

## Research Article

# Tendencies of Technologies and Platforms in Smart Cities: A State-of-the-Art Review

**Pablo Chamoso , Alfonso González-Briones , Sara Rodríguez, and Juan M. Corchado**

*BISITE Digital Innovation Hub, University of Salamanca. Edificio Multiusos I+D+I, 37007 Salamanca, Spain*

Correspondence should be addressed to Pablo Chamoso; [chamoso@usal.es](mailto:chamoso@usal.es)

Received 22 June 2018; Accepted 2 August 2018; Published 14 August 2018

Academic Editor: Ramón Sanchez

Copyright © 2018 Pablo Chamoso et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Technology is starting to play a key role in cities' urban sustainability plans. This is because new technologies can provide them with robust solutions that are of benefit to citizens. Cities aim to incorporate smart systems in their industrial, infrastructural, educational, and social activities. A Smart City is managed with intelligent technologies which allow improving the quality of the services offered to citizens and make all processes more efficient. However, the Smart City concept is fairly recent. The ideas that it encompasses have not yet been consolidated due to the large number of fields and technologies that fit under this concept. All of this led to confusion about the definition of a Smart City and this is evident in the literature. This article explores the literature that addresses the topic of Smart Cities; a comprehensive analysis of the concept and existing platforms is performed. We gain a clear understanding of the services that a Smart City must provide, the technology it should employ for the development of these services, and the scope that this concept covers. Moreover, the shortcomings and needs of Smart Cities are identified and a model for designing a Smart City architecture is proposed. In addition, three case studies have been proposed: the first is a simulator to study the implementation of various services and technologies, the second case study to manage incidents that occur in a Smart City, and the third case study to monitor the deployment of large-scale sensors in a Smart City.

## 1. Introduction

Over the past five years, the term Smart City (SC) has become a very popular concept around the world; it relates to cities that implement the latest technologies to obtain benefits in a wide range of areas. At present, cities of all sizes are including SC proposals in their urban sustainability programmes. This concept is commonly misrelated to energy efficiency alone. Although energy efficiency is a very important aspect of a SC, the whole idea of a SC is not focused solely on energy or buildings. SC encompasses the entire human ecosystem: it focuses on providing social benefits, economic growth, and creating new opportunities.

Therefore, SC touches upon many different scientific fields as it intends to provide cities with intelligence. Being a fairly new concept, SC requires further research and proposals that will consolidate the technologies and ideas involved in the development of a SC. This will allow resolving all doubts with regard to the definition of the SC concept. Below, we examine the prominent definitions of a SC found in the literature.

The idea of a SC first appeared in 1993 when Singapore city presented itself as an “intelligent city”, in [1]. Between 2000 and 2010, the concept of a “digital city” emerged and was closely related to the idea of a SC, although there are some nuances between the two concepts. In [2] a digital city was defined as an open, complex, and adaptable system, based on a computer network and urban information resources, that make up the city's virtual digital space. In [3], on the other hand, two different definitions for the concept of a digital city were proposed: (i) a city that is being transformed or redirected through digital technology; (ii) a digital representation or reflection of some aspects of a real or imagined city. Digital cities are therefore a precedent for what is meant by SCs.

In 2007, Giffinger et al. [4] published a document that presents one of the first definitions for the term SC as it is understood today. In addition, it already pointed to the ambiguity of this concept. The authors present a SC as a city that prospectively performs its activity in the industrial, educational, citizen participation, and technical infrastructure fields, combining them intelligently to serve its citizens.

However, it is not until 2010 that the interest in SCs began to grow exponentially, with the number of definitions increasing considerably. The IBM company published its own definition of the concept; Harrison et al. [5] specifically defined a SC as an instrumented, interconnected, and intelligent city. Instrumented referred to the collection and integration of real data in real time from the use of sensors, applications, personal devices, and other resources. Interconnected referred to the integration of all such data into a computing platform that provides a set of services. Finally, intelligent referred to the complex elements, such as analytical calculations, modeling, optimization, and visualization of services for better operational decisions.

Another known definition is the one provided by Hancke et al. [6], where the authors look at a SC as a city that operates in a sustainable and intelligent way, thanks to the cohesive integration of all its infrastructure and citizen services and to the use of intelligent devices for monitoring and control. All this allows guaranteeing a SC's sustainability and efficiency.

To understand the importance of SCs, it is necessary to understand the scope of this concept; it addresses multiple areas of citizen's daily life. Dirks et al. [7] proposed a SCs framework in the following areas: transportation, energy, education, healthcare, building, physical infrastructure, food, water, and public safety.

Not only do the definitions vary but also the different approaches to the dimensionality of SCs vary. The oldest of all, presented in [8], indicates that the key dimensions of a SC are information technology (IT) in education, IT in infrastructure, IT in economics, and quality of life. More recently, Rudolf et al. [9] proposed economics, mobility, environment, people, and government/administration as key dimensions. Eger et al. [10] defined that the key aspects of a SC are technology, economic development, employment growth, and increasing the quality of life of its citizens.

However, the inclusion of intelligence in each of the city's subsystems individually is not sufficient to create a SC. A SC cannot be considered intelligent unless it functions as an organic ensemble [11], so that it matches the definition presented in [6].

The final beneficiary in the vast majority of SC definitions is the citizen. A recent SCs review [12] concluded that their main focus is "people first and foremost".

The aim of this paper is to propose a model for the design of SC architectures. The main novelty is that it offers a series of services that can be reused by other cities and allows the creation of a catalog of services that cities can "clone", taking advantage of development, without having to program common functionalities.

We first look at how to provide the architecture with specific and generic characteristics. This will allow easily making changes in the system, for example, in (i) message sending; (ii) storage or data collection modules, without affecting the architecture of the SC, as well as the technological support capable of storing and processing all information in a connected and distributed way; or (iii) the services that must provide different areas that benefit all its citizens to promote the adoption of this concept by local governments. In this sense, we describe the technical solutions to adopt (design

of Internet of Things (IoT) services, cloud storage system, messaging protocols, and related technologies), discussing the appropriate SC environment.

The rest of the article is structured as follows. Section 2 presents a review of the different services that a SC may offer. Section 3 reviews proposals, developments, and technologies applied in this area. Moreover, we study the proposals which use performance analysis in these systems. Section 4 outlines our architecture proposal and the technologies that have been used in the development stage. Section 5.1 presents the evaluation of the proposed architecture and three case studies are proposed. The Smart City platform as a simulator for the implementation of various services and technologies, incident management, and large-scale sensor deployment. Finally, Section 6 presents the conclusions drawn from this work.

## 2. Smart City Architecture Conceptualization

In this section a specific SC architecture is conceptualized; it considers all the functionalities that such architecture must incorporate in order to make a city a SC. It studies the basic services that a SC architecture must provide in the different areas and the technologies that must be available (those that are capable of obtaining data through any sensor, of transmitting and communicating data between devices and the system, of transforming data into useful information, and of information analysis for decision-making). Existing platforms are also examined in this respect to see to what extent they comply with these requirements. This will allow identifying the defects of these platforms and to indicate how they can be improved to achieve the degree of intelligence required in a SC.

Given the different definitions of the SC concept and its scope, it is necessary to establish the types of services that SCs should provide to their citizens. This will help clarify their objectives and to understand the requirements that a city must comply with in order to be considered a SC. This knowledge will allow evaluating different aspects of a city and verify whether they can be classified as smart.

Despite the lack of a fully shared and globally accepted definition, it is possible to describe the most common characteristics of a SC, what services can improve its degree or level of "intelligence", and its most important aspects. The definition of the concept of "intelligence" must be more specific in order to be able to evaluate intelligence. Different authors have agreed on three main aspects when defining intelligence [13, 14]:

- (i) Effectiveness refers to the ability of a city to effectively provide public and private services, such as citizens (students, workers, and elderly people), companies, or nonprofit organizations. In other words, a SC is not in itself intelligent, but by the public value it creates for people
- (ii) Environmental benefits refers to improving the quality of the environment in large cities. One of the main pillars of SC is to prevent environmental degradation. To this end, it is necessary to conduct major studies

regarding energy consumption, air and water pollution, or traffic regulation. Therefore, a SC must focus on these solutions to preserve environmental quality

- (iii) Innovation means that a SC must apply cutting-edge technologies to improve the quality of its main components, so that better services are provided. Technology is therefore a central aspect of a city's intelligence

Thus, the intelligence of a city's components can improve if they are transformed into effective and innovative tools that are not harmful for the environment. This provides public value. However, these three aspects are not enough to increase public value. As proposed in [15], creating public value must be the ultimate goal of a smarter city and requires that all projects and initiatives be targeted at citizens. The concept of public value is complex and includes several dimensions [16]:

- (i) Creating both economic and social values, which are difficult to unite and sometimes enter into conflict with each other
- (ii) Creating value for different stakeholders, which may have different expectations that are not always compatible with each other
- (iii) Creating value regarding the different dimensions of life in the city, which also implies understanding what the real needs and priorities are

Therefore, in order to create public value in a SC, it is necessary to pool a large set of variables through a well-defined overall framework. The framework must be capable of providing for the needs, expectations, and perceptions of citizens with regard to what they expect from a SC in their daily lives. In fact, in most cases of SCs, the benefits are not defined, measured, and reported, so that although they produce improvements in the daily lives of citizens, they are often not informed in advance [17]. To solve this inconvenience, it is necessary that the services that this general framework offers are published.

SC technology can be applied to a wide range of aspects of daily city life. Thus, it is necessary to identify the services that are frequently offered to society in a structured way, classifying them according to their domain. In existing works such as the one published in [18], a classification according to domain is presented:

- (i) Natural resources and energy:

- (a) smart grids: services that enhance the experience from the use of electricity grids that take into account consumer habits, sustainable, affordable and secure distribution, and affordable and safe use of electricity grids [19, 20]
- (b) lighting: street lighting with streetlights that offer features such as air pollution control or Wi-Fi connectivity that allow you to incorporate software for reducing consumption based on a variety of criteria [21]

- (c) renewable energies: exploitation of natural resources that are regenerative or inexhaustible, such as heat, water, or air [22, 23]
- (d) waste management: collection, recycling, and deposit of waste using methods that prevent negative effects on the environment or inadequate waste management [24, 25]
- (e) water management: analysis and management of the quantity and quality of water used in agriculture, municipal, or industrial purposes [26]
- (f) food and agriculture: such as the use of wireless sensor networks for harvest management and knowledge of the conditions in which plants grow [27]

- (ii) Transport and mobility:

- (a) city logistics: the improvement of the logistics of cities by efficiently integrating business needs with traffic conditions and geographical and environmental issues [28]
- (b) mobility information: distribution and use of dynamically selected information, both prior to the completion of the journey and during the journey, with the aim of improving traffic and transport efficiency, as well as ensuring high quality travel experience [29]
- (c) mobility of people: use of different innovative and sustainable ways to provide transport to people in cities, such as the development of public transport modes and green-powered vehicles, all supported by advanced technologies and the proactive behavior of citizens [30]
- (d) services exposing district information models: which are domain-specific models that include data models of Building Information Models (BIM), Geographic Information Systems (GIS), and System Information Models (SIM) [31, 32]

- (iii) Smart building:

- (a) facilities management: cleaning and maintenance of urban facilities [33]
- (b) construction services: use of services such as power grids, lifts, fire safety systems, and telecommunications [34, 35]
- (c) housing quality: aspects related to the quality of life in residential buildings, such as comfort, lighting, heating, ventilation, and air conditioning. This category includes everything related to increasing the level of satisfaction of people in their home life [36]

- (iv) Daily life:

- (a) entertainment: ways to stimulate tourism, provide information on leisure events, proposals for free time and nightlife [37]

- (b) hospitality: the ability of a city to accommodate students, tourists, and other nonresidents by offering appropriate solutions to their special needs [38]
  - (c) pollution control: control of emissions and water waste from the use of different devices. Decision-making to improve the air, water, and environmental quality in general [39]
  - (d) public security: protection of citizens and their belongings based on the active involvement of public organizations, the police, and even citizens themselves. Collection and monitoring of information for crime prevention [40, 41]
  - (e) health: prevention, diagnosis, and treatment of diseases supported by information and communication technologies [42, 43]
  - (f) welfare and social inclusion: improving the quality of life by stimulating social learning and participation. Certain groups of citizens require special attention, such as the elderly and persons with disabilities [44]
  - (g) culture: dissemination of information on cultural activities and motivating citizens to get involved in them
  - (h) management of public spaces: care, maintenance, and active management of public spaces to improve the attractiveness of the city and solutions that provide visitors with information in places of tourist attraction in a city [45]
- (v) Government:
- (a) e-governance: digitization of public administration through the management of documents and formalities using digital tools, in order to optimize work and provide fast and new services to citizens [46]
  - (b) E-democracy: use of information and communication systems for the management of votes [47]
  - (c) transparency: allowing citizens to access official documents in a simple way and decreasing the chances of abuse of authorities who may use the system for their own interests or withhold relevant information from authorities [48]
- (vi) Economy and society:
- (a) innovation and entrepreneurship: measures to promote innovation systems and urban entrepreneurship, for example, by using incubators [49]
  - (b) cultural heritage management: the use of digital systems can provide visitors of cultural heritage sites with new experiences. Asset information management systems are used for performing maintenance in historic buildings [50]
  - (c) digital education: extensive use of methodologies and digital tools in schools [51, 52]
  - (d) human capital management: policies that improve human capital investments and attract and retain talent, avoiding the flight of human capital, known popularly as the brain drain of [53, 54]

Although there are more classifications, the one presented above is one of the most complete in terms of domains and subdomains where a SC may be able to provide services. Given the great diversity of existing services, which is increased by the evolution of technology almost every day, the proposed architecture must be used as a basis for deploying new services, capable of supporting the heterogeneous set of technological solutions.

### 3. Technological Support

This section describes different technologies that have been used in SC-related developments. These technologies can be used at different levels to provide optimal solutions to specific problems. Below we present platforms that support SCs and are based on the presented technologies.

*3.1. Heterogeneous Device Networks.* In the field of SCs, IoT is presented as a tool that provides a series of specific services which give low level support to different applications offered to citizens [55].

The adoption of the IoT concept has grown substantially since its appearance. This is due to technological advances, such as wireless communications and the great standardization of low-power communication protocols, which make it possible to obtain sensor data almost everywhere and at any time.

Furthermore, the main objective of IoT is to interconnect all the things and to guarantee that all those things are intelligent. In [56], IoT is presented as a paradigm that allows things to communicate in people's environment, through the Internet as if they were computers. On the other hand, [57] is defined as things or objects such as devices, sensors, actuators, and smartphones that are capable of interacting with each other and cooperating with intelligent components to achieve common goals.

What seems obvious is that when talking about certain services that a SC must offer, the concept of IoT plays a relevant role, especially in the collection of information from the environment through sensors, as well as in the execution of certain actions through actuators.

Another concept that is closely related to IoT is M2M (machine to machine). M2M is especially used in SCs, where the connection is not only between technology and people but also between the machines and where any object can become part of the network. Such a large number of elements can be connected, so some kind of mechanism is needed to support both the volume of connections and the type of communication. This is the purpose of the M2M concept which manages data in a stable and reliable way, besides unifying the different standards [58]. Thus, M2M allows implementing

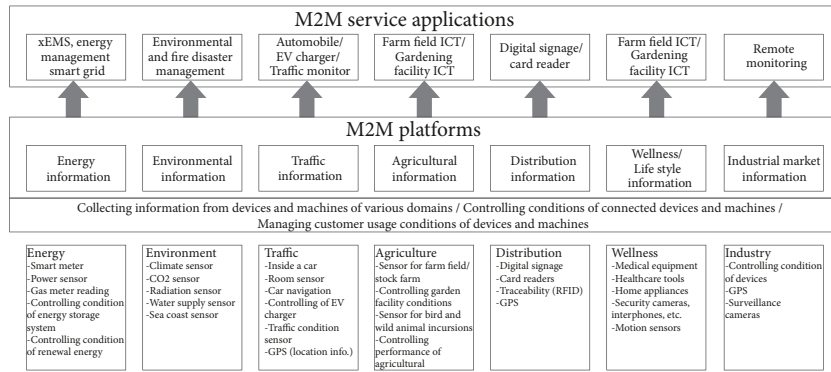


FIGURE 1: M2M services schema.

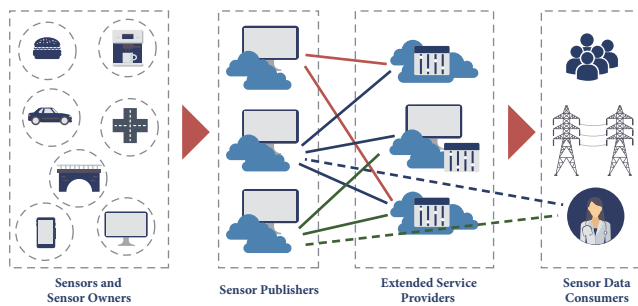


FIGURE 2: Sensing as a service model (SenaaS).

and combining services, despite the great amount of possible connections. The services that M2M can offer are presented in [59] and a summary can be seen in Figure 1.

Among the technologies that can be part of an M2M system, forming heterogeneous sensor networks are ZigBee, Bluetooth, Wi-Fi, WiMax, PLC, GSM/GPRS, 6LoWPAN, EnOcean, or Z-Wave allowing the connection of any object regardless of its nature integrating it into the platform. However, the number of technologies is steadily increasing [60].

3.2. *Sensing as a Service.* With the expansion of IoT, the number of objects connected to Internet will increase and more sensors will be available for use [27]. Currently, the use of some of the sensors is restricted to the objects in which they are embedded and are operated by platforms that follow the sensing as a service (SenaaS) [27] model. This model can be used even if the sensor is not a physical object and it is not necessary for it to have an associated hardware; it may take a software system as a sensor that is capable of obtaining and providing data. The SenaaS model can bring benefit to the public from the collected data. This is because the owner of a sensor, whether a private or a public person or organization, can make the sensor data available to the public [61].

As can be seen in Figure 2, the SenaaS model consists of four conceptual layers [27]:

- (i) Sensor and sensor property layer: a sensor represents the concept of an entity (software or hardware) that

detects or measures physical property, records it, or transmits it properly. The sensors can be classified into four categories:

- (a) personal or domestic: all entities that do not belong to public or private organizations will be equipped with sensors in the future
- (b) private organizations or places: entities belonging to a private organization or to the above category should have the right to make the decision to publish these sensors or not
- (c) organizations or public places: entities that belong to public places and public infrastructures are found in this category. Depending on government policy, sensor data will or will not be published
- (d) providers of commercial sensor data: which deploy and manage sensors on public or private property depending on strategy and demand, taking into account legal terms. Permission from the owner is required for the deployment of both public or private sensors and it is possible that an economic agreement be made between both parties. The source of income for these suppliers is based on the publication of sensor data. The owner of the sensor decides whether or not to publish a sensor in the cloud. If you decide to publish a sensor you will need to register with a sensor publisher. Sensor owners may define restrictions, conditions, and some benefit from sensor publishers, whether economic or not

- (ii) Sensor publisher layer: detection of available sensors, communication with sensor owners, and obtaining permission to publish services in the cloud are the main functions of a sensor publisher. Sensor publishers depend entirely on payments to sensor owners and consumers of these data
- (iii) Extended service provider layer: provides added value to sensor publishers, providing new services for data consumers. Sensor publishers themselves can also be

service providers. Service providers can communicate with multiple sensor publishers by acquiring data to help the data consumer. Added value should be included in the payments

- (iv) Sensor data consumer layer: those applications that use the data provided by sensor publishers or service providers. All communications with the consumers of data will be between the publishers of sensors or with the suppliers of the sensor.

In other words, this model allows managing all the entities of the city that wish to be incorporated as data sources (sensors), from the data collection itself to its use by applications that provide the final services to citizens.

**3.3. Cloud Computing.** The infrastructure of a SC architecture must allow the communication of its information to the connected objects or entities, allowing obtaining information collected by the different heterogeneous entities. Since the number of entities connected to the infrastructure may vary as well as the volume of the generated data, the computing infrastructure must be able to adapt to storage, volume, and speed of response needs.

To provide a solution to needs where these factors are unknown, the CC (Cloud Computing) technology appears [62]. However, the concept of CC is quite broad; it refers both to the applications that are offered in the form of services over the Internet and to hardware and software systems that provide those services in data centers. The main advantage for the user of this type of solution is that CC makes it easy to adapt computational resources to respond to peak demand, without having to have an infrastructure to provide them, transforming capital expenditures into operational costs [63].

**3.4. Information Persistence.** We have looked at how a distributed infrastructure can be obtained. This infrastructure must be able to adapt the resources required to provide different services in real time, regardless of their computational requirements. Now it is necessary to determine how information should be processed. In this regard, much of the information may be discarded because it is considered irrelevant or even out of date after some time has passed since that information is no longer relevant. New technologies have emerged in recent years to provide a solution to the problems involved in handling large amounts of data. NoSQL databases are one of the solutions to information persistence [64].

NoSQL databases, in comparison to relational databases, offer another set of very varied and more storage-oriented features for large volumes of information [65]. There are different models of NoSQL databases that offer different features and depending on the need will be more or less useful for the management, exploitation, and relationship of these volumes of data. Currently there are different models in which you can classify the NoSQL databases [66]:

- (i) Key-value databases: these are the simplest NoSQL databases; they store a tuple containing a key and its value that is completely opaque to the database. This type of database is used when data is accessed by

means of a primary key and can perform simple operations such as storing, querying, and deleting the tuple. The main advantage of these databases is that their performance is very high despite their simplicity and they offer very simple scalability

- (ii) Column-oriented databases: column-oriented databases store data in columnar families, which would be comparable to a table in a relational database. Each of these column families contains one row with many columns associated with a key. In addition, rows do not contain the same columns. This type of database can resemble a slightly more sophisticated key type value, in which the value is no longer opaque within the database and has a defined form. The biggest advantage of this type of database is the ability to analyze large-scale data in real time
- (iii) Graph-oriented databases: these database models are based on graph theory; that is, they use nodes and their relationships to represent a set of stored data. The nodes store the entities while the different relationships are modeled by the arcs that join them; they may be strong or weak depending on whether they are directly connected or not. It should be noted that networks can be directed or not; thus patterns between nodes can be found. Currently, this type of database is mainly used in recursive searches with many levels, recommendation engines, or social networks.
- (iv) Document-oriented databases: document-oriented databases are based on storing documents rather than a key-value pair. A document is a tree-shaped data structure that can consist of maps, collections, and scalar values that can be stored in XML-type text files (Extensible Markup Language, translated as “extensible markup language”), JSON (JavaScript Object Notation), BSON (Binary JSON), and so on. It is not necessary for documents to be identical; each of them can store different information, although they should be similar to each other. In this type of database the documents can be examined to obtain characteristics by consulting their properties. The main advantage of these systems is the high horizontal scalability of the data, making it a very effective database type for storing large volumes of data.

Therefore, all kinds of architectures are available and we can choose those that best adapt to the needs of the data model. Each architecture offers a different structure to organize the information; however, these databases allow decoupling the way information is organized from the structure of the information itself. Therefore, in cases where the structure of the information received may vary (as is the case of this paper), the use of NoSQL database models is the most appropriate.

**3.5. Information Analysis.** The massive generation of digital data in current general networks and in particular cities has opened the door to new technological trends that seek added value through the analysis of these digital data.

One of the main reasons for the constant generation of digital data today is that mobile devices are becoming increasingly similar to desktop computers, with high-speed Internet connection and geolocation systems.

All this means that all the actions made on mobile phones, computers, and connected objects, practically every element of a city, leaves a trace. All of this generates large volume of information that is constantly increasing and accumulating while waiting to be analyzed.

In addition, some of the new applications that emerged recently have had a great impact on society, such as open data sources, social networks, the emergence of IoT and, above all, SCs; they have favoured the generation and availability of information on day-to-day aspects of everyday life, but which were unthinkable years ago.

The analysis of this information has opened the door to new opportunities, which can bring great benefits: cost savings, increased profits, job creation, etc. Analyses can therefore help improve the management of businesses or cities, allowing making predictions in a range of fields or detecting cause-and-effect relationships between data that could not be understood up until now. Therefore, Big Data is a crosscut that many areas of knowledge must consider.

Large amounts of information are generated nowadays (as in the examples provided in [67], where it is stated that CERN, the European Organization for Nuclear Research, in Switzerland, generates 1PB of data per day and the world's largest social network, Facebook, generates 500TB per day); the analysis of this data requires the use of computer sets with large processing capacities, running highly optimized software.

However, when conducting analysis a certain level of intuition is also required; this is where the human factor or artificial intelligence plays an important role. Certain filters are capable of reducing the volume of data for analysis by eliminating those that are clearly of no interest.

Researchers at the Computer Science and Artificial Intelligence Laboratory, at the Massachusetts Institute of Technology, have recently proposed a new system called Data Science Machine [68], a software based on artificial intelligence that is capable of finding patterns in the relationships between data and of making predictions from these patterns, better than most human experts and in a much shorter period of time. One of the fundamental factors for extracting knowledge autonomously from a large dataset is the identification of the variables that are necessary to acquire the knowledge sought for.

But there are more problems inherent in data that complicate the process of acquiring knowledge, such as the deconstructing of data from heterogeneous open and closed sources, the wide variety of formats (such as text, images, or videos), the quality of such data, or the veracity of the data. In this regard, different scientists indicate that when analyzing the information, there are five complications that are important to bear in mind, the so-called 5Vs model of Big Data proposed in [69]:

- (i) Volume: this is the increase in the volume of data that must be analyzed nowadays. Terabytes to petabytes can be handled

- (ii) Speed: there is a noticeable increase in the speed at which data is generated and distributed. Generated data is streamed; i.e., it is generated, distributed, and consumed in real time
- (iii) Variety: significant increase in the heterogeneity of data sources due to factors such as the large number of data sources, the need to process different sources at different levels of structure, and the diversity of distribution formats; we used to know their structure and how to process them. All this has changed as we now have unstructured data such as texts or images
- (iv) Accuracy: increased uncertainty about the quality of the available data; incorrect data leads to incorrect conclusions. It is important to establish the veracity of the data sources with which one works in order to avoid losses
- (v) Value: strategic decisions for improvement should be taken. To this end, the value of a data source has to be determined. Determining it a priori, without knowing the sources, is highly complicated

Big Data thus emerged as a response to these five complications that previously existing processing and analysis technologies did not face.

In the early definitions of the Big Data concept, such as the one proposed by Manyka et al. [70], Big Data is about datasets with a size that is beyond the capabilities of typical databases to be captured, stored, managed, and analyzed.

According to the definition of IBM, Big Data is a way of dealing with the processing or analysis of large volumes of information that by their unstructured nature cannot be analyzed in an acceptable time using traditional BI (Business Intelligence) processes and tools [71].

The definition has been broadened and the Big Data concept is defined in [72] according to three main features: (i) the data itself; (ii) the analysis of the data; (iii) the presentation of the results of the analysis.

Thus, what began with only one large set of data, now encompasses that data, its analysis, and interpretation and presentation. Thus, Big Data's goal is to extract useful knowledge from the data. It is not something trivial as has been mentioned, since it presents problems such as the availability of data in streaming, its large size, the low level of structure, the variety of formats, or its quality.

Until the Big Data appeared, the data came from traditional information systems, but with these new analysis techniques, new types of data are available that can be analyzed: documents, digitized images, data from payment transactions or currency exchange, among others. This internal data can be used in conjunction with external data, for example, from open data sources or social networks.

However, there are different problems and difficulties in implementing Big Data. According to a survey conducted by "The Data Warehouse Institute" in 2013, presented at [73], only 12 percent of Big Data companies were found to be successful. Sixty-four percent of companies reported moderate success, while 24 percent of companies reported having failed with Big Data.

The different problems that may arise are closely linked to the five complications mentioned above, which must always be taken into account. Some of the reasons for this failure are (i) the immense complexity of integrating data sources, (ii) poor data quality, (iii) real-time management of the generated data, (iv) lack of staff with the right skills, and (v) the choice of an incorrect architecture.

Therefore, it is extremely important to take into account these aspects: the choice of the personnel involved in the project, the architectural basis of the project (database type, location of the infrastructure, use of third-party services, etc.), the evaluation of the data sources to be incorporated (it is necessary to verify the quality before the integration and after the integration process), coexistence with the company's data (partner cost, maintenance, etc.), integration of the company's data, and so on.

Massive data analysis management techniques are adaptations of high-volume data analysis techniques, optimized to improve your execution time performance.

Common techniques include the following: A/B testing, association rules [74], classifiers item [75], data mining, self-organized maps, k-means, decision trees, support vector machines, genetic algorithms, predictive learning, machine learning [76], and Linear regression [77].

In short, the most commonly used techniques in data analysis since years ago have been incorporated into solutions capable of handling large datasets in order to take full advantage of today's large processing capacity, for example, thanks to distributed environments. The use of these techniques provides a level of intelligence capable of extracting advanced knowledge from the information held in a city.

*3.6. Existing Platforms.* Throughout this section the different existing solutions created to support SC-oriented services are presented. The vast majority of them use technologies such as those previously presented.

When describing the different existing platforms, it is important to differentiate between platforms oriented to the use only on the part of a set determined by city managers (specific platforms) and generic platforms oriented to provide services in a general way, which can be used by any city that wishes to do so. However, they can also be classified into public access platforms or private platforms, depending on whether any developer can use the platform to use its services or not, in which case the business model involves selling the private platform to different municipalities that decide to use its services when transforming their city into a SC.

- (i) **Sentilo** is the piece of architecture that will isolate the applications that are developed to exploit the information “generated by the city” and the layer of sensors deployed across the city to collect and broadcast this information. The project began in 2012 with the Barcelona City Council and was used to put Barcelona at the forefront of SCs [78]. Nowadays, although it originated exclusively oriented towards the city of Barcelona, it has evolved and is being used by other cities such as Terrassa or Reus. It is also an open-source software, which provides the source code

through its own repository. It is therefore a generic and public system

- (ii) **SmartSantander** project proposes an urban-scale experimental research facility on which different applications are supported and typical SC services are deployed. The project is not intended to be limited solely to the city of Santander but is intended to extend to other cities such as Belgrade, Guildford, or Lübeck. The project is private; however, it offers a free information access system for developers to use to make new applications. It is also adaptable to new cities, although privately, so it is considered specific to a closed group of certain cities
- (iii) **IBM Intelligent Operation Center** is a private platform, owned by the IBM company, which is located in different cities around the world, such as Rio de Janeiro. It offers an environment that provides different default tools but can be customized on demand. It is therefore a private and specific system, since it requires adaptation and maintenance by the owner company [79]
- (iv) **CitySDK** project aims to provide a programming structure for deploying SCs systems, which has been tested in 8 cities in Europe: Amsterdam, Barcelona, Helsinki, Istanbul, Lamia, Lisbon, Manchester, and Rome, involving more than 5 private companies in collaboration with 5 universities. They allow the integration of new cities, but only from the use of the API they propose. It is therefore a private and general system
- (v) **Open Cities** is a platform that allows users to use (read and write access) the data stored on it in order to be used by developers to offer services in cities. It is a private system of free use of a generic nature (not oriented to specific cities)
- (vi) **i-SCOPE** is a platform that provides three types of services to SCs [80]: (i) improving the inclusion and mobility of citizens with routing systems and signalling of barriers in the city; (ii) optimizing energy consumption; (iii) environmental control. However, it is a private project, already completed and specific, of which the cities in which it is implemented are unknown.
- (vii) **People** is a platform that provides services, generally open-source, for the community to use and share those they develop, always oriented towards SCs. For example, they have services in cities such as Bilbao, Bremen, or Thermi. It is a public project but can only be used in specific environments.
- (viii) **IoT Open platforms** is an initiative that provides a set of libraries, technical documentation, web services, and protocols in an open way for use by the entire developer community. Among the tools they offer is VITAL-OS Smart City Platform, for example, which provides a set of visual tools to develop applications with reduced cost and effort. They use specific city-specific datasets.



TABLE 1: Existing SC-oriented platforms.

Name	+1 city	Open source	SC services	Open data
Sentilo	Yes	Yes	Not open source	Yes
SmartSantander	Yes	No	-	Yes
IBM IOC	Yes	No	-	No
CitySDK	Yes	No	-	Yes
Open Cities	No	Yes	No	Yes
i-SCOPE	-	No	-	No
People	Limited	Yes	Yes	No
IoT Open platforms	-	Yes	Only IoT-based	Yes

In conclusion, as we can observe in Table 1, it can be seen that there are numerous service-oriented platforms that can be delivered in SCs; however, there are few solutions that have explicitly emerged directly oriented to support SCs. Most of them are private, although they allow a free access or use (in exchange for giving the information) and only Sentilo allows developers to download the code to be able to replicate in other cities freely, although it does not provide any high-level service by default.

There is therefore a lack of a generic and open platform that includes a set of high-level services that developers can use to take greater advantage of the available information and knowledge of the city that is encapsulated in the platform.

#### 4. Proposed Architecture

This section describes the proposed global architecture for managing the services offered in a SC, regardless of the characteristics of the city. This is so, since one of the objectives is that any city can deploy its own platform following the model of the proposed architecture. To achieve this, architecture has to be self-adaptive. This is the only way to ensure that, regardless of the size of the city, as many services can be deployed as needed in any of the areas presented in Section 2.

The design of the architecture, therefore, should make it possible to fill those shortcomings of the current platforms. The design of this architecture is focused on the fact that it can be used in more than one city, can be open-source platform, and provides high-level open data SC services.

The proposed architecture is based on two main parts: (i) a set of existing technological solutions, the combination of which is capable of providing the services necessary to manage information autonomously (obtaining) and (ii) a distributed system adaptable to the needs of the city that provides the basic functionalities and tools required by any SC (such as representation of high-level information, analysis or services). More specifically, the architecture has been structured in 4 main layers, as we can see in Figure 3:

- (i) Physical support layer is the physical layer, the lowest level, which by far is the lowest level of the physical layer. On the one hand, it provides the necessary processing capacity and, on the other hand, allows device connectivity by making use of different standards of communication

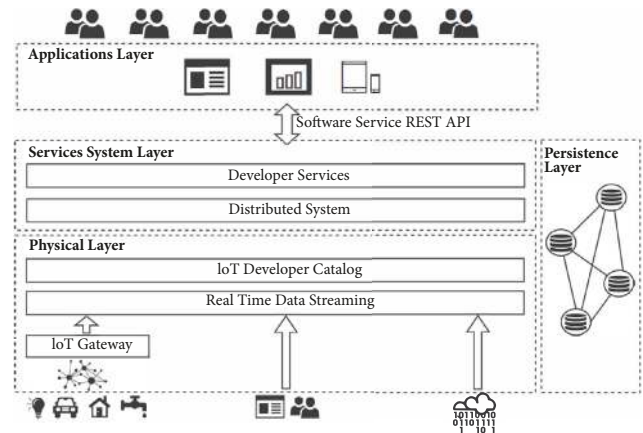


FIGURE 3: Proposed architecture.

- (ii) Persistence layer is the layer that allows representation and storage of the information you need to store in the system
- (iii) Services system layer, together with the following layer, provides the which differentiates the present architecture for SCs from the existing ones. In this layer, agents are deployed and structured to provide services
- (iv) Application layer is the most important layer of the architecture, in which the different applications that developers create and offer to all citizens are deployed in an open or private way. All applications must be provided when acquiring the system, by default it offers an app for user control and permissions, a dashboard for monitoring adaptive sensors, and a tool for sending and resolving incidents, so that anyone who purchases the software can use them and even extend the functionality according to the needs of the city in which the platform is deployed.

To better understand how the architecture works, the layers that make it up are presented separately, following the logic that the lower layers provide lower level support to the upper layers, so that they can offer all of their functionality as the highest level of service.

*4.1. Physical Support Layer.* The distributed infrastructure of the architecture offers a support that allows to manage multiple resources and instantiate services guaranteeing their availability. In many SC scenarios (status of lighting, containers, traffic lights, or traffic lights) it is essential to transmit information in real time, so streaming techniques must be used. This streaming concept is composed of three main components:

- (i) **Producer:** this is a software based system that connects to the data source. Producers publish possible data to a streaming system after obtaining it from the source, transforming it into the desired format and optionally filtering, grouping, or enriching it
- (ii) **The streaming system:** it obtains the data published by producers, stores them, and delivers them reliably to consumers
- (iii) **Consumer:** this component consists of streaming processing engines that subscribe to the data streaming itself provides and manipulate analyze that data to look for alerts and additional knowledge

The integration of the information coming from each of the possible existing producers, city infrastructures, citizens and applications, and open data and external databases is done in a different way. For information associated with city infrastructures and citizens and applications, REST-based technology is used, while to obtain information from external databases and other data warehouses, connectors have been used to communicate directly and obtain the information.

Most sensors and actuators thanks to the use of a gateway, which acts as a gateway between the information received by the sensor and the platform, can communicate using HTTPS connections, adapting it to the integration mechanism of the REST-based platform. However, it is proposed to use Apache Kafka to provide computing capacity for real-time data streaming because it is distributed, scalable, fast, and robust, more than desirable features in the field of SCs [81, 82]. Compared to traditional queuing systems such as RabbitMQ or ActiveMQ, Kafka stands out for its simple scalability and higher speed. Due to the computing deficiency that may exist when deploying a Kafka client on the infrastructure itself, Kafka REST API appears, which allows the infrastructure to easily publish and subscribe to Kafka topics, making the architecture much more agile. Any device with HTTP/HTTPS capability can communicate directly with Kafka. Kafka is able to handle more than 1,000 messages per second.

Communication with producers from databases, data warehouses, or distributed file systems is in many cases not freely modifiable and the only way to access them is through the use of pre-designed connectors that allow connection with Kafka of legacy data stores, such as databases or data warehouses and file systems, such as HDFS (Hadoop Distributed File System). This connection eliminates the need to consume a custom intermediate application. This is achieved using the Kafka Connect5 tool, which provides a reliable connection to the most common data warehouses and integrates them into Kafka very quickly for large volumes of data. Through

this module you can connect platforms that support different SCs following the present architecture, achieving a horizontal scalability of the system.

Kafka needs the use of an orchestrator (coordinator for distributed systems) and ZooKeeper has been chosen for several reasons. In addition to meeting all of Kafka's needs, it is selected for its easy integration with Apache Software Foundation technology. As an orchestrator, it is a recognition service to meet new Kafka cluster members so that the tasks (load) are distributed among them.

It has been detailed how sensors can be integrated into the architecture, making use of the Kafka REST API, but if you want to define this API to add new infrastructure, or you want to access an existing infrastructure and do not know how to do it, you need to define a method to use the information.

To this end, the Hypercat specification, proposed in [83], is included as a lightweight hypermedia catalog format based on JSON to consult and represent web-based resource catalogs (URIs). It has a clear orientation to IoT technology support by SCs and industries, to the extent that it defines itself as the specification that allows IoT customers to discover information about IoT assets over the web, while allowing developers to create applications that run on multiple systems or servers. Its main objectives are safety and interoperability of the IoT. Catalog resources are described by a list of triple declarations similar to the RDF (Resource Description Framework) to provide information on the format and semantics of the URL. The platform must serve the catalog through an accessible URL, a web address that returns the catalog in a JSON document with all its elements. The following is an example of a catalog.

*4.2. Persistence Layer.* The persistence layer should be seen as a cross-sectional level of the architecture, as it supports all other existing layers. Computationally supported by Layer 1, it supports the persistence of information from both that same layer and the top layers it must store.

The great disadvantage when dealing with systems, where information can be oriented to the large number of services that can be found in a SC and that the sources for each of them vary, is that the information will not follow a uniform structure and the system of persistence itself is not able to adapt to this information with traditional methodologies. In addition, there is a large volume of data.

Mainly due to these reasons, NoSQL-type databases are used [66], having also decided to use a document-oriented database architecture. This is mainly due to the fact that, in this type of database, documents do not need to be identical; each of them can store a different type of information, so the information can follow a less strict structure than in the other databases NoSQL.

In addition, in this type of databases documents can be examined to obtain characteristics through queries to their properties, so that the highest level services themselves can easily discover which structure they should adapt to consult the information. Another of the main advantages of these systems is the high horizontal scalability of the data, being a very effective type of database to store large volumes of data, as is the present case.

Going deeper into the existing options of document-oriented NoSQL databases, the most commonly used are currently MongoDB and Apache CouchDB, whose differences are not highly significant.

*4.3. Services System Layer.* This layer of architecture is very important as it is where the main differentiating feature resides in the design of the platforms for the management of SCs proposed so far. This layer is intended to offer both generic functionalities, which any SC-oriented platform should have, as very specific, where the expert knowledge that is produced can be reused, as well as the mechanisms of reasoning or AI used for this purpose.

The architecture must provide high-level services thanks to the analysis and processing processes of the data captured in the SC. This should provide the ability to regulate lighting status, container filling status, traffic light operation status, or traffic management, among others.

In order to be able to work with the vast amount of data coming from the physical layer, as mentioned above, it is necessary to develop a distributed system that uses the Apache Spark framework. This development environment makes it possible to use the Scala language for supervised automatic learning tasks with the scalable MLlib and ML packages from Spark in a Spark cluster.

- (i) Binary classification: predictions of whether a traffic jam will occur or not (1/0) on a given motorway
- (ii) Regression problem: prediction of the duration of a traffic jam (in minutes) after a traffic jam has occurred

The distributed system must allow the modeling process to be carried out. Training and evaluations must be conducted on relevant test datasets and precision metrics. Data from the producers, city infrastructures, citizens and applications, open data, and external databases are imported in the form of resilient distributed datasets (RDDs) and data frames are defined according to the scheme. This provides obtaining data concerning permissions (number of cars per lane, time, and event data at locations in the area) and filtering out damaged or nonrelevant information. Once this filtering process has been completed, Spark provides a better understanding of the data through exploration and visualization using third-party applications.

The distributed system using Spark allows you to use applications such as Jupyter or Apache Zepellin to create and share documents that contain live code, equations, visualizations, and narrative text. Uses include data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

Other applications that can be part of the architecture's functionality to provide high-level basic services could be Microstrategy or Tableau applications, which focus on business analysis and intelligence.

*4.4. Application Layer.* This last layer of applications is part of the architecture; however, it is not part of the platform, but rather makes use of it. It refers to the applications that integrate using the platform as a support, making use of the previous layer.

These applications depend directly on the developers that integrate them; however, certain generic applications common to all SCs are developed, which can be used by developers to duplicate and modify their functionality to adapt it to their needs. These generic applications are

- (i) Simulation environment: it allows to evaluate algorithms, simulate applications, and demonstrate their operation before being implemented in a real environment
- (ii) Incident management system: it allows administrations to keep a record of incidents and notifications and display them in real time, for example, on maps of the city itself
- (iii) Infrastructure monitoring system: it allows visualizing information in real time of the sensors displayed in the city, determining what information is to be obtained and its form of visualization, as well as registering the infrastructure through the web interface.

These common services allow them to be freely duplicated and modified by developers without the need to develop similar functionality again. In addition, work has been done on SC-oriented services that have required a large research component and have been integrated into the platform.

To summarize, the main novelty of the proposed system is that it allows new cities that want to use this architecture to reuse a series of services by easily adapting them to their characteristics and needs. The modular design of the architecture makes this adaptation easy. The modular design allows the technologies used to be decoupled so that new technologies can be included without involving a modification in the rest of the system.

## 5. Case Studies

This section describes three case studies that have been carried out using the proposed platform as a technological basis in order to verify its correct functioning. The three case studies have been chosen because they deal with "generic" cases that any Smart City could include, so that it can reuse the code to adapt it to its needs.

More specifically, the first case study is a simulation scenario that allows testing the connectivity of different sensors through different communication technologies, as well as the algorithms that use their information. The second case study presents a system for managing incidents that users can report from their mobile devices and the different authorities in charge of solving them or a centralised authority will visualize these incidents on a dashboard to solve them more quickly, offering a better service to citizens. Finally, the third case study presents a web platform that allows the management of sensors (registrations, cancellations, and modifications) deployed.

*5.1. Simulation Environment.* To enable the validation of the proposal presented in this paper, a software and hardware simulation environment has been developed, as shown in Figure 4. Simulation environments have played an important

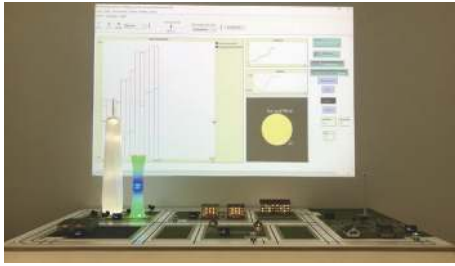


FIGURE 4: Simulation environment (software and hardware).

role in SCs for years [84, 85]. In our case, the simulation software environment has been developed using NetLogo, a multiagent programmable modeling environment. This allows for collaborative SCs simulations and participation under the SaaS methodology. The software environment together with the Services System Layer provides the features that differentiate our proposal for SCs from the existing ones. Services System Layer deploys the distributed system adaptable to the needs of each city providing the ability to simulate any high-level services that are deployed. In addition, the environment provides a functional middleware with dynamic visualization techniques facilitating the creation of new apps.

The hardware environment has been developed as a modeling mock-up. This mock-up has sets of nodes based on NodeMCU (an open-source, programmable, low cost, WI-FI enabled, and Arduino-like hardware), which perform the function of IoT gateway, allowing the control of different services, such as regulating street lamps, traffic lights, energy consumption, and security services.

To simulate a SC, the urban area in Salamanca, called “Las Pajas”, has been recreated in a mock-up, as part of the ecoCASA project (<https://ecocasa.usal.es>, grant ID RTC-2016-5250-6). The deployment of sensors in the real infrastructure has been carried out in the frame of the IOTEC project (<https://iotec.usal.es>, grant ID 0123\_IOTEC\_3.E). This area consists of different types of buildings such as single-family homes, office buildings, and large buildings. The simulations have focused on three objectives: (i) scalability, (ii) technologies, and (iii) services.

In scalability, message transmission has been evaluated using Kafka, a durable, distributed streaming platform. On a couple of nodes Kafka can sustain the generation of up to 2,000 messages per second, partitioning and ordering data, and keeping the durability and ordered partitioning of data. This is considered to be high performance and only the top few companies have higher requirements than this. The nodes are automatically deployed by the architecture according to the size of the city (this directly affects the number of sensors connected) and the deployment of nodes is done when the total capacity is at 80 percent of the performance, so that the availability of resources to meet all requests is always guaranteed.

In technologies, with the simulation environment, it has been verified that it is easy to integrate and use different information transmission technologies with the proposed

platform. More specifically, Bluetooth, ZigBee, Wi-Fi, 6LoWPAN, Z-Wave, and WirelessHART have been tested.

In services, the developed architecture allows the deployment of a large number of services, which can be obtained or incorporated from different software entities. These services are monitored and controlled from the dashboard provided by the developed NetLogo-based environment. Web and mobile app versions of the dashboard have also been developed.

5.2. *Management of Incidents.* All SCs must have a system that offers a range of core services such as the following:

- (i) Citizen participation: citizens must be able to access services that allow them to participate in different aspects of their city. Therefore, utilities or applications that can be freely used to provide information in different aspects should be offered
- (ii) Information management: administrations or service providers must be able to access the information stored on the platform with a set of permissions. Thus, there must be applications with restricted access that offer, for example, a dashboard for the control and visualization, in an intuitive way so that it does not require specialized technical personnel of this information
- (iii) Real-time notifications: whenever you work with real-time information systems, there may be a need to perform actions, triggered by the type of information handled, immediately. In the event that such action is to be performed by a human being, the best way to let them know or remind them is through a system of notifications
- (iv) Visualization of the information: the information that is presented, both to the citizen and to the administration staff, must be presented using different methodologies. The most common way for SCs when the information is associated with a geographic component is to present it in a structured way according to its location

In the first part of this first case study, it is intended to apply all these basic elements in a system that provides the necessary functionality to demonstrate the possibilities of citizen participation, manage incidents from a dashboard that allows information to be viewed intuitively, and issue notifications to the managing authority.

When it comes to reaching citizens and allowing them to report incidents from anywhere in the city, it has been determined that the best way is to create a mobile application so that, from their own devices, they can easily report the incidents they decide on at any time.

The use of the mobile device facilitates the sending of incidents, as it allows autocompletion of fields such as location, from the user’s GPS position, or the use of the device’s camera to justify or specify some aspect of the incident itself. To this end, an app has been developed for Android devices and another for iOS devices with the Xamarin tool, which

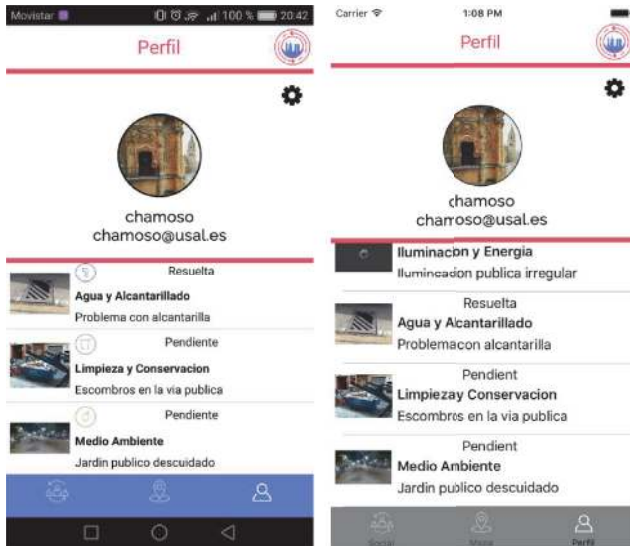


FIGURE 5: Incident reporting app.

allows maintaining an application for both operating systems with a single version of the code.

A web application (HTML5, CSS, JavaScript, Bootstrap, and Socket.io as the main development techniques) has also been developed to monitor and manage all the content, allowing the visualization and resolution of incidents, proposing the best solution to the human supervisor according to the type of incident. When a developer indicates that he or she wishes to obtain a copy of the project, the content of the web page is included within each node associated with the case study, so you can easily modify the entire code to suit your design and functional needs.

One of the developments of the proposed framework is the mobile application designed for citizens who can report any problems that they encounter in the city. This mobile app is available for iOS and Android, the two most used mobile operative systems today. In Figure 5 we can see two screenshots of the app. Users can create their own profile and manage it. When users report an incident, they help the corresponding administration to verify it. The reliability of every profile is measured by an incorporated metric based on previously correctly reported incidences.

Users can send a report with the help of just two simple clicks; if the user is reporting from the location of the incident, the mobile obtains the location automatically. In addition, users can attach a graphic evidence to accompany the written text describing the incident; this can be done because the app provides access to the phone's camera or gallery.

These reports can be received as notifications in the browsers of the corresponding organizations; this depends on the category that the user has marked: water and sewerage, lighting and energy, cleaning and conservation, environment, pedestrians and cyclists, health, safety, traffic and roads, transport, or urban planning. However, there may be a global organization profile that would include all incidents regardless of their categories.

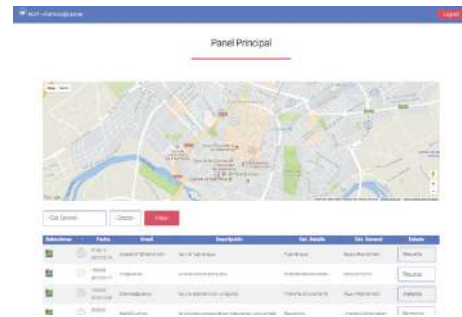


FIGURE 6: Notifications dashboard.

The dashboard displays the reported incidents on the city map (Figure 6) with an icon that shows its category; in addition there is a map that presents their status:

- (i) Pending: when the incident has been received from the citizen, but has not yet been attended
- (ii) Attended: when the corresponding notice has been sent to the maintenance entity responsible for solving the type of problem reported
- (iii) Resolved: when the reported incident has been resolved

The citizen who reported incidents can track their status from their profile on the app.

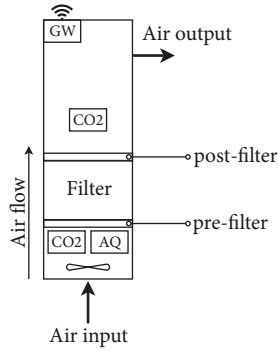
The system provides telephone or email contacts of the entities that are responsible for attending the different types of incidents; in this way the incident can be solved quickly.

The service provided by this application, therefore, allows the city's government entities to attend the needs of its citizens. On the other hand, citizens are provided with a tool that helps them manifest the problems or inconveniences that they find in their city, moreover it is very convenient as they can operate it from their smartphone.

Therefore, this app gives an important role to citizens who, with its use, can participate in improving their city and contribute to a faster resolution of the different problems, since citizens are the most likely to encounter an incident in the city.

**5.3. Deployed Sensors Monitoring.** One of the most necessary and used services in SCs involves the deployment of certain sensors in the city, their integration with the platform, and the representation of the information they capture. Some of the relevant information directly associated with the sensors are about their location. This leads to the need to present to the end user of the application a map with each of the sensors, both to locate and monitor them and to manage them (additions, deletions, and modifications).

This case study aims to develop a real-time measurement system to help absorb CO<sub>2</sub> to bring pollution to more environmentally friendly levels. This is done in a way that is compatible with the natural system, but has the advantage that it can be installed in areas where there is the highest concentration of emissions and it is not possible to plant trees (as protected areas of World Heritage cities). It also allows the

FIGURE 7: CO<sub>2</sub> collector device.FIGURE 8: CO<sub>2</sub> capture monitoring software.

measurement of other air quality parameters (temperature, relative humidity, atmospheric pressure, smoke, and CO<sub>2</sub> and NO<sub>2</sub>) in order to be consulted and to feed intelligent systems to manage the SC.

To this end, a number of devices have been deployed, such as the one shown in Figure 7, all with the same functional characteristics but in different locations, in one city (Salamanca is the city that has been used as a pilot). The devices consist of an air inlet, a fan, two CO<sub>2</sub> sensors (one after the inlet and one right at the start, CO<sub>2</sub> in the image), an air quality sensor (AQ in the image), a Wi-Fi gateway (GW in the image), a prefilter, a filter, and a postfilter.

The system works as follows:

- (i) The outside air is attracted by a fan
- (ii) A series of sensors measure air quality (humidity, pressure, etc.)
- (iii) A CO<sub>2</sub> probe measures its concentration in the air
- (iv) A prefilter prevents the entry of particles
- (v) An impregnated amine filter absorbs the CO<sub>2</sub>
- (vi) A postfilter prevents the impregnated amine from being expelled to the outside
- (vii) A CO<sub>2</sub> measuring probe measures its concentration at the outlet
- (viii) A wireless device collects and sends data from sensors and probes
- (ix) The SC platform collects the information

The device includes a power socket that can be connected to the grid as well as microgeneration systems for weak (renewable), solar, wind, or kinetic energies. The SC2 project will also include a versatile design adaptable to different types of street furniture (street lamps, traffic lights, mupis, or city buses).

The gateway that transmits the information to the platform is based on the Raspberry Pi 3, as mentioned above, which receives the information from the sensor and sends it via Wi-Fi via the Kafka REST API, so that the measurements are available in real time.

The designed system allows registering any kind of sensor. The user can define what type of information they want to be

viewed, such as the values collected by the sensor, times at which the information was received, the sensor identifier, or MAC address.

When registering a sensor, a description of it has to be provided (required by Hypercat) and its location has to be specified. Since this case study focuses on the monitoring of fixed sensors deployed in a city, their location cannot change over time nor the information about them.

Once this is done, the platform is ready to collect the information provided by the sensor. In the described case study, information is provided in an open way, so no identification/password is required to access it. Like on open data platforms, citizens or developers can download the information to use it in other applications.

The developer has a number of options to choose from on how they want to represent the information on the map. In the designed case, the value of CO<sub>2</sub> before and after it passes through a filter is shown on the heatmap (Figure 8). Also, the location of sensors on the map is indicated with markers.

The application can easily be replicated by any developer. Therefore, an open data service can easily be established by using the information gathered from the sensors deployed in any city. This can be achieved by defining the necessary settings on communication type, a specific communication protocol, and security.

## 6. Conclusion and Future Work

The architecture model presented in this work follows a multilayer platform designed to provide different services that can be deployed in a SC. These services cover a wide range of areas, such as employability, leisure, energy management, or healthcare.

The architecture is designed in such way that it adapts to the specific needs of each city, enabling cities of any size to use only the resources they need. The utilization model has three possibilities: (i) deploying the solution developed in a proprietary environment, where you have physical control of all the computing resources available in the cluster; (ii) deploying the solution in an external cloud environment that charges for the services used; (iii) using a hybrid system according to different needs, which in most cases would be

the best option, but depends on each scenario in which you intend to deploy.

The novelty of this architecture lies in the abstract layer design; none of the previous platforms included layers in their architectures. The use of layers makes it possible to change the technology used to execute a particular process within the system, without having to make changes in the rest of layers or modules. In addition, the development of a specified service layer (Services System Layer) allows reuse of common functionalities in basic services for all SCs. An open data system could be created to allow cities to access the knowledge acquired by all the cities on these common functionalities.

The architecture has been validated with the three case studies presented. These case studies have been selected so that they can be reused in any city to manage incidents, visualize data, or perform tests in simulation environments.

As future work, the aim is to expand the number of generic case studies so that any city can reuse them and thus save time in developing common functionalities in Smart Cities.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research has been partially supported by the European Regional Development Fund (ERDF) within the framework of the Interreg program V-A Spain-Portugal 2014-2020 (PocTep) under the IOTEC project grant 0123\_IOTEC\_3\_E and by the Spanish Ministry of Economy, Industry and Competitiveness and the European Social Fund under the ecoCASA project grant RTC-2016-5250-6.

## References

- [1] T. M. Heng and L. Low, "The intelligent city: Singapore achieving the next lap: Practitioners forum," *Technology Analysis & Strategic Management*, vol. 5, no. 2, pp. 187–202, 1993.
- [2] Q. Li and S. Lin, "Research on digital city framework architecture," in *Proceedings of the International Conferences on Info-Tech and Info-Net, ICII 2001*, pp. 30–36, Beijing, China, November 2001.
- [3] T. Ishida, H. Ishiguro, and H. Nakanishi, "Connecting digital and physical cities," in *Digital Cities II: Computational and Sociological Approaches*, vol. 2362 of *Lecture Notes in Computer Science*, pp. 246–256, Springer, Berlin, Germany, 2002.
- [4] R. Giffinger, C. Fertner, H. Kramar, R. Kalasek, Pichler-Milanovic, and N. Meijers, *Smart Cities: Ranking of European Medium-Sized Cities*, Centre of Regional Science (SRF), Vienna University of Technology, Vienna, Austria, 2007.
- [5] C. Harrison, B. Eckman, R. Hamilton et al., "Foundations for smarter cities," *IBM Journal of Research and Development*, vol. 54, no. 4, pp. 1–16, 2010.
- [6] G. P. Hancke, B. D. C. de Silva, and G. P. Hancke Jr., "The role of advanced sensing in smart cities," *Sensors*, vol. 13, no. 1, pp. 393–425, 2013.
- [7] S. Dirks and M. Keeling, *A Vision of Smarter Cities: How Cities Can Lead The Way into A Prosperous and Sustainable Future*, IBM Global Business Services, Somers, NY, USA, 2009.
- [8] A. Mahizhnan, "Smart cities: The Singapore case," *Cities*, vol. 16, no. 1, pp. 13–18, 1999.
- [9] G. Rudolf, C. Fertner, H. Kramar, R. Kalasek, N. Pichler-Milanovic, and E. Meijers, "Smart cities-ranking of european medium-sized cities," Tech. Rep., Vienna Centre of Regional Science, 2007.
- [10] J. M. Eger, "Smart growth, smart cities, and the crisis at the pump a worldwide phenomenon," *I-WAYS-The Journal of E-Government Policy and Regulation*, vol. 32, no. 1, pp. 47–53, 2009.
- [11] R. Moss Kanter and S. S. Litow, "Informed and interconnected: A manifesto for smarter cities," Tech. Rep. 09-141, Harvard Business School General Management Unit Working Paper, 2009.
- [12] M. N. K. Boulos, A. D. Tsouros, and A. Holopainen, "Social, innovative and smart cities are happy and resilient: insights from the who euro 2014 international healthy cities conference," *International Journal of Health Geographics*, vol. 14, no. 1, article 3, 2015.
- [13] S. Alawadhi, A. Aldama-Nalda, H. Chourabi et al., "Building understanding of smart city initiatives," in *EGOV 2012: Electronic Government*, vol. 7443 of *Lecture Notes in Computer Science*, pp. 40–53, 2012.
- [14] R. Carli, M. Dotoli, R. Pellegrino, and L. Ranieri, "Measuring and managing the smartness of cities: A framework for classifying performance indicators," in *Proceedings of the 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013*, pp. 1288–1293, Manchester, UK, October 2013.
- [15] M. H. Moore, *Creating Public Value: Strategic Management in Government*, Harvard University Press, 1995.
- [16] J. Benington and M. H. Moore, *Public value: Theory and practice*, Palgrave Macmillan, Basingstoke, UK, 2011.
- [17] B. Baccarne, P. Mechant, and D. Schuurman, "Empowered Cities? An Analysis of the Structure and Generated Value of the Smart City Ghent," in *Smart City*, Progress in IS, pp. 157–182, Springer International Publishing, 2014.
- [18] P. Neirotti, A. De Marco, A. C. Cagliano, G. Mangano, and F. Scorrano, "Current trends in smart city initiatives: Some stylised facts," *Cities*, vol. 38, pp. 25–36, 2014.
- [19] T. Sousa, H. Morais, J. Soares, and Z. Vale, "Day-ahead resource scheduling in smart grids considering Vehicle-to-Grid and network constraints," *Applied Energy*, vol. 96, pp. 183–193, 2012.
- [20] L. Gomes, P. Faria, H. Morais, Z. Vale, and C. Ramos, "Distributed, agent-based intelligent system for demand response program simulation in smart grids," *IEEE Intelligent Systems*, vol. 29, no. 1, pp. 56–65, 2014.
- [21] J. F. De Paz, J. Bajo, S. Rodríguez, G. Villarrubia, and J. M. Corchado, "Intelligent system for lighting control in smart cities," *Information Sciences*, vol. 372, pp. 241–255, 2016.
- [22] J. Viitanen and R. Kingston, "Smart cities and green growth: Outsourcing democratic and environmental resilience to the global technology sector," *Environment and Planning A*, vol. 46, no. 4, pp. 803–819, 2014.

- [23] A. González-Briones, J. Prieto, F. De La Prieta, E. Herrera-Viedma, and J. M. Corchado, "Energy optimization using a case-based reasoning strategy," *Sensors*, vol. 18, no. 3, 2018.
- [24] P. Chamoso, J. F. De Paz, S. Rodríguez, and J. Bajo, "Smart Cities Simulation Environment for Intelligent Algorithms Evaluation," in *ISAmI 2016: Ambient Intelligence- Software and Applications – 7th International Symposium on Ambient Intelligence (ISAmI 2016)*, vol. 476 of *Advances in Intelligent Systems and Computing*, pp. 177–187, Springer International Publishing, 2016.
- [25] A. González-Briones, P. Chamoso, H. Yoe, and J. M. Corchado, "GreenVMAS: Virtual organization based platform for heating greenhouses using waste energy from power plants," *Sensors*, vol. 18, no. 3, 2018.
- [26] S. Barthel and C. Isendahl, "Urban gardens, Agriculture, and water management: Sources of resilience for long-term food security in cities," *Ecological Economics*, vol. 86, pp. 224–234, 2013.
- [27] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Sensing as a service model for smart cities supported by internet of things," *Transactions on Emerging Telecommunications Technologies*, vol. 25, no. 1, pp. 81–93, 2014.
- [28] K. Nowicka, "Smart city logistics on cloud computing model," *Procedia-Social and Behavioral Sciences*, vol. 151, pp. 266–281, 2014.
- [29] S. Hörold, C. Mayas, and H. Krömker, "Towards paperless mobility information in public transport," in *Human-Computer Interaction: Design and Evaluation*, vol. 9169 of *Lecture Notes in Computer Science*, pp. 340–349, Springer International Publishing, 2015.
- [30] T. Jeske, "Floating car data from smartphones: What google and waze know about you and how hackers can control traffic," in *Proceedings of the BlackHat Europe*, pp. 1–12, Amsterdam, Netherlands, 2013.
- [31] A. Krylovskiy, M. Jahn, and E. Patti, "Designing a smart city internet of things platform with microservice architecture," in *Proceedings of the 3rd International Conference on Future Internet of Things and Cloud (FiCloud '15)*, pp. 25–30, Rome, Italy, August 2015.
- [32] X. Liu, X. Wang, G. Wright, J. C. P. Cheng, X. Li, and R. Liu, "A state-of-the-art review on the integration of Building Information Modeling (BIM) and Geographic Information System (GIS)," *ISPRS International Journal of Geo-Information*, vol. 6, no. 2, article 53, 2017.
- [33] S. U. M. Tobi, D. Amaratunga, and N. M. Noor, "Social enterprise applications in an urban facilities management setting," *Facilities*, vol. 31, no. 5/6, pp. 238–254, 2013.
- [34] L. Hernandez, C. Baladron, J. M. Aguiar et al., "A survey on electric power demand forecasting: Future trends in smart grids, microgrids and smart buildings," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1460–1495, 2014.
- [35] T. A. Nguyen and M. Aiello, "Energy intelligent buildings based on user activity: a survey," *Energy and Buildings*, vol. 56, no. 1, pp. 244–257, 2013.
- [36] C. Wilson, T. Hargreaves, and R. Hauxwell-Baldwin, "Smart homes and their users: a systematic analysis and key challenges," *Personal and Ubiquitous Computing*, vol. 19, no. 2, pp. 463–476, 2014.
- [37] U. Gretzel, M. Sigala, Z. Xiang, and C. Koo, "Smart tourism: foundations and developments," *Electronic Markets*, vol. 25, no. 3, pp. 179–188, 2015.
- [38] C. Koo, F. Ricci, C. Cobanoglu, and F. Okumus, "Special issue on smart, connected hospitality and tourism," *Information Systems Frontiers*, vol. 19, no. 4, pp. 699–703, 2017.
- [39] P. Kumar, L. Morawska, C. Martani et al., "The rise of low-cost sensing for managing air pollution in cities," *Environment International*, vol. 75, pp. 199–205, 2015.
- [40] Z. Khan, Z. Pervez, and A. Ghafoor, "Towards cloud based smart cities data security and privacy management," in *Proceedings of the 7th IEEE/ACM International Conference on Utility and Cloud Computing, UCC 2014*, pp. 806–811, London, UK, December 2014.
- [41] A. González-Briones, G. Villarrubia, J. F. De Paz, and J. M. Corchado, "A multi-agent system for the classification of gender and age from images," *Computer Vision and Image Understanding*, 2018, In Press.
- [42] A. Solanas, C. Patsakis, M. Conti et al., "Smart health: a context-aware health paradigm within smart cities," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 74–81, 2014.
- [43] A. D. Black, J. Car, C. Pagliari et al., "The impact of ehealth on the quality and safety of health care: a systematic overview," *PLoS Medicine*, vol. 8, no. 1, Article ID e1000387, 2011.
- [44] A. Hussain, R. Wenbi, A. L. Da Silva, M. Nadher, and M. Mudhish, "Health and emergency-care platform for the elderly and disabled people in the Smart City," *The Journal of Systems and Software*, vol. 110, pp. 253–263, 2015.
- [45] J. Borràs, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: a survey," *Expert Systems with Applications*, vol. 41, no. 16, pp. 7370–7389, 2014.
- [46] T. Clohessy, T. Acton, and L. Morgan, "Smart city as a service (SCaaS): A future roadmap for e-government smart city cloud computing initiatives," in *Proceedings of the 7th IEEE/ACM International Conference on Utility and Cloud Computing, UCC 2014*, pp. 836–841, London, UK, December 2014.
- [47] S. M. Ali, C. A. Mehmood, A. Khawja et al., "Micro-controller based smart electronic voting machine system," in *Proceedings of the 2014 IEEE International Conference on Electro/Information Technology, EIT 2014*, pp. 438–442, Milwaukee, Wis, USA, June 2014.
- [48] M. Janssen and J. van den Hoven, "Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy?" *Government Information Quarterly*, vol. 32, no. 4, pp. 363–368, 2015.
- [49] S. Y. Perng, R. Kitchin, and D. Mac Donncha, "Hackathons, entrepreneurship and the passionate making of smart cities," 2017.
- [50] B. Ridel, P. Reuter, J. Laviolle, N. Mellado, N. Couture, and X. Granier, "The revealing flashlight: Interactive spatial augmented reality for detail exploration of cultural heritage artifacts," *Journal on Computing and Cultural Heritage*, vol. 7, no. 2, article 6, 2014.
- [51] T. Kim, J. Y. Cho, and B. G. Lee, "Evolution to smart learning in public education: a case study of korean public education," in *Open and Social Technologies for Networked Learning*, vol. 395 of *IFIP Advances in Information and Communication Technology*, pp. 170–178, Springer, Berlin, Germany, 2013.
- [52] T. Di Mascio, P. Vittorini, R. Gennari, A. Melonio, F. De La Prieta, and M. Alrifai, "The learners' user classes in the TER-ENCE adaptive learning system," in *Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies, ICALT 2012*, pp. 572–576, Rome, Italy, July 2012.



- [53] M. Kavaratzis, "From city marketing to city branding: towards a theoretical framework for developing city brands," *Place Branding and Public Diplomacy*, vol. 1, no. 1, pp. 58–73, 2004.
- [54] P. Chamoso, A. Rivas, S. Rodríguez, and J. Bajo, "Relationship recommender system in a business and employment-oriented social network," *Information Sciences*, vol. 433–434, pp. 204–220, 2018.
- [55] R. Minerva, A. Biru, and D. Rotondi, "Towards a definition of the internet of things (iot)," *IEEE Internet Initiative*, no. 1, 2015.
- [56] O. Said and M. Masud, "Towards internet of things: Survey and future vision," in *Proceedings of the International Journal of Computer Networks 5(1)*, vol. 5, 2013.
- [57] A. Iera, C. Floerkemeier, J. Mitsugi, and G. Morabito, "The Internet of things," *IEEE Wireless Communications Magazine*, vol. 17, no. 6, pp. 8–9, 2010.
- [58] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. D. Johnson, "M2M: from mobile to embedded internet," *IEEE Communications Magazine*, vol. 49, no. 4, pp. 36–43, 2011.
- [59] S. Okuya, "M2M and big data to realize the smart city," *NEC Technical Journal*, vol. 7, no. 2, pp. 67–71, 2012.
- [60] H. Okkonen, O. Mazhelis, P. Ahokangas et al., "nترنت-of-things market, value networks, and business models: state of the art report," *Computer Science and Information Systems Reports* 39, 2013.
- [61] A. Zaslavsky, C. Perera, and D. Georgakopoulos, "Sensing as a service and big data," 2013, <https://arxiv.org/abs/1301.0159>.
- [62] P. Mell, T. Grance et al., *The Nist Definition of Cloud Computing*, 2011.
- [63] M. Armbrust, A. Fox, R. Griffith et al., "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.
- [64] R. Hecht and S. Jablonski, "NoSQL evaluation: a use case oriented survey," in *Proceedings of the International Conference on Cloud and Service Computing (CSC '11)*, pp. 336–341, Hong Kong, China, December 2011.
- [65] J. Pokorny, "NoSQL databases: a step to database scalability in web environment," *International Journal of Web Information Systems*, vol. 9, no. 1, pp. 69–82, 2013.
- [66] A. Moniruzzaman and S. A. Hossain, "Nosql database: New era of databases for big data analytics-classification, characteristics and comparison," 213, <https://arxiv.org/abs/1307.0191>.
- [67] A. Mukherjee, "Iot and the rise of smart machines - prosperity or doomsday," in *Proceedings of the International Conference on Digital Libraries (ICDL) 2016: Smart Future: Knowledge Trends that will Change the World*, *The Energy and Resources Institute*, pp. 116–125, 2016.
- [68] J. M. Kanter and K. Veeramachaneni, "Deep feature synthesis: Towards automating data science endeavors," in *Proceedings of the IEEE International Conference on Data Science and Advanced Analytics, DSAA 2015*, Paris, France, October 2015.
- [69] D. Beulke, "Big data impacts data management: The 5 vs of big data," 2011, <http://davebeulke.com/big-data-impactsdata-management-the-five-vs-of-big-data/>.
- [70] J. Manyika, M. Chui, B. Brown et al., "Big data: The next frontier for innovation, competition, and productivity," 2011.
- [71] P. J. Jamack, *Big Data Business Intelligence Analytics*, IBM Developer Works Technical Library, 2012.
- [72] J. Gantz and D. Reinsel, "The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the far east," in *IDC iView: IDC Analyze the future*, pp. 1–16, 2012.
- [73] P. Russom, "Managing big data," TDWI BEST PRACTICES REPORT 4, 2013.
- [74] G. Piatetsky-Shapiro, "Discovery, analysis and presentation of strong rules," in *Knowledge Discovery in Databases*, pp. 229–248, 1991.
- [75] S. Suthaharan, "Big data classification: problems and challenges in network intrusion prediction with machine learning," *ACM SIGMETRICS Performance Evaluation Review*, vol. 41, no. 4, pp. 70–73, 2014.
- [76] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, *Machine Learning: An Artificial Intelligence Approach*, Springer Science & Business Media, 2013.
- [77] G. A. F. Seber and A. J. Lee, *Linear Regression Analysis*, vol. 936, John Wiley & Sons, New York, NY, USA, 2012.
- [78] T. Bakıcı, E. Almirall, and J. Wareham, "A smart city initiative: the case of Barcelona," *Journal of the Knowledge Economy*, vol. 4, no. 2, pp. 135–148, 2013.
- [79] L. Zhuhadar, E. Thrasher, S. Marklin, and P. O. de Pablos, "The next wave of innovation—Review of smart cities intelligent operation systems," *Computers in Human Behavior*, vol. 66, pp. 273–281, 2017.
- [80] R. De Amicis, G. Conti, D. Patti, M. Ford, and P. Elisei, *I-Scope-Interoperable Smart City Services through an Open Platform for Urban Ecosystems*, na, 2012.
- [81] G. Wang, J. Koshy, S. Subramanian et al., "Building a replicated logging system with apache kafka," *Journal of Proceedings of the VLDB Endowment*, vol. 8, no. 12, pp. 1654–1655, 2015.
- [82] R. Ranjan, "Streaming big data processing in datacenter clouds," *IEEE Cloud Computing*, vol. 1, no. 1, pp. 78–83, 2014.
- [83] M. Blackstock and R. Lea, "IoT interoperability: A hub-based approach," in *Proceedings of the 2014 International Conference on the Internet of Things, IOT 2014*, pp. 79–84, Cambridge, Mass, USA, October 2014.
- [84] S. Latré, P. Leroux, T. Coenen, B. Braem, P. Ballon, and P. Demeester, "City of things: An integrated and multi-technology testbed for IoT smart city experiments," in *Proceedings of the 2nd IEEE International Smart Cities Conference, ISC2 2016*, Trento, Italy, September 2016.
- [85] B. Braem, S. Latré, P. Leroux, P. Demeester, T. Coenen, and P. Ballon, "City of Things: a multidisciplinary smart cities testbed for IoT, big data and living labs innovation," in *Network Design and Optimization for Smart Cities*, vol. 8 of *Series on Computers and Operations Research*, pp. 279–292, World Scientific, 2017.

