frontiers | Frontiers in Sustainable Cities

TYPE Hypothesis and Theory
PUBLISHED 21 September 2022
DOI 10.3389/frsc.2022.846987

OPEN ACCESS

EDITED BY
Giuseppe Nicosia,
University of Catania, Italy

REVIEWED BY
Gauthier Picard,
Université de Toulouse, France
Davide Quaglia,
University of Verona, Italy

*CORRESPONDENCE
Souvik Sengupta
souvik.sengupta@iex.ec

SPECIALTY SECTION
This article was submitted to Smart Technologies and Cities, a section of the journal Frontiers in Sustainable Cities

RECEIVED 31 December 2021
ACCEPTED 19 August 2022
PUBLISHED 21 September 2022

CITATION
Simonet-Boulogne A, Solberg A, Sinaeepourfard A, Roman D, Perales F, Ledakis G, Plakas I and Sengupta S (2022) Toward blockchain-based fog and edge computing for privacy-preserving smart cities. *Front. Sustain. Cities* 4:846987. doi: 10.3389/frsc.2022.846987

Toward blockchain-based fog and edge computing for privacy-preserving smart cities

Anthony Simonet-Boulogne[1], Arnor Solberg[2], Amir Sinaeepourfard[3], Dumitru Roman[4], Fernando Perales[5], Giannis Ledakis[6], Ioannis Plakas[6] and Souvik Sengupta[1]*

[1]iExec Blockchain Tech, Lyon, France, [2]Tellu, Asker, Norway, [3]IEEE Member, Oslo, Norway, [4]SINTEF, Trondheim, Norway, [5]JOT Internet Media, Madrid, Spain, [6]UBITECH, Athens, Greece

The rapid development of Smart Cities is aided by the convergence of information and communication technologies (ICT). Data is a key component of Smart City applications as well as a serious worry. Data is the critical factor that drives the whole development life-cycle in most Smart City use-cases, according to an exhaustive examination of several Smart City use-cases. Mishandling data, on the other hand, can have severe repercussions for programs that get incorrect data and users whose privacy may be compromised. As a result, we believe that an integrated ICT solution in Smart Cities is key to achieve the highest levels of scalability, data integrity, and secrecy within and across Smart Cities. As a result, this paper discusses a variety of modern technologies for Smart Cities and proposes our integrated architecture, which connects Blockchain technologies with modern data analytic techniques (e.g., Federated Learning) and Edge/Fog computing to address the current data privacy issues in Smart Cities. Finally, we discuss and present our proposed architectural framework in detail, taking into account an online marketing campaign and an e-Health application use-cases.

KEYWORDS

Blockchain, Smart Cities, Fog/Edge computing, internet-of-things (IoT), smart contract

1. Introduction

Recent advances in cloud and edge computing, networking and big data processing drive an increasing interest in IoT and Smart City applications. If implemented properly, Smart Cities could make an important turn in privacy and data sovereignty by processing data locally, hence avoiding the concentration of private and sensitive data in the hands of a few cloud providers.

The so-called *Cloud Computing Continuum* which combines fog and edge computing cloudlets (Dolui and Datta, 2017) with fiber optical and 5G networks implies that urban areas are already filled with computing resources and connected devices such as IP cameras, traffic lights, and all sorts of sensors. The Cloud Continuum, however, is highly fragmented; IoT networks, cloud computing and network providers are all deployed and exploited within silos that prevent their discovery and complicates their integration in

real-world applications. These silos also tend to create frictions for new actors wishing to enter a market (e.g., renting some local virtualized servers, monetising data streams, etc.) and threatens the privacy of the end users. As ambitious application scenarios such as e-Health, e-Governance, public services, mobility and transportation will require sharing more and more data among different parties, transparency, traceability and accountability are paramount.

This work proposes a vision of a new privacy-preserving platform and framework for building Smart City applications. Our design is based on three pillars which address the aforementioned challenges; first, a *computing and data marketplace* will provide smooth interactions between consumers and providers of computing resources, data and devices from the Cloud Continuum, and bring it all within one namespace; second, a *Blockchain-based computing infrastructure* will provide trust, traceability and auditability in data processing activities; third, *Privacy-preserving processing* techniques such as Federated Learning and Trusted Execution Environments (TEE) will ensure data privacy and confidentiality from end to end.

The framework we propose in this article is based on the most recent software research and hardware components and is provided as *Platform-as-a-Service* (PaaS). First, a Domain Specific Language (DSL) is used for describing high-level application requirements; second, a software platform selects the most appropriate resources from the Computing Continuum for satisfying the requirements; then the platform provisions the resources and deploys the application services; finally, the platform monitors both the resources and the application, adapting them automatically to infrastructure and environment changes. Further, we describe two industrial use-cases and use them as drivers throughout our work, from the design to the evaluation of our approach. By giving developers access to extremely diverse resources while automating application life cycle management, we believe that our solution will be a major enabler for future Smart City and IoT applications.

The rest of this article is organized as follows; Section 2 studies the state of the art related to infrastructures, platforms, and technologies for Smart City applications; Section 3 defines our driving industrial use cases; Section 4 describes the architecture of our solution into details, while Section 5 provides concluding remarks and hints at future works related to the implementation of our approach.

2. Background study

Benefiting from recent advances in distributed systems, networking and IoT, Smart Cities are expected to enhance the quality of life of urban citizens and propel cities toward a more sustainable future. The availability of a variety of IoT is actively contributing to the development of Smart Cities, but the tremendous amount of data that needs to be generated, curated, classified, analyzed, and transferred to provide context-aware real-time services to inhabitants creates specific challenges (Allam and Dhunny, 2019; Lee et al., 2020). Scalable ICT infrastructure management, successful monetisation of data, applications and computing resources, and the development of hybrid privacy-enhancing technologies are three critical criteria for developing advanced privacy-preserving Smart City ecosystems (Bélissent, 2010; Pahuja, 2021). In particular, performing data processing in distant cloud data centers present many limitations in terms of performance, resilience, energy efficiency and privacy (Perera et al., 2017). Edge and Fog Computing (Bonomi et al., 2014; Shi and Dustdar, 2016) have emerged as a promising infrastructure solution for latency-sensitive and geographically distributed applications. By placing many small cloud points of presence (or *cloudlets* to the edge of the Internet, Fog/Edge Computing brings the users and their IoT devices closer to the processing resources. As a result data transfers are kept to a minimum, reducing the network latency and saving precious bandwidth. In certain models fog and edge devices have limited processing capacity (e.g., in the case when Single-Board Computers are used) and cannot support massive data processing (Ghobaei-Arani et al., 2020); in others models, enterprise-grade servers are geographically distributed in strategic locations (e.g., in urban areas) to improve cost, scale and reliability without sacrificing performance (Church et al., 2008; Valancius et al., 2009). Furthermore, the Computing Continuum, by combining the centralized cloud with servers deployed at the edge of the Internet and in-between (*fog*), presents a suitable option for latency-sensitive situation-awareness applications because these are typically sensitive to network delays and connection instability (Masip-Bruin et al., 2016).

Famous computer scientist Prof John McCarthy envisioned a futuristic computing environment where processing power, applications, and data might be structured as a public utility and sold through a metered business model at the dawn of the contemporary computation age. According to Prof John McCarthy, such an ecosystem might enable scalable access to virtually limitless computer resources as well as a wide range of IT services (Rajaraman, 2014). Selling processing power, applications, and data eventually became a reality and a market that is largely dominated by a few large IT companies. As a consequence, individuals and small/medium businesses face massive obstacles in penetrating this market. More recently, several independent marketplaces have appeared, considering that open, rule-based, uniformly governed, and federated marketplaces are required to break free of the current oligopoly (Perera et al., 2015; Maciel et al., 2019). Nonetheless, a significant issue that remains for any marketplace is providing verified, reputable, and relevant resources to meet gain the trust of buyers. Because they can ensure the security, transparency, auditability, and the traceability of all deals and transactions

among peers, blockchain technologies (Johansen, 2018; Rejeb et al., 2020; Waleed et al., 2021) are viewed as a major enabler to meet the need of decentralized marketplaces.

One of the essential requirements for developing a sustainable Smart City is adaptability, for which artificial intelligence (AI) and machine learning (ML) technologies can play a significant role (Ullah et al., 2020). In the context of smart cities and AI, data is one of the most critical components. However, data raise two significant concerns; firstly, handling data without breaching the privacy of their owners can be challenging, since it constantly flows from one computing system to another; second, the security and performance of AI/ML applications can be harmed by fabricated or corrupted data (Li et al., 2017; Sambasivan et al., 2021; Xu et al., 2021). Therefore, maintaining a high level of security throughout the whole computing chain is paramount to safeguard the data, and can only be achieved through comprehensive infrastructure management and auditability (Al Nuaimi et al., 2015). In light of this, three different main ICT architectural reference models and their related data management and flow architectures within the Smart City paradigm have been proposed started from Decentralized-to-Centralized (DC2C-ICT) Sinaeepourfard et al. (2017) to Distributed-to-Centralized (D2C-ICT) (Sinaeepourfard, 2017; Sinaeepourfard et al., 2020), and hybrid (Sinaeepourfard et al., 2020, 2021; Alamouti et al., 2022) ICT architectures, as represented in Figure 1. With the help of the proposed three main ICT reference frameworks, the authors showed how efficiently storing, managing, processing, analyzing data, and software services management can be done within and across the city/cities within the scope of the Smart City.

Developing Smart City ICT reference architecture models walks through diverse strategic design views. Those views explain how the ICT architecture models may connect technologies, organizations, and information in the unified architecture design (Schieferdecker et al., 2017). As shown in Figure 1, we finally realize that designing ICT architecture in Smart Cities (Schieferdecker et al., 2017; Sinaeepourfard et al., 2020) may propose through two different strategic ICT design views from "bottom-top" or "top-bottom."

Besides developing a robust and sustainable ICT architecture and their data management infrastructure, building strong data privacy and the security mechanism is essential. Therefore, many technological innovations have emerged to support hardware-level and software-level data privacy-preserving and confidential computing to ensure the highest level of data security and privacy. For example, hardware-based Trusted Execution Environments (TEE) offer an isolated execution environment inside the main processor for securely executing applications and data even in a compromised system (Valadares et al., 2021). It ensures that the code and data loaded inside are protected in terms of confidentiality and integrity. In addition, TEE provides security features such as isolated execution, the integrity of applications running inside, secure data storage, and

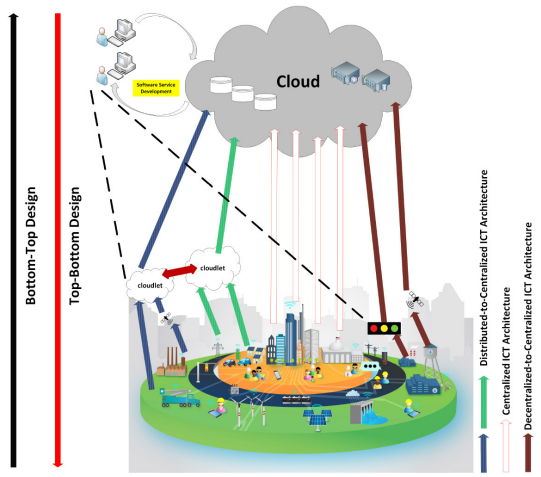

FIGURE 1
Different ICT architectures in Smart City for data management, processing, and software services development, extract main idea from Sinaeepourfard (2017), Sinaeepourfard et al. (2020).

key management and cryptographic algorithms as an isolated runtime environment (Ning et al., 2017). In comparison, technologies like secure multiparty computation, homomorphic encryption, federated learning, differential privacy, and zero-knowledge proof support application-level data privacy and security (Danezis et al., 2015). In addition to that, federated learning (Konečnỳ et al., 2016) enables the facilities to keep the private data to the owners, which helps significantly reduce the data flow in the Smart City domain and helps to utilize the network resources within the Smart City paradigm adequately.

Ensuring the highest level of data privacy and security is certainly essential for making a more advanced and sustainable Smart City. Considering this fact and focusing on the different application scenarios, we propose an architectural solution for developing a privacy-preserving computation framework for managing big data within the scope of the Smart City domain.

3. Use-case descriptions

In this section, by a thorough discussion of two different Smart City application scenarios, we compiled the main requirements for developing a privacy-preserving computing ecosystem within the scope of the Smart City domain.

3.1. Use-case 1: Smart cities services optimization based on innovative marketing data pipelines

Smartness is a more user-friendly concept in marketing than the more elitist term "intelligence," which is often limited to fast thinking and attentive to input. However, other interpretations

argue that "smart" also includes the term "intelligence" because smartness can only be fulfilled when an intelligent system adjusts to the needs of its users. Significantly, in the case of digital marketing Smart City application scenario depends on having access to relevant data that allows for the creation of innovative services aligned with both population interests and city resource management. Within the scope of the Smart City application domain, the information gathered from the various online marketing apps and digital online stores can be used to analyze user behavior and interests. In addition, developing new data pipelines and providing insights regarding population interest based on location and time will directly impact the short-medium term. Retail businesses, for example, may alter their stock levels based on local preferences and consumers (both on a regular and *ad-hoc* basis), security resource allocation based on young population behavior, and sports service offerings based on demand, among other things.

Depending on the set of variables which describes the impact of the campaigns; for examples, clicks, impressions, conversion rate, and click-through rate, it is possible to ideally monitor the relevance of the campaigns depending on the category in almost real-time. Besides that, in any Smart City, many connected IoT devices capture various events for generating data and create a massive data flow over the Smart City. Furthermore, data from citizens' personalized IoT devices (e.g., smartwatches, health monitoring devices) also increase data flow. Implicitly, analyzing those data could potentially improve the market campaigning and help the market managers to promote their business through the public area connected devices' (e.g., Digital billboards). Finally, adopting the augmented reality/virtual reality (AR/VR) technology in Smart City's supermarket improves the marketing experience for the customers. Also, it helps to run the business smoothly and effectively (Yi, 2020).

In the Smart City scenario, in addition to the existing pipeline process for downloading marketing statistics in batch mode, the development of a new edge middleware component called the *Redirector* is required. The *Redirector* assigns a unique identifier per click. It is also responsible for ensuring complete traceability of user actions. Due to the system's complexity, there are several requirements and critical resources that must be met in order to provide a positive user experience, including; 24/7 operational, real-time data acquisition, audibility and verification of data, and edge processing to assign the identifier less than 200 ms to avoid affecting the internet user experience.

Based on the user engagement with the online real-time marketing campaign management system, we can find some key participants/actors participating in the system to run the campaign successfully. Besides the data contributors (Smart City's citizens), the following vital actors are mainly responsible for improving the overall marketing management system in Smart City. Below, we are going to describe their role and functionalities briefly:

- **Campaign managers** will define the applications needed to monitor the performance of the campaigns based on the available data coming from the daily (batch) ingestion of the campaigns' performance statistics as well as the stream (real-time) data collected by the Redirector. In addition, they will use the orchestrator and scheduler to update the reporting services depending on their needs, typically eight times per day.
- **Data Scientists** will also define data pipelines but generate new insights to enrich the information related to the keywords, such as category classification, seasonality and profit prediction. In this sense, they will exploit the resource provisioning to adapt the resources to the experiment requirements (processing time, data volume and so on). Also, the orchestrator synchronizes the processing steps and the scheduler when the models are uploaded to productive conditions.
- **IT Managers** will deploy the Redirector component to process and collect data at the edge as well as exploit all the data pipeline management resources: (i) Security layer to control user access to the confidential data as well as access to the data repository, (ii) Monitoring to evaluate the performance of the processing steps and identify bottlenecks and downtimes due to the resources allocation inefficiencies. The IT managers will also manage the automatic adaptation of the resources to implement constraints to optimize the financial costs. This resource allocation will be carried out based on the existing infrastructure, which in our case combines Azure and Google Cloud.

In this way, the coordinated system combines the specific needs and constraints to define and deploy a data pipeline in productive environments and the flexibility to be adapted to the specific requirements of the main stakeholders of the business case. In addition, the combination of the two data collection pipelines enables the generation of innovative services to be applied in a Smart City environment. Also, it demands the creation of new business models based on data monetising for third parties that can be directly implemented utilizing the marketplace functionalities. The conceptual diagram of the life-cycle management of analytical pipelines for market campaigning has been presented in Figure 2.

3.2. Use-case 2: Digital health system for smart cities

The health business case represents an effective Smart City service that allows the citizens to be protected and cared for in terms of their welfare and health while living at home. This includes featuring care services to allow the elderly to live at

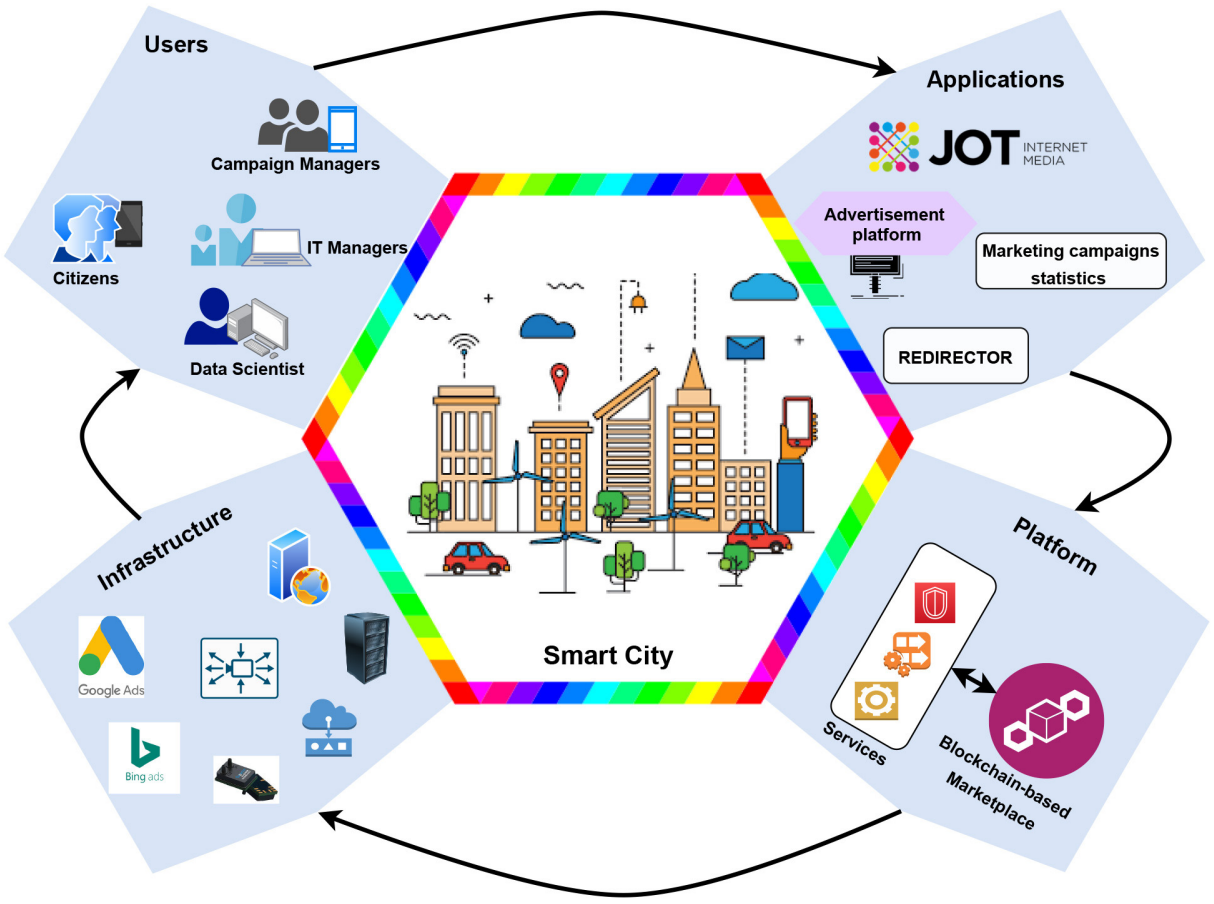

FIGURE 2
Life-cycle of analytical pipelines of marketing campaigns management.

home as long as possible and allow health personnel to do remote follow up of citizens that suffer from chronic diseases to enable the patients to stay at home to the extent possible during treatment, allowing them to live close to a normal daily life. In these scenarios, data management needs to be strictly secured and controlled to ensure privacy and protect sensitive data.

It is significant to have optimized data management and pipelines in order to gather data from sensors and devices (e.g., welfare sensors, medical devices etc.) and provide the correct data to the right stakeholder, whether it is the doctor, care providers, next of kin. The data may also be provided to relevant city services such as the fire department in case of a smoke detector alarm in the home of the elderly etc. The information needs to be timely and need to be provided securely and reliably.

Moreover, the data collected has potentially significant value, for example, in predicting and understanding health parameters' evolution and generating new insights. Therefore, there is great value potential in getting the technology and mechanisms in place to take full advantage of data, improve workflows, and create new services using connected devices and sensors. The digital health use case both controls sensors that are deployed for remote supervision, such as bed sensors, motion sensors, sensors for indoor and out-door location, video-based supervision, so on and various types of medical devices and specific sensors supporting the care and wellness for the specific patient (e.g., blood pressure meter, scales, glucose meter, medicine reminder). In addition, the system needs to integrate with other systems, for instance, to provide information or alarms to response centers, caregivers, physicians, next of kin etc., and to feed information to medical systems such as an Electronic Health Record (EHR) Systems (Kalra, 2006). Figure 3 depicts the typical data flows of this use-case.

Several pipelines may be applied to a citizen. For example, one pipeline may originate from medical devices automatically sending data as the citizen do their measurements. That information is stored and reviewed regularly (e.g., daily basis) by a doctor or a nurse to follow up on the patient. Another pipeline could be related to more critical alarms, e.g., leaving

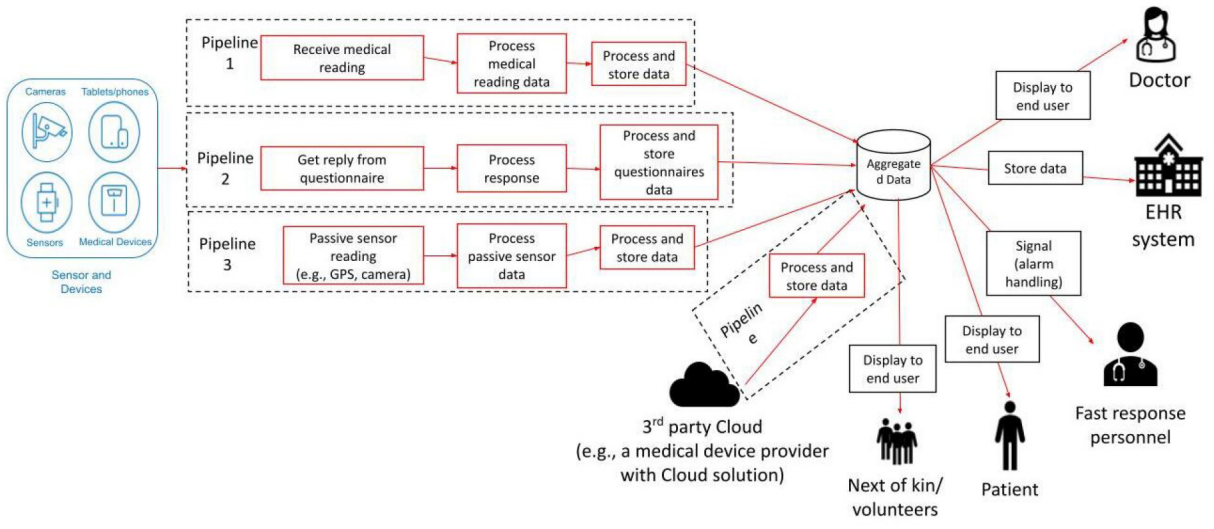

FIGURE 3
Typical data flows of a e-Health system in a Smart City environment.

the bed or a room during night alarm from camera-based sensors, which should be sent to response centers that need to urgently check it out through camera-based supervision or by sending out personnel. A third pipeline may be video-based supervision setting up a secure video stream to watch the patient's room.

To build and manage such systems is a complex task, and there is a set of roles that need to gain proper tool support to enable them to manage their tasks, in particular:

- **DevOps Engineers:** full-stack developers and system operators (in a DevOps set-up) develop, test, deploy and operate data pipelines. It has the ability to describe a big data (BD) pipeline potentially using a graphical interface, including reuse of sub-components and previously defined pipelines, and to perform simulation and tests. Manages which pipelines are online and their quality (metrics), and how it affects the overall system (infrastructure).
- **System Architect:** System architects specify the overall end to end data pipelines at the technical level. It has the ability to describe a BD pipeline potentially using a graphical interface, including the reuse of sub-components and previously defined pipelines. Responsible for end-to-end control over the pipelines and how they are applied.
- **Product Managers:** Provides domain expertise (e-health) for the product (e.g., what kinds of data flows are needed in the pipeline). Puts forward requirements on what is needed for a client or product. Managing product development and evolution. Take part in testing. Has some technical skills.
- **Business Developer** - Domain expert with technical background. It has the ability to describe a BD pipeline using a graphical interface, including the reuse of sub-components and previously defined pipelines.

The envisioned coordinated system combines the specific needs and constraints to define and deploy data pipelines in production environments and the flexibility to adapt to the specific requirements for this use case. In particular, at the application level, this coordinated system facilitates the development of deployment and operation of data pipelines to help care providers remotely supervise and follow their patients and support them while staying at home during treatment and care as well as supporting patients and elderly to accomplish self-care living at home. At the IoT/Edge level, the system gathers data from medical devices and cameras and does local processing, such as filtering, encryption, anonymization and local storage, for security and privacy reasons. At the Cloud level, the system manages and orchestrates data pipelines for a set of SaaS applications for patients and health personnel and integrates and exchanges data to third party services such as national Electronic Health Record Systems, response centers, and big data analytic tools that for example can predict diseases or pandemic evolution. The envisioned coordinated system also enables handling the scaling of sensors and devices needed to scale to potentially hundreds of thousands of homes.

Four essential functional requirements have been summed up after a thorough investigation of the two use-cases above. Those functional requirements are — 1. Adequate computing and network infrastructure management to offer better smart service experiences for the Smart City residents; 2. Fair, a traceable and transparent market mechanism for monetising resources (e.g., computing, data, apps); 3. Ensures the highest

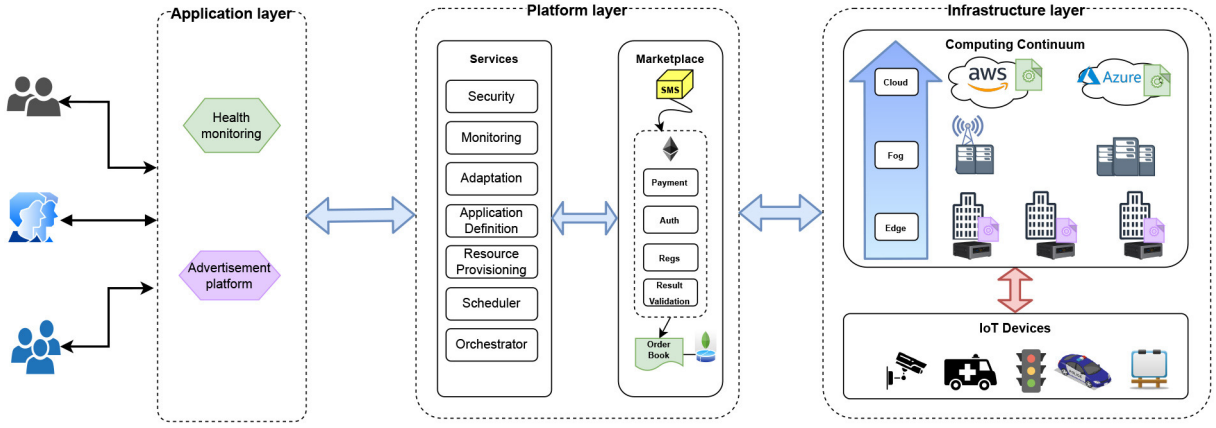

FIGURE 4
Proposing architectural framework for managing the Big Data pipelines in Smart Cities.

level of trustability to authenticate data, application and computing resources; 4. Finally, enable the privacy-preserving computation to protect the participants' confidentiality and restrict unauthorized data access. Significantly these four essential requirements build the foundation for proposing a new architectural framework in the Smart City domain. The architectural framework has been thoroughly described in the next section.

4. Proposed architecture

We found potential challenges in Smart Cities by studying the use-cases. In order to overcome these concerns, we proposed a three-layer computer architecture, as shown in Figure 4. The proposed architecture splits into three layers: 1. *Application*, 2. *Platform*, and 3. *Infrastructure*. The *Application* layer is acting as the gateway between the end-users and our envisioned computing framework. End-users interact with the proposed computing framework by using smart applications (e.g., e-Health app, e-Governance app) that reside in the *Application* layer. Most importantly, in this architectural framework, one of our main contributions is to develop the *Platform* layer by integrating the Blockchain-based resource marketplace and different computing services (e.g., resource provisioning, scheduling, orchestrating, security and privacy provisioning etc.). Taking the advantages of Blockchain technology and modern privacy-enhancing techniques (e.g., Trusted Execution Environment, Homomorphic Encryption, Federated Learning), this *Platform* layer enables a secure and privacy-preserving computing framework for executing various smart applications over the resources of the *Infrastructure* layer. By summarizing the functionalities of each layer, a detailed description of the individual layer has been presented below.

4.1. Application layer

In our proposed computing framework, the *Application layer* is acting as the gateway for the end-users (e.g., subscribers). This layer comprises several smart applications that various use-case owners have built to provide various intelligent services in the Smart Cities paradigm. However, the deployment of smart applications (e.g., e-Health apps, e-Governance apps, Digital marketing) in a heterogeneous ecosystem (e.g., Smart Cities) is challenging. We have already explored the requirements for execution of the respective applications and their accompanying Big Data pipelines in the prior Section 3 by thoroughly outlining two prospective use-cases. However, meeting all of the application's requirements for the operating system (OS), libraries, and other resources (such as data and computation) is complex. Container-based application deployment is one of the most popular techniques for dealing with this problem. Moreover, container technology is one of the most popular ways of delivering applications in our envisioned architectural framework because of its ability to encapsulate packages and dependencies. On the other hand, different components of the *Platform layer* assist in on-boarding apps by selecting appropriate resources, and those components are in charge of effectively managing the life-cycle of deployed applications. We will describe how the *Platform layer* components can help to efficiently deploy and run different apps in our proposed framework in the next part by thoroughly explaining the functionality of the different components of the *Platform layer*.

4.2. Platform layer

Depending on the requirements of the use-cases and following the Smart City paradigm, we have realized that it

is essential to build a composite middleware (*Platform* layer) for developing an intelligent and privacy-preserving computing framework. Notably, this *Platform* layer can ensure two different aspects in our envisioned computing architecture. First, the *Services* block is in charge of the secure and proper execution of various smart applications. Whereas, the *Marketplace* helps the application owners to rent the appropriate resource(s) for executing some tasks and offering the services among the end-users.

4.2.1. Services block

In the suggested architecture, the platform layer allows the deployment, execution and management of the smart applications and big data pipelines over the resources provided by the infrastructure layer. To support these actions, the following building blocks are required:

- **Application Definition:** The first step toward deploying and managing a smart application is the proper definition of the application and the parts composing it. As the applications are distributed, we consider using a detailed model that allows the definition of applications as graphs composed of interconnected but independent components. In addition to using such a model, a domain-specific language (DSL) is used to abstractly express details about the application placement and behavior. The DSL includes soft and complex requirements, adaptation policies, and SLOs to be achieved.
- **Scheduler:** The scheduler is responsible for identifying the needs of the smart application and proceeding with the deployment. The scheduler requests or chooses the appropriate deployment environment for each component based on the application graph and any constraints defined using the application definition model or DSL.
- **Resource Provisioner:** Resource provisioner is responsible for providing the proper environment for the execution of the application components. In general, usage of Virtual Machines or containers for the applications is suggested, so using a Cloud provider (e.g., AWS EC2, Azure) API or container management (e.g., Kubernetes) API is required. For the proposed architecture, we go one step further by suggesting the usage of a Blockchain-based resource marketplace that provides a decentralized trusted execution infrastructure for the application.
- **Orchestrator:** Orchestrator is a service that uses a control loop that has a primary responsibility to perform the deployment procedure of all application components, constantly ensuring the proper functioning of the components and the whole application, and manage the maintenance of it.
- **Monitoring:** Monitoring process aims to collect and utilize available monitoring data regarding resource utilization from the underlying infrastructure (e.g., compute, memory, network) and the behavior from the deployed smart application by tailored application-level metrics (e.g., throughput, active users). Although both push-based and pull-based approaches for collecting metrics can be used, we consider the pull-based approach for such a distributed deployment of the smart apps. A pull-based approach (e.g., by using Prometheus[1]) is more accessible to implement in a secure and trusted way by using encryption over HTTP and better controlling the authenticity of the monitoring nodes.
- **Adaptation:** Adaptation process is responsible for maintaining the desired state of the application based on the description of the application, using the DSL or using custom policies. Notably, that is achievable by enforcing specific rules of action when the hard and soft constraints are not achieved.
- **Security:** In the Smart City domain, many small and large resource providers participate in building a large-scale ICT system, where ensuring network security is a necessary requirement. In our proposed framework, the Security process is dedicated to offering high network security and helps to build a secure overlay network over the infrastructure layer. Also, this process block is responsible for controlling user access to confidential data and access to the private data repositories.

4.2.2. Marketplace

Core services of the Marketplace block are developed based on the iExec Marketplace for cloud computing resources (Fedak et al., 2018). This block builds a Blockchain-based decentralized marketplace for software appliances (e.g., Apps, VMs, containers), computing resources (cloud, fog, and edge), and data generators (e.g., sensors, actuators, transportation services, health monitoring sensors). Therefore, monetisation of these resources and software appliances is possible through our proposed framework. Moreover, the Marketplace block helps the application owners to find the appropriate decentralized trusted computing resources for securely running their applications.

This component enables equal opportunities for individuals and small/medium/large companies to monetise their resources (e.g., computing, data and apps); thus, it scale-up the overall infrastructure and makes it highly diverse. The underlying Blockchain technology makes the marketplace entirely traceable and transparent to all its participants; every task execution and corresponding transaction are permanently recorded on the Blockchain nodes. Market mechanisms consider the reputation of the resource providers and their resources: the reputation of resource providers is built based on their past operations and the quality of services they offered in the past. The Blockchain-based marketplace ensures open and fair competition among the

1 https://prometheus.io/

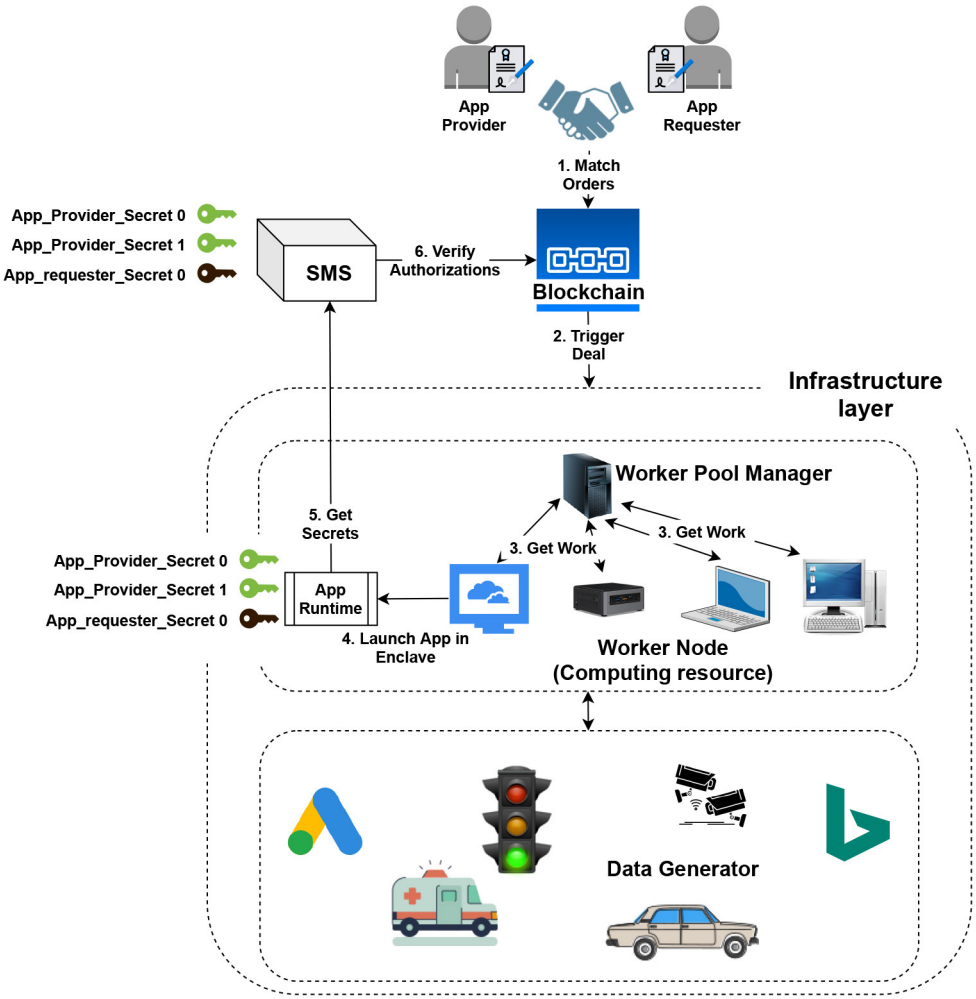

FIGURE 5
Work-diagram: For deal-based resource selection and secure App execution in trusted environment.

various resource providers, which helps consumers choose the appropriate resources and facilitate better resource governance facilities for the overall system.

One of the main objectives of this Blockchain-based Marketplace is to bring more trust among the various participants (e.g., Application/Service developer/owner, Application/Service requester, Resource Provider, etc.) of the Smart City paradigm. Setting up strict rules and policies could drive Smart Cities toward a trusted computing ecosystem. Similarly to the iExec Marketplace, trust is ensured by combining smart contracts and Trusted Execution Environments (TEE). Notably, the Secret management Service (SMS) is one of the pivotal components which helps develop a robust, trusted computing framework in our envisioned ecosystem. It is the secure intermediary between use-case/application owners and task executioners. The SMS is responsible for storing all keys (i.e., secrets) to remote applications running inside identified and trusted enclaves. Significantly, our envisioned platform is able to enforce the different participants to abide by the set rules and policies for executing the tasks in a Smart City domain. The secret provisioning policy of the SMS is developed based on the iExec provided on-chain Access Control Lists (i.e., PoCo). PoCo smart contracts define simple Access Control Lists (ACL) rules, where individuals have ownership of on-chain objects which have been exposed through the Marketplace (e.g., resources, dataset, etc.). In Figure 5, we have presented the functional diagram for selecting the resources from the Blockchain-based Marketplace. The diagram shows how the *Deal* is being done once the request for resources has been matched and the task is being executed in a TEE-based system.

Importantly, in our Blockchain-based marketplace the deal marks the beginning of the task execution. Once the deal has been fixed between the requester and provided then, the

appropriate resources (e.g., computing resources, data, etc.) has been chose to execute the user requested applications. Indeed the SMS is responsible for making sure that only the authorized actors (e.g., application) access the specific sensitive data which is already residing inside the enclave of TEE. Finally, after successfully executing the corresponding the tasks of that authorized applications, the result is being return to the given requester. By following all these steps SMS along with TEE and Smart Contract helps to build a trustworthy, privacy-preserving computing platform in Smart City domain.

4.3. Infrastructure layer

Thoroughly investigating two use-cases and considering the Smart City scenarios, we realized that besides the networking devices and functions, two different kinds of elements are responsible for developing the *Infrastructure layer* in our envisioned architectural framework. First, different types of heterogeneous computing resources are participating in Smart Cities. Notably, all of those computing resources have been provided by private or public cloud providers, small and medium fog/edge providers, and individuals. Most importantly, combining all of these computing resources build the *Computing Continuum* in our envisioned architecture. Besides the computing resources, IoT devices (e.g., personalized healthcare devices, smartwatch), various web of things elements, i.e., pay-per-click advertising platforms (e.g., Google Ads, Bing Ads, Etc.), public digital infrastructure devices (e.g., digital display boards, traffic lights) are playing pivotal roles in successfully running the smart applications in a Smart City paradigm by generating a massive amount of data. In our proposed architecture, these IoT devices, web of things elements, and public digital infrastructure devices are considered as the *Data Generator*.

For satisfying the DSL requirements of the smart applications and corresponding Big Data pipelines; the *Platform layer* components, more precisely the *Services* and the Blockchain-based *Marketplace* combinedly help to find the appropriate and suitable resources (i.e., computing and data) from the *Infrastructure layer*. Our proposed Blockchain-based *Marketplace* not only helps to find adequate resources but also can provide the TEE-enabled computing resources for ensuring secure, trusted and auditable tasks execution mechanisms to develop a robust and sustainable privacy-preserving trustworthy computational framework in Smart Cities.

5. Concluding remarks

In this paper, we outlined the importance of data as a key enabler as well as a major source of concern in designing and deploying Smart City applications. Through example use cases in the area of Smart City applications (including Smart Cities services optimisation based on innovative marketing data and digital healthcare), we showed that data is an essential element that drives the entire application development lifecycle. At the same time, mishandling the data in such scenarios can potentially have disastrous consequences in terms of unreliable data and data privacy. Therefore, data processing scalability and data authenticity and confidentiality were identified as critical requirements in designing and deploying trustworthy Smart City applications in this context. We proposed an architectural framework based on combining modern technological paradigms and novel peer-to-peer distributed networks to address such requirements. In the proposed architectural framework, we argued for the need to connect Blockchain technologies with modern data analytic techniques (e.g., Federated Learning) and Edge/Fog computing as the baseline for ensuring data processing scalability and data authenticity and confidentiality in Smart Cities applications.

The proposed architectural framework is exemplified in the context of Smart Cities applications in this paper; however, its applicability goes beyond the domain of Smart Cities, in areas where data processing over heterogeneous and untrusted resources across the Computing Continuum is needed. In fact, elements of the proposed architectural framework are currently being implemented in a EU funded H2020 project[2] — a research and innovation project for creating a novel paradigm for Big Data pipeline processing over heterogeneous resources encompassing the Computing Continuum, covering the complete lifecycle of managing Big Data pipelines, from data pipelines discovery, design, simulation, to provisioning, deployment and adaption.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

AS-B and SS have equal and significant contributions. ASo, ASi, DR, FP, GL, and IP contributed to describing the use-cases and some parts of the envisioned architecture. All authors contributed to the article and approved the submitted version.

2 https://datacloudproject.eu

Funding

This work received partial funding from the projects DataCloud (H2020 101016835), BigDataMine (NFR 309691), and SINTEF SEP-DataPipes.

Conflict of interest

Authors AS-B and SS are employed by iExec Blockchain Tech. Author ASo is employed by Tellu. Author ASi was employed by Aker Solution and currently holding IEEE Membership. Author DR is employed by SINTEF. Author FP is employed by JOT Internet Media. Authors GL and IP are employed by UBITECH.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Al Nuaimi, E., Al Neyadi, H., Mohamed, N., and Al-Jaroodi, J. (2015). Applications of big data to smart cities. *J. Internet Serv. Appl.* 6, 1–15. doi: 10.1186/s13174-015-0041-5

Alamouti, S. M., Arjomandi, F., and Burger, M. (2022). Hybrid edge cloud: a pragmatic approach for decentralized cloud computing. *IEEE Commun. Mag.* 1–28. doi: 10.1109/MCOM.001.2200251

Allam, Z., and Dhunny, Z. A. (2019). On big data, artificial intelligence and smart cities. *Cities* 89, 80–91. doi: 10.1016/j.cities.2019.01.032

Bélissent, J. (2010). *Getting Clever About Smart Cities: New Opportunities Require New Business Models*. Cambridge, MA: Cambridge University.

Bonomi, F., Milito, R., Natarajan, P., and Zhu, J. (2014). "Fog computing: a platform for internet of things and analytics," in *Big Data and Internet of Things: A Roadmap for Smart Environments* (Springer), 169–186.

Church, K., Greenberg, A. G., and Hamilton, J. R. (2008). "On delivering embarrassingly distributed cloud services," in *HotNets* (Calgary, CA: Citeseer), 55–60.

Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J. -H., Metayer, D. L., Tirtea, R., et al. (2015). Privacy and data protection by design-from policy to engineering. *arXiv preprint arXiv:1501.03726*. doi: 10.48550/arXiv.1501.03726

Dolui, K., and Datta, S. K. (2017). "Comparison of edge computing implementations: fog computing, cloudlet and mobile edge computing," in *2017 Global Internet of Things Summit (GIoTS)* (Geneva: IEEE), 1–6.

Fedak, G., Bendella, W., and Alves, E. (2018). *Iexec: Blockchain-based decentralized cloud computing*. Technical report. Available online at: http://iex.ec/wp-content/uploads/pdf/iExec-WPv3

Ghobaei-Arani, M., Souri, A., and Rahmanian, A. A. (2020). Resource management approaches in fog computing: a comprehensive review. *J. Grid Comput.* 18, 1–42. doi: 10.1007/s10723-019-09491-1

Johansen, S. K. (2018). *A Comprehensive Literature Review on the Blockchain as a Technological Enabler for Innovation*. Mannheim: Department of Information Systems, Mannheim University.

Kalra, D. (2006). Electronic health record standards. *Yearb Med. Inform.* 15, 136–144. doi: 10.1055/s-0038-1638463

Konečnỳ, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., and Bacon, D. (2016). Federated learning: strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*. doi: 10.48550/arXiv.1610.05492

Lee, R., Jang, R. -Y., Park, M., Jeon, G.-Y., Kim, J. -K., and Lee, S. -H. (2020). "Making iot data ready for smart city applications," in *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)* (Busan: IEEE), 605–608.

Li, W., Song, H., and Zeng, F. (2017). Policy-based secure and trustworthy sensing for internet of things in smart cities. *IEEE Internet Things J.* 5, 716–723. doi: 10.1109/JIOT.2017.2720635

Maciel, P. D., Verdi, F. L., Valsamas, P., Sakellariou, I., Mamatas, L., Petridou, S., et al. (2019). "A marketplace-based approach to cloud network slice composition across multiple domains," in *2019 IEEE Conference on Network Softwarization (NetSoft)* (Paris: IEEE), 480–488.

Masip-Bruin, X., Marín-Tordera, E., Tashakor, G., Jukan, A., and Ren, G. -J. (2016). Foggy clouds and cloudy fogs: a real need for coordinated management of fog-to-cloud computing systems. *IEEE Wireless Commun.* 23, 120–128. doi: 10.1109/MWC.2016.7721750

Ning, Z., Zhang, F., Shi, W., and Shi, W. (2017). "Position paper: challenges towards securing hardware-assisted execution environments," in *Proceedings of the Hardware and Architectural Support for Security and Privacy*, 1–8.

Pahuja, N. (2021). "Smart cities and infrastructure standardization requirements," in *Solving Urban Infrastructure Problems Using Smart City Technologies* (Elsevier), 331–357.

Perera, C., Liu, C. H., and Jayawardena, S. (2015). The emerging internet of things marketplace from an industrial perspective: a survey. *IEEE Trans. Emerg. Top. Comput.* 3, 585–598. doi: 10.1109/TETC.2015.2390034

Perera, C., Qin, Y., Estrella, J. C., Reiff-Marganiec, S., and Vasilakos, A. V. (2017). Fog computing for sustainable smart cities: a survey. *ACM Comput. Surveys* 50, 1–43. doi: 10.1145/3057266

Rajaraman, V. (2014). Johnmccarthy–father of artificial intelligence. *Resonance* 19, 198–207. doi: 10.1007/s12045-014-0027-9

Rejeb, A., Keogh, J. G., and Treiblmaier, H. (2020). How blockchain technology can benefit marketing: six pending research areas. *Front. Blockchain* 3, 3. doi: 10.3389/fbloc.2020.00003

Sambasivan, N., Kapania, S., Highfill, H., Akrong, D., Paritosh, P., and Aroyo, L. M. (2021). ""Everyone wants to do the model work, not the data work": data cascades in high-stakes AI," in *proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15.

Schieferdecker, I., Tcholtchev, N., Lämmel, P., Scholz, R., and Lapi, E. (2017). "Towards an open data based ict reference architecture for smart cities," in *2017 Conference for E-Democracy and Open Government (CeDEM)* (Krems: IEEE), 184–193.

Shi, W., and Dustdar, S. (2016). The promise of edge computing. *Computer* 49, 78–81. doi: 10.1109/MC.2016.145

Sinaeepourfard, A. (2017). *Hierarchical Distributed Fog-to-Cloud Data Management in Smart Cities*. Universitat Politècnica de Catalunya.

Sinaeepourfard, A., García Almiñana, J., Masip Bruin, X., and Marín Tordera, E. (2017). "Fog-to-cloud (f2c) data management for smart cities," in *Proceedings of 2017 Future Technologies Conference (FTC): 29-30 November 2017, Vancouver, Canada* (Vancouver, BC: The Science and Information (SAI) Organization), 162–172.

Sinaeepourfard, A., Krogstie, J., and Petersen, S. A. (2021). *Zen Data Management and Monitoring. Requirements and architecture*. Trondheim: SINTEF Academic Press.

Sinaeepourfard, A., Krogstie, J., and Sengupta, S. (2020). Distributed-to-centralized data management: a new sense of large-scale ICT management of smart city iot networks. *IEEE Internet Things Mag.* 3, 76–82. doi: 10.1109/IOTM.0001.1900038

Ullah, Z., Al-Turjman, F., Mostarda, L., and Gagliardi, R. (2020). Applications of artificial intelligence and machine learning in smart

cities. *Comput. Commun.* 154, 313–323. doi: 10.1016/j.comcom.2020.02.069

Valadares, D. C. G., Will, N. C., Spohn, M. A., de Souza Santos, D. F., Perkusich, A., and Gorgonio, K. C. (2021). "Trusted execution environments for cloud/fog-based internet of things applications," in *11th International Conference on Cloud Computing and Services Science CLOSER* (Prague), 111–121.

Valancius, V., Laoutaris, N., Massoulié, L., Diot, C., and Rodriguez, P. (2009). "Greening the internet with nano data centers," in *Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies* (Rome), 37–48.

Waleed, M., Latif, R., Yakubu, B. M., Khan, M. I., and Latif, S. (2021). T-smart: trust model for blockchain based smart marketplace. *J. Theor. Appl. Electron. Commerce Res.* 16, 2405–2423. doi: 10.3390/jtaer16060132

Xu, Q., Arafin, M. T., and Qu, G. (2021). "Security of neural networks from hardware perspective: a survey and beyond," in *2021 26th Asia and South Pacific Design Automation Conference (ASP-DAC)*, (Tokyo: IEEE), 449–454.

Yi, S. (2020). Reform of sports health supermarket based on ar technology. *Front. Sport Res.* 2, 020511. doi: 10.25236/FSR.2020.020511