Revista Ciência Agronômica, v. 51, Special Agriculture 4.0, e20207755, 2020
Centro de Ciências Agrárias - Universidade Federal do Ceará, Fortaleza, CE
www.ccarevista.ufc.br

Artigo Científico
ISSN 1806-6690

Survey on connectivity and cloud computing technologies: State-of-the-art applied to Agriculture 4.0

Pesquisa sobre tecnologias de conectividade e computação em nuvem: Estado da arte aplicado à Agricultura 4.0

Rafael Simionato[1], José Rodrigues Torres Neto[1], Carla Julciane dos Santos[1], Bruno Silva Ribeiro[1], Fernando Cesar Britto de Araújo[1], Antonio Robson de Paula[1], Pedro Augusto de Lima Oliveira[1], Paulo Silas Fernandes[1] and Jin Hong Yi[1]*

ABSTRACT - In recent years, agriculture has faced many challenges, from a growing global population to be fed, the work power evasion in the sector, to sustainability requirements and environmental constraints. To satisfy the increasingly demanding stakeholders, the agricultural sector has looked for new ways to tackle these issues. In this context, Information and Communications Technologies (ICTs) have been applied to help the agricultural sector overcome these challenges. This article investigates how two ICTs - connectivity and cloud computing - can leverage and traverse other ICTs, such as Internet of Things and artificial intelligence, enabling the entire productive sector to be supported by decision-making systems, which in turn are based on data-driven models. Moreover, a successful case study on how cloud computing has helped one of SiDi's biggest customers - a global company - improve its operational performance by obtaining insights from its data is presented.

Key words: Telecommunication. Serverless. Data lake. Data analysis. Internet of Things.

RESUMO - Nos últimos anos, a agricultura tem enfrentado muitos desafios, desde uma crescente população global a ser alimentada, a evasão da força de trabalho no setor, até requisitos de sustentabilidade e restrições ambientais. Para satisfazer as partes interessadas cada vez mais exigentes, o setor agrícola tem procurado novas formas de lidar com essas questões. Nesse contexto, as Tecnologias da Informação e Comunicação (TICs) têm sido aplicadas para ajudar o setor agrícola a superar esses desafios. Este artigo investiga como duas TICs - conectividade e computação em nuvem - podem alavancar e explorar outras TICs, como Internet das Coisas e inteligência artificial, permitindo que todo o setor produtivo seja apoiado por sistemas de tomada de decisão, que por sua vez são baseados em modelos orientados por dados. Além disso, é apresentado um estudo de caso de sucesso sobre como a computação em nuvem ajudou um dos maiores clientes da SiDi - uma empresa global - a melhorar seu desempenho operacional ao obter insights a partir de seus dados.

Palavras-chave: Telecomunicação. Sem servidor. Lago de dados. Análise de dados. Internet das Coisas.

DOI: 10.5935/1806-6690.20200085
Editores do artigo: Professor Daniel Albiero - daniel.albiero@gmail.com e Professor Alek Sandro Dutra - alekdutra@ufc.br
*Author for correspondence
Received for publication 06/01/2020; approved on 20/10/2020
[1]Departamento de Pesquisa e Desenvolvimento, SiDi, Campinas-SP, Brasil, r.simionato@sidi.org.br (ORCID ID 0000-0001-9392-6598), jr.neto@sidi.org.br (ORCID ID 0000-0002-1834-9978), carla.j@sidi.org.br (ORCID ID 0000-0002-8763-8853), bruno.ribeiro@sidi.org.br (ORCID ID 0000-0002-4846-8409), fernando.cesar@partner.sidi.org.br (ORCID ID 0000-0002-1902-6919), robson.p@sidi.org.br (ORCID ID 0000-0002-4243-8196), p.lima@sidi.org.br (ORCID ID 0000-0002-3447-7574), paulo.silas@sidi.org.br (ORCID ID 0000-0001-9338-5226) and j.yi@sidi.org.br (ORCID ID 0000-0003-4459-1036)

INTRODUCTION

In the past years, the agricultural sector has faced great challenges: 1) how to feed an ever-growing population (9.1 billion people by 2050) (FOOD AND AGRICULTURE ORGANIZATION, 2009); 2) how to deal with rural populations abandoning agricultural land, seeking better economic prospects in urban areas; and 3) how to render agricultural activities sustainable with a low impact on the environment.

Aside from these challenges, agriculture has to deal with numerous uncertainties classified into four categories (ESTESO; ALEMANY; ORTIZ, 2018; LEZOCHE *et al.,* 2020): product (shelf-life, deterioration rate, lack of homogeneity, food quality, and food safety); process (harvesting yield, supply lead time, resource needs, and production); market (demand and market prices); and environment (weather, pests and diseases, and regulations).

In this context, the concept of Industry 4.0 has evolved to address issues from the realm of agriculture, which in turn has given rise to Agriculture 4.0, the incorporation of Information and Communications Technologies (ICTs) into agriculture to mitigate these risks.

The main ICTs used to digitalize Agriculture 4.0 can be divided into four major classes:

- Sensors present in equipment (combines and tractors); sensors in the livestock; sensors in the field to collect data on soil nutrients, moisture, and health and also on plant health and yield;

- Data collection by these sensors to obtain a real-time picture of the farm/livestock and the environment, enabling the creation of a historical database;

- Intelligence generation and aggregation from this large volume of data in a way that allows farmers to improve their decision-making system by obtaining insights from this data;

- Automation of robotics with real-time information based on the insights provided by the collected data.

The digitalization of the agricultural supply chain structures enables farmers to monitor and visualize potential risks in real-time, which allows for a more flexible and timely response, thus improving the robustness and resilience of the farm. However, the process of digitalization can be expensive for small farmers. Therefore, the Farmer Producer Organizations (FPO) provide service to their members, allowing small farmers to have access to cutting-edge technologies. One example of FPO service is Cocamar (COCAMAR, 2020), which worked in cooperation with IBM in a hackathon (IBM, 2020) to develop a tool for visual recognition of pests in the soybean plantation (MARATONA BEHIND THE CODE, 2020). This project was also developed alongside the IT developers' community from Latin America.

The communication between various endpoints within the agricultural structure is an important aspect of digitalization. Currently, there are numerous technologies that can help famers achieve connectivity between their properties and the cloud-based systems. This article investigates some of these connectivity technologies and provides an in-depth examination of the cloud systems enabled by them, thus providing creativity and expansion of Agriculture 4.0 projects.

In this context, this research aimed to convey how these ICTs can and should benefit from connectivity and cloud computing to achieve better results for agricultural challenges while reducing the cost for the farmer/producer by answering the following research questions:

- How have communication technologies been employed for the connectivity of monitoring devices and agricultural machinery in Agriculture 4.0?

- How can cloud computing enhance the efficiency of the production processes in Agriculture 4.0?

CONNECTIVITY AND NETWORK IN AGRICULTURE 4.0

The advances in connectivity and networks to connect devices are a crucial factor for Agriculture 4.0. As we move into the digital age, the growing evidence of this is observed from connections, data feedback, and agricultural information systems, which characterize the process of digitalization in farms (FIELKE *et al.*, 2020).

In farming, digitalization promotes the use of technologies that facilitate the agribusiness in operational and strategic decision-making. These technologies help increase productivity, reduce costs, and increase efficiency (MEDEIROS, 2020), which in turn improve the production quality, reduce the environmental impact, allow better planning of agricultural inputs, and provide support to commercialization (EMBRAPA; SEBRAE; INPE, 2020). The convergence of different technologies, such as geotechnology, precision agriculture, and Internet of Things (IoT), provides scenarios for digital transformation in rural properties and establishes the smart farming concept (EMBRAPA, 2018).

To obtain the benefits of digital transformation, countryside connectivity needs to be provided, which includes two main aspects: Internet access and the coverage of rural properties for access to devices and machines. Of the more than 5 million rural properties in Brazil alone, only 29% have access to the Internet. These properties occupy 351.3 million ha (IBGE, 2017).

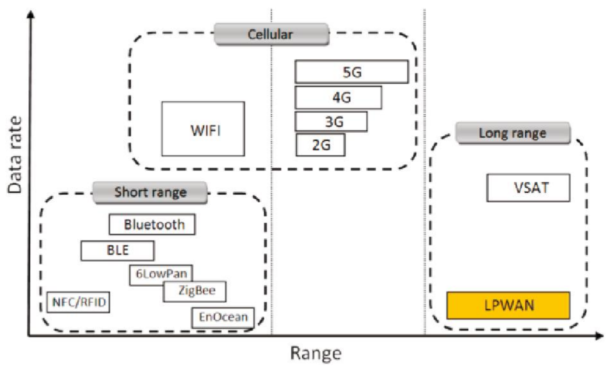
To provide Internet access to rural properties, in addition to conventional accesses, such as fiber optics and radio, which may be considered impractical due to the distances between farms, access alternatives based on wireless, terrestrial, and satellite communication technologies can be used. This section focuses on field coverage with terrestrial wireless access technologies and long-range connectivity for large stretches of land, thus enabling precision agriculture and monitoring of agricultural and livestock variables. In addition to the extensive use of robotics that generate more data for existing systems in the countryside, the use of data processing technologies, such as algorithms and artificial intelligence, facilitate in the decision-making processes.

In the following section, long-range wireless access technologies based on the evolution of IoT and cellular technology, such as the 5th Generation (5G) connectivity, are presented. As they are wireless solutions, the use of the spectrum and regulation applied to the service will also be considered.

Low Power Wide Area Networks (LPWAN)

The IoT has become one of the most promising ways to increase agricultural production, enhance precision, and help producers in managing their farms. Nowadays, there is no unified connection technology that meets everyone's needs. Each country and the producers therein have their specific characteristics of relief, land dimensions, business model, problems to be solved, and bureaucratic restrictions. These characteristics influence the model to be adopted, as there are applications where a high range is more important than a high data transmission rate. Figure 1 presents the relationship of data rate and range compared with other technologies.

Figure 1 - Required data rate vs. range capacity of radio communication (MEKKI *et al*., 2019)



LPWANs are networks with low energy consumption and low transmission rates, although their main characteristic is their high territorial reach (MEKKI *et al*., 2019). They are widely used in applications where sensors are located in remote areas, with no possibility of power recharge or electrical connection and with low quantities of data transmitted during the day. Hence, the equipment must have a lifecycle of 2 or more years, and a connectivity model that fits the specific needs on hand should be carefully chosen considering the hardware and software components to be used (ARJONA *et al*., 2018).

Numerous LPWAN technologies have emerged in the licensed and unlicensed frequency bandwidth (MEKKI *et al*., 2019). This topic mainly focuses on the leading technologies, such as SigFox and Lora. NB-IoT will be discussed in the next topic.

A - SigFox

SigFox, which is based on an Ultra Narrow Band technology, was first developed in France in 2010, with a vision to connect everything. It is now present in more than 70 countries (Figure 2).

Using radio technology to connect devices, the network operates in the unlicensed frequency range, utilizing the 902-MHz range in the United States; the 920-MHz range in South America, New Zealand, and Australia; and the 868-MHz range in Europe. It features extremely low battery consumption, a 100-bps data transfer rate, and a maximum distance of 40 km in open areas, in addition to very low noise levels (RAZA; KULKARNI; SOORIYABANDARA, 2019).

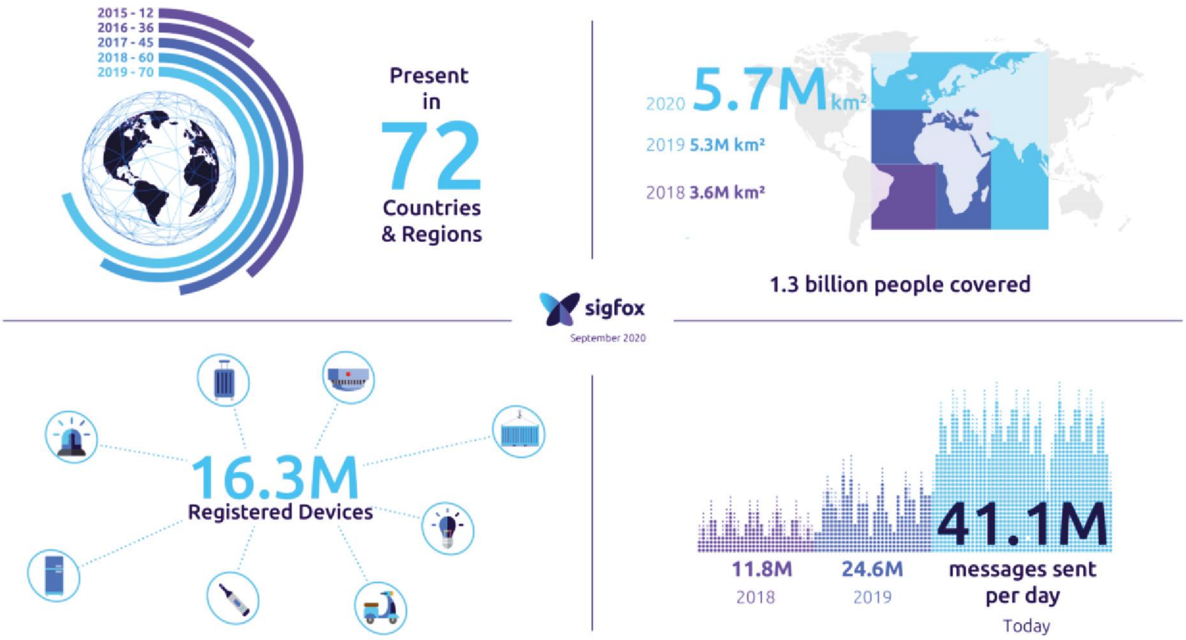
Gomez *et al.* (2019) conducted a study on this specific area and provided a more detailed description of the SigFox technology and its architecture.

B - Long-Range (LoRa)

Like SigFox, LoRa (Long-Range) was developed in France and standardized by the LoRa Alliance. It uses unlicensed Industrial, Scientific, and Medical bands and a frequency range of 868 MHz in Europe and 915 MHz in North America (MEKKI *et al*., 2019). LoRa actuates in the physical layer, whereas LoRaWAN actuates in the logic layer. LoRaWAN is an open-standard network protocol designed to wirelessly connect everything into the Internet. It has a maximum value of 50-kbps data

Figure 2 - SigFox overview[1]. SigFox is present in 72 countries and regions around the world and has more than 16.3 million devices connected

[1]https://www.sigfox.com/en/sigfox-story



transmission rate and up to 20-km distance in rural areas. By using LoRaWAN, each message transmitted by an end device is received by all the base stations in the range (MEKKI *et al*., 2019). Figure 3 presents the differences in the use of LoRaWAN in different regions.

A study conducted by Subashini; Venkateswari and Mathiyalagan (2019) provides a more detailed analysis of the architecture and specifications of LoRa and LoRaWAN globally.

Conversely, a study by the Semtech Corporation (SEMTECH CORPORATION, 2019) described the main solutions using LoRaWAN in the context of Agriculture 4.0.

SigFox and LoRaWAN share a similar market and serve projects that require long distance, low data transmission rate, low energy consumption, and high scalability. However, in choosing the best alternative and cost benefit, even the smallest details that differentiate each project can be crucial. Raza; Kulkarni and Sooriyabandara (2019) provide an overview of the LPWAN technologies.

Through the study by Klerkx; Jakku and Labarthe (2019), the comparative studies on LPWAN in a German farm and the local issues therein can be possibly understood.

4G Long-Term Evolution (LTE) and Private Networks

The evolution of cellular networks and their access technologies began in the 1st Generation (1G) in the 1980s, with telephony in an analog network, then in the 2nd Generation (2G) in the 1990s, with telephony in a digital network. The advent of broadband and smartphones in the 3rd Generation (3G) in the 2000s has led to the emergence of services in the 4th Generation (4G) in the 2010s and now, with greater speed and less latency, the 5th Generation (5G) (BERNARDOCKI *et al.*, 2020). Cellular networks have strongly standardized specifications worldwide by the 3rd Generation Partnership Project (3GPP), which is in line with the International Telecommunication Union.

The connectivity of devices began to emerge globally with 2G and 3G, but it was consolidated only with 4G, particularly for the mass market (low-cost devices, low data volume, and extreme coverage) (BERNARDOCKI *et al.*, 2020). With advanced LTE modulation techniques in 4G, the LTE Release 13 (3GPP, 2016) was rolled out, which further evolved with the LTE Release 14 (3GPP, 2017).

NB-IoT and LTE-M coexist in 4G and are required in 5G, which ensure the continuity of these technologies in

Figure 3 - Regional specification (LORA ALLIANCE, 2015). It shows the specific technical information of the LoRaWAN technology in six distinct areas, namely, Europe, North America, China, Korea, Japan, and India

	Europe	North America	China	Korea	Japan	India
Frequency band	867-869MHz	902-928MHz	470-510MHz	920-925MHz	920-925MHz	865-867MHz
Channels	10	64 + 8 + 8	In definition by Technical Committee	In definition by Technical Committee	In definition by Technical Committee	In definition by Technical Committee
Channel BW Up	125/250kHz	125/500kHz				
Channel BW Dn	125kHz	500kHz				
TX Power Up	+14dBm	+20dBm typ (+30dBm allowed)				
TX Power Dn	+14dBm	+27dBm				
SF Up	7-12	7-10				
Data rate	250bps-50kbps	980bps-21.9kbps				
Link Budget Up	155dB	154dB				
Link Budget Dn	155dB	157dB				

the cellular network evolution. According to Bernardocki *et al.* (2020), these technologies have the following main characteristics:

- **NB-IoT:** 200-kHz bandwidth and data transmission rate of up to 250 kbps, with a battery life that can last for years. The most common use cases are sensors and meters;

- **LTE-M:** 1.4-MHz bandwidth and data transmission rate of up to 1 Mbps, with greater complexity and cost than NB-IoT, but with less latency, greater positioning accuracy, mobility in connected mode, and even voice support. However, its battery life is shorter than that of NB-IoT due to the greater complexity of the modem. The main applications are connected vehicles and trackers.

The combination of access technology and spectrum used defines the coverage scope, which meets the connectivity requirements. For the countryside, where a long-range coverage is required, the frequency range employed is a determining factor for technical and economic feasibility. In the licensed spectrum that is available for cellular networks, there are bands known as sub-GHz, namely, the 450-MHz and 700-MHz bands. These bands have a range of approximately 30 km, and the bandwidth is dependent on access technology and low latency. There are evolutions in NB-IoT that already cover more than 100 km (BERNARDOCKI *et al.*, 2020).

For Brazil, as the ecosystem in the 450-MHz band did not develop, operators chose the 700-MHz band as the most efficient to provide agribusiness services. It is worth noting that the 700-MHz band is also available in other countries in Latin America (TELE.SÍNTESE, 2020).

From the viewpoint of technological coverage, in Brazil, 88% of the municipalities have 4G (TELECO, 2020); however, only over 11% of the total area of the country has coverage (CARAM, 2020). The challenge of the extension of cellular coverage is that 41% of the national territory is destined to agribusiness (IBGE, 2017). Thus, investments linked to new business models are necessary and should be in line with cellular network operators, rural producers, and the ANATEL national regulator to achieve feasible spectrum usage costs. These investments will enhance the productivity and innovation in a sector that contributed 21.4% of the 2019 gross domestic product in Brazil.

The business model approach addressed by traditional cellular network operators is service based on private network. Private network is a cellular network implanted in a building, in a determined area (campus), and in an extensive area that serves a single organization. As it is a cellular network, it has the network elements necessary for its operation: core network, radio access network, spectrum, and edge computing. For agribusiness, private network represents the infrastructure that ensures connectivity in the field, employing a large number of sensors, actuators, and machines for data collection and processing, thus enabling an intelligent decision-making process (BERNARDOCKI *et al*., 2020).

Traditional cellular network operators focus on the end consumer and massive market. Moreover, although their business operating structures are not adopted to meet this type of supply to agribusiness, there have been initiatives to model possibilities for serving private networks according to the deployment mode, which can be from dedicated (exclusive infrastructure) to different levels of hybrid solutions (infrastructure that is partly exclusive, partly shared, or even virtualized) (5G ACIA, 2019). For these different deployment modes, the traditional cellular network operator must evaluate the technical and economic feasibility, plan agreements with the ecosystem, and create confidence in farmers with regard to the service-level agreement, coverage, cost, security, and data privacy (ANALYSIS MASON, 2020).

Depending on the company's nature of business, a licensed, unlicensed, and even shared spectrum can be employed in the private network, which provides opportunities for new players to provide this service. These new players offer various levels of specialization and can be classified into traditional suppliers of cellular network elements, specialized wireless network suppliers, cloud suppliers, and system integrators, each of which has its strengths and weaknesses with respect to different elements of the value chain. Both traditional cellular network operators and these new players must have business and operating models suitable for agribusiness to develop new channels for this market (ANALYSIS MASON, 2020).

In the countryside, this connectivity demand only existed for a few years, which restricted the improvement of competitiveness in this sector. Due to the delay in the provision of coverage in this sector, one of the leading machine manufacturers in the market, for example, did not have new products to launch in Brazil. However, through technological partners, the company developed a variant solution of cellular technology that combines two characteristics: cellular network with 4G LTE technology, which is used in the 250-MHz band, and private network, in which the ANATEL regulator released this band for this type of application (private limited service). It also rendered a feasible business model for the manufacturer, where its distribution chain made its machines and coverage available to its customers. In this case, the challenge is to invest in this infrastructure, which has two different approaches: for large producers, the investment capacity approach, productivity, and return gains are well defined, and for small producers with low investment capacity, the approach can be asset sharing through cooperatives or producer associations, including investment as a service option.

The 250-MHz LTE variant solution, as well as the LoRa and SigFox technologies (which operate at 900 MHz), are proprietary solutions which, therefore, limit the amount of solutions in the ecosystem. However, these solutions promote a competitive environment and can be utilized in addition to traditional cellular networks.

Figure 4 presents the main differences between SigFox, LoRa, and NB-IoT.

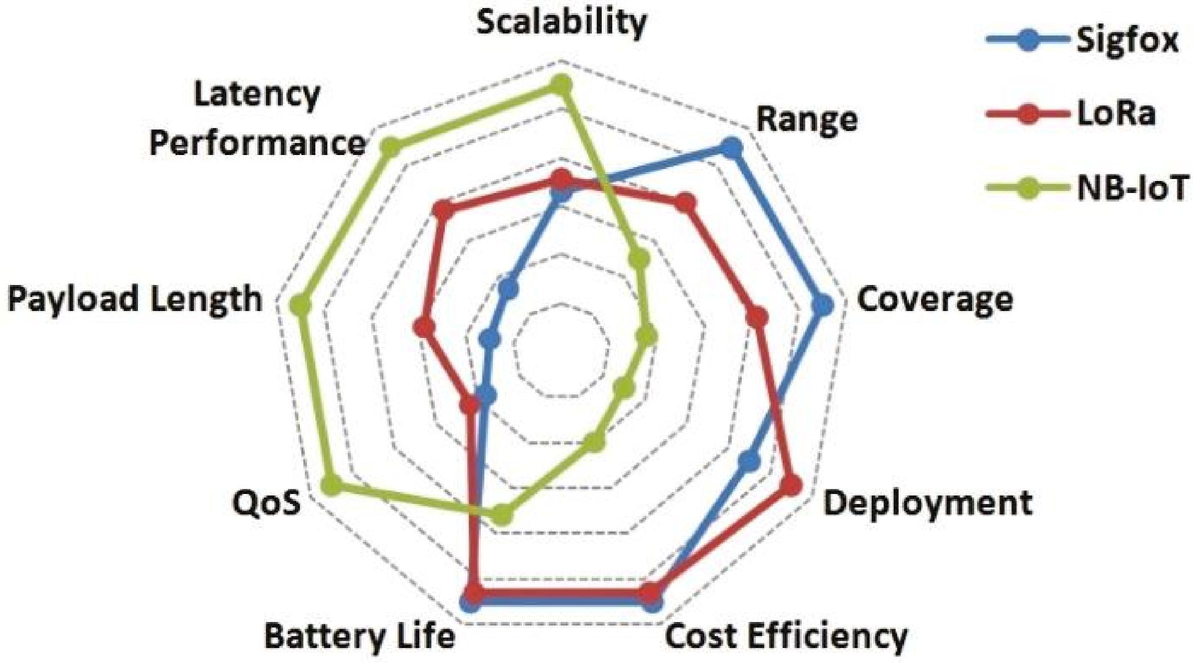
5G Connectivity

Despite the benefits of the abovementioned technologies, the 5G cellular network projects are in the works and will be revolutionary. In addition to the connection of services for humans, as is the case with the previous technologies, 5G also targets services for thousands of connections between devices to deploy the IoT. Unlike the previously mentioned technologies, 5G has a holistic design that provides understanding of both wireless and wired scenarios. Wireless communication is standardized by the 3GPP, whereas wired communication is supported by the Internet Engineering Task Force (IETF) (FITZEK *et al*., 2020).

In this context, the IoT and 5G have provided multiple possibilities of information exchange and communication between devices in a powerful way (LIU *et al.*, 2020). In Agriculture 4.0, such technologies exhibit high scalability, flexibility, individualization, etc., benefiting everything from small companies to large cooperatives (GIAMBENE; ADDO; KOTA, 2019). These characteristics are related to a new generation of agriculture, namely, Agriculture 4.0. For example, applications using 5G can monitor livestock or a set of agricultural machines and equip the combine to automatically order the carrier at the exact time of the current harvest and provide instructions on where to deliver the harvest (GAGLIORDI, 2018; GIAMBENE; ADDO; KOTA, 2019).

5G is a wireless broadband technology that shows promise in improving the speed and coverage of the 4G technology by providing low latency in wireless communication (SHAFI *et al*., 2017; LIU *et al*., 2020). Although the 4G technology is efficiently used by farms in different precision farming solutions, the 5G technology is expected to provide more reliable speed, more adequate bandwidth to large plantation areas, support for precision farming, and real-time connectivity (GIAMBENE; ADDO; KOTA, 2019). For example, during task development in the field, not all equipment may need to support the new technology. Only one tractor is needed to act as a connection and task hub for the other equipment.

Figure 4 - Respective advantages of SigFox, LoRa, and NB-IoT in terms of IoT factors (MEKKI *et al*., 2019)



Moreover, to synchronize and optimize the task in real-time, harvesters can exchange information and details with each other.

However, when using the 5G technology, numerous infrastructure challenges may be encountered in achieving maximum connectivity in a large farm and increasing its yield owing to its low population density. This infrastructure is important to support the use of the 5G technology in rural areas (GAGLIORDI, 2018; GIAMBENE; ADDO; KOTA, 2019). The deployment of the 5G infrastructure in rural areas enables the use of new sensors, the large-scale collection of agricultural data, and the application of new data analysis and machine learning techniques in decision-making and data protection (LIU *et al.*, 2020).

The 5G technology supports three categories of scenarios of larger mobile broadband, from which Agriculture 4.0 can benefit (SHAFI *et al*., 2017):

- Enhanced mobile broadband (eMBB): It increases the current data rate performance by providing a more seamless user experience. In other words, the eMBB technology offers user coverage with high mobility and high data rates;

- Ultra-reliable and low latency communications (URLLC): This technology supports applications that strictly require the characteristics of reliability, latency, and availability, for example, applications that use unmanned aerial vehicles to perform tasks (TORRES NETO *et al.*, 2015; TORRES NETO *et al.*, 2014; GIAMBENE; ADDO; KOTA, 2019), remote medical assistance (NETO *et al.*, 2017), intelligent networks (ROCHA FILHO *et al*., 2018), public protection (VÖLK *et al.*, 2020), and disaster relief (VÖLK *et al*., 2020), among others;

- Massive machine-type communications (mMTC): It promotes data generation, information exchange, and performance without human intervention. This means that communication is only between machines through wired or wireless networks. This technology can provide connectivity and network communication to billions of machines in device-to-device (TORRES NETO *et al*., 2017; TORRES NETO *et al*., 2019) or vehicle-to-vehicle (ROCHA FILHO *et al*., 2020) scenarios.

To enable complete operation of the 5G technology, there are several capacity and service quality requirements to achieve high coverage. The transmission rate requires optimal conditions for a guaranteed data delivery, whereas low latency requires shorter transmission time intervals and low power consumption when there is no data to transmit. To achieve the maximum capacity of the 5G technology, the following aspects must be met (SHAFI *et al*., 2017):

- Increased bandwidth: The deployment of the 5G technology provides a macro layer in the microwave bands for the control plane, and the microlayer in the millimeter wave band contains user-plane traffic, unlike the networks currently offered, where most frequencies are below 3 GHz;

- Massive multiple-input, multiple-output (MIMO) antenna arrays at the base station: Large-scale antenna arrays at the base station provide higher 5G frequencies, making it possible to overcome the loss of paths and to provide spatial multiplexing gain. Moreover, the arrays can form groups of dipoles to achieve the desired gain;

- Advances in MIMO: Simultaneous transmission to numerous users depends on the location, deployed methods, and spatial flows that can be supported by the base station. This problem can be solved by using 2D matrices and multiuser pre-coding;

- Network densification: 5G antennas provide a narrower beamwidth than the current sectorial antennas, which reduce the interference levels. Downloading network traffic to small cells is also possible;

- New waveforms: Although the orthogonal frequency-division multiple access in LTE provides high data transmission rates, the packet size for the mMTC is still small. 5G requires a new radio interface that is capable of providing a new multiple access scheme.

CLOUD COMPUTING IN AGRICULTURE 4.0

The 4th industrial revolution in the agricultural domain is currently underway, and it is all about data (LEZOCHE *et al*., 2020). This data generated in the field is provided by IoT devices and reaches the cloud *via* a proper telecommunication technology using lightweight protocols, as described in the first part of this work. Once in the cloud, aside from all the infrastructure technologies providing unlimited storage and computing power, the entry data is combined with intelligent systems that are capable of predicting situations, such as machinery maintenance.

Although achieving this kind of data pipeline maturity may appear easy nowadays, some key aspects need to be observed:

- Since the 1940s, numerous AI algorithms and statistical techniques have been discussed (HAENLEIN; KAPLAN, 2019). However, at that time, no computing power was available to promote such a data revolution;

- The amount of data generated by sensors applied to agriculture can quickly escalate, and maintaining the infrastructure on-premises to store and process all the gathered data may render the task too expensive. As discussed in the work of Mark and Griffin (2016), each plant generates approximately 0.5 KB of data. Considering the entire US corn plantation in 2015 (roughly 89 million acres), it corresponds to a total of 1.3 PB of gathered data;

- The necessary profile of a data scientist (DAVENPORT; PATIL, 2012) is not an off-the-shelf profile. It mostly requires skills from different areas of expertise.

In this section, we describe the state-of-the-art for cloud computing technologies. In addition, we describe a case study on the success of SiDi R&D and how the technology was employed to enable a data pipeline to empower the customer service business area of one of its biggest customers - a global company - with a wide vision of the currently developing trends across Brazilian cities, thus providing better results.

The following remarks need to be taken into consideration:

- Despite the fact that this case study is related to another sector, the essentials can be applied to other sectors, including the agricultural sector;

- Although the concepts may be applicable to most cloud service providers, as the attendance platform was implemented in the Amazon Web Services (AWS) cloud infrastructure, the examples and references are mostly focused on this player.

The Evolution of Computing Power

A fundamental concept for understanding the evolution of computer power, provided by the cloud computing paradigm, is elasticity. Al-Dhuraibi *et al.* (2018) defined elasticity as the ability of the system to allocate and remove resources “on the fly” so as to adapt to the load variation in real-time. These resources can be understood as processing and storage.

There are two types of elasticity: horizontal and vertical. Horizontal elasticity consists of the addition or removal of the instances of computing resources associated with an application. Conversely, vertical elasticity consists of the increase or decrease of computing resources, such as CPU time, cores, memory, and network bandwidth.

The elastic computing power in the cloud was provided by the software innovation called virtualization. Al-Dhuraibi *et al.* (2018) stated that virtualization makes it possible to simultaneously run multiple operating systems and multiple applications on the same server.

It creates an abstract layer that hides the complexity of the working environments of both the hardware and software. The cloud computing paradigm allows the deployment and quick scaling of workloads through the rapid provisioning of virtualized resources. Such a deployment is performed through virtual machines that enable a much higher capacity utilization and provide a much easier way to instantiate new workloads.

One example of service that exhibits computing elasticity is AWS Elastic Beanstalk[2]. It is used by SiDi R&D to deploy and scale services by adjusting the right computing power required by the clients.

A - On-Premises Versus On-Demand

To better understand the state-of-the-art cloud computing, it is important to go back in time and observe a number of key points. Before the advent of cloud computing, organizations would maintain their own on-premises data centers to process medium-to-large workloads. Although there is nothing wrong with maintaining servers in-house, some drawbacks must be observed:

- One of the disadvantages involves the thousands of dollars disbursed upfront for the entire asset;

- Notwithstanding the high investment, setting up servers in-house is time consuming, from the purchase order until a significant quantity of bytes can be processed by the business;

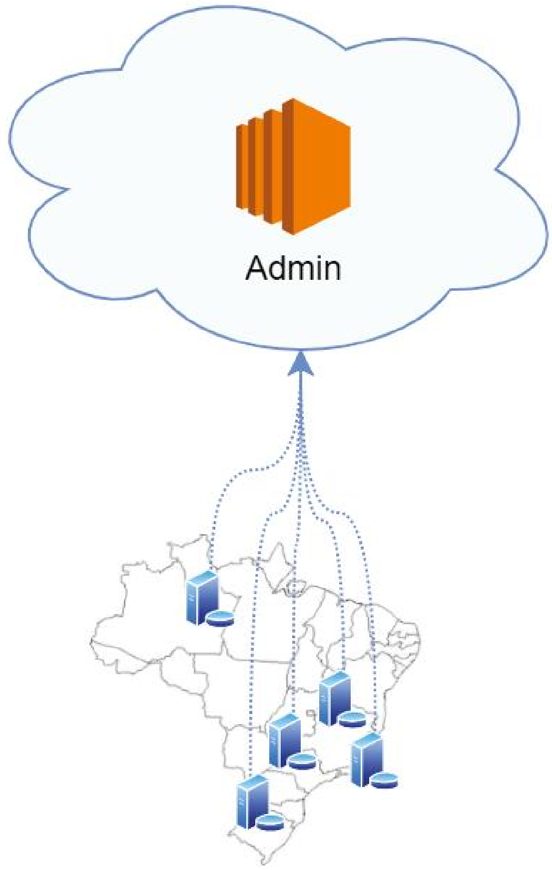
- In certain cases, the remote support for on-premise data centers is a very difficult task owing to a non-homogeneous environment that has different hardware platforms, different operating systems, different software from the various departments competing for hardware resources, etc.

These are only some of the possible drawbacks that highlight the stark contrast in efficiency when compared with the on-demand resources in the cloud. Instead, of waiting for a few days to conduct an experiment, any organization can launch and terminate virtual machines just by clicking a button. Moreover, in case more capacity is required (either computing power or memory), virtual machine resources can be immediately deployed. This is reflected in Andy Jassy's - AWS CEO - famous quote: "Invention requires two things: the ability to try a lot of experiments, and not having to live with the collateral damage of failed experiments."

In the case study on the customer service attendance platform, the first step in the digital transformation journey lies on this subject. One of the requirements of the attendance platform is that it should run on-premises in each service center across the country and that it should not depend on Internet access. In addition, it should eventually sync the service center data to the admin server in the cloud, as seen in Figure 5.

[2]https://aws.amazon.com/elasticbeanstalk

Figure 5 - Original deployment of the attendance platform



After a few days into production, as expected, all the drawbacks previously mentioned were observed, as were many others. Thus, the SiDi R&D team proposed lifting the entire solution to the cloud due to only a few deployment adjustments (Figure 6). Without any development effort, both the computing and database layers could be detached into different resources with specific purposes: Amazon EC2[3] instances for computing and Amazon Relational Database Service (RDS)[4] for database.

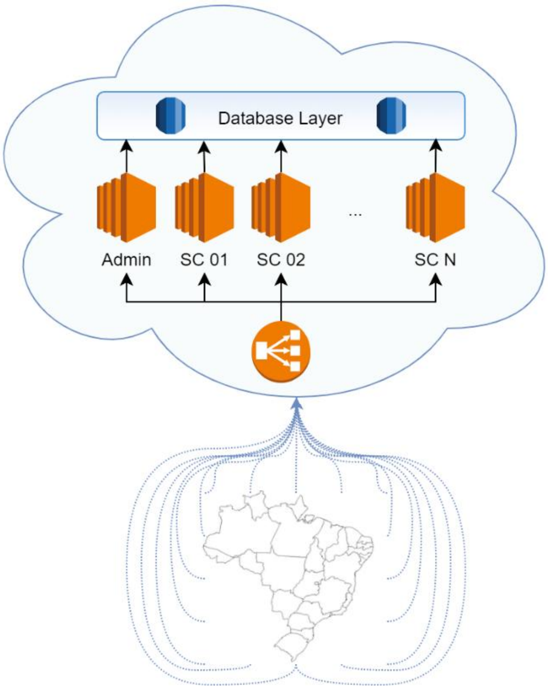
After the application of the abovementioned strategy, a number of benefits were immediately noticed:

- Minimal investment to set up new solution server for a new service center;

[3]https://aws.amazon.com/ec2

[4]https://aws.amazon.com/rds

Figure 6 - Attendance platform lifted to the cloud



- Homogeneous environment that facilitates remote support;

- Accurate backup process for all generated data;

- Better performance, which reduces hardware expenditure.

Another strategy is the formation of farmer associations where the computing infrastructure may be maintained and run on their own on-premise data centers.

In the next section, the state-of-the-art of computing and how, by using it, SiDi R&D was able to shift the attendance platform are described.

B - Serverless Services

Serverless computing is the next evolution in cloud computing (PAVAN *et al*., 2020). In this paradigm, the development team is free from server management tasks as they are entirely handled by the service provider. The available solutions vary depending on the abstraction level presented to the developers, from an initial level where defining the software stack is possible to a final level where they only have to implement business rules as single methods executed by requests or triggered by events.

The first level is commonly observed in the big service providers in the form of Containers as a Service (CaaS) (CAREY, 2020). In these solutions, aside from the implementation of the business rules, the developers also need to perform data source integration, task management, and security and network setup. Conversely, the service provider is responsible for container orchestration and scalability.

There are intermediary levels where the developers do not need to worry about containers, clusters, or software stack management; they can work using popular development tools and languages but still deploy their solution in a bundle in which the service provider is able to scale as needed. These services are known as Platform as a Service (PaaS) (VIOLINO, 2019).

Finally, in the last level of abstraction, the developers only need to worry about the application code, whereas all the rest is handled by the service provider. The same applies to the intermediary level, where the developers can usually choose from popular languages and tools. This level also offers automatic, managed, transparent, and fine-grained scaling capability and is usually referred to as Function as a Service (FaaS) (ROBERTS, 2018).

Among the benefits of serverless computing, apart from the freedom to handle only the application code, there are also cost reduction and shorter cycles both in the developmental and operational tasks. In addition, serverless applications exploit the growing number of managed services made available in the cloud ecosystem. Such services are used in the application to handle different tasks, such as message delivery, authentication, queues, data storage, and manipulation.

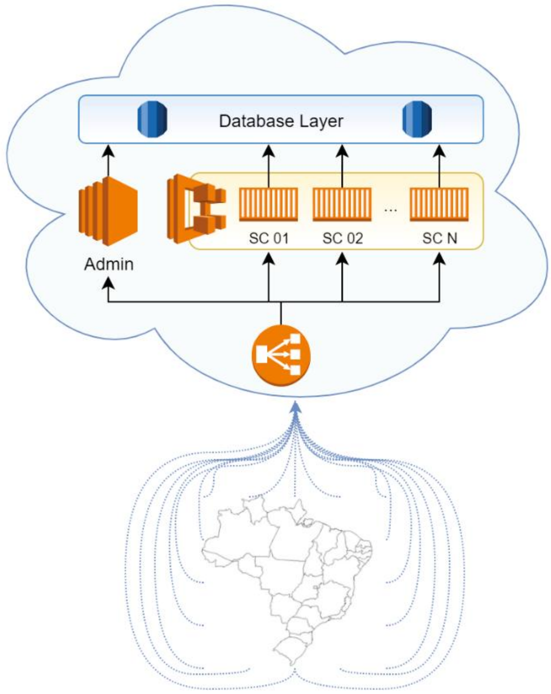
However, the use of such services has numerous drawbacks, such as not being able to control important aspects like service downtime, usage limits, and forced upgrades of application programming interface. Finally, one important aspect that needs to be considered when shifting into serverless implementation is vendor lock-in. While changing the service provider when running serverless applications in CaaS is easier, it could be harder in FaaS.

In the case study on the attendance platform of customer service centers, the second step in the digital transformation journey lies on this subject. After lifting the entire solution to the cloud, all the computing resources required to run the attendance platform were under the control of SiDi R&D; thus, maintaining these resources became easier. However, as the fleet of service centers quickly increased, the management of each individual instance became difficult, generating a scale in price since the solution considered one Amazon EC2 instance per service center.

To overcome this problem, SiDi R&D started using the Amazon Elastic Container Service[5] (Amazon ECS) with AWS Fargate[6] to orchestrate a fleet of containers in serverless mode, thus replacing the Amazon EC2 instances (Figure 7).

[5]https://aws.amazon.com/ecs
[6]https://aws.amazon.com/fargate

Figure 7 - Attendance platform shifted to serverless



In addition to facilitating the deployment process and management of each service center, it was possible to only allocate ¼ of the resources to the containers used by each Amazon EC2 instance, thus reducing the total computing costs.

Data and Storage

In the agricultural context, Big Data provides new opportunities in Smart Farming, not only in primary production but also in providing predictive insights into farming operations, making real-time operational decisions, and redesigning business processes for game-changing business models (WOLFERT *et al*., 2017).

The data must be stored in a way that it can be accessed easily so that it can be useful. In the cloud context, this can be achieved by using storage and database managed services that enable entities to save their resources (data and files) off-site, making these resources available through the Internet.

Databases and the evolution of storage managed services in the cloud are not different from the evolution of computing power observed in the past. When developing software in the cloud, storage solutions vary from on-premises to serverless. Considering that the application needs a database, in case the data handled by the application is too sensitive, it could require the database to run internally on its own on-premises servers. However, if it requires high scaling capabilities, serverless solutions that are capable of worldwide petabyte scaling are available. Among those options, a range of services to address different application requirements exists.

A - Cloud Storage Services

Cloud storage service providers, such as AWS[7] and Google Cloud Services[8], commonly provide object-based storage solutions, which have some of the following characteristics:

- The resources are stored as an object–a container that includes the data in the resource, any associated metadata, and a unique identifier;

- The objects can be organized in containers called buckets;

- The objects can be versioned.

The providers are responsible for guaranteeing availability, so users can access their data when needed; durability, through replication, ensuring that the object will not be lost; and security, by providing encryption mechanisms, access control list, and bucket policies.

As previously mentioned, cloud-based solutions are cost effective as they provide elastic allocation for on-demand resources. Cloud storage managed services allow cost management through lifecycle management: objects can be stored in different storage classes according to the specific policies of the customer and how this object is frequently accessed. In general, this management can be manual or automated.

Lifecycle management is possible through the selection of different storage classes, which are designed for different use cases:

- **General-purpose storage class:** It is designed to offer low latency and high throughput for frequently accessed objects;

- **Infrequent access storage class:** It is designed to offer cheap storage price for objects that are not frequently accessed;

- **Archive storage class:** It is designed to offer cheaper storage price for data archiving.

[7]https://aws.amazon.com/s3/
[8]https://cloud.google.com/storage

B - Cloud Databases

Cloud database managed services are cloud computing services that allow users to access and use databases in the cloud infrastructure without purchasing or performing database administrative tasks, such as provisioning and data backup.

The key benefits of using a cloud database are as follows:

- Cost savings: The organizations do not need to purchase the infrastructure for the database; they are only charged for the resources they consume;

- Scalability: It is easy to scale database clusters according to the organization's requirements;

- Disaster recovery: Cloud database managed services provide features that enable database deployment and backup in different regions, which enhances the availability and durability of the company's data;

- Security: The services typically provide data encryption at rest and in transit.

These managed services provide at least two types of database: Structured Query Language (SQL) database, which is a relational database, and NoSQL databases[9]. Cloud service providers, such as AWS, typically provide database services either as a fully managed service (which still demands a server) or as a serverless solution.

Relational databases

Relational databases organize data in tables, a predefined schematized set of columns and rows, and the relationship between them using common fields known as foreign keys. These databases can be accessed using SQL, which was standardized by the American National Standards Institute and is supported by the most popular relational database engines.

Relational databases have some important aspects:

- Data integrity and consistency: It is provided by the primary key, foreign keys, and a set of data constraints that must be respected in any transaction;

- Transaction: It is a set of SQL statements that must be executed in a sequence as a single unit of work;

- Atomicity: All the statements in a single transaction are successfully executed; if any statement fails, the entire transaction is voided, and none of the statements are executed;

- Durability: It ensures that all the changes made in a successful transaction are permanent.

[9]https://www.ibm.com/cloud/learn/nosql-databases

The AWS cloud service provider provides a range of relational database engines (Oracle Database, Microsoft SQL Server, MySQL, PostgreSQL, MariaDB, Amazon Aurora[10]) as fully managed services branded as Amazon RDS. Built for the cloud and equipped with hardware-level optimizations, Amazon Aurora is a relational database compatible with MySQL and PostgreSQL "up to five times faster than standard MySQL databases and three times faster than standard PostgreSQL databases." It also relies on a serverless version, which can automatically handle scaling.

NoSQL databases

This type of database is designed to provide flexibility in the schemas used to structure, store, and retrieve data.

NoSQL database can manage the following data models:

- Key-value structure: The data is stored using a key (usually a string), and the data itself is stored as primitive types or as complex objects. This data model replaces the rigidity of the relational schemas and allows developers to modify the structure of the stored object without losing the integrity of the database;

- Document-based: The data is stored in a semi-structured way, such as JSON (JavaScript Object Notation)-like documents, which provides flexibility to store varying attributes within a document;

- Column-based: The data is stored in rows, and each row can contain a variable number of columns, thus providing great flexibility;

- Graph-based: The data is stored in a graph structure, where the nodes are the data and the edges are the relationship among the data.

This type of database is a good choice when a very large amount of unstructured/semi-structured data must be supported without a predefined schema.

AWS provides fast and flexible solutions for NoSQL databases[11]: Amazon DynamoDB as a key-value and document-based database; Amazon Keyspaces as a column-based database; and Amazon Neptune as a graph-based database.

Data Analysis and Insights

The International Data Corporation published a white paper, *The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things* (TURNER *et al.*, 2014), concluding that in 2014, the digital universe was growing by 40% annually, and in

[10]https://aws.amazon.com/rds/aurora

[11]https://aws.amazon.com/products/databases/

2020, it is expected to reach approximately 44 ZB or 44 trillion GB. This is mainly due to the increasing number of smart devices and connected things.

As mentioned previously, the 4th industrial revolution in the agricultural realm is all about data, and cloud computing has certainly enabled the transformation of data into valuable information. However, the collection and storage of data in itself are insufficient without a data scientist. A data scientist, which is an important figure in this landscape, uses the available data and tools to provide important insights into the business.

One of the most important tools used by a data scientist is a data lake (MILOSLAVSKAYA; TOLSTOY, 2016). According to Khine and Wang (2018), data lake is a relatively new concept that attracts more attention from business enterprises than from academic institutions. However, this does not mean that the topic is any less relevant. Rather than perceiving data lake as a new marketing label for the traditional big data concept (MILOSLAVSKAYA; TOLSTOY, 2016.), Fang (2015) defined data lake as "A methodology enabled by a massive data repository based on low-cost technologies that improves the capture, refinement, archival, and exploration of raw data within an enterprise. A data lake contains the mess of raw unstructured or multi-structured data that for the most part has unrecognized value for the firm."

In data lakes, data essentially comes from the producers and is stored in a proper repository until use for analysis to provide important business insights. When designing a data lake, a solution architect must carefully handle two points: data ingestion and storage. Due to the unpredictability of the data production rate (sensor readings on an oil rig, product sales, voting on an interactive TV show, etc.), concepts such as elasticity, redundancy, and availability are indicated to address all the incoming events without losing any bytes. It should be noted that, with persisting data, durability is also indicated.

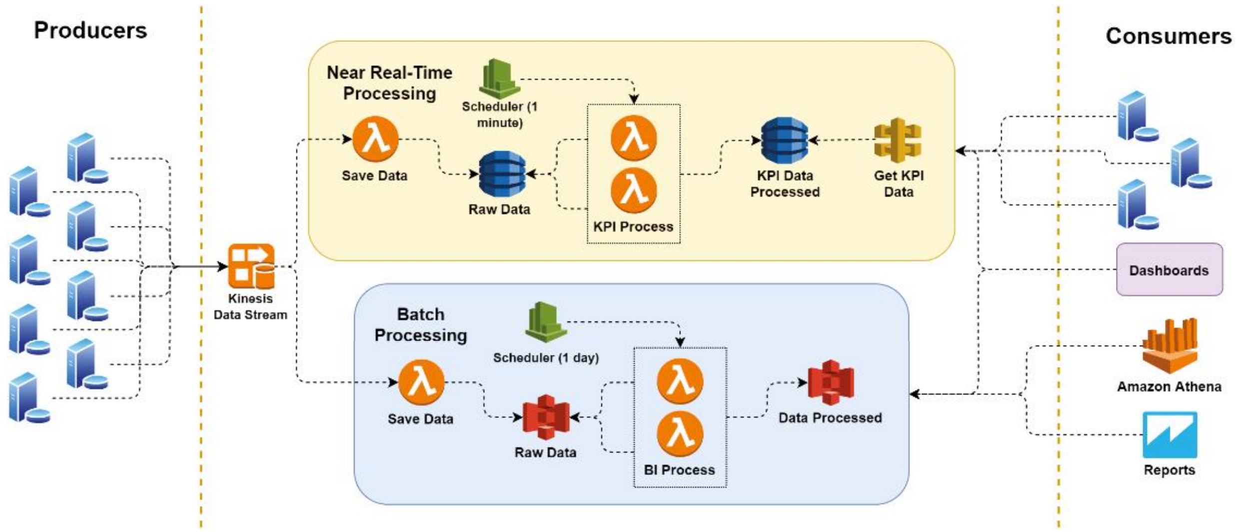
In the study case on the attendance platform of customer service centers, the third step in the digital transformation journey lies on this subject. The creation of a data lake for the attendance platform provides a near real-time view of the service center administrators and helps them fix bottlenecks and provide better service. Moreover, it provides the executive team with a wide perception from the business intelligence (BI) reports, thus enabling them to drive the business from the data. Figure 8 illustrate the core elements in this data lake implementation.

In this case, the producers are the service centers across Brazil, which in total produce a few thousand utterance events per minute. In front of both structures (near real-time and batch processing), is Amazon Kinesis Data Streams[12], which is a fully managed service designed to be "a massively scalable and durable real-time data streaming service." It captures all the data produced by the fleet of service centers.

Then, AWS Lambda[13], which is a serverless computing service, plays a significant role in the consumption and proper storage of data. It is also employed when the "Scheduler" triggers either the Key Process Indicator (KPI) process (every each other minute)

[12]https://aws.amazon.com/kinesis/data-streams
[13]https://aws.amazon.com/lambda

Figure 8 - Attendance platform data lake for near real-time and batch processing



or the BI process (once a day). Each of the processes has its own business rules implemented in the supported programming languages.

For the persistence layer, the architecture contemplates two separate storage types used according to each process:

- Near real-time processing persists the data using Amazon DynamoDB[14], which is a serverless database service offering a single-digit millisecond performance on any scale. This database was chosen mainly because in this part of the process, computing is designed to be triggered every minute, and it must deliver the currently calculated KPIs as fast as possible whenever required by any business dashboard or service center. Particularly, for this part of the process, the raw data is dropped at the end of the day, and the data growth is therefore not a concern, nor are the resulting costs, because with a continuous increase, the costs could be rapidly scaled.

- Batch processing persists the data using Amazon S3[15], which is a serverless object storage service designed to achieve high availability and durability, natively integrated with a myriad of AWS products (e.g., Amazon Athena[16], a serverless service that enables querying data in Amazon S3 using standard SQL). It is a very cost-effective storage solution with varying prices for different types of access, enabling a long lifecycle for raw data. Accessing information in this type of storage is nowhere near the Amazon DynamoDB single-digit millisecond access time; however, this is not a problem as this part of the computing process is designed to be triggered only once a day. Moreover, at the end of this pipeline are the BI reports created using Amazon QuickSight[17], which is a fully managed service for the creation of dashboards and reports.

IoT Backend Services and Edge Computing

According to O'grady *et al.* (2019), an average farm may quickly generate large data volumes from satellites, drones, and IoT sensors. Processing such a large volume of data in a way that it is easily interpreted requires the use of a variety of analysis techniques, from image processing to machine learning. Cloud computing technologies are attractive solutions for data analysis, as mentioned in the previous sections.

Since not every location where an IoT sensor is deployed has access to the Internet, assuming that a cloud backend is always available is impractical. In addition, even in cases in which Internet is available, a high latency may occur, and for some applications, such a high latency is not tolerated. Thus, a near real-time response is required.

To overcome these limitations, a new paradigm called edge computing was proposed. It is an extension of the cloud computing paradigm and moves the computing (and storage) resources to the proximity of the devices/users from the core to the edge of the network (Figure 9).

The objective of edge computing is to support a wide range of IoT applications and thus allow data processing anywhere along the IoT–edge–cloud continuum. This distributed computing structure has inherent advantages and also provides system scalability to allow the infrastructure to process large-scale data generated by IoT devices that are geographically distributed. With edge computing, data processing is performed closer to the physical location of the IoT device, and only digests or exceptions are sent to the cloud for more processing or storage (NETO *et al., 2017*; TORRES NETO *et al.*, 2019; QIAO *et al.*, 2020).

Among the numerous cloud provider offerings for IoT backend services, the AWS IoT Core is the top-ranking (AWS-IC, 2020). It forms the backbone for IoT deployments to guarantee secure connection to IoT devices and handle the data at scale. The AWS IoT Core allows the devices to securely connect to the AWS cloud and to each other. This allows the devices to route, process, and act on the data and messages coming from these devices. Moreover, with AWS IoT Core, it is possible to deploy and develop applications that can interact with IoT devices even when offline.

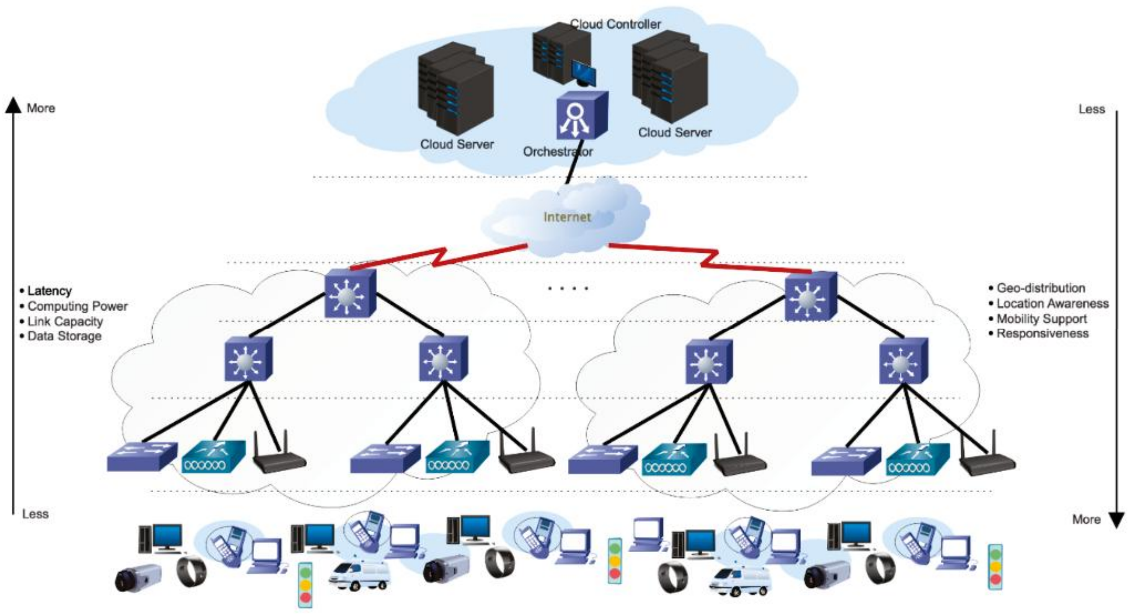
AWS IoT Greengrass (AWS-GG, 2020) is the major cloud platform for edge computing. It seamlessly extends AWS to edge devices so they can locally act on the data they generate while still using the cloud for management, analytics, and durable storage. AWS IoT Greengrass enables the connected devices to run AWS Lambda functions, Docker containers, or both and to execute predictions based on machine learning models, keep device data in sync, and securely communicate with other devices, even when offline. It comprises a large ecosystem of services meant to quickly bring IoT solutions to customers worldwide. At SiDi R&D, we have expertise in the integration of these services and development of tailored solutions to our customers while leveraging the right services for each case.

TRENDS FOR AGRICULTURE 4.0

The trends of ICTs in Agriculture 4.0 in the face of existing challenges, which mainly involve the increase

[14]https://aws.amazon.com/dynamodb
[15]https://aws.amazon.com/s3
[16]https://aws.amazon.com/athena
[17]https://aws.amazon.com/quicksight

Figure 9 - IoT-Edge-Cloud architecture (QIAO *et al*., 2020)



in agricultural production without expansion of the planted area, provide new opportunities for the use of innovations. These trends indicate that the agricultural sector requires new ICTs to manage data, information, and knowledge at all stages of the production chain in a new infrastructure where the physical and digital worlds are totally interconnected.

The advancements in science and technology have significantly contributed to the improvement in global food production. The productive capacity of agriculture has increased between two-and-a-half fold and threefold in the last 50 years (FAO, 2017). This has leveraged the growth of global food production to keep pace with the population growth. In addition to the growing demand for foodstuffs, food production faces other challenges that make the context even more complex, such as climate change, which restricts the capacity of natural resources, namely, water and soil.

Innovation is important to ensure that future generations have access to quality food and nutrients. For that to happen, the way we produce food needs to be changed. Increasing productivity alone is insufficient. Thus, a more comprehensive approach that involves sustainable production and consumption is required to guarantee food security for future generations (EMBRAPA, 2015).

To guarantee a promising future, advanced digital technologies must be involved in the agricultural production process to enable these technological innovations to promote an intensive agriculture and massive knowledge, with high levels of productivity and sustainability, reduced costs, and better working conditions in the field.

The search for optimizing the use of natural resources will cause intense monitoring and automation in the future farms. Sensors spread across the cropland and connected to the IoT platforms in the cloud will generate the big data that needs to be filtered, stored, and analyzed. The human workforce will not be able to handle such an amount of data; thus, better algorithms through computational intelligence and cognitive computing techniques for assistance are required in the analysis process. After the analysis, the cycle is ended using remote controls for tractors and agricultural implements that, equipped with intelligent processing and location systems, only intervene as needed to optimize costs and production, reduce the environmental impact, and upgrade logistics, thus expanding the business horizon (ESQUERDO, 2014).

With connected agriculture, companies and independent producers, as well as governments themselves, will be able to perform real-time monitoring of the machine performance for the storage and transportation of crops, where automatic data transmission over data networks is essential.

In this interconnected environment, where the generation of knowledge and mobility as well as the growing offer of mobile applications are an expanding market, Agriculture 4.0 is expected to help every farmer reap the benefits of this technology.

In this scenario, Agriculture 4.0 can certainly be cited within the context of the following four macro trends (REDAÇÃO AGRISHOW, 2016):

1 - Digital mindset and open innovation

In the acceleration of the digital transformation of agriculture, cultural aspects cannot be left out, and the power of communication and perception to change behaviors should also not be overlooked.

Numerous individuals from the younger generations, with a natural propensity to use technology, tend to work in agriculture, which has been attracting considerable attention. This scenario can significantly collaborate with the digital transformation and provide a more widespread and open form of innovation. Currently, there are several initiatives in this direction.

2 - Big data and predictive models

The increased connectivity in the field will drive the growth of big data and provide access to digital technologies anywhere in the world. However, the first condition for the digital transformation of the big data industry is that the raw material must exist, in this case: data. To create larger databases, companies have used new technologies, such as satellites, drones, IoT, and sensors, on major machines to enable the capture and transformation of production data into digital solutions (MASSRUHÁ, 2014).

The next step is to know what to do with the data. In this sense, agribusiness is well matured. One of the first experimental statistics was born with Ronald Fisher in 1919 with agricultural research at the Rothamsted Research experimental center in Harpenden, England. Over time, a lot of data and models have been generated. Given the new technological resources, the generation of data and models will now be further accelerated.

3 - Customer experience and perception

By adopting a higher level of transparency and traceability throughout the agribusiness production chain, including the production, distribution, and stock sectors, a more dynamic model of trust and price, which will benefit the customer, can be generated. In the end, much of this transformation pillar involves better data management and a critical consumer culture.

If better customer experience is implemented in the agricultural value chain, producers, suppliers, and consumers will also benefit, which can threaten the traditional cooperative model. Aware of this, cooperatives have been making great strides in terms of customer experience.

4 - New business models

Platform models, ranging from e-commerce to the market, are being generated to transform traditional business models, along with innovative digital tools for market research, sales, and distribution. In the scenario of Covid-19 pandemic, where the food security of the nations will be under a high level of scrutiny, new digital business models can contribute to the generation of shorter supply chains. Disruptive innovations will occur outside the front door. This is because it is in the processes of production financing, purchasing, and marketing where the entire links that can be totally transformed are observed, eliminating players, modifying business models, and redesigning the complex agricultural chain in its entirety.

In addition to being complex and inefficient as well as involving hundreds of entities, these processes are practically common to all cultures; therefore, they bring about one of the factors that significantly attract digital transformation: the volume and scalability of the solutions.

CONCLUSIONS

1. The transformation sectors have gone through several revolutions, from mechanical production based on steam-powered systems to the use of electricity and automation of the stages, already making use of computer resources. In recent years, digital transformation has motivated changes in production processes based on artificial intelligence, big data, and IoT. All sectors, both public and private, are undergoing adaptations, and agribusiness is no different;
2. In this field, the innovation involved pillars, such as soil analysis, livestock remote monitoring, crop monitoring and real-time production, sensors connected to communication networks, biotechnology, tractors equipped with autonomous systems, and connected drones, among many other solutions that generate a huge amount of data. The collection and treatment of data are great propellers of this new phase as they enable a more accurate monitoring of weather conditions, diseases and pests, resource management, logistic improvements, storage solutions, and design of new business models;
3. Countryside connectivity-Internet access and the coverage of farmland with connected machinery and devices–represents one of the many challenges of the

massive adoption of ICT in the agricultural sector. However, great efforts have been exerted to overcome this issue, and the results are reflected in the developed technologies, such as SigFox, Lora, 4G, and 5G, each with its own characteristics meeting the specific requirements for the croplands;

4. Cloud computing services provide a great deal of solutions–data storage, computing power, and data analysis tools–with high scalability and flexibility, thus allowing the extraction of meaning from the collected data and application of the insights to change the decision-making process in agriculture;
5. Connectivity and cloud computing link together many other ICTs that can improve the production process in agriculture, such as artificial intelligence, IoT, and data analysis. These technologies have been successfully employed in other domains, and agriculture is adopting a digital mindset that is open for innovation and new ventures.

ACKNOWLEDGMENTS

All the other authors and I, John (Jin Hong) Yi, in the position of SiDi's Superintendent - one of the main R&D centers in Latin America and avant-garde in disruptive technologies -, would like to thank Professor Daniel Albiero, Agricultural Engineering, Ph.D. for his support and for extending the invitation to participate in this article. A special thanks to the Universidade Federal do Ceará (UFC) and Revista Ciência Agronômica (RCA), the fifth scientific journal with the greatest impact factor in the area of Agricultural Sciences, for all their support during the development of this article.

REFERENCES

3rd GENERATION PARTNERSHIP PROJECT-3GPP. Release 13, [S.I.], 2016. Available: <https://www.3gpp.org/release-13>

3rd GENERATION PARTNERSHIP PROJECT-3GPP. Release 14, [S.I.], 2017. Available: <https://www.3gpp.org/release-14>

5G ALLIANCE FOR CENNECTED INDUSTRIES AND AUTOMATION-5G ACIA. 5G Non-Public Networks for Industrial Scenarios, [S.I.]. 2019. Available: <https://www.5g-acia.org/fileadmin/5G-ACIA/Publikationen/5G-ACIA_White_Paper_5G_for_Non-Public_Networks_for_Industrial_Scenarios/WP_5G_NPN_2019_01.pdf>

AL-DHURAIBI, Yahya *et al.* "Elasticity in Cloud Computing: State of the Art and Research Challenges," in **IEEE Transactions on Services Computing**, v. 11, n. 2, pp. 430-447, 1 March-April 2018, doi: 10.1109/TSC.2017.2711009.

ANALYSIS MASON. Webinar Private LTE/5G networks: opportunities for operators, [S.I.]. 2020. Available: <https://www.analysysmason.com/events-and-webinars/webinars/private-lte5g-networks-an-iot-growth-opportunity-for-operators>

ARJONA, R *et al*. An Experimental End-to-End Delay Study of a Sub-1GHz Wireless Sensor Network with LTE Backhaul. In: 2018 **IEEE Global Communications Conference (GLOBECOM)**. IEEE, 2018. p. 1-7.

AWS-IC. Amazon Web Services, "AWS IoT Developer Guide," 2020. Available: <https://docs.aws.amazon.com/iot/latest/developerguide/what-is-aws-iot.html>. Accessed October 2020.

AWS-GG. Amazon Web Services, "AWS Greengrass Developer Guide," 2020. Available: <https://docs.aws.amazon.com/greengrass/latest/developerguide/what-is-gg.html>. Accessed October 2020.

BERNARDOCKI, P *et al.* E-Book Cobrindo o campo com IoT Celular, São Paulo-SP, Ericsson. 2020. Available: <https://digital.futurecom.com.br/o-futurecom/e-book-cobrindo-o-campo-com-iot-celular>

CARAM, V. Live Painel Telebrasil 2020 Workshop 1: Futuras Demandas por Espectro, 08/09/20, ANATEL presentation. Available: <https://www.telesintese.com.br/anatel-apenas-14-do-territorio-brasileiro-tem-cobertura-3g-e-4g>

CAREY, S. "What is CaaS? Simpler container management." Available: <https://www.infoworld.com/article/3567202/what-is-caas-simpler-container-management.html>

COCAMAR. **Cocamar participa de maratona digital**. 2020. Available: <https://www.cocamar.com.br/noticia/Cocamar_participa_de_maratona_digital/5412> Accessed at October 29, 2020.

DAVENPORT, T. H.; PATIL, D. J. Data scientist: The Sexiest Job of the 21st Century. **Harvard Business Review,** v. 90, n. 5, p. 70-76, 2012.

EMBRAPA. Embrapa em números. Brasília, DF, 2015. 138p. Available: <https://www.embrapa.br/embrapa-em-numeros>

EMBRAPA. Visão 2030 - O Futuro da Agricultura Brasileira, Brasília-DF. 2018. Available: <https://www.embrapa.br/visao-2030>

EMBRAPA; SEBRAE; INPE. Pesquisa Agricultura Digital no Brasil, Campinas-SP. 2020. Available: <https://www.embrapa.br/agropensa/produtos-agropensa>

ESQUERDO, J. C. D. M.; CRUZ, S. A. B.; MACÁRIO, C. G. do N.; ANTUNES, J. F. G.; SILVA, J. dos S. V. da; COUTINHO, A. C. Tecnologias da informação aplicadas aos dados geoespaciais. In: MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A.; LUCHIARI JUNIOR, A.; ROMANI, L. A. S. (Ed.). Tecnologias da informação e comunicação e suas relações com a agricultura. Brasília, DF: Embrapa, 2014. Cap. 8. p. 139-156.

ESTESO, A.; ALEMANY, M M; ORTIZ, A. Conceptual framework for designing agri-food supply chains under uncertainty by mathematical programming models.

International Journal of Production Research, v. 56, n. 13, p. 4418-4446, 2018.

FANG, H. 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, 2015. p. 820-824.

FAO - ORGANIZATION OF THE UNITED NATIONS with inputs from International Food Policy, Research Institute (IFPRI) and Organization of Economic Cooperation and Development (OECD). Food and Agriculture Organization of the United Nations Rome, 2017.

FIELKE, S; TAYLOR, B; JAKKU, E. Digitalisation of agricultural knowledge and advice networks: A state-of-the-art review. **Agricultural Systems**, v. 180, p. 102763, 2020.

FITZEK, F HP *et al.* On the need of computing in future communication networks. In: **Computing in Communication Networks**. Academic Press. p. 3-45, 2020.

FOOD AND AGRICULTURE ORGANIZATION. How to Feed the World in 2050. In: **Executive Summary-Proceedings of the Expert Meeting on How to Feed the World in 2050**. Rome, Italy: Food and Agriculture Organization, 2009.

GAGLIORDI, N. **How 5G will impact the future of farming and John Deere's digital transformation**. 2018. Available at: <https://www.zdnet.com/article/how-5g-will-impact-the-future-of-farming-and-john-deeres-digital-transformation/>. Accessed at 29 October 2020.

GIAMBENE, G; ADDO, E O; KOTA, S. 5G Aerial Component for IoT Support in Remote Rural Areas. In: **2019 IEEE 2nd 5G World Forum (5GWF)**. IEEE, 2019. p. 572-577.

GOMEZ, C *et al*. A Sigfox energy consumption model. **Sensors**, v. 19, n. 3, p. 681, 2019.

HAENLEIN, M; KAPLAN, A. A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. **California Management Review**, v. 61, n. 4, p. 5-14, 2019.

IBM. <MARATHON/> behind the code 2020 [S.I][2020]. Available: < https://maratona.dev/en >. Accessed at 2020, October 27.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA–IBGE. Censo Agropecuário 2017, Rio de Janeiro-RJ. 2017. Available: <https://censos.ibge.gov.br/agro/2017/>

KHINE, P P; WANG, Zhao Shun. Data lake: a new ideology in big data era. In: ITM web of conferences. EDP Sciences, 2018. p. 03025.

KLERKX, L; JAKKU, E; LABARTHE, P. A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. **NJAS-Wageningen Journal of Life Sciences**, v. 90, p. 100315, 2019.

LEZOCHE, M *et al*. Agri-food 4.0: a survey of the supply chains and technologies for the future agriculture. **Computers in Industry**, v. 117, p. 103187, 2020.

LIU, Y *et al*. From Industry 4.0 to Agriculture 4.0: Current Status, Enabling Technologies, and Research Challenges. **IEEE Transactions on Industrial Informatics**, 2020.

LORA ALLIANCE. **A technical overview of LoRa® and LoRaWAN™**. [S.I] [2015]. Available: <https://www.tuv.com/media/corporate/products_1/electronic_components_and_lasers/TUeV_Rheinland_Overview_LoRa_and_LoRaWANtmp.pdf>. Accessed at: 2020, Aug. 28.

MARATONA BEHIND THE CODE. Desafio 01 Cocamar: Maratona Behind the Code, 2020. Available: < https://github.com/maratonadev-br/desafio-1-2020>. Accessed at 27 october 2020.

MARK, T; GRIFFIN, T. Defining the barriers to telematics for precision agriculture: Connectivity supply and demand. 2016.

MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A.; MOURA, M. F. Os novos desafios e oportunidades das tecnologias da informação e da comunicação na agricultura (AgroTIC). In: MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A.; LUCHIARI JUNIOR, A.; ROMANI, L. A. S. (Ed.). Tecnologias da informação e comunicação e suas relações com a agricultura. Brasília, DF: Embrapa, 2014. Cap. 1. p. 23-38.

MEDEIROS, J. Produtor Digital. Live AGROtic 2020, Qual o Perfil do Agricultor Digital, 15/set/2020, CNA/SENAR presentation.Available:<https://www.eventos.momentoeditorial.com.br/live-agrotic-2020>

MEKKI, K *et al*. A comparative study of LPWAN technologies for large-scale IoT deployment. **ICT express**, v. 5, n. 1, p. 1-7, 2019.

MILOSLAVSKAYA, N; TOLSTOY, A. Big data, fast data and data lake concepts. **Procedia Computer Science**, v. 88, n. 300-305, p. 63, 2016.

NETO, J. R. *et al*. INCA: Um sistema healthcare flexível baseado no paradigma fog computing e publish/subscribe. In: **Anais do I Workshop de Computação Urbana**. SBC, 2017.

O'GRADY, M.D.; LANGTON, D., O´HARE, G.M.P. Edge computing: A tractable model for smart agriculture?, **Artificial Intelligence in Agriculture**, v. 3, 2019, p. 42-51, ISSN 2589-7217, https://doi.org/10.1016/j.aiia.2019.12.001.

PAVAN, L. V. *et al*. "A Study of Serverless Architecture: An Overview" (IJRASET) SSN: 2321-9653; IC Value: 45.98; **SJ Impact Factor: 7.429** v. 8 Issue VI June 2020- Available: <www.ijraset.com>.

QIAO, YUANSONG & SENHAJI HAFID, ABDELHAKIM & AGOULMINE, NAZIM & KARAMOOZIAN, AMIR & TAMAZIRT, LOTFI & LEE, BRIAN. Edge Computing and Distributed Intelligence. **Springer Handbook of Internet of Things**, 2020.

RAZA, U; KULKARNI, P. Mahesh Sooriyabandara Low Power Wide Area Networks: An Overview **IEEE Communications Surveys & Tutorials**. 2017, v. 19, Issue: 2

REDAÇÃO AGRISHOW. Produção agrícola conectada com o universo digital: entenda a tendência da Agricultura 4.0. 2016. Avaliable: <http://agrishow.com.br/blog/producao-agricola-conectada-com-o-universo-digital-entendatendencia-da-agricultura-4-0/>.

ROBERTS, Mark, "Serverless Architectures." Available: <https://martinfowler.com/articles/serverless.html>. Accessed on October 2020.

ROCHA FILHO, G. P. *et al*. Enhancing intelligence in traffic management systems to aid in vehicle traffic congestion problems in smart cities. **Ad Hoc Networks**, v. 107, p. 102265, 2020.

ROCHA FILHO, G. P. *et al.* Um sistema de controle neuro-fog para infraestruturas residenciais via objetos inteligentes. In: **Anais Principais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos**. SBC, 2018.

SEMTECH CORPORATION. **LoRaDevices. Smart Agriculture: Real World Solutions**. [S.I] [2019]. Available: <https://info.semtech.com/hubfs/Semtech-UseCaseEBook-SmartAg-2020-web.pdf?hsLang=en-us>. Accessed at: 2020, Aug. 28.

SHAFI, M *et al.* 5G: A tutorial overview of standards, trials, challenges, deployment, and practice. **IEEE Journal on Selected Areas in Communications**, v. 35, n. 6, p. 1201-1221, 2017.

SUBASHINI, S.; VENKATESWARI, R.; MATHIYALAGAN, P. A study on LoRaWAN for wireless sensor networks. In: **Computing, Communication and Signal Processing**. Springer, Singapore, 2019. p. 245-252.

TELECO, Cobertura de Redes 4G no Brasil, [S.I.]. 2020. Available: <https://www.teleco.com.br/4G_cobertura.asp>

TELE.SÍNTESE. Relatório Campo Digital, São Paulo-SP. 2020. Available: <https://www.eventos.momentoeditorial.com.br/wp-content/uploads/2020/09/CampoDigital_completo.pdf>.

TORRES NETO, J. R. *et al.* Towards the Use of Unmanned Aerial Vehicles for Automatic Power Meter Readings. In: **2015 IEEE International Conference on Computer and Information Technology**; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing. IEEE, 2015. p. 379-386.

TORRES NETO, J.; GUIDONI, Daniel Ludovico; VILLAS, Leandro. A new solution to perform automatic meter reading using unmanned aerial vehicle. In: **2014 IEEE 13th international symposium on network computing and applications**. IEEE, 2014. p. 171-174.

TORRES NETO, J. R. *et al*. Exploiting offloading in IoT-based microfog: experiments with face recognition and fall detection. **Wireless Communications and Mobile Computing**, v. 2019, 2019.

TORRES NETO, J. R. *et al.* Performance evaluation of unmanned aerial vehicles in automatic power meter readings. **Ad Hoc Networks**, v. 60, p. 11-25, 2017.

TURNER, V; GANTZ, J F.; REINSEL D; MINTON, S. The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things. IDC white paper sponsored by EMC Corporation, 2014.

VIOLINO, B. "What is PaaS? Platform as a service explained." Available: <https://www.infoworld.com/article/3223434/what-is-paas-software-development-in-the-cloud.html>

VÖLK, F *et al*. Emergency 5G Communication on-the-Move: Concept and field trial of a mobile satellite backhaul for public protection and disaster relief. **International Journal of Satellite Communications and Networking**, 2020.

WOLFERT, S *et al*. Big data in smart farming–a review. **Agricultural Systems**, v. 153, p. 69-80, 2017.