devices and the backhaul which transit then data to the cloud. However, the right balance between cloud and edge/fog computing is required (Badidi, 2020) based on available resources and whether or not the task is sensitive.

Taneja et al. used a strategy Edge-Fog-Cloud to develop a detection system of lameness for cattle. The data from the pedometer is transmitted to the Fog node by means of a Long-Range proprietary protocol at 433 MHz on a distance of 2 km. Fog node stores in local database, preprocess and aggregates them. Fog node communicates with IBM Watson IoT Platform with MQTT protocol. Arriving data are picked up and stored in Cloudant NoSQL JSON database in IBM cloud. A mobile application synchronizes data with PouchDB, its local database via the REST API of Cloudant database when an Internet connection is available (Taneja et al., 2020). Alonso et al. presented Global Edge Computing Architecture (GECA), a modular tiered architecture (IoT Layer, Edge Layer, Business Solution Layer) to monitor dairy and feed grain state in real-time. In this architecture, a Distributed Ledger Technologies provides security from IoT Layer to Business Solution Layer. In the IoT layer, a set of agents call oracles to verify incoming data and afterward calculate hash of data with SHA-256 which is stored in the blockchain to verify the non-alteration of data. In parallel data is encrypted with the RSA algorithm and then sent to the Edge layer. The Edge Layer is responsible of the preprocessing of data and filters out data transmitted to the cloud. It enables also various data analyses. In the business Solution Layer, final storage, authentication, analysis for decision making is achieved. It provides also a knowledge base and APIs (Alonso et al., 2020).

Edge-Edge is a paradigm in which devices interact to collaborate, exchange, and process data. The deployment of the 5G network allows the interconnection between UAVs and UGVs/ agricultural machinery (Tang et al., 2021). This high throughput network will allow to developping new collaboration between UAVs/ UGVs and agricultural machinery, for example, a drone will provide information to a harvester to avoid a non-desirable area of the field or avoid obstacles. A fleet of drones can also collaborate to coordinate their operations on the field between them of course subject to availability in rural areas, a transmission network with sufficient bandwidth and short-latency or capabilities to communicate between them in direct connection or in a mesh network. (Tang et al., 2021).

Four cooperation strategies have been identified, two of which use the cloud, namely Fog-Cloud and Edge-Cloud. The other two remaing, do not involve the cloud; namely, Fog-Edge, and Edge-Edge cloud. The first two strategies complement the cloud to help us to address issues relating to production data and trade secrets, network congestion, and response times. The other two strategies do without the cloud and therefore assume that the devices/ vehicles have sufficient capacity to perform the processing. Despite these cooperation strategies between different levels of processing, some questions remain unanswered: How to store all the raw data when the data is so important that it would take colossal means to process it? What about security? How to organize the distribution of tasks between the edge, the fog, and the cloud? How to ensure operation and/ or treatment when network connections are intermittent or faulty? How to improve the maintainability of these architectures? These are the questions that the new trends that we describe in the next paragraph attempt to answer.

5. New trends

In this section, we present two emerging architectures not based on the batch or/and real-time architectures or their derivatives. Afterward, we describe Osmotic and Dew computing as two new paradigms, which allow us to respectively choose where the processing must be achieved and improve the user experience. New trends are additional elements that allow enriching the analysis of Section 4 in order to address the third research question.

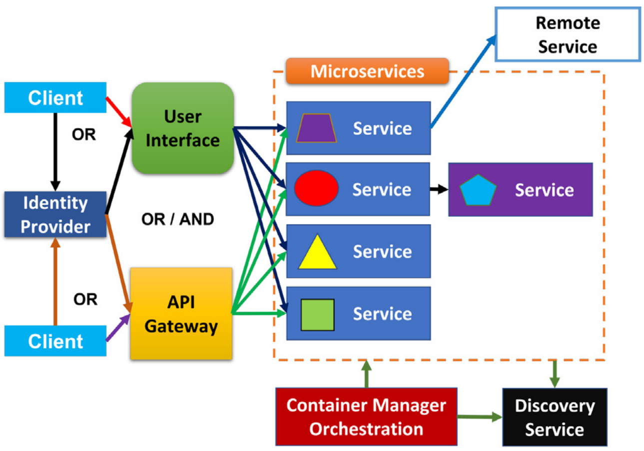

Fig. 6. Microservices Architecture General Scheme.

Table 13
Pros and Cons of Microservices Architecture.

Pros	Cons
- Fractionating of monoliths facilitates the maintainability and scalability of low coupled microservices. - The discovery of micro-services allows the development new applications more easily than with monoliths. - More resilient, when a microservice is down, all others continue to function.	- Need to find microservice adapted with needs. Fraction complex monolith is not easy.

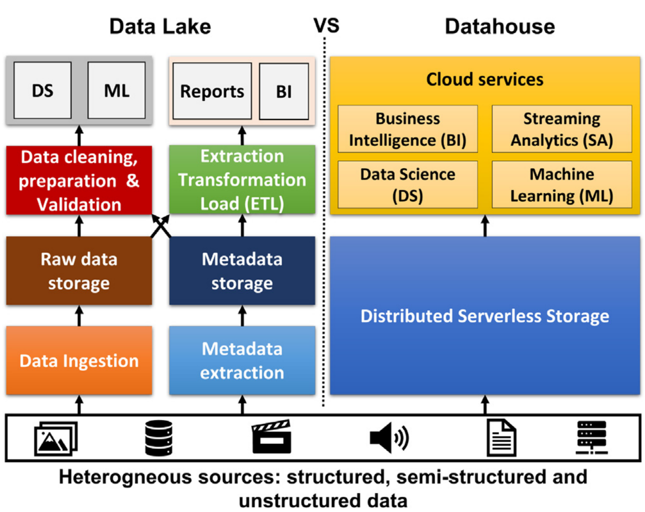

Fig. 7. Data lake and Lakehouse General Scheme.

The **Microservices Architecture** (MA) is a new system software design pattern that divides complex monolithic application in micro services dedicated for a single function. Microservice addresses defects of monolithic applications in which improving of service performance needs multiple deployment; a change in a function can affect all the monolith due to high dependencies between components; all the monolith uses a sole technology stack and development standards which limits possibilities to solve problems of physical heterogeneity.

The advantages of this architecture are using a lightweight communication mechanism to interact between services with a minimal overload (Sun et al., 2017). The design proposed by (Sun et al., 2017) is composed of 8 microservices (Geo, Security, Tenant, Devices, Big Data, Automation, AI, and Application) and a core service coordinating. These services provide respectively: (1) Geo, a GIS layer to render data; (2) Security, user/group/role management, access control, administration, and authentication mechanism; (3) support for multiple IoT applications with a single core; (4) device plugins and communication protocols for sensing and actuating; (5) scalable persistence to store data; (6) process, analyze events and notify appropriate participant; (7) Artificial intelligence tools for IoT big data; (8) components to interact with client interfaces; (9) support for data exchanging by message with the devices. Authors argue that their approach is more flexible, scalable and platform-independent. Fig. 6 Table 13.

Bixio et al. proposed a stream processing architecture event-driven based on proxy, adapter, and data processing microservices. This architecture extends the IoT platform Senseioty and using the Java OSGi framework (Bixio et al., 2020).

The **Data Lake Architecture** (DLA) (Fang, 2015; Miloslavskaya and Tolstoy, 2016) enables the storage of large volumes of data of all types: raw data in its native format, structured, semi-structured, in a cost-effective manner. In this architecture, data is stored in its native format until it needs to process them by engines (Miloslavskaya and Tolstoy, 2016), which allows a fast transformation and refinement of stored data regardless of the amount of data stored. The architecture makes it possible to consume all types of data (logs, web services, database, files, etc.); different ingestion systems consume the data and then stored it in data repository. Once the data is stored, query systems can query the data lake. This architecture is considered in the corporate world as an evolution of existing architectures. The advantage of the Data Lake architecture is that it can easily and inexpensively store large amounts of data. It is particularly well suited to storing data in a typical format. In Enterprise Data Lakes are used; in addition to, data warehouses. Data lakes are, however, unsuitable for assessing data quality, data can be placed in data lakes without content control, and performance is also poorer than on specially designed and optimized infrastructures. The Lakehouse is a variant of the Data Lake where storages of data are generally achieved with Hadoop in the data lake is replaced by a distributed storage such as Amazon S3, Azure Blob Storage, Google Cloud Storage, and analysis are directly achieved by infrastructure managed by Cloud Service Providers such as Amazon Athena, EMR, or Databricks, Google Data proc, Azure HDInsight. The Fig. 7 provides a comparison between data lake and gatehouse structure.

It crucial in agriculture to explore datasets from different sources. The data lake is indicated to manage the complexity of agricultural ecosystems and centralized all data sources to find new correlations. (Madera et al., 2017). A data lake provides views based on metadata. It is nevertheless necessary to have advanced analysis tools for predictive modeling and statistical analysis. López et al. used a data lake to achieve the fusion of data from different domains in smart the agriculture context (López et al., 2020). Gallinucci et al. (Gallinucci et al., 2019; Gallinucci et al., 2020) present an innovative architecture 3 tiers architecture, called Mo.Re.Farming (MOnitoring and REmote system for a more sustainable FARMING) based on a data lake using Apache Hadoop and storing structured, semi-structured, and unstructured raw data, and in which subsequent processing and enrichment activities are separated. An Operational Data Store (ODS) using PostgreSQL with PostGIS to stores structured and detailed data and address limitations of big data solutions in properly handling continuous field geographic data. Finally, a spatial cube enables Spatial OnLine Analytical Processing (SOLAP). Neves et al. described an architecture in which raw data is stored in a datalake. Then, ETLs transforms data to be storable in a database. The data is enriched thanks to a knowledge base and its exploration by data mining algorithms (machine learning). The result of processing is filtered

Table 14
Pros and Cons of Datalake/DataHouse.

Pros	Cons
- Store the data in its raw form without transforming them immediately.	- Availability of results depend of the ingesting speed by processing services.
- Allow store massive low-value data without investing energy to transform and store them in a database.	- Data analysis by sampling does not give exact results but is estimated.
- Provides a solution to situations where the volume of data is so large that it can no longer be processed immediately	- Data House may be limited by the services offered by cloud providers for data analysis.

Table 15
Pros and Cons of Osmotic Computing.

Pros	Cons
- Micro Element (microservice + micro dataset) easy to migrate between fog and cloud.	- All datasets are not decomposable in micro dataset.

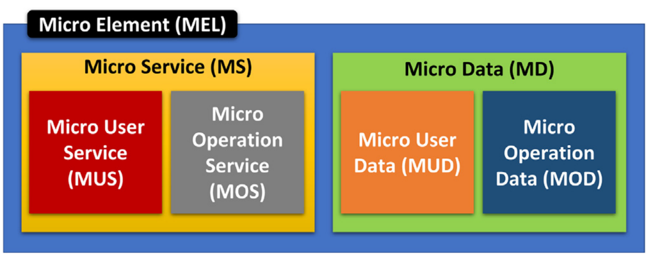

Fig. 8. Micro Element Structure.

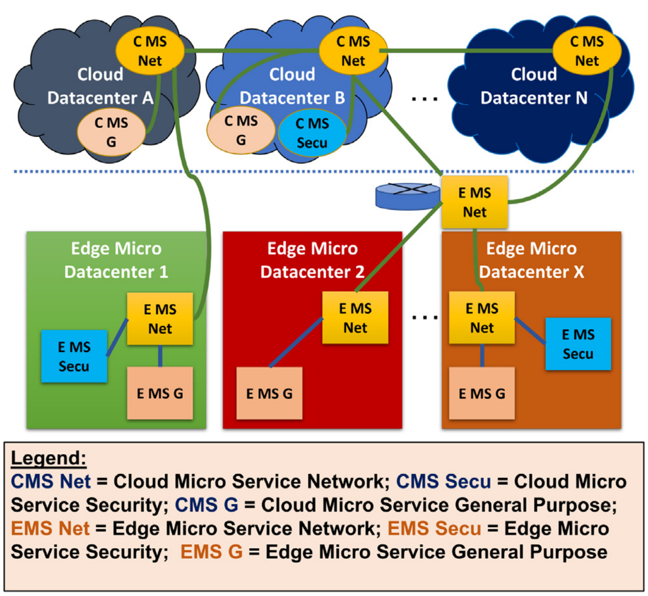

Fig. 9. Osmotic Computing General Scheme.

to improve the quality of structured data (Neves and Cruvinel, 2020). Table 14.

Osmotic Computing (OC) (Villari et al., 2016) is a new paradigm inspired by the chemical osmosis process that corresponds to a dynamic and bidirectional flow of microservices between cloud and edge. OC exploits container-based solution to allows an automatic deployment of portable, mobile, and cross-platform microservices between Edge and cloud levels (Villari et al., 2016). Osmotic computing introduces the concept of Micro Elements (MELS) which decouples user data and applications in Micro Services (MS) i.e. a docker container and Micro Data (MD) i.e. an entity self-explicative in JSON. MS associates one operating system (Micro Operation Service) with an application (Micro User Service) while MD associates a microservice configuration (Micro Operational Data) and User data (Micro User Data). These MELS can be deployed on Microcontrollers (MCU) or Multiprocessor (MPU) (Villari et al., 2017). Table 15.

The bidirectional migration of microservices between Edge and Cloud must, on one hand, avoid application breakdown and QoS degradation and on the other hand manage them dynamically, in high heterogeneously physical resources context, in the function of infrastructure and applications requirements (Villari et al., 2016). Carnevale et al. have applied osmotic computing to the Internet of Things by means of a distributed multi-agent system. Each agent is self-orchestrated, works independently, and manages the workflow as a composition of MELs. It monitors the overloading state of microservices by means of response time metric and decides to relocate them to another agent based on a Deep Reinforcement Learning algorithm or Time Series Analysis (Carnevale et al., 2019). Figs. 8 9.

In an IoT context, OC allows to deploy lightweight micro services at edge level while complex micro services are deployed at fog/cloud level, and balance load between edge, fog, and cloud. (Maksimović, 2018). Morshed et al. proposed to use OC to distribute Deep Learning across edge, cloud, and mobile edge in a holistic way (Morshed et al., 2017). However, Kaur et al. in their Osmotic Computing applications survey have identified the need of standardization in terms of infrastructure deployment and micro-services distribution. The orchestration is crucial to manage efficient services. Security remains an important challenge because the service migration is supported by different layers (Kaur et al., 2020).

Dew Computing (DC) (Skala et al., 2015) allows to further improve response times by pushing from Central cloud to end-users, computing applications, data, and low-level services. Client microcomputers are used to store a part of the data locally in the background and to limit access to the cloud, reduce network dependency and drastically reduce processing cost (Skala et al., 2015). Dew computing is the additional piece of cloud computing. It is mainly composed of a wide range of heterogeneous devices and varied equipment ranging from smartphones to smart sensors (Wang, 2016). DC is highly and effectively capable in terms of scalability and ability to perform sophisticated operations and to process numerous applications and tools. Additionally, the equipment of DC is ad hoc programmable and self-adaptive. They have the qualifications to running the process within another process in a distributed way without a focal communication network (Skala et al., 2015). Applications running in the on-premises computers provide services to users and/or devices independently of the cloud but collaborating with cloud services (Wang, 2016). DC can provide access web fraction without Internet connection (WiD), Storage in dew has a cloud copy (STiD), Local database has a cloud backup (DBiD), Software ownership and settings have a cloud copy (SiD), SDK and projects have a cloud copy (PiD), On-premises computer settings and data have a cloud copy (IaD), Other services (DiD) (Wang, 2016). The Fig. 10 presents the dew computing in the general scheme Cloud-Fog-Edge Computing. Table 16.

Rajakaruna et al. presented a dew architecture based on a drone to retrieve and process data, manage WSN, and play the role of dew server. The drone communicates with sensors, and actuators with BLE protocol, collect, store data, and then when the drone is at the

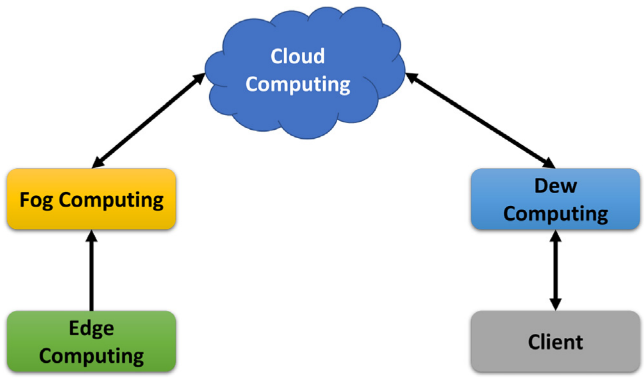

Fig. 10. Dew Computing General Scheme.

Table 16
Pros and Cons of Dew Computing.

Pros	Cons
- Allows access to a local copy of data when the connection is unavailable. - Improve the reliability and the false tolerance.	- Replication of data is bandwidth-consuming. - Difficult to exploit if bandwidth is insufficient.

docking station it sends data to the cloud (Rajakaruna et al., 2018). Grovers et al. described a reliable and fault-tolerant architecture at 4 levels (edge, dew, fog, and cloud) in which sensed data is replicated at edge, fog and cloud level in order to take over the application's control when a server is failed. In their architecture, dew servers are closed and linked with sensors producing data. The fault tolerance is ensured by mobile agents working as a resource exchanging the application and link-state information between us, and the network monitoring agent (Grover and Garimella, 2018).

The **Blockchain** is a distributed digital ledger of transaction distributed maintained by a network of multiple computing nodes. This ledger can be deployed among the IoT nodes network (Bermeo-Almeida et al., 2018). In the blockchain, transactions namely blocks are managed by a specific software platform ensuring the data transmission, processing and storage, and its representation in a human-readable form allowing a consistent view and a consensus between the participants (Kamilaris et al., 2019). Different mechanisms of consensus whose two main ones are the "Proof of Work (PoW)" and the Proof of Stake (PoS). The PoW requires the solving of difficult computational tasks before validating transactions and the adding of the block in the blockchain. In this approach "miners" are in competition to be the first and obtain the rewards, which has an impact on the environment, need expending large a amount of computer and energy, and involves a risk of centralization. While the PoS approach, "validators" are randomly selected with a probability which depends on the amount of stake held. At the end of the validation process, it earns a fee. Other less used consensus mechanisms exist such as (1) Proof of Elapsed Time (PoET) in which each node generates a random wait time and goes to sleep for that specified duration; (2) Simplified Byzantine Fault Tolerance (SBFT), an improvement of Practical Byzantine Fault Tolerance (PBFT) specifically designed for blockchain in which each new block is maintained by a delegation of nodes with increasing authority. Each one uses the internal time to decide when actions must be done; (3) Proof of Authority (PoA) in which approved accounts process to the automated validation of transaction and blocks. Table 17.

Table 17
Pros and Cons of Blockchain.

Pros	Cons
- Data distributed (Alonso et al., 2020). - Immutable, durable, verifiable, secure, and transparent (Alonso et al., 2020). - Transactions P2P at low cost.	- Energy consumption for the complex signature verification process can be important. - Not adapted to store images, video.

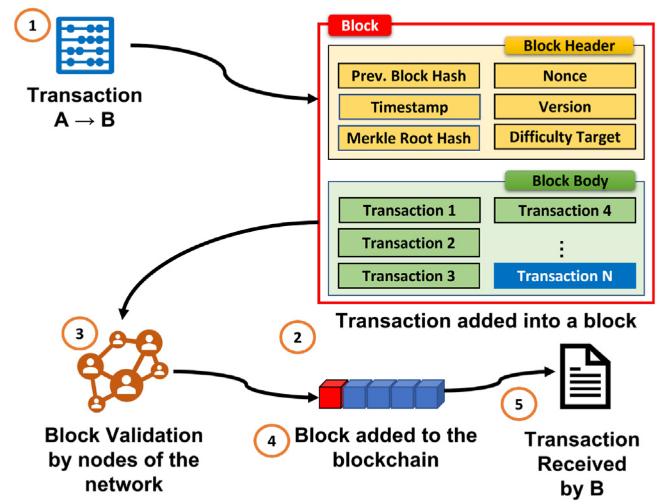

Fig. 11. Blockchain General Scheme.

The Fig. 11 shows the blockchain general scheme.

The block chain is mainly used in Agriculture to make the data of the supply chain transparent and open (Bermeo-Almeida et al., 2018) and ensure the complete traceability of the food chain from the fork to the plate. The block chain allows to record information about: (1) Transactions between provider and farmer as well as information relating to the crops, material and chemical products; (2) The farm, cultivation practices and management, animals feeding, and complementary information such as weather conditions, animals welfare, diseases, treatment, etc; (3) Information about factory and its equipment, the processing method, batch numbers but also financial transactions with producers and distributors; (4) Warehousing, storage conditions (temperature, humidity), methods of transport, transit time, and all financial transactions between the distributors and retailers; (5) food items information such as quantity available, quality, expiration date, time spent on the shelf or in the stock (Bermeo-Almeida et al., 2018; Kamilaris et al., 2019). The Fig. 12 shows an example of blockchain applied to an agri supply chain.

To a lesser extent, secured data storage, remote monitoring, and automation. The blockchain address some challenges of IoT such as decentralization, data anonymization, and security. Moreover, it allows faster and efficient operations, to improve reliability and scalability (Bermeo-Almeida et al., 2018).

The analysis of new trends shows that: (1) **Micro service architecture** allows decomposing monoliths in microservices lowly coupled which makes it easier to maintain it while allowing other services to continue operating. Furthermore, this type of architecture is more resilient because if one of the services is down, the other services due to the weak coupling can continue to operate at least in a degraded mode. (2) **Data Lake/DataHouse** propose a

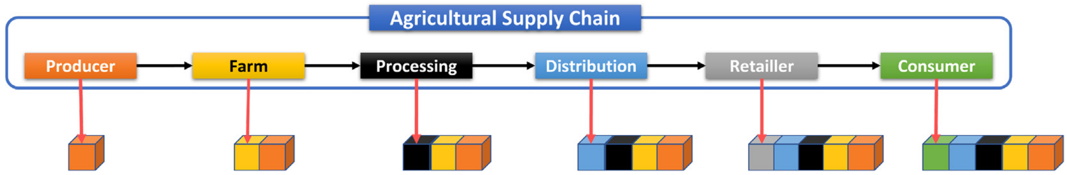

Fig. 12. Supply chain based on a blockchain.

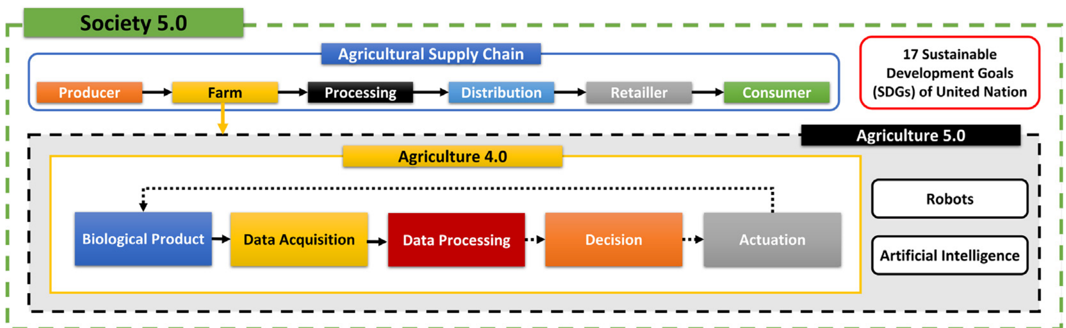

Fig. 13. Integration of the Agriculture 5.0 in the context of the Society 5.0.

new approach Load Transform Extract (LTE) where data are firstly stored in their original format, which are then transformed in order to extract information. This paradigm is particularly well adapted when the amount of data is so important that process all data is too costly. In this case, data can be sampled in order to obtain information. This paradigm is also well adapted if we want to conserve also raw data or complete a generic architecture, for example, to store data that will be processed in batch processing. (3) **Osmotic Computing** attempts to propose a solution to the repartition of workload between fog and cloud in decomposing treatments in microelements composed of a microservice associated with a micro dataset. The osmotic computing could also be associated with the micro service architecture to allow the distribution of instances of microservices at different levels of the network according to their respective load. (4) **Dew Computing** aims to replicate data near sensors or users to continue to store data or allows to continue to consult data when connection is intermittent. It allows improving the reliance of architectures on connection interruptions. (5) **Blockchain** provides an answer to authentication and security problems by making it possible in particular to verify that the data has not been altered or compromised. Nevertheless, it is not possible to store large amount of data such as high definitions images, or videos in the blockchain but hashes of datasets allowing to verify their authenticity well.

6. Towards Agriculture 5.0

According Myklevy et al., the world must improve the amount of food produced by 70% by 2050 to produce global food needs for a population (Mykleby et al., 2016) of 9.7 billion according to the Food and Agriculture Organization of the United Nations (FAO) (Zhang, 2016). To overcome these problems and contribute to achieve the second objective of 17 Sustainable Development Goals (SDGs) of the United Nations (UN) with a timeframe in the range 2015 to 2030, the concept of Agriculture 5.0 has been born (Martos et al., 2021). Agriculture 5.0 aims to increase production sustainably while consuming fewer resources and taking care of the environment. This next wave of agricultural revolution will imply the use of robots integrating machine learning to compensate for the shortage of workers. Farm robots are drastically increasing productivity in improving the human labor workforce and can also harvest a more important volume faster than a human. Nevertheless, these early technologies are still too expensive for most farmers especially small farms (Saiz-Rubio and Rovira-Más, 2020). Fig. 13 show the coupling between Agriculture 4.0 and Agriculture 5.0 and their integration in the context of the agri-food supply chain, the Society 5.0 and 17 Sustainable Development Goals (SDGs) of United Nations (See Fig. 13).

7. Conclusion

Our review is boosted by four research questions dectitaed as follow: (1) Which storage and processing architectures are best suited to Agriculture 4.0 applications and address its particularities? (2) Can generic architectures meet the needs of Agriculture 4.0 application cases? (3) What are the horizontal valuation possibilities that allow the transition from research to industrialization? (4) What are the vertical valuation possibilities to move from algorithms trained in the cloud to embedded or autonomous products?.

The analysis of the literature shows that a multitude of architecture coexists. Nevertheless, the Lambada and Kappa architectures seem to emerge as generic architectures. These must generally be accompanied by complementary architectural components to address specific needs and be part of a storage and processing strategy in which the cloud architecture is a component of the chain or may also and more rarely be absent.

The traditional centralized cloud computing will continue to remain an important part of computing systems (Ai et al., 2018), for sciences even if other paradigms appear. Indeed, cloud, fog, and edge computing complementary interact with each other to form a mutually beneficial and interdependent service continuum. Some functions are naturally more suitable or advantageous at a level than another in function of requirements in response time, computing, or latency tolerance. However, the cloud cannot be completely replaced by fog and edge computing because some computation-intensive tasks can only be processed at the cloud level, which has the computing power and storage capacities (Wang et al., 2020). In Agriculture 4.0, this is particularly the case for the processing of satellite images, the training of artificial intelligence algorithms such as Deep Convolutional Neural Network (DCNN).

New trends make it possible to address various problems: (1) The Data lake/Data House offers a more economical alternative to massive cloud storage in databases. In this paradigm, all data are stored in a state and transform only when they are to be exploited. This approach is particularly interesting on one hand when all data are not exploited and on the other hand when a decision or an action is not expected immediately. Data lake also allows the fusion of agriculture data from various origins in different formats and granularity. (2) The blockchain provides solutions in particular to the security problems, the possibility of distributing data storage and ensuring the traceability of transactions in agrifood supply chains (3) As the literature has shown, Dew Computing can be placed in two different places in the network either as close as possible to the sensors to allow processing to continue during transmission interruptions or as close as possible to users in order to have a local copy of the data in order to be able to consult them offline. It should be noted, however, that for the second option, there are other means of local caching at the device level, for amounts of data of a few mega as those offered by Progressive Web Apps (PWA) by example. (4) Osmotic computing provides a solution to the question of how to distribute the load between the different processing levels (edge, fog, cloud). It uses the concept of microelement associating a microservice and its micro dataset. In addition, osmotic computing can also be associated with micro-service architectures. (5) The microservice architecture offers the possibility of decoupling the monolithic architectures into weakly coupled microservices. These services can be more easily associated, maintained, or evolved independently. The combination of these microservices makes it easier to develop new services for the end-user that are also easier and faster to evolve according to technological developments and needs.

In addition, at the network level, the 5G network offers new possibilities in terms of Wireless Sensors and Actuators Network (WSAN), communication between machines, UAVs, and UGVs. Moreover, the coupling with MEC opens the field of processing close of end-users. The SDN/NFV Architecture allows to facilitate the design, and to improve the flexibility of network. Software-defined networking (SDN) allows decoupling transmission of data and network control functionality while Network function virtualization (NFV) abstracts transfer network and related network functions (Friha et al., 2021).

Two trends in the use of processing architecture coexist, on the one hand, users of a paid or open source IoT platform, and on the other hand, users who develop specific architectures to implement particular use cases. From the point of view of transferability, we understand that it is easier for ready-made chargeable infrastructures and that it can be limited for turnkey open-source infrastructures where the type of license adopted may pose a problem. However, the sustainability of paid infrastructure is conditional on the development granted by the company that manages them and on its financial health. The development of architecture based on paid software bricks is facilitated but its durability is conditioned by the availability and the maintenance of these software bricks. As for transferability, it is linked to the acquisition of ad hoc licenses. The development of architecture using open source software bricks from foundations such as Apache Foundation makes it possible not to be limited by licenses but is dependent on developments and maintenance carried out by the community of developers. These software bricks can be abandoned by the community, the company that sponsors them, or the foundation that hosts them. The development of a sustainable architecture would go through an emancipation of software bricks which would make it possible to easily change them on the one hand when one of them disappears or if a new more efficient software brick appears.

The deployment of 5G and satellite Internet will bring in a new player, which are the telecommunications companies that will be able to provide processing capacities and services as close as possible to users at the level of 5G antennas, which will impact processing architectures. The problem will then arise of interoperability between the networks of sensors and actuators with these new high-speed, low-latency networks. The new networks offered by the telecommunications companies will make it possible to offer new services or even to decouple the software from the hardware, which will make it possible to make the sensors and actuators interchangeable. This should make it possible to reduce the cost of the equipment and make it accessible to developing countries or areas not covered by traditional LPWAN and 3GPP networks. The impact of these new networks will have to be reviewed in the future to identify the new trends offered by 5G and satellite Internet.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have interfered with overall quality of the work reported in this paper.

Acknowledgment

This research is partially funded by Infortech and Numediart institutes. The authors would like to express their gratitude to Dr. Meryem Elmoulat for accepting to edit the writing of this paper and to Mr. Fabrice Nolack Fote for his help in the elaboration of the conceptual framework.

References

Agency, E.G., 2020. Power-efficient positioning for THE Internet of Things – White Paper. European GNSS Agency. https://doi.org/10.2759/26127.
Ai, Y., Peng, M., Zhang, K., 2018. Edge computing technologies for internet of things: a primer. Digital Commun. Networks 4, 77–86. https://doi.org/10.1016/j.dcan.2017.07.001.
Alonso, R.S., Sittón-Candanedo, I., García, Óscar, Prieto, J., Rodríguez-González, S., 2020. An intelligent edge-iot platform for monitoring livestock and crops in a dairy farming scenario. Ad Hoc Netw. 98,. https://doi.org/10.1016/j.adhoc.2019.102047 102047.
Amazon, 2021a. Amazon dynamodb. url:https://aws.amazon.com/fr/dynamodb/.
Amazon, 2021b. Amazon web services. url: https://aws.amazon.com/.
Ampatzidis, Y., Partel, V., Costa, L., 2020. Agroview: Cloud-based application to process, analyze and visualize uav-collected data for precision agriculture applications utilizing artificial intelligence. Comput. Electron. Agricul. 174,. https://doi.org/10.1016/j.compag.2020.105457 105457.
Andriamandroso, A.L.H., Lebeau, F., Beckers, Y., Froidmont, E., Dufrasne, I., Heinesch, B., Dumortier, P., Blanchy, G., Blaise, Y., Bindelle, J., 2017. Development of an open-source algorithm based on inertial measurement units (imu) of a smartphone to detect cattle grass intake and ruminating behaviors. Comput. Electron. Agricult. 139, 126–137. https://doi.org/10.1016/j.compag.2017.05.020.
Apache Software Foundation, A., 2021a. Cassandra. url: https://cassandra.apache.org.
Apache Software Foundation, A., 2021b. Druid. url: https://druid.apache.org.
Assis, M.R., Bittencourt, L.F., 2016. A survey on cloud federation architectures: identifying functional and non-functional properties. J. Network Comput. Appl. 72, 51–71. https://doi.org/10.1016/j.jnca.2016.06.014.
AT&T, P., 2021. At&t continues to fuel growth of the internet of things with launch of new developer-friendly managed service. url: https://about.att.com/story/m2x_data_service_for_enterprise_developers.html.
Ayaz, M., Ammad-Uddin, M., Sharif, Z., Mansour, A., Aggoune, E.H.M., 2019. Internet-of-things (iot)-based smart agriculture: toward making the fields talk. IEEE Access 7, 129551–129583. https://doi.org/10.1109/ACCESS.2019.2932609.
Badidi, E., 2020. Qos-aware placement of tasks on a fog cluster in an edge computing environment. J. Ubiquitous Syst. Pervasive Networks 13, 11–19. https://doi.org/10.5383/JUSPN.13.01.002.
Bermeo-Almeida, O., Cardenas-Rodriguez, M., Samaniego-Cobo, T., Ferruzola-Gómez, E., Cabezas-Cabezas, R., Bazán-Vera, W., 2018. Blockchain in agriculture: A systematic literature review, in: International Conference on Technologies and Innovation, Springer. pp. 44–56. doi:10.1007/978-3-030-00940-3_4.
Bixio, L., Delzanno, G., Rebora, S., Rulli, M., 2020. A flexible iot stream processing architecture based on microservices. Information 11, 565. https://doi.org/10.3390/info11120565.

Blynk, 2021. Blynk iot platform: for businesses and developers. url: https://blynk.io.
Botta, A., De Donato, W., Persico, V., Pescapé, A., 2016. Integration of cloud computing and internet of things: a survey. Future Generat. Comput. Syst. 56, 684–700. https://doi.org/10.1016/j.future.2015.09.021.
Byers, C.C., 2017. Architectural imperatives for fog computing: Use cases, requirements, and architectural techniques for fog-enabled iot networks. IEEE Commun. Mag. 55, 14–20. https://doi.org/10.1109/MCOM.2017.1600885.
Carnevale, L., Celesti, A., Galletta, A., Dustdar, S., Villari, M., 2019. Osmotic computing as a distributed multi-agent system: the body area network scenario. Internet of Things 5, 130–139. https://doi.org/10.1016/j.iot.2019.01.001.
Cisco, C., 2018. Cisco global cloud index: Forecast and methodology, 2016–2021. Cisco, San Jose.
Codeluppi, G., Cilfone, A., Davoli, L., Ferrari, G., 2020. Lorafarm: A lorawan-based smart farming modular iot architecture. Sensors 20. https://doi.org/10.3390/s20072028. url: https://www.mdpi.com/1424-8220/20/7/2028.
Conesa-Muñoz, J., Valente, J., Del Cerro, J., Barrientos, A., Ribeiro, A., 2016. A multi-robot sense-act approach to lead to a proper acting in environmental incidents. Sensors 16, 1269. https://doi.org/10.3390/s16081269.
Corp, P.H., 2020. Sensorcloud. url: https://sensorcloud.com/.
Debauche, O., El Moulat, M., Mahmoudi, S., Boukraa, S., Manneback, P., Lebeau, F., 2018. Web monitoring of bee health for researchers and beekeepers based on the internet of things. Proc. Comput. Sci. 130, 991–998. https://doi.org/10.1016/j.procs.2018.04.103.
Debauche, O., Mahmoudi, S., Andriamandroso, A.L.H., Manneback, P., Bindelle, J., Lebeau, F., 2019. Cloud services integration for farm animals' behavior studies based on smartphones as activity sensors. J. Ambient Intell. Humanized Comput. 10, 4651–4662. https://doi.org/10.1007/s12652-018-0845-9.
Debauche, O., Mahmoudi, S., Elmoulat, M., Mahmoudi, S.A., Manneback, P., Lebeau, F., 2020. Edge ai-iot pivot irrigation, plant diseases, and pests identification. Proc. Comput. Sci. 177, 40–48. https://doi.org/10.1016/j.procs.2020.10.009.
Debauche, O., Mahmoudi, S., Mahmoudi, S.A., Manneback, P., Bindelle, J., Lebeau, F., 2020. Edge computing and artificial intelligence for real-time poultry monitoring. Proc. Comput. Sci. 175, 534–541. https://doi.org/10.1016/j.procs.2020.07.076.
Debauche, O., Mahmoudi, S., Mahmoudi, S.A., Manneback, P., Bindelle, J., Lebeau, F., 2020. Edge computing for cattle behavior analysis. In: 2020 Second International Conference on Embedded & Distributed Systems (EDiS), IEEE, pp. 52–57. https://doi.org/10.1109/EDiS49545.2020.9296471.
Debauche, O., Mahmoudi, S., Mahmoudi, S.A., Manneback, P., Lebeau, F., 2020. A new edge architecture for ai-iot services deployment. Proc. Comput. Sci. 175, 10–19. https://doi.org/10.1016/j.procs.2020.07.006.
Debauche, O., Mahmoudi, S.A., De Cock, N., Mahmoudi, S., Manneback, P., Lebeau, F., 2020. Cloud architecture for plant phenotyping research. Concurrency and Computation: Practice and Experience 32,. https://doi.org/10.1002/cpe.5661 e5661.
Debauche, O., Mahmoudi, S.A., Mahmoudi, S., Manneback, P., 2018. Cloud platform using big data and hpc technologies for distributed and parallels treatments. Proc. Comput. Sci. 141, 112–118. https://doi.org/10.1016/j.procs.2018.10.156.
Debauche, O., Trani, J.P., Mahmoudi, S., Manneback, P., Bindelle, J., Mahmoudi, S.A., Guttadauria, A., Lebeau, F., 2021. Data management and internet of things: a methodological review in smart farming. Internet of Things 14,. https://doi.org/10.1016/j.iot.2021.100378 100378.
Díaz, M., Martín, C., Rubio, B., 2016. State-of-the-art, challenges, and open issues in the integration of internet of things and cloud computing. J. Network Comput. Appl. 67, 99–117. https://doi.org/10.1016/j.jnca.2016.01.010.
Drakos, A., Protonotarios, V., Manouselis, N., 2015. aginfra: a research data hub for agriculture, food and the environment. F1000Res. 4. doi:10.12688/f1000research.6349.2.
El-Sayed, H., Sankar, S., Prasad, M., Puthal, D., Gupta, A., Mohanty, M., Lin, C.T., 2017. Edge of things: the big picture on the integration of edge, iot and the cloud in a distributed computing environment. IEEE Access 6, 1706–1717. https://doi.org/10.1109/ACCESS.2017.2780087.
Elijah, O., Rahman, T.A., Orikumhi, I., Leow, C.Y., Hindia, M.N., 2018. An overview of internet of things (iot) and data analytics in agriculture: benefits and challenges. IEEE Internet Things J. 5, 3758–3773. https://doi.org/10.1109/JIOT.2018.2844296.
Estrada, R., Ruiz, I., 2016. Big data smack: A guide to apache spark. Mesos, Akka, Cassandra, and Kafka.
Fan, D., Gao, S., 2018. The application of mobile edge computing in agricultural water monitoring system. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing, p. 012015.
Fang, H., 2015. Managing data lakes in big data era: What's a data lake and why has it became popular in data management ecosystem, in: 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), IEEE. pp. 820–824. doi:10.1109/CYBER.2015.7288049.
Farooq, M.S., Riaz, S., Abid, A., Abid, K., Naeem, M.A., 2019. A survey on the role of iot in agriculture for the implementation of smart farming. IEEE Access 7, 156237–156271. https://doi.org/10.1109/ACCESS.2019.2949703.
Feng, X., Yan, F., Liu, X., 2019. Study of wireless communication technologies on internet of things for precision agriculture. Wireless Pers. Commun. 108, 1785–1802. https://doi.org/10.1007/s11277-019-06496-7.
Fernandez, R.C., Pietzuch, P.R., Kreps, J., Narkhede, N., Rao, J., Koshy, J., Lin, D., Riccomini, C., Wang, G., 2015. Liquid: Unifying nearline and offline big data integration., in: CIDR, pp. 1–8. url: http://hdl.handle.net/10044/1/23433.
Ferrag, M.A., Shu, L., Yang, X., Derhab, A., Maglaras, L., 2020. Security and privacy for green iot-based agriculture: review, blockchain solutions, and challenges. IEEE Access 8, 32031–32053. https://doi.org/10.1109/ACCESS.2020.2973178.
Fote, F.N., Roukh, A., Mahmoudi, S., Mahmoudi, S.A., Debauche, O., 2020. Toward a big data knowledge-base management system for precision livestock farming. Proc. Comput. Sci. 177, 136–142. https://doi.org/10.1016/j.procs.2020.10.021.
Friha, O., Ferrag, M.A., Shu, L., Maglaras, L.A., Wang, X., 2021. Internet of things for the future of smart agriculture: a comprehensive survey of emerging technologies. IEEE CAA J. Autom. Sinica 8, 718–752. https://doi.org/10.1109/JAS.2021.1003925.
Gallinucci, E., Golfarelli, M., Rizzi, S., 2019. A hybrid architecture for tactical and strategic precision agriculture. In: Ordonez, C., Song, I.Y., Anderst-Kotsis, G., Tjoa, A.M., Khalil, I. (Eds.), Big Data Analytics and Knowledge Discovery. Springer International Publishing, Cham, pp. 13–23. https://doi.org/10.1007/978-3-030-27520-4_2.
Gallinucci, E., Golfarelli, M., Rizzi, S., 2020. Mo. re. farming: A hybrid architecture for tactical and strategic precision agriculture. Data Knowl. Eng. 129,. https://doi.org/10.1016/j.datak.2020.101836 101836.
García, L., Parra, L., Jimenez, J.M., Lloret, J., Lorenz, P., 2020. Iot-based smart irrigation systems: an overview on the recent trends on sensors and iot systems for irrigation in precision agriculture. Sensors 20, 1042. https://doi.org/10.3390/s20041042.
Garcia Lopez, P., Montresor, A., Epema, D., Datta, A., Higashino, T., Iamnitchi, A., Barcellos, M., Felber, P., Riviere, E., 2015. Edge-centric computing: Vision and challenges. doi:10.1145/2831347.2831354.
Giebler, C., Stach, C., Schwarz, H., Mitschang, B., 2018. Braid, in: Proceedings of the 7th International Conference on Data Science, Technology and Applications, pp. 294–301. doi:10.5220/0006861802940301.
Google, 2021. Firebase. url: https://firebase.google.com/.
Granell, C., Miralles, I., Rodríguez-Pupo, L.E., González-Pérez, A., Casteleyn, S., Busetto, L., Pepe, M., Boschetti, M., Huerta, J., 2017. Conceptual architecture and service-oriented implementation of a regional geoportal for rice monitoring. ISPRS Int. J. Geo-Inform. 6. url: https://www.mdpi.com/2220-9964/6/7/191, doi:10.3390/ijgi6070191.
Grover, J., Garimella, R.M., 2018. Reliable and fault-tolerant iot-edge architecture, in: 2018 IEEE sensors, IEEE. pp. 1–4. doi:10.1109/ICSENS.2018.8589624.
Guardo, E., Di Stefano, A., La Corte, A., Sapienza, M., Scatà, M., 2018. A fog computing-based iot framework for precision agriculture. J. Internet Technol. 19, 1401–1411. https://doi.org/10.3966/160792642018091905012.
Gupta, M., Abdelsalam, M., Khorsandroo, S., Mittal, S., 2020. Security and privacy in smart farming: challenges and opportunities. IEEE Access 8, 34564–34584. https://doi.org/10.1109/ACCESS.2020.2975142.
Hausenblas, M., 2014. Internet of things architecture (iot-a) home page. url: https://github.com/mhausenblas/iot-a.info.
Iaksch, J., Fernandes, E., Borsato, M., 2021. Digitalization and big data in smart farming–a review. J. Manage. Anal. 8, 333–349. https://doi.org/10.1080/23270012.2021.1897957.
IBM, 2015. Ibm watson iot platform. url: https://internetofthings.ibmcloud.com/.
IBM, 2021. Ibm cloud. url: https://www.ibm.com/cloud.
Influxdata, 2021. Infludb cloud. url: https://www.influxdata.com/products/influxdb-cloud/.
Integra Souces, I., 2021. Iot solution development services. url: https://www.integrasources.com/iot-page/.
Jayaraman, P.P., Yavari, A., Georgakopoulos, D., Morshed, A., Zaslavsky, A., 2016. Internet of things platform for smart farming: experiences and lessons learnt. Sensors 16, 1884. https://doi.org/10.3390/s16111884.
KaaIoT, 2021. Ubidots. url: https://docs.kaaiot.io/KAA/docs/current/Welcome/.
Kamilaris, A., Fonts, A., Prenafeta-Boldύ, F.X., 2019. The rise of blockchain technology in agriculture and food supply chains. Trends Food Sci. Technol. 91, 640–652. doi:10.1016/j.tifs.2019.07.034.
Kaur, A., Kumar, R., Saxena, S., 2020. Osmotic computing and related challenges: a survey. In: 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), IEEE, pp. 378–383. https://doi.org/10.1109/PDGC50313.2020.9315757.
Kazim, M., Liu, L., Zhu, S.Y., 2018. A framework for orchestrating secure and dynamic access of iot services in multi-cloud environments. IEEE Access 6, 58619–58633. https://doi.org/10.1109/ACCESS.2018.2873812.
Khanna, A., Kaur, S., 2019. Evolution of internet of things (iot) and its significant impact in the field of precision agriculture. Comput. Electron. Agricul. 157, 218–231. https://doi.org/10.1016/j.compag.2018.12.039.
Kodati, S., Jeeva, S., 2019. Smart agricultural using internet of things, cloud and big data. Int. J. Innov. Technol. Exploring Eng. (IJITEE) 8, 3718–3722. https://doi.org/10.35940/ijitee.J9671.0881019.
Kreps, J., 2014. Questioning the lambda architecture. Online article, July 205.
Lakhe, B., 2016. Practical Hadoop migration: how to integrate your RDBMS with the Hadoop ecosystem and re-architect relational applications to NoSQL. Apress.
Liu, Y., Ma, X., Shu, L., Hancke, G.P., Abu-Mahfouz, A.M., 2020. From industry 4.0 to agriculture 4.0: current status, enabling technologies, and research challenges. IEEE Trans. Industr. Inf. 17, 4322–4334. https://doi.org/10.1109/TII.2020.3003910.
López, I.D., Grass, J.F., Figueroa, A., Corrales, J.C., 2020. A proposal for a multi-domain data fusion strategy in a climate-smart agriculture context. Int. Trans. Oper. Res. https://doi.org/10.1111/itor.12899.
Luis Bustamante, A., Patricio, M.A., Molina, J.M., 2019. Thinger. io: an open source platform for deploying data fusion applications in iot environments. Sensors 19, 1044. https://doi.org/10.3390/s19051044.

Mach, P., Becvar, Z., 2017. Mobile edge computing: a survey on architecture and computation offloading. IEEE Commun. Surveys Tutorials 19, 1628–1656. https://doi.org/10.1109/COMST.2017.2682318.
Madera, C., Laurent, A., Rouge, T.L., Miralles, A., 2017. How can the data lake concept influence information system design for agriculture? In: 11th European conference dedicated to the future use of ICT in the agri-food sector, bioresource and biomass sector (EFITA 2017), pp. 181–182.
Maksimović, M., 2018. The role of osmotic computing in internet of things. In: 2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH), IEEE, pp. 1–4. https://doi.org/10.1109/INFOTEH.2018.8345538.
Manyika, J., Chui, M., 2015. By 2025, internet of things applications could have $11 trillion impact. Insight Publications.
Martos, V., Ahmad, A., Cartujo, P., Ordoñez, J., 2021. Ensuring agricultural sustainability through remote sensing in the era of agriculture 5.0. Appl. Sci. 11, 5911. https://doi.org/10.3390/app11135911.
Marz, N., Warren, J., 2013. Big Data: Principles and best practices of scalable real-time data systems. Manning.
Maureira, M.A.G., Oldenhof, D., Teernstra, L., 2011. Thingspeak–an api and web service for the internet of things. World Wide Web.
Meehan, J., Zdonik, S., Tian, S., Tian, Y., Tatbul, N., Dziedzic, A., Elmore, A., 2016. Integrating real-time and batch processing in a polystore. In: 2016 IEEE High Performance Extreme Computing Conference (HPEC), IEEE, pp. 1–7. https://doi.org/10.1109/HPEC.2016.7761585.
Meola, A., 2021. Smart farming in 2020: How iot sensors are creating a more efficient precision agriculture industry. url: https://www.businessinsider.com/smart-farming-iot-agriculture?IR=T.
Microsoft, 2021a. Azure. url: https://azure.microsoft.com.
Microsoft, 2021b. Azure iot. url: https://azure.microsoft.com/en-us/overview/iot/.
Miloslavskaya, N., Tolstoy, A., 2016. Big data, fast data and data lake concepts. Proc. Comput. Sci. 88, 300–305. https://doi.org/10.1016/j.procs.2016.07.439.
Misra, N., Dixit, Y., Al-Mallahi, A., Bhullar, M.S., Upadhyay, R., Martynenko, A., 2020. Iot, big data and artificial intelligence in agriculture and food industry. IEEE Internet Things J. https://doi.org/10.1109/JIOT.2020.2998584.
Mongo, 2021. Mongodb atlas. url: https://www.mongodb.com/en-us/cloud/atlas.
Morshed, A., Jayaraman, P.P., Sellis, T., Georgakopoulos, D., Villari, M., Ranjan, R., 2017. Deep osmosis: holistic distributed deep learning in osmotic computing. IEEE Cloud Comput. 4, 22–32. https://doi.org/10.1109/MCC.2018.1081070.
Munir, A., Kansakar, P., Khan, S.U., 2017. Ifciot: Integrated fog cloud iot: a novel architectural paradigm for the future internet of things. IEEE Consumer Electron. Mag. 6, 74–82. https://doi.org/10.1109/MCE.2017.2684981.
Mykleby, M., Doherty, P., Makower, J., 2016. The New Grand Strategy: Restoring America's Prosperity, Security, and Sustainability in the 21st Century. St. Martin's Press.
Navarro, E., Costa, N., Pereira, A., 2020. A systematic review of iot solutions for smart farming. Sensors 20, 4231. https://doi.org/10.3390/s20154231.
Navarro-Hellín, H., Martinez-del Rincon, J., Domingo-Miguel, R., Soto-Valles, F., Torres-Sánchez, R., 2016. A decision support system for managing irrigation in agriculture. Comput. Electron. Agricul. 124, 121–131. https://doi.org/10.1016/j.compag.2016.04.003.
NECTEC, 2020. Netpie - network platform for internet of everything. url: https://netpie.io.
Neves, R.A., Cruvinel, P.E., 2020. Model for semantic base structuring of digital data to support agricultural management. In: 2020 IEEE 14th International Conference on Semantic Computing (ICSC), IEEE, pp. 337–340. https://doi.org/10.1109/ICSC.2020.00067.
Nkamla Penka, J.B., Mahmoudi, S., Debauche, O., 2021. A new kappa architecture for iot data management in smart farming. Proc. Comput. Sci.. In press.
Oracle, 2021. Mysql. url: https://www.mysql.com.
Persico, V., Pescapé, A., Picariello, A., Sperlí, G., 2018. Benchmarking big data architectures for social networks data processing using public cloud platforms. Future Gener. Comput. Syst. 89, 98–109. https://doi.org/10.1016/j.future.2018.05.068.
Pesonen, L.A., Teye, F.K.W., Ronkainen, A.K., Koistinen, M.O., Kaivosoja, J.J., Suomi, P. F., Linkolehto, R.O., 2014. Cropinfra–an internet-based service infrastructure to support crop production in future farms. Biosyst. Eng. 120, 92–101. https://doi.org/10.1016/j.biosystemseng.2013.09.005.
Petcu, D., Di Martino, B., Venticinque, S., Rak, M., Máhr, T., Lopez, G.E., Brito, F., Cossu, R., Stopar, M., Šperka, S., et al., 2013. Experiences in building a mosaic of clouds. J. Cloud Comput.: Adv., Syst. Appl. 2, 1–22. https://doi.org/10.1186/2192-113X-2-12.
Popescu, D., Stoican, F., Stamatescu, G., Ichim, L., Dragana, C., 2020. Advanced uav–wsn system for intelligent monitoring in precision agriculture. Sensors 20, 817. https://doi.org/10.3390/s20030817.
Radoglou-Grammatikis, P., Sarigiannidis, P., Lagkas, T., Moscholios, I., 2020. A compilation of uav applications for precision agriculture. Comput. Netw. 172,. https://doi.org/10.1016/j.comnet.2020.107148 107148.
Rajakaruna, A., Manzoor, A., Porambage, P., Liyanage, M., Ylianttila, M., Gurtov, A., 2018. Lightweight dew computing paradigm to manage heterogeneous wireless sensor networks with uavs. arXiv preprint arXiv:1811.04283.
Ray, P.P., 2017. Internet of things for smart agriculture: technologies, practices and future direction. J. Ambient Intell. Smart Environ. 9, 395–420. https://doi.org/10.3233/AIS-170440.
Ren, J., Guo, H., Xu, C., Zhang, Y., 2017. Serving at the edge: a scalable iot architecture based on transparent computing. IEEE Network 31, 96–105. https://doi.org/10.1109/MNET.2017.1700030.
Rodriguez, M.A., Cuenca, L., Ortiz, A., 2018. Fiware open source standard platform in smart farming – a review. In: Camarinha-Matos, L.M., Afsarmanesh, H., Rezgui, Y. (Eds.), Collaborative Networks of Cognitive Systems. Springer International Publishing, Cham, pp. 581–589. https://doi.org/10.1007/978-3-319-99127-6_50.
Roman, R., Lopez, J., Mambo, M., 2018. Mobile edge computing, fog et al.: a survey and analysis of security threats and challenges. Future Gener. Comput. Syst. 78, 680–698. https://doi.org/10.1016/j.future.2016.11.009.
Roukh, A., Fote, F.N., Mahmoudi, S.A., Mahmoudi, S., 2020. Big data processing architecture for smart farming. Proc. Comput. Sci. 177, 78–85. https://doi.org/10.1016/j.procs.2020.10.014.
Roukh, A., Fote, F.N., Mahmoudi, S.A., Mahmoudi, S., 2020. Wallesmart: cloud platform for smart farming, in. In: 32nd International Conference on Scientific and Statistical Database Management, pp. 1–4. https://doi.org/10.1145/3400903.3401690.
Ruan, J., Jiang, H., Zhu, C., Hu, X., Shi, Y., Liu, T., Rao, W., Chan, F.T.S., 2019. Agriculture iot: Emerging trends, cooperation networks, and outlook. IEEE Wirel. Commun. 26, 56–63. https://doi.org/10.1109/MWC.001.1900096.
Saiz-Rubio, V., Rovira-Más, F., 2020. From smart farming towards agriculture 5.0: a review on crop data management. Agronomy 10, 207. https://doi.org/10.3390/agronomy10020207.
Sallah, A.H.M., Tychon, B., Piccard, I., Gobin, A., Van Hoolst, R., Djaby, B., Wellens, J., 2019. Batch-processing of aquacrop plug-in for rainfed maize using satellite derived fractional vegetation cover data. Agric. Water Manage. 217, 346–355.
Scott, J., 2015. Zeta architecture: Hexagon is the new circle. an enterprise architecture solution for scale and efficiency. url: https://www.oreilly.com/ideas/zeta-architecture-hexagon-is-the-new-circle.
Sethi, P., Sarangi, S.R., 2017. Internet of things: architectures, protocols, and applications. J. Electr. Comput. Eng. 2017. https://doi.org/10.1155/2017/9324035.
Shafi, U., Mumtaz, R., García-Nieto, J., Hassan, S.A., Zaidi, S.A.R., Iqbal, N., 2019. Precision agriculture techniques and practices: From considerations to applications. Sensors 19, 3796. https://doi.org/10.3390/s19173796.
Sharofidinov, F., Muthanna, M.S.A., Pham, V.D., Khakimov, A., Muthanna, A., Samouylov, K., 2020. Agriculture management based on lora edge computing system. In: Vishnevskiy, V.M., Samouylov, K.E., Kozyrev, D.V. (Eds.), Distributed Computer and Communication Networks. Springer International Publishing, Cham, pp. 113–125.
Shi, W., Cao, J., Zhang, Q., Li, Y., Xu, L., 2016. Edge computing: vision and challenges. IEEE Internet Things J. 3, 637–646. https://doi.org/10.1109/JIOT.2016.2579198.
Shi, X., An, X., Zhao, Q., Liu, H., Xia, L., Sun, X., Guo, Y., 2019. State-of-the-art internet of things in protected agriculture. Sensors 19, 1833. https://doi.org/10.3390/s19081833.
Siciliani, T., 2015. Big data using lambda architecture. url: https://dzone.com/articles/lambda-architecture-big-data.
Sigrimis, N., Arvanitis, K., Ferentinos, K., 2002. Macqu: An open scada system for intelligent management and control of greenhouses, in: 2002 ASAE Annual Meeting, American Society of Agricultural and Biological Engineers. p. 1. doi:10.13031/2013.9618.
Singh, K.N., Behera, R.K., Mantri, J.K., 2019. Big data ecosystem: review on architectural evolution. Emerging Technol. Data Mining Inform. Secur. 335–345. https://doi.org/10.1007/978-981-13-1498-8_30.
Sipos, G., Turilli, M., Newhouse, S., Kacsuk, P., 2013. A european federated cloud: Innovative distributed computing solutions by egi, in: EGU General Assembly Conference Abstracts, pp. EGU2013–8690.
Skala, K., Davidovic, D., Afgan, E., Sovic, I., Sojat, Z., 2015. Scalable distributed computing hierarchy: cloud, fog and dew computing. Open J. Cloud Comput. (OJCC) 2, 16–24. https://doi.org/10.19210/1002.2.1.16.
SQLite, 2021. Sqlite. url: https://www.sqlite.org.
Sun, L., Li, Y., Memon, R.A., 2017. An open iot framework based on microservices architecture. China Commun. 14, 154–162. https://doi.org/10.1109/CC.2017.7868163.
Talavera, J.M., Tobón, L.E., Gómez, J.A., Culman, M.A., Aranda, J.M., Parra, D.T., Quiroz, L.A., Hoyos, A., Garreta, L.E., 2017. Review of iot applications in agro-industrial and environmental fields. Comput. Electron. Agricul. 142, 283–297. https://doi.org/10.1016/j.compag.2017.09.015.
Taneja, M., Byabazaire, J., Jalodia, N., Davy, A., Olariu, C., Malone, P., 2020. Machine learning based fog computing assisted data-driven approach for early lameness detection in dairy cattle. Comput. Electron. Agricul. 171,. https://doi.org/10.1016/j.compag.2020.105286 105286.
Taneja, M., Jalodia, N., Byabazaire, J., Davy, A., Olariu, C., 2019. Smartherd management: a microservices-based fog computing–assisted iot platform towards data-driven smart dairy farming. Software: Practice Experience 49, 1055–1078. https://doi.org/10.1002/spe.2704.
Tang, Y., Dananjayan, S., Hou, C., Guo, Q., Luo, S., He, Y., 2021. A survey on the 5g network and its impact on agriculture: challenges and opportunities. Comput. Electron. Agricul. 180,. https://doi.org/10.1016/j.compag.2020.105895 105895.
The PostgreSQL Global Development Group, P., 2021. Postgresql: The world's most advanced open source relational database. url: https://www.postgresql.org/.
Tran, T.X., Hajisami, A., Pandey, P., Pompili, D., 2017. Collaborative mobile edge computing in 5g networks: new paradigms, scenarios, and challenges. IEEE Commun. Mag. 55, 54–61. https://doi.org/10.1109/MCOM.2017.1600863.
Triantafyllou, A., Sarigiannidis, P., Bibi, S., 2019. Precision agriculture: a remote sensing monitoring system architecture. Information 10, 348. https://doi.org/10.3390/info10110348.

Tzounis, A., Katsoulas, N., Bartzanas, T., Kittas, C., 2017. Internet of things in agriculture, recent advances and future challenges. Biosyst. Eng. 164, 31–48. https://doi.org/10.1016/j.biosystemseng.2017.09.007.

Ubidots, 2021. Ubidots. url: https://ubidots.com/.

Uehara, M., 2017. Mist computing: Linking cloudlet to fogs, in: International Conference on Computational Science/Intelligence & Applied Informatics, Springer. pp. 201–213. doi:10.1007/978-3-319-63618-4_15.

Valecce, G., Strazzella, S., Grieco, L.A., 2019. On the interplay between 5g, mobile edge computing and robotics in smart agriculture scenarios. In: Palattella, M.R., Scanzio, S., Coleri Ergen, S. (Eds.), Ad-Hoc, Mobile, and Wireless Networks. Springer International Publishing, Cham, pp. 549–559. https://doi.org/10.1007/978-3-030-31831-4_38.

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

Villari, M., Celesti, A., Fazio, M., 2017. Towards osmotic computing: Looking at basic principles and technologies, in: Conference on Complex, Intelligent, and Software Intensive Systems, Springer. pp. 906–915. doi:10.1007/978-3-319-61566-086.

Villari, M., Celesti, A., Fazio, M., Puliafito, A., 2014. Alljoyn lambda: an architecture for the management of smart environments in iot. In: 2014 International Conference on Smart Computing Workshops, IEEE, pp. 9–14. https://doi.org/10.1109/SMARTCOMP-W.2014.7046676.

Villari, M., Fazio, M., Dustdar, S., Rana, O., Ranjan, R., 2016. Osmotic computing: a new paradigm for edge/cloud integration. IEEE Cloud Comput. 3, 76–83. https://doi.org/10.1109/MCC.2016.124.

Wang, X., Han, Y., Leung, V.C., Niyato, D., Yan, X., Chen, X., 2020. Edge AI: Convergence of Edge Computing and Artificial Intelligence. Springer Nature. https://doi.org/10.1007/978-981-15-6186-3.

Wang, Y., 2016. Definition and categorization of dew computing. Open J. Cloud Comput. (OJCC) 3, 1–7. https://doi.org/10.19210/1002.3.1.1.

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

Yang, S., 2017. Iot stream processing and analytics in the fog. IEEE Commun. Mag. 55, 21–27. https://doi.org/10.1109/MCOM.2017.1600840.

Zhai, Z., Martínez, J.F., Beltran, V., Martínez, N.L., 2020. Decision support systems for agriculture 4.0: Survey and challenges. Comput. Electron. Agricul. 170,. https://doi.org/10.1016/j.compag.2020.105256 105256.

Zhang, Q., 2016. Precision agriculture technology for crop farming. Taylor & Francis. doi:10.1201/b19336.

Zhou, Y., Zhang, D., Xiong, N., 2017. Post-cloud computing paradigms: a survey and comparison. Tsinghua Sci. Technol. 22, 714–732. https://doi.org/10.23919/TST.2017.8195353.