2017/18

AN ANALYSIS OF THE EMERGING DIGITAL TECHNOLOGIES AND THEIR APPLICATION IN THE SMART CITY CONTEXT

JAGADISH BALAJI SELVAM - 873302

SUPERVISOR: DAVIDE CHIARONI

POLITECNICO DI MILANO

SCHOOL OF INDUSTRIAL AND INFORMATION ENGINEERING

MASTER OF SCIENCE (MSC.) IN ENERGY ENGINEERING

ACADEMIC YEAR 2017-2018

(This page has been intentionally left blank)

Acknowledgement

First and foremost, I would like to express thanks to my parents for their unconditional love and support throughout my life and through my hardship in the last two years. A very big thank you to the two of you for believing in me and giving me the strength to reach for the stars and chase after my mighty dreams. I would also like to render my thanks to my grandparents, uncle, aunts, and cousins who have always rooted for my success.

I would not have been able to complete this thesis successfully without the guidance of Francesca Capella and Davide Perego. I would love to express my deepest gratitude to them for their exceptional guidance, patience and for providing me with a supportive rapport. I would also like to extend my thanks to the whole of Politecnico di Milano and all of its staffs for providing me support throughout my duration of the study.

I would like to thank all my beloved friends in PoliMi and back home, especially, Divya, Julia, Jelena, Gioia, and Siddhant who stood by me during my times of hardships and supported and encouraged me to get through it. Each of you have made my life a wonderful experience.

This thesis marks the end of my University life and the beginning of my real-life journey.

Finally, I would like to leave the remaining space in loving memory of an aspiring young doctor, my beloved friend Joel Rajkumar (27th April 1995 – 13th July 2015), who will not be there anymore to celebrate with my successes and support my failures.

Contents

List of Figures

List of Tables

Abstract

With the urban population growing by 60 million each year due to rapid urbanization it is crucial to mitigate the issues that might come along with it such as overcrowding, increased energy consumption and so on. Smart Cities could help in addressing these issues as they aim to make use of all the interconnected information available today to better understand and control its operations and optimize the use of limited resources. This is achieved collectively using smart technologies such as the Internet of Things (IoT), Big Data Analytics, Artificial Intelligence (AI), Machine Learning (ML), and Blockchain Technology. Where Internet of Things help in sensing and generating data while Big Data Analytics helps in processing the data and creating value from it which can be used for a multitude of applications. Artificial Intelligence and Machine Learning help in automation of time and labor-intensive processes and also helps in prediction which is useful in Analytics. Blockchain provides safe, private and secure ways for data transaction.

For a thorough analysis the Individual components of each technology were studied – For IoT, architectures and different models of communication; for Big Data, its characteristics, value creation journey and data software platform architecture, the functional and non-functional requirements and the methods of integrating AI and ML into the platform; and for Blockchain Technology, the different categories, architecture, consensus models and smart contracts. For a broader scrutiny, the plausibility of intertwining between the technologies was explored based on interoperability, scalability, privacy, security, and safety. Upon analyzing the technologies and their various applications, it can be concluded that the individual components of each technology influence the effectiveness of the Smart City and so does the interplay between each technology. From a broader perspective, these technologies are certainly anchoring the creation of a highly functional smart city. However, as with the evolution of any technology, more investigation is needed for an in-depth understanding of these technologies and thus for enhancing the development of a Smart City.

Keywords: Internet of Things, Big Data Analytics, Artificial Intelligence, Machine Learning, Blockchain Technology, Smart City, urbanization, architecture

Abstract (Italian)

Con una popolazione urbana in crescita di 60 milioni all'anno a causa della rapida urbanizzazione, risulta di fondamentale importanza attenuare i possibili problemi ad essa correlati, come ad esempio sovraffollamento e aumento del consumo di energia. Parte della risposta a questi problemi potrebbe essere rappresentata dalle Smart Cities, realtà basate sull'utilizzo delle informazioni interconnesse disponibili, al fine di ottimizzare la gestione della città stessa e l'utilizzo delle risorse di cui dispone. Questo risultato viene raggiunto attraverso tecnologie come Internet of Things (IoT), Big Data Analytics, Artificial Intelligence (AI), Machine Learning (ML) e Blockchain Technology. Internet of Things aiuta a rilevare e generare dati, elaborati poi attraverso Big Data Analytics, andando così a creare un valore aggiunto effettivamente monetizzabile. intelligenza artificiale e apprendimento automatico aiutano invece nell'automazione dei processi a uso intensivo di tempo e di lavoro, facilitando anche la stesura di previsioni, utilizzabili poi nella fase di analisi. Blockchain invece fornisce modi sicuri, privati e affidabili per la transazione dei dati.

Al fine di condurre un'analisi più approfondita, sono stati inoltre studiati i componenti individuali di ciascuna tecnologia: per IoT, architetture e diversi modelli di comunicazione; per i Big Data, le sue caratteristiche, il percorso di creazione di valore e l'architettura della stessa piattaforma software di dati, oltre ai requisiti funzionali e non funzionali e i metodi di integrazione di AI e ML nel software; per la Blockchain Technology, invece, sono stati presi in esame le diverse categorie, l'architettura, i modelli di consenso e i contratti intelligenti. Al fine di presentare un'analisi più ampia, è stata poi esplorata la plausibilità dell'intreccio tra le suddette tecnologie, analizzando l'interoperabilità, la scalabilità, la privacy, la sicurezza e l'affidabilità di tale possibile accoppiamento.

In conclusione, analizzando le tecnologie e le loro varie applicazioni, si evince come i componenti di ciascuna tecnologia influenzano l'efficacia della Smart City, sia presi singolarmente che adottando un loro utilizzo congiunto. Si può quindi affermare con certezza che queste tecnologie stanno preparando il campo alla creazione di una Smart City altamente funzionale. Tuttavia, come d'altronde per l'evoluzione di qualsiasi tecnologia, sono necessarie ulteriori indagini al fine di una più approfondita comprensione delle stesse, e quindi per migliorare poi lo sviluppo della Smart City.

Keywords: Internet of Things, Big Data Analytics, Artificial Intelligence, Machine Learning, Blockchain Technology, Smart City, urbanizzazione, architettura

Methodology

For acquiring a basic understanding of the topic involved, firstly, reports and review articles which explained the concepts of Smart Cities, Internet of Things, Big Data Analytics, Machine Learning, Artificial Intelligence and Blockchain Technology were studied. Next, review articles which discussed the mechanisms of the technologies were shortlisted based on the year of publication (focusing more on recent publications) and relevant to the topics at hand.

Using these review articles, the original papers, journals and web articles were backtracked, and the mechanisms were studied in detail. Next, articles which discussed the application of the different technologies in a Smart City were selected and compared to understand the different approaches possible. Then, papers describing and discussing the perspective of the application of the various technologies in the energy sector of a Smart City were studied. Papers corresponding to each technology were analyzed holistically for two different aspects:

1) For the individual components of the technology and how it would contribute to the building of a Smart City

2) For the interplay between these technologies and the benefits and limitations that are put forward by them in the context of a Smart City.

Out of these papers, other commonly mentioned factors including the benefits, challenges and limitations were also selected to be used in the thesis.

Finally, from the numerous journals, reports, review papers and web articles concerning the technological concepts involved in the thesis, only those which are relevant to the topic of the thesis were focused on. All the collected information was made to fit into pre-formed outline for the thesis. Any lacking information was obtained by backtracking from the selected articles or textbooks. Images were chosen based on how clearly, they convey the intended message without portraying excessive information.

Introduction

Modern-day technology has been growing exponentially over the last few years and so has been the incorporation of them in the day-to-day lives. As per the estimation of the United Nations global population would double by the year 2050 and the urban population is growing by 60 million each year. With these rates of urbanization, it would be inordinately difficult to keep the systems of livelihood, functioning with the already existing system. This demand drives the need for new and innovative methods to handle the complexities of urban living. Few of the critical issues that the large-scale urbanization is overcrowding, increased energy consumption, strenuous resource management, and poor environmental protection measures and these need to be tackled in a well-organized manner. An acceptable and insightful solution would be to turn the cities into Smart Cities. A Smart City would essentially be, a conglomeration of concepts which could help in accommodating the problem of urbanization.

Since 1990's the term Smart City has been surfacing in various works of literature but only in 2010, it gained popularity due to IBM's Smarter Cities Challenge. Right after the global financial crisis, IBM sent experts to various cities which they had targeted for their technology offer for city infrastructure and local governments. They proposed solutions which would eventually make the cities "smarter and more effective". According to IBM, a Smart City is defined as "one that makes optimal use of all the interconnected information available today to better understand and control its operations and optimize the use of limited resources". And Cisco defines it to be those cities which adopt "scalable solutions that take advantage of information and communication technology (ICT) to increase efficiencies, reduce costs, and enhance the quality of life."

From the time of inception of the concept of Smart Cities the spurts of familiarization driven by the various projects can be categorized into three waves. Large technology companies drove the ***first wave*** of Smart Cities as these companies focused on the technological component which is the key element to their conception of smart cities. They targeted systems which laid the foundation of a big city such as energy, water, and transport. The critical outcomes that a Smart City focuses on are resource efficiency, improved decision making, and so on. The definition of a Smart City has been strongly debated ever since.

The conception of the 'Smart City' was influenced by the broader aspects of city living and aspects such as governance, education, and inclusion. This eventually brought citizens into the picture and their engagement noticeably happened to be the ***second wave*** of smart

cities. Citizens were reached out by local authorities, especially in Europe, through digital platforms, open data portals, civic crowdfunding, hackathons, innovation competitions, co-design and living labs and more.

The various companies started making use of the city as a platform to establish their own markets during the ***third wave***. Previously the Smart City was conceived as a reimagining of city services through the technological transformation of large-scale traditional infrastructure. But transforming incumbent infrastructure and processes have proven to be extremely difficult. In these circumstances, companies are disrupting old business models and bypassing old systems, often delivering directly to citizens, or in this context, consumers.

Whereas, according to the Directorate General for Internal Policies EU, Smart cities are defined along six dimensions as shown in the figure.

Figure 1 - Six Dimensions of a Smart City

Despite the several bemused ways of defining what a Smart City is, at its core, Smart Cities are supposed to aid the conception and interdependence of human capital, social capital and Information, and Communication Technology (ICT) infrastructure so as to generate appreciable and more sustainable economic development and a better quality of life.

The ultimate purpose of Smart Cities is to improve the quality of life by making use of the various streams of data and the different digital Smart Technologies. Real-time data are comprehensive and provide agencies the access to scrutinize the various events as they occur and allow them to understand the changes in the patterns of demand and supply. Therefore, this allows them to act on it considerably faster which makes the solution relatively low in cost.

The nature and economics of the infrastructure of a city are subjected to change due to the use of Smart Technologies. As they eventually reduce the cost involved with the collection of information about the usage patterns and with a revolutionary amount of data sources in their disposal, city governments, employers and residents will be able to figure out new and innovative methods to optimize existing systems and also create new ones.

Governments are constantly faced with implications of the digital transformation of urban life that is under place. There are several steps taken by them to protect their citizens and their data in the newly created environment. They go through a lot of pressure just to enhance the positive impact of the technologies and to safeguard the people from the negative effects of it.

Some smart solutions in addition to responding to demand also involve the public in shaping it. Some of the beneficial uses are that they advise the people to use transit during off-peak hours, and also to change them to use less energy and water. They also contribute considerably to the healthcare system as they promote preventive self-care. All of the uses collectively optimize the city to be more liveable and also makes it a more productive place for businesses to operate.

Globally, tech companies are making use of the universality of digital technology to achieve rapid scaling and by doing so they are eventually transforming numerous cities around the world. There are many start-ups which are built upon digital connectivity and infrastructure, such as taxis, food deliveries, laundry, travel planning, accommodation, and so on.

According to Economist and energy visionary Jeremy Rifkin, the Third Industrial Revolution would be based on the convergence of communication and technology and would promote a radical sharing economy, and it has already begun to take place. He states that China and Europe are already forerunning the revolution as they have begun to make radical changes to their policies and institutional frameworks.

The strategies for Smart Cities that exist today are not suitable enough to provide guidance to the governments on the issues discussed. The function of a Smart City can be essentially dissected into three layers. Each layer has a technology which is integrated into it and plays a pivotal role.

The first layer is formed by the different sensors and devices which are capable of generating data across the city which are connected using high-speed communication networks to transmit the generated data and also open data portals. The data generated consist of variables such as the flow of traffic, consumption of energy, quality of air and many other

aspects of daily life which are granted ease of access. This idea of connecting the different sensors and accessing data generated by them is the fundamental basis of the technology of the Internet of Things. The mechanism, possibilities, and the applications are discussed in detail in Chapter 1.

The second layer deals with the creation of value from the data that is generated by the first layer. The raw data generated is analysed by the Big Data Analytics Platform and suitable outcomes are derived from it which are then translated into alerts, insights, and actions of the system linked with it. Furthermore, Machine Learning and Artificial Intelligence are integrated within the Big Data Analytics Platform to help in automating and increasing the efficiency of the analytics and predictions that are carried out. These technologies, their functionalities and use cases are collectively discussed in detail in Chapter 2.

Figure 2 - Three layers of Smartness
Source: McKinsey Global Institute

The third layer is constituted by the numerous application that is used by the public to access the data. A key criterion for success among these applications is the ability to adapt and change their behaviour according to the needs of the users, usually placing the individual user in the driving seat and providing them more transparent information which would allow them to choose better. However, few of the main concerns with regards to data is its safety, security, and privacy, all of which are constantly subjected to the risk of being hacked, stolen or tampered with. A novel solution to tackle this problem could be the implementation of Blockchain Technology into the different layers of the Smart City which would help in

securing the data due to its almost impossible to tamper with mechanism making it trustworthy. This has been detailly discussed in Chapter 3.

1. Internet of Things

1.1 Overview

The IOT is an influential topic that has been in the limelight over the past few years which has been known for its immense potential that can be readily exploited by the different players of the technology industry. The ripples of this tech-wave have been seen to have a strong influence in the engineering fields, policies, innovation and development projects, etc. The technology provides new capabilities that were not possible in that past by taking advantage of the development in computing power, miniaturization of electronics and interconnection between networks and integrating them into a wide range of sensors, systems, and networked products. There has been a collective growth in technology and market trends, the confluence of which has put this technology in the limelight. Owing to the same reason it has become relatively easy to produce smaller devices at a cheaper cost and use them to establish the interconnected networks. The various contributive factors are:

- Ubiquitous connectivity – The emergence of low-cost and high-speed network connectivity on a global level through licensed and unlicensed wireless services and technology help in connecting almost everything.
- Global adoption of IP-based networking – IP has revolutionized the global standards for networking as it provides a well-structured and widely used platform of tools and software that can be integrated into a wide spectrum of devices in a cheap and easy manner.
- Economics of Computing – Moore's Law holds true in this case as the computing power grows constantly as the price points and power consumption is decreasingly correspondingly.
- Miniaturization – With time the size of the devices has been decreasing considerably and this has enabled them to be integrated into even tiny objects. Along with the results of Moore's Law, this has greatly contributed to the futuristic development of small and cheap sensors in devices, which can pave the way to may IoT applications.
- Development in Data Analytics – Analysis methods of the data obtained through the different sensors have gotten a boost with new and efficient algorithms and with the rapid growth of computing power, storage of data and service connectivity to the cloud.
- Leap in Cloud Computing Capabilities – Cloud Computing allows small and distributed devices to interact with powerful back-end and control capabilities as it allows them

to process, manage and store data remotely through networked computing resources.

The extensive implementation of IoT devices assures transformation of everyday life as it is known into a more energy efficient and secure one. New age products like Internet-enabled appliances, energy management devices and home automation components act as the bridge to close the gap between the conventional home and a 'Smart home'. Similarly, systems such as networked vehicles, real-time info from the sensors embedded in roads and pathways, and intelligent traffic systems seem to bring the idea of ultra-connected 'Smart cities' closer to reality. This could help with the reduction of consumption of energy and minimize congestion. IoT is also transforming the agriculture sector, industries, energy production, and distribution by increasing the quantity and quality of the information available along its value chain using sensors connected to networks.

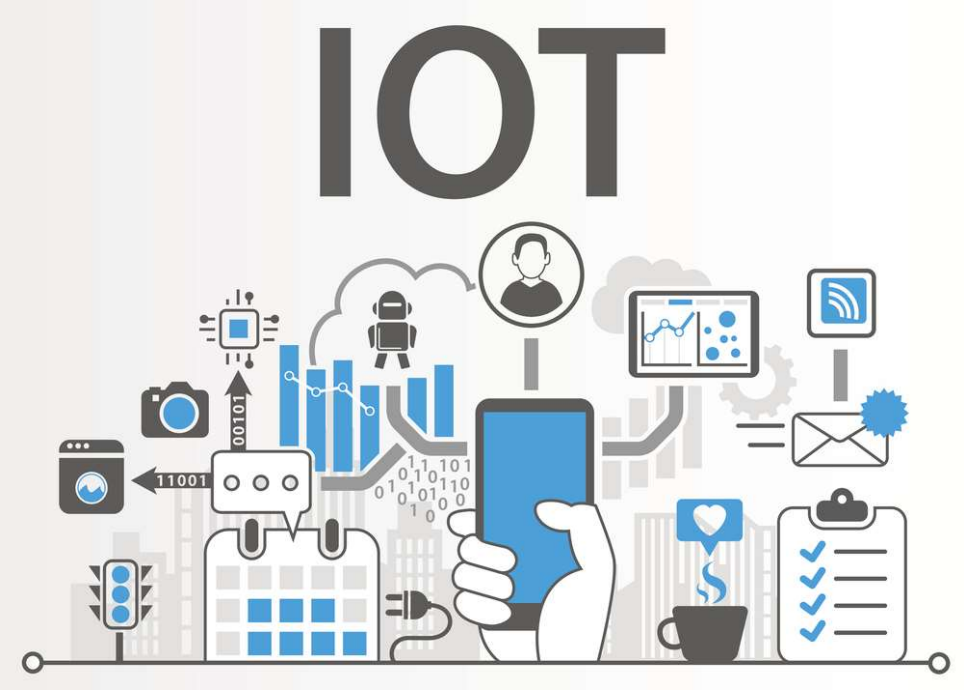


Figure 3 - Internet of Things
Source: Flexware Innovation

There are several research organization and private companies that have released detailed projections as to how the impact of IoT on the Internet and on the economies of the world in the next decade. Cisco, being one amongst them, predicts that there would be more than 24 billion Internet-connected devices by 2019; while Morgan Stanley forecasts at 75 billion connected devices by 2020 and Huawei projects about 100 billion network-connected devices by 2025. In terms of economics, McKinsey Global Institute suggests that the financial impact of IoT on the global economy may be as much as $3.9 to $11.1 trillion by 2025.

Though the dissimilarity between the projections make the numbers questionable, they do not fail to clearly and strongly showcase the growth and developmental influence the technology bestows.

1.1.1 Smart Cities

IoT has been increasingly used in the setting of Smart Cities in the recent past. There have been numerous initiatives and entrepreneurial programs by governments around the globe to promote the concept of Smart City. This is primarily carried out by experimental IoT applications which aim to improve the quality of services, conservation of energy and water, relieve traffic congestion and the quality of life. Cities are a target-rich environment for such applications as they are made of complex infrastructure and largely concentrated populations. Being the engines for global economic growth and with the increased urbanization in developing economies, where 60 percent of the world's population – about 4.7 billion people are estimated to live by 2025, cities have the most to gain.

According to the 2015 report of *the Internet of Things: Mapping the value beyond the hype* from McKinsey Global Institute, the estimated potential economic impact of the Internet of Things on Smart Cities could easily exceed $1.7 trillion per year by 2025. A single major chunk of this impact is expected to be associated with public health applications where water and air monitoring would lead to improved health outcomes. The value generation is estimated to be around $700 billion a year by 2025. However, transportation applications when collectively considered would have a greater impact of up to $808 billion a year by 2025. Traffic applications such as management of real-time traffic flow, smart meters and better and efficient use of public transportation could be worth more than $570 billion a year globally. While about $235 billion could be valued using autonomous vehicles which would reduce traffic accidents, consumption of fuel and carbon emissions.

1.2 IoT Architecture

There exist multiple definitions for IoT, IoT platforms, IoT architectures, and the IoT things. In order for a deeper and uniform understanding for the future of IoT, it is essential to define how IoT and its architecture is perceived to be in its current state. Researchers have been debating over how the smartness of each city can be defined. Conventionally the architecture was composed of three layers namely, the sensor layer, network layer, and the application layer. While the latest of these proposed architectures targeting the security and privacy concerns is made up of five different layers, where there is a processing and business layer in addition to the three conventional layers.

1.2.1 Sensor Layer

The Internet of Things as we know it today encompasses a wide spectrum of "things" which include cyber-physical devices, devices, end-points, entities, and human entities. Each of

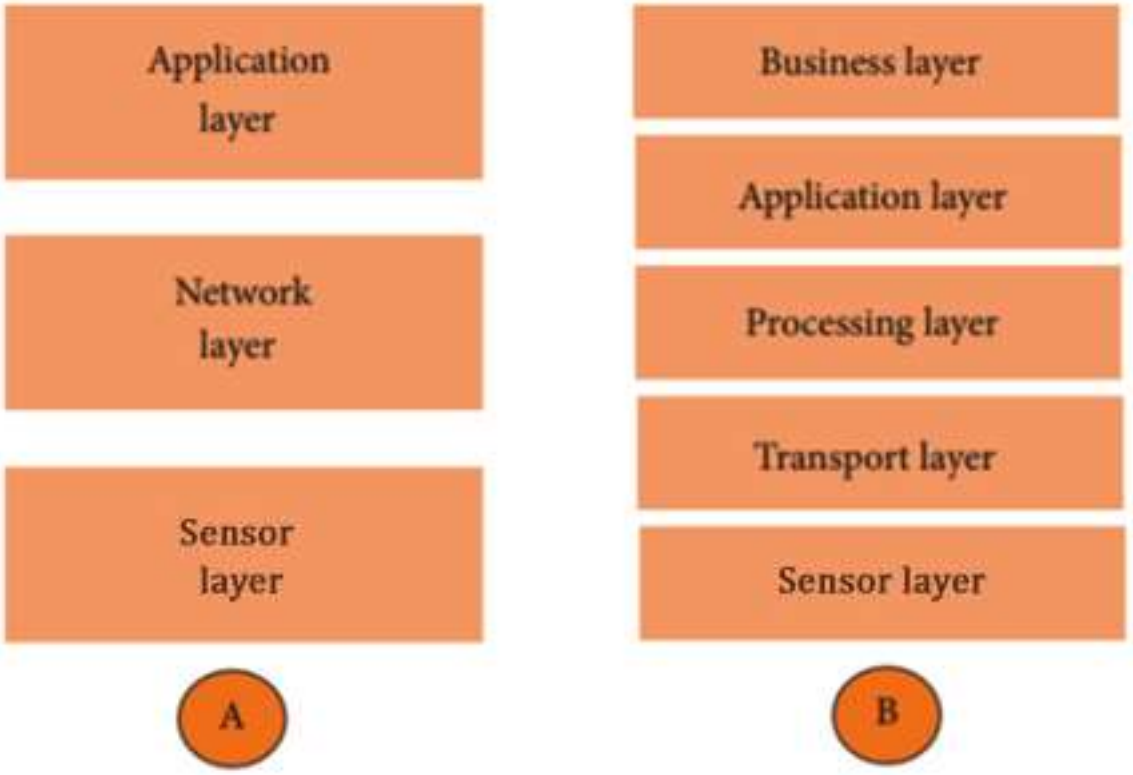


Figure 4 - Architecture of IoT (A: Three Layers) (B: Five Layers)
Source: Sethi, P., & Sarangi, S. R. (2017). Internet of Things: Architectures, Protocols, and Applications

these things regardless of the domain to which they belong to have an attribute in common, which is their individual identity as a physical device/object. These devices/objects usually possess some degree of computational power, either integrated into them or directly attached in the form of sensors. Therefore, this layer is known as the sensor layer.

The sensors function like the eyes, ears, and nose of human beings as it identifies and collects information from the devices. The sensors range from RFID, 2D Barcodes to actuators and controllers and these are chosen based on what the application requires. There are three categories into which sensors can be divided: in situ technical sensors, remote technical sensors and human sensors. Where in situ sensors generally measure in the vicinity of the sensor itself, while remote technical sensors are usually measuring from a distance and human sensors are collected by human-generated measurements. However, it is difficult to distinguish these categories as they usually overlap with each other.

The plethora of data which is gathered by these sensors are generally regarding the location, changes in the environment, air quality, and composition, motion, vibration, etc. This is the prime reason why most threats are related to the sensors and why attackers usually focus on replacing these sensors with their own to use the information that can be obtained for personal benefits. It is prominently accepted that the implementation of these sensors should be in a form of sensor nodes, where the same power source, communication, and processing unit are shared by a number of sensors. This enhances the possibility to collect diverse information from multiple sensors in the same location.

In the case of smart cities, since a huge number of sensors will have to be installed it is vital to anticipate their size and energy demand and plan accordingly. Geomatics, which deals with the collection, analysis, and interpretation of data relating to the earth's surface, plays a noteworthy role in the sensors and sensor systems for smart cities. In order to provide complete data from the sensor measurements which would then help in making decisions, it is vital to determine the spatial component along with the quantitative and qualitative components. Further, the data can also be displayed to the end user in terms of their spatial components.

1.2.2 Network Layer

The layer which bridges the sensor layer and the application layer is the network layer. It is also called as transmission layer as it carries and transmits the data gathered through the sensors of the physical objects. This can be either wireless or wire-based and is also responsible for the connection between smart things, network devices and networks to each other.

Taking into account the feasibility and cost-effectiveness, established technologies that already exist in today's world are given priority. The most commonly used technologies are 3rd Generation (3G), LongTerm Evolution (LTE), Wireless fidelity (Wi-Fi), Worldwide Interoperability for Microwave Access (WiMAX), ZigBee, Z-Wave, Dash7, Near Field Communication (NFC), Radiofrequency identification (RFID), 6LoWPAN and Satellite communications. Based on the kind of area that the networks will have to cover they are classified into on the area they cover, they are usually divided into Home Area Network (HAN), Wide Area Network (WAN) and Narrow-Area Network (NAN).

The monitoring and control systems that are usually found in a household are generally connected using HAN which includes technologies that allow short distance data transmissions such as ZigBee, Dash7, Wi-Fi or network technologies such as Ethernet. WAN is used for transmission over distances longer than that of HAN, usually for communication

between service providers and end users. While NAN is widely used to connect customers' premises to infrastructure substations in smart grids.

The interconnection between the uniquely identifiable embedded computing like devices within the Internet infrastructure is a key feature of IoT. This offers advanced connectivity of devices, systems, and services that extend further than machine-to-machine communications (M2M) and encompasses a variety of domains and applications. As per this concept, the sensor nodes should be energy efficient and be connected wirelessly to form a Wireless Sensor Network (WSN) which would be a way to avoid the high costs and issues that arise from the network cable installations. Due to the security issues concerning integrity and information authenticity that is transmitted in the network, it is highly sensitive to attacks and threats.

Classification	Name of the technology	Maximum Range
Home Area Network (HAN)	RFID, 6LoWPAN, Bluetooth, Z-Wave, Zigbee, NFC	10 cm – 150 m
Narrow Area Network (NAN)	WiMAX, Wi-Fi, Dash7	50 m – 50 km
Wide Area Network (WAN)	3G, LTE, Satellite Communication	5 km – Worldwide

Table 1 - Communication Network Range

1.2.3 Application layer

All the applications in which IoT is used or is deployed in are collectively called as the application layer and it utilizes the Device-to-Application layer gateway model for communication. These applications of IoT include smart homes, smart cities, smart health, animal tracking, etc. It is responsible for providing the various services to the applications which might be different for each application but will depend on the information that the sensor collects. There are numerous issues in the application layer which raises security concerns.

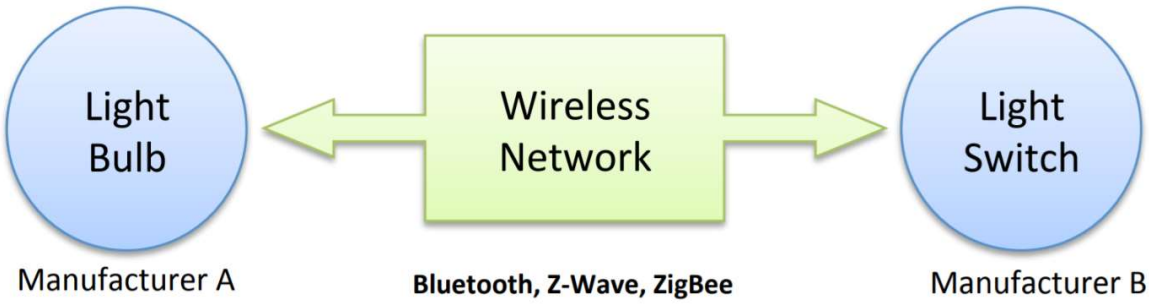
As applications focus on creating smart homes, smart cities, power system monitoring, demand-side energy management, coordination of distributed power storage, and integration of renewable energy generators, the last layer which is the application layer, is the layer in which the information is received and processed.

1.3 IoT – Models for Communication

It is beneficial to understand how the different IoT devices can be networked to communicate with each other and the technicalities involved in it. A guide document which provides details for setting up the architecture of networking of smart devices was released by the Internet Architecture Board (IAB) in March 2015. The different frameworks and its characteristics as mentioned by the guide are discussed below.

1.3.1 Device-to-Device Communications

In this model, two or more devices connect with each other directly and communicate eliminating the need for an intermediary application server. This is achieved through IP networks or the Internet. However, the devices often make use of protocols such as



Source: Tschofenig, H., et. al., Architectural Considerations in Smart Object Networking. Tech. no. RFC 7452. Internet Architecture Board, Mar. 2015. Web. <https://www.rfc-editor.org/rfc/rfc7452.txt>.

Figure 5 - Device-to-Device Communication Model

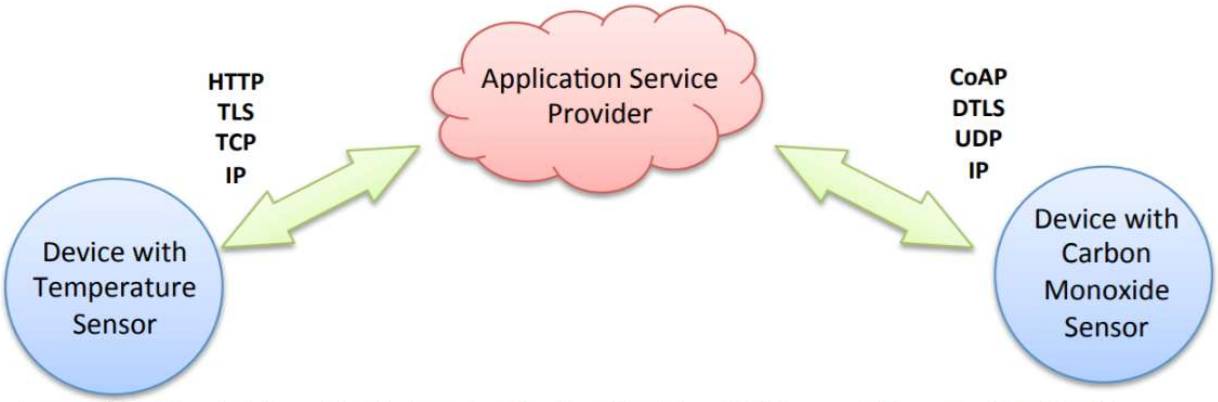
Bluetooth, Z-wave or Zigbee for the device-to-device communication as shown below.

The devices which follow this model have a specific protocol which they adhere to in order to facilitate communication and exchange of the messages between them. This kind of a model is predominantly used in systems which involve the exchange of considerably smaller data packets of information between each other. For instance, in a Smart Home, the residential IoT devices such as light bulbs, switches, thermostats, air conditioners, door locks, and so on usually exchange a small and limited amount of information between each other.

An IETF Journal article states, “these devices often have a direct relationship, they usually have built-in security and trust [mechanisms], but they also use device-specific data models that require redundant development efforts [by device manufacturers]”. Which implies that the different manufacturers and service providers will have to invest in research and development activities to custom build models to suit their products rather than relying on

the standard formats available in the market. This allows the manufacturer or service provider to increase the quality and the features provided by them to a significant level, from the consumers perspective this forces them to use devices from the same manufacturer or using the same protocol as it becomes increasingly hard to seamlessly connect with other devices if that is not the case. Though the incompatibility issue prevails, consumers learn that the products from the same manufacturer or service provider works relatively well and causes fewer troubles.

1.3.2 Device-to-Cloud Communication



Source: Tschofenig, H., et. al., Architectural Considerations in Smart Object Networking. Tech. no. RFC 7452. Internet Architecture Board, Mar. 2015. Web. <https://www.rfc-editor.org/rfc/rfc7452.txt>.

Figure 6 - Device-to-Cloud Communication Model

In a model of this sort, the data generated by the device is exchanged with the Internet cloud service provider possibly through an application. This model usually utilizes the pre-existing communication mechanism such as Wi-Fi or the wired Ethernet connections to connect the device to the IP network and eventually connect with the cloud service.

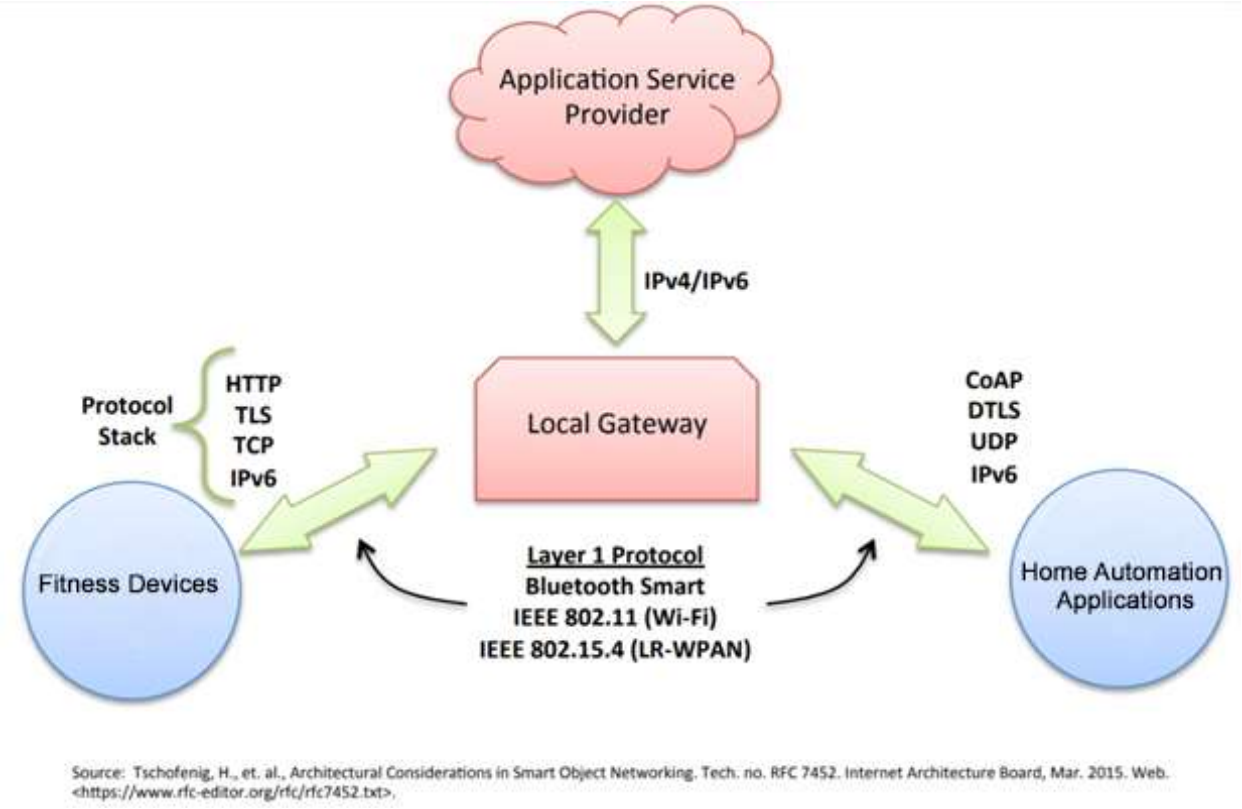
This model is prevalently used by manufacturers such as Nest Labs in their product Learning Thermostat which transmits the data to a cloud where the home energy usage can be analyzed. Also, the user will be connected to the Cloud through a Smartphone or Web interface and can monitor and control the thermostat remotely.

Like for the Device-to-device model, the issue of interoperability still exists for this model as well, where there might be challenges arising when trying to integrate devices from different manufacturers. If proprietary data protocols are used to operate the cloud and the device, the user might be restricted from using alternate service providers This is usually referred to as “vendor lock-in”, which usually encompasses other features such as the ownership and access to the data of the user. However, devices that are designed for a specific platform are growing being able to be integrated.

1.3.3 Device-to-Gateway Model

The device-to-device model or as commonly known as the device-to-application-layer-gateway model the devices connect to the cloud service using an ALG service. This ALG service uses an application software which serves as a local gateway and helps in establishing the intermediary connection between the device and the cloud service and also secures and translates the data or the protocol.

This model is being used in several ways in consumer devices and in most cases, the local gateway device is usually a smartphone or a tablet which has an installed app that relays data from the device to the cloud.



Source: Tschofenig, H., et. al., Architectural Considerations in Smart Object Networking. Tech. no. RFC 7452. Internet Architecture Board, Mar. 2015. Web. <https://www.rfc-editor.org/rfc/rfc7452.txt>.

Figure 7 - Device-to-Gateway Communication Model

Another growingly popular form of a device with such a model is the "hub" devices which are increasingly used in home automation systems. The "hub" would serve as the gateway between the IoT devices at home and between the cloud and help in relaying information. The option of interoperability comes along with the hubs as they can have the transceivers for more than one family of devices such as Zigbee and Z-Wave in one package and hence can be more convenient for the consumers. However, this model is usually used to integrate new smart devices into a legacy system which has devices that are not interoperable. Though it could serve as an effective model, the development of the application-layer-

gateway software and systems increases the complexity and cost of the overall system. The pros and cons of this model are still unfolding.

1.3.4 Back-End Data-Sharing Model

The back-end data-sharing model is a system that allows users to analyze the data of the Smart device stored on the cloud service along with data from several other sources. This sort of a model facilitates the data to be shared with third parties upon the user's grant of access. This is very similar to the device-to-cloud communication model where there is traditionally only one device, and, in this model, there can be multiple connected devices and the data can be shared from the cloud.

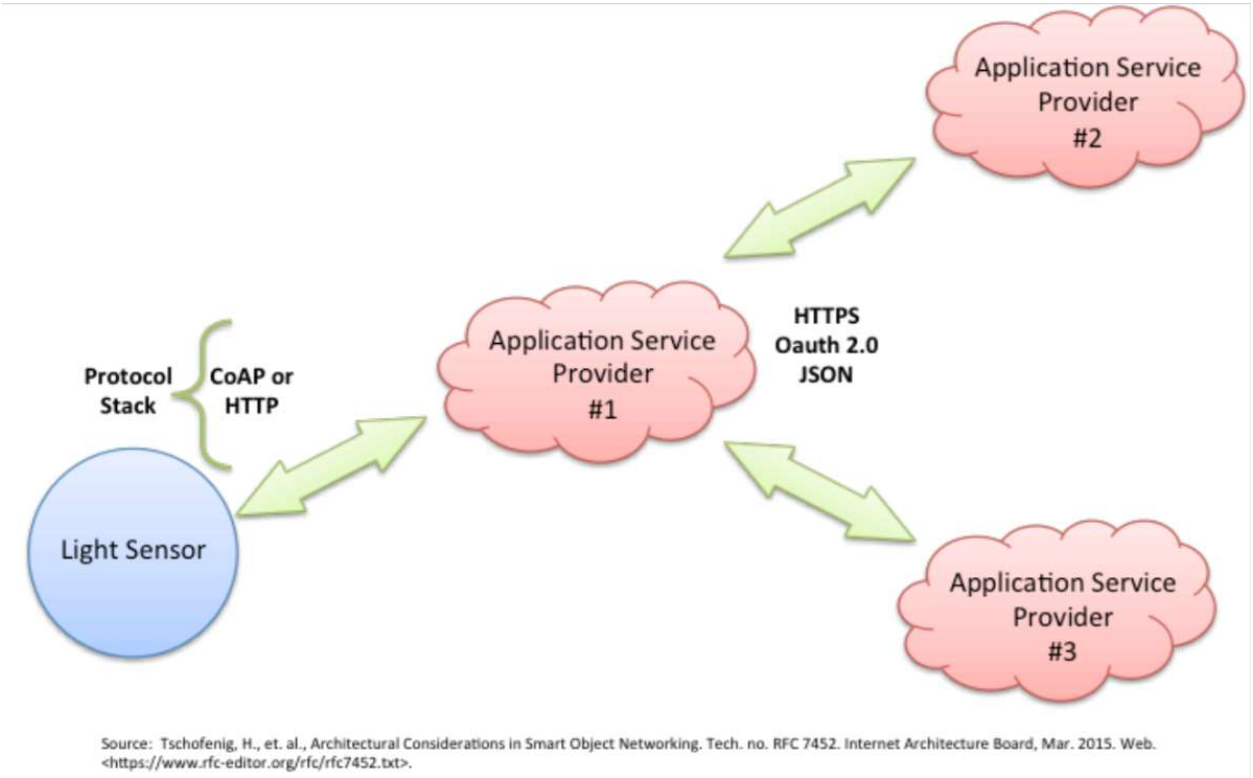


Figure 8 - Back-End Data-Sharing Model

For instance, let's consider a corporate office that has numerous IoT devices which help in collecting and in the analysis of the data on consumption of energy and utilities. A traditional device-to-cloud model would let the data produced by each IoT device sit as a stand-alone data silo while an effective back-end data-sharing architecture would enable the user to access the entire spectrum of data forms all the devices in the office through the same framework. This also enables and allows data portability and ease of use breaking the traditional stand-alone data silo barriers. To effectively functionalize the interoperability and

facilitate the access to all the data an API (Application Programming Interface) is usually necessary.

This model is an efficient way to facilitate interoperability among the different back-end systems and it can only be as effective and functional as the IoT system design which holds it together. Closed system designs cannot be entirely overcome by Back-end data-sharing models.

1.4 Smart City Applications

1.4.1 Smart Transport and Mobility Tracking

Figure 9 - Smart Transportation
Source: Quantzig

The need for transportation management systems that are efficient has seen an increase which is proportional to the fleeting growth in the number of cars for avoiding traffic congestions and optimization of the flow of traffic, especially at intersections. As the Transport system in a Smart City could have a varied range of mobile and fixed sensors, that are connected wirelessly or through wires which increases the potential of the system thus making it an Intelligent Transport System (ITS) service. This service would comprise of features such as car parking aids, adaptive vehicle navigation, road incident detection, congestion avoidance, human driver monitoring, speed control via smart interaction with roadside controls and overall better driving safety.

The use of traffic lights are the normative means to regulate the flow of traffic. As such conventional systems are not dynamic since they cannot adapt to the varying interval times and are often fixed to one given interval. This affects the average fuel usage during traffic congestions due to the recurring fitful navigation along with increased carbon emissions. A scheme which could be more adaptive with regards to the inflow of traffic could be desirable and this would require an algorithm which would take into account of the number of cars approaching and dynamically change the intervals accordingly. Such an intelligent system would require a dedicated infrastructure and methods to detect the required information. An alternative solution where the vehicles will be wirelessly interconnected with each other

and make use of a decision-making algorithm for collision avoidance and entirely eliminating the need for traffic regulation systems.

One of the predominantly used and a cheap method for vehicle sensing is the use of induction loop detectors that are usually buried in the roads. These loop detectors which will be able to detect metals and thereby consider them to be vehicles. Information such as the kind of the vehicle, its speed and/or other parameters can be obtained from these and can be used effectively for traffic flow management. Methods like the Weight in Motion (WIM) which uses sensors (e.g. piezoelectric systems, capacitive mats, bending plates, load cells, and optical WIM) embedded on the road surfaces to determine the weight of a vehicle that crosses it. Such methods usually demand infrastructure and hence come bearing a high installation and maintenance cost.

A more sophisticated but a less intrusive approach is the use of video cameras. Due to the easiness in their installation and low maintenance, they are widely preferred. In the cases where a surveillance system is already functional, they can be used for intelligent transportation applications. However, a major setback with the computer vision-based methods is the dependency of their performance on environmental conditions, such as lighting, occlusions, and weather. Considering the fact that vehicular distance can be detected using the headlights of other vehicles, street lamps and traffic lights can also be used for vehicle sensing in night-time scenarios.

Mobility data is also recorded by the integrated sensors in addition to the vehicle detection data through the use of a system such as M-Atlas which focuses on the trajectory of a moving object. The trajectory is learned by reconstructing the timestamped location data from the moving object using the data that is sensed from various contexts including Global Navigation Satellite System (GNSS) which tracks navigation devices that are hand-held or vehicular, call detail records from mobile phones with the help of GSM carriers and providers, time-stamped location records that can be found online on social networks or other such services, and so on.

The recording of an individual's mobility data is allowed to be performed across the entire urban network through GNSS technologies. In Italy, for insurance purposes, information which provides single trajectories with a spatial scale of 2 km and a time scale of 30 seconds are recorded for 3% of the whole population of the vehicles. The start or stop of the engine is always recorded in one datum and information such as speed position, the direction of motion and quality of GNSS is recorded in each datum. Though the spatial resolution of such GNSS data is considerably poor it could still be used reconstruct in real-time the individual

trajectory dynamics on the road network. The regulation and planning of flow of traffic can be aided by the analytical pattern detection in spatiotemporal mobility.

1.4.2 Smart Environment

If the city is not well organized and equipped to face the stress of rapidly growing urbanization, chaos quickly takes place. The quality of life, the security, the effectiveness of citizens' services, the economic development and attractiveness, and the quality of the environment may decrease quickly affecting the life of the community.

In the case a city is not organized properly it will be faced with a chaotic situation of urbanization that rapidly grows and it would affect the factors such as life quality, economic development, and attractiveness, citizen services effectiveness as well as the deterioration of environmental quality which negatively affects the life of the community. The continuous monitoring of the quality of water, air, humidity, and other parameters could help in forming an effective smart environment thereby serving as a solution to preserve the environment of the Smart City. This would allow keeping constant tabs on the levels of the different parameters and thereby understanding if a serious threat was underway or if a mitigation process is successfully effective. For such monitoring, a plethora of sensors need to be installed in outdoor locations such as rivers, parks, etc. The use and installation of wireless sensors on vehicles or buildings could be beneficial in monitoring the pollution in a Smart City.

As technology develops the numbers of sensors that are used for such monitoring has increased steeply in order to obtain credible results. With the advent and affordability of 3G and 4G systems, the possibility and extent of acquiring sensor data have been extended in recent years. The issue regarding the effect of the electromagnetic fields produced due to such networks of sensors generally varies from Hz to GHz and has been hotly debated. Light pollution caused due to the lighting systems in the areas around the big cities and the increased consumption of energy, pollution due to noise which majorly affects the people residing in cities and the management of waste in rapidly developing cities are some of the few important problems that are commonly faced. The advancement in technologies should also be used to find a remedy for such a situation.

1.4.3 Smart Grid (SG)

The International Energy Agency has estimated that globally there would be more than two-thirds of an increase in the demand for electricity by the year 2035. This kind of an increase

could result in a heavy stress on the existing power infrastructure which is outdated and overstressed. There is a high amount of unreliability in the power grid due to the lack of efficient monitoring, fault diagnostic and automation techniques and such unreliable conditions extend to water, gas, heat and other infrastructures of the city as well. However, the solutions are predominantly found and applied only for power related issues which include a unidirectional flow of the current to the customers from the generating stations, estimation and prediction of the supply based on previous data, slower response times due to mechanical switches and centralized schemes of generation.

As the name suggests Smart grid aims to tackle the challenges that were discussed above and contribute to making the world power systems more secure, reliable, efficient, flexible and sustainable. In other words, it can be explained as the use of automated control, modern communication infrastructure, monitoring and measuring technologies and advanced energy management based on optimization of demand, availability of the grid and so on to define a modern infrastructure with a relatively high efficiency and reliability.

The Smart Grid plays a major part in the energy sector and it can be classified into three types of networks as follows: a Home Area Network (HAN), a Neighborhood Area Network (NAN) and a Wide Area Network (WAN). While HAN is the primary layer, it consists of smart devices, home appliances, electric vehicles as well as the renewable energy sources and manages the on-demand power requirements of the consumers. It is widely deployed and used in residential setups, industrial plants and also in commercial buildings where the electrical appliances are connected with the smart meters. NAN, which is also known as the Field Area Network (FAN), aids the communication between the substations for distribution and the electrical devices in the field which facilitates the power distribution. It also connects, collects the data regarding metering information from multiple HANs and transmits it to the data collectors which connect NANs to a WAN. WAN is the third layer of a Smart Grid and it functions as the crux for the communication between gateways or points of aggregations. It facilitates the communication among power transmission systems, bulk generation systems, renewable energy sources and control centers

1.4.3.1 Mechanism: Integration of the IoT into a SG

The smart grid is prevalently known and adopted for its capabilities of information sensing, transmission, and processing, and IoT plays a considerable role in the construction of the grid. The overall efficiency of a SG is improvised as the IoT technology provides access to real-time, highspeed and two-way data sharing across various applications and also real-time interactive control of the various devices connected to the network. The application of the IoT's can be classified into three different types, first would be the monitoring of the

state of the equipment's in which the IoT smart device is deployed. Second would be to collect information from the equipment through the use of various communication technologies and third would be to control the SG through the dedicated interfaces.

IoT devices usually comprise wireless sensors, RFIDs, M2M (machine-to-machine) devices, cameras, infrared sensors, laser scanners, GPSs, and various data collection devices. These devices enhance the information sensing capabilities of a SG by playing an important role in deploying infrastructure for sensing of data and for transmission in the SG, network construction assistance, safety and security management, maintenance, the collection of information, user interaction, and so on. Moreover, the flow of information, power, and distribution in a SG are collectively enabled by the use of IoT.

The existing architectures of the SG contribute to the management of the entire power grid by mainly focusing on the power distributors needs. Through the use of General Packet Radio Service (GPRS) or other mobile networks, the smart meters network of the consumers are accessed. The existence of other smart home infrastructures in the homes of the consumers has not yet been accounted for into the existing SG architectures. There are architectures which do consider such situations, but they are not scalable for larger infrastructures. Specific protocols for combining IoT and SG will have to formulate as the existing ones only focus on the individual technologies.

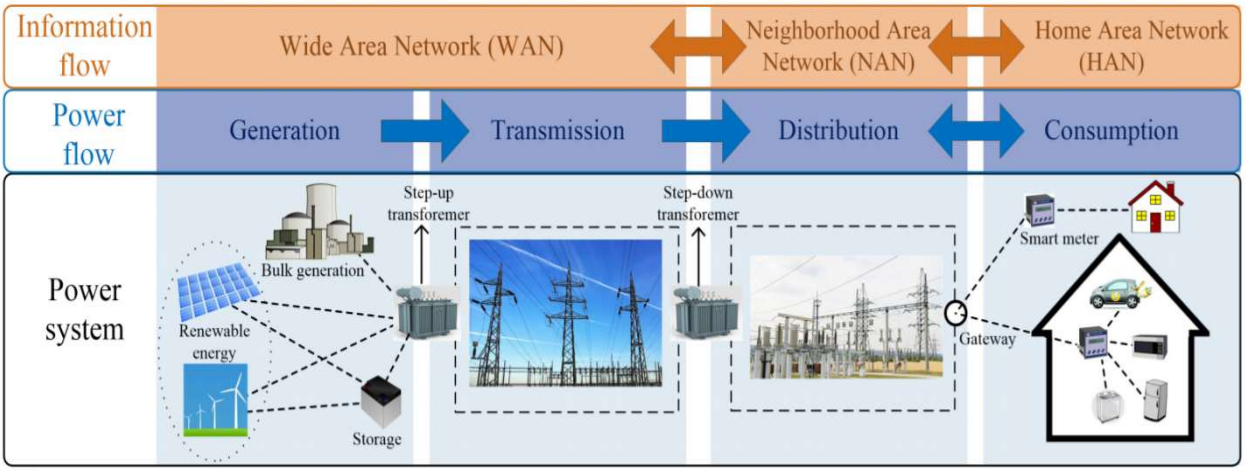


Figure 10 - Smart grid (SG) architecture presenting power systems, power flow and information flow.
Source: Saleem, Y., Crespi, N., Rehmani, M. and Copeland, R. (2018). Internet of Things-aided Smart Grid: Technologies, Architectures, Applications, Prototypes, and Future Research Directions

Power generation, transmission, distribution, and utilization are the four main subsystems of a SG. IoT application to each of these subsystems would help in the betterment of it and improve the overall effectiveness. In power generation, IoT can be used for the monitoring and controlling of energy consumption, units, equipment, gas emissions, and pollutants discharge, power use/production prediction, energy storage and power connection, as well as for managing distributed power plans, pumped storage, wind power, biomass power and

photovoltaic power plants. In power transmission, IoT can be used to monitor and control the substations, the transmission lines and also for ensuring the safety of the tower. In power distribution, IoT can be used for automation of distribution and to manage the operations and types of equipment. In power utilization IoT can be used for smart homes, automatic meter reading, charging and discharging of electric vehicles, for collecting information about the energy consumption of appliances, controlling of power load, monitoring, and management of energy efficiency, management of power demand and multi-network consumption.

The application of IoT based on the three layers of the SG are discussed as follows:

1) Home Area Networks (HANs): HANs usually have the topology of a star or a mesh and prefer the use of ZigBee, Bluetooth, Wi-Fi and wired technology for communication. A wide range of smart devices like the home gateway, smart meters sensor and actuator nodes, smart appliances, and electric vehicles fall under HAN. The smart meters are connected to the home gateways and regularly transfer consumption data of the home appliances using it. HANs has the functions of commissioning and control, wherein commissioning involves identification of the new devices and managing them, and control allows for communication between the connected smart devices and also carries out operations for the various SG layers with increased reliability. Its two-way communication allows for demand response management services. The forward communication direction allows for the collection of consumption information from the various IoT devices by the home gateways and then transmitting it from the consumer side to the NAN and then to be forwarded to a utility center. While the backward communication direction allows the home gateway to receive

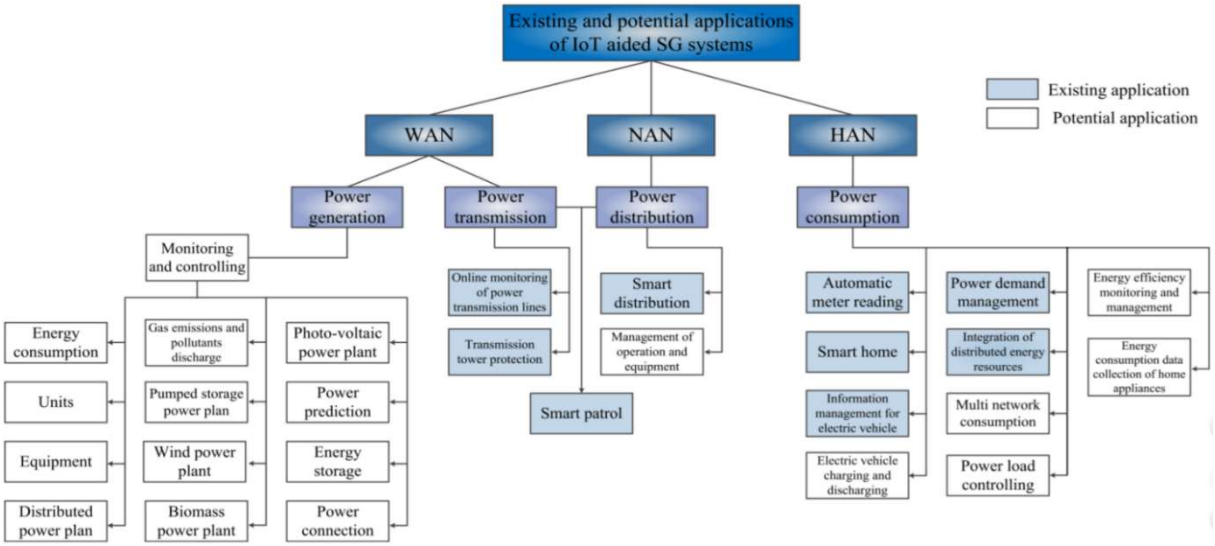


Figure 11 - Existing and potential applications of IoT-aided SG systems classified into WAN, NAN and HAN. Source: Saleem, Y., Crespi, N., Rehmani, M. and Copeland, R. (2018). Internet of Things-aided Smart Grid: Technologies, Architectures, Applications, Prototypes, and Future Research Directions

the dynamic electricity pricing from the NAN and then transmit it to the different IoT devices which would initiate the required actions.

2) Narrow Area Networks (NANs): The NANs are required of to cover a radius of a thousand meters and hence there should not be any interference in the channels of communication between the smart meters and the data aggregation points. The data collected from the smart meter through the HAN network is transmitted to the NAN gateway which is then transmitted to utility companies by using a public or a private WAN. The NAN comprises of a HAN and a NAN gateway where the latter serves as a single access point to the multiple HAN gateways connected. The transmission of the data from HAN gateways can be through either wired (e.g., PLC, DSL) or wireless (e.g., cellular, mobile broadband wireless access or digital microwave technology) communication technologies.

3) Wide Area Networks (WANs):

The WAN serves as a facilitation structure for the communication between network gateways, NANs, distributed grid devices, utility control centers, and substations. It is made up of a core network and a backhaul network both of which are interconnected. The core network handles low latency and high data rate communication using fiber optics or cellular communications. While the backhaul networks handle the broadband connections and devices for monitoring to NANs using wired (e.g., optical networks, DSL), wireless (e.g., cellular network, mobile broadband wireless access) or hybrid fiber-wireless networks.

1.4.3.2 Existing Applications of IoT-aided SG systems

In this section, the existing applications of IoT-aided SG systems in the literature are discussed.

A. HAN Applications

1) *Smart Home:* IoT technology has a predominant role in the SG for the realization of smart homes and appliances such as lighting control, temperature monitoring, detection of fire, home security systems, smart TVs, refrigerators and washing machines. A smart home consists of sensor and actuator nodes for environmental monitoring that transmit the surveillance data to the control unit of the house. The control unit allows users to monitor and control the appliances at anytime and anywhere, remotely. Hence, the smart home is an important element of the SG to understand real-time interaction between users and the grid, enhance service quality and upgrade the capacity of integrated grid services, and to meet users' energy demands in a most efficient way possible. A broad use of smart home services is associated with optimizing the daily power consumption. For example, even before arriving home, users can switch on air conditioners and heaters to enjoy their

Figure 12 - Smart Home
Source: Intermedia

preferred home environment without the wait. Moreover, electrical appliances that are power intensive like washing machines could be turned on at midnight, when there is cheaper electricity cost. The control unit additionally utilizes surveillance data to identify suspicious activities and notify users to take necessary actions. IoT Technology makes these functions possible and understandable. Viswanath et al. have formulated and developed a system and testbed for smart home - a house with one living and three bedrooms, fitting for six to nine people. Based on dynamic pricing, this system controls the energy and functions as an energy management system. Thus, the usage of home appliances during peak hours could be prevented. The authors developed An Android application was also developed by the authors for consumers to have remote access. Similarly, Shah et al. suggested a model for security enhancement in the smart home by implementing Reed Solomon Codes for error detection and correction.

The IoT is applicable to different aspects of the SG in smart homes, such as, in a smart home's sensor LAN protocol to manage smart appliances, multi-meter reading, data collection of power consumption (including electricity, water, and gas), load monitoring as well as control user interaction with smart appliances. NANs also can be connected to the IoT technology by linking a group of smart homes in a neighborhood through a NAN to form a smart community. Smart homes in the smart community can hence share outdoor surveillance camera data to identify any accident or suspicious activity and notify the relevant police stations and emergency centers autonomously. A Smart City could be developed from extending the concept of the smart community. A comprehensive surveillance system could be created in a Smart City to manage and detect various activities within an entire city or even a country.

2) *Information Management System for Electric Vehicles:* Eco-friendly transportation can be attained through Electric Vehicles (EVs) that reduce carbon dioxide emissions. This is an

interesting platform for IoT aided SG systems. The charging system for EV contains a power supply system, charging equipment and a monitoring system. The output and management of electricity is the responsibility of the power supply system while the charging equipment charges and discharges the EVs and includes both AC and DC chargers. AC chargers cater slow charging and are typically applied in the home. Rapid EV charging requires DC chargers, typically used at public charging stations. Both types of chargers also provide a billing function. The monitoring system is in charge of real-time monitoring of the charging system and its security. IoT technology has an instrumental role in this monitoring system, generating an information management system that combines different components of the charging system. An example would be IoT Technology allowing power supply and real-time monitoring systems to transfer information to the information management system. The latter then passes information to control station for appropriate actions to occur. EVs are equipped with GPS, enabling the IoT to aid drivers in more efficient management of batteries by locating closest, most ideal and least crowded charging station and providing parking and traffic details.

The auto manufacturer Nissan has partnered with ENEL, one of Europe's largest power companies, to generate a 'Vehicle-to-Grid' (V2G) system. The companies indicate this

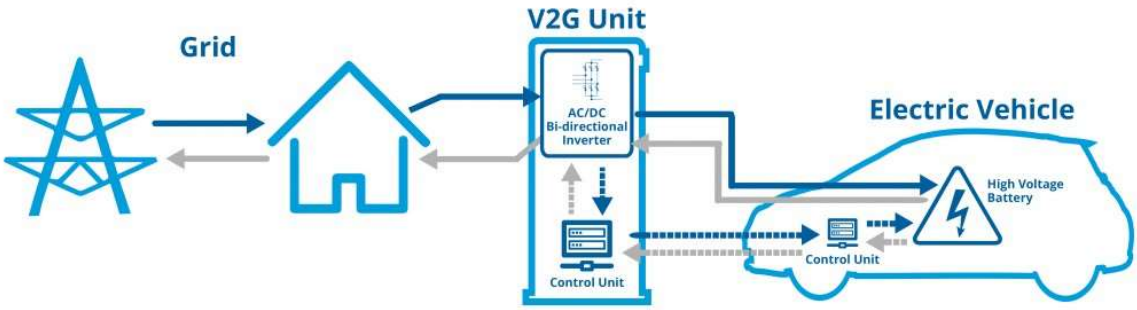


Figure 13 - Vehicle to Grid
Source: Cenex

system "to allow drivers and energy users to operate as individual energy hubs with the ability to use, store and return excess energy to the grid". The first trials for the technology will occur in Denmark, with Germany, the Netherlands and some other European countries ensuing if the trials are fruitful.

Generally, the operational efficiency of the power grid gets affected due to the daily load demand fluctuations and regulation of the voltage and frequency from the power grid which often result in high cost and heavy energy waste. The concept of V2G helps with this by utilizing the large amounts of energy stored in the Electric Vehicles as a buffer for the renewable energy and the power grid thereby providing a service similar to the Ancillary service market. This way the power demand during the peaks can be satisfied through this

mechanism and the grid operation costs for regulation of voltage and frequency can be scrapped. And the EV and its owners would be able to benefit from this by storing energy when the price is low (off-peak hours) and can feed it back to the grid when the price is high (peak hours).

However, the EV's will not be able to freely access the grid at any time as it would cause serious damage is the charging demand of the EV's happens to be huge during the peak-load periods of the grid. As for the vehicles, in addition to providing ancillary services for the grid, they should satisfy the daily routine driving requirements. Therefore, it is necessary to investigate the V2G technology to coordinate the charging/discharging behaviors between vehicles and grid so that it will not affect the power grid operation and constrain the normal use of automobiles. This system can be made possible only when ample amount of data is available for managing the network which will be facilitated by the IoT sensors and using the HAN network which will be connected to the cloud.

3) *Automatic Meter Reading:* The traditional manual collection of power consumption information on site at specific time intervals, inevitably resulted in inadequacies with respect to accuracy and timeliness. IoT allows Automatic Metering Infrastructure (AMI) or remote meter reading systems based on WSN, PLC and Optical PLC (OPLC) by using public or private communication networks. An AMI system is one of the most prominent functions of the SG, gathering the real-time electricity consumption data with high reliability, processing this data thus enabling real-time monitoring, statistics and power consumption analysis. This system favors the timeliness, efficiency, and accuracy in the power consumption data. Through this system, IoT technology could benefit users to save money by altering their electricity usage based on the analysis of their power consumption.

GreenPocket, a German company, aids its customers in capitalizing on smart meter data by maximizing their customer relationships. It provides utilities and business customers with a highly intuitive IoT software solution to display the significance of smart meter data by offering modular white label standard software as SaaS. The software visualizes, assesses and interprets smart meter data by employing intelligent algorithms. The software also offers useful functions for utilities to improve their services and customer dialogue and engagement. End users get critical insights into their energy consumption that can help to improve their energy efficiency to reduce costs. This, in turn, creates an opportunity for them to contribute to a more sustainable future by reducing their CO2 emissions.

A smart metering system allows the water, electric and gas utilities a continuous consumption reading and recording in time intervals or, at least, daily reporting, monitoring,

and billing. This allows the utility to gather interval data, time-based demand data, outage management, service interruption, service restoration, quality of service monitoring, distribution network analysis, distribution planning, peak demand, demand reduction, customer billing, and work management. In recent years, the advances in Information and Communication Technologies sector have permitted the evolution from the simplest Automated Meter Reading systems to Advanced Meter Infrastructures (AMI). These allow demand-response management, enabling customers to make informed decisions and consumption prediction.

AMI systems include sensors, hardware, software, communications, consumption displays and controllers, customer systems, data mining software, meter data management software, and business systems. They provide two-way communications with the meter, allowing sending commands from the utility to the smart meter for multiple purposes, including monitor real-time values and change the frequency of readings among others. The network between the smart meters and the utility center allows the collection and distribution of information to customers, suppliers, utility companies, and service providers.

4) *Integration of Distributed Energy Resources (DERs):* Renewable energy generators, such as solar cells, photovoltaic cells, and wind turbines, are gradually incorporated into today's power grid. They have recently gained great attention in SG studies due to climate changes and environmental pressures. Renewable energy has a positive impact on the global environment by generating electricity without carbon emissions. Decreasing the annual increase in greenhouse gas emissions leads to limiting the Earth's increasing temperature. Over the past few years, governments and organizations have installed a substantial number of solar cells and wind turbines to fulfill part of their power requirements. Power generation patterns of renewable energy sources (solar and wind), which are disseminated over the grid, are intermittent in nature and influenced by location and climate, so they are major challenges for the predictability and reliability of the power supply. These issues are addressed using the seamless interoperability and connectivity provided by the IoT technology. Moreover, the IoT technology uses sensors to collect real-time weather data which aids in forecasting energy availability in the near future. A Kalman Filter-based state estimation and discrete-time linear quadratic regulation method for managing the state deviations is proposed in, using IoT aided SG systems. It utilizes IoT technology and WSNs to sense, estimate and control the states of DERs. Furthermore, a web of things-based SG architecture is proposed in for the remote monitoring and controlling of renewable energy sources.

5) *Power Demand Management:* Power demand management, or demand-side energy management, is defined as "the change in the energy consumption profiles of consumers according to the time-varying electricity prices from utility companies". It is used to reduce the consumer's electricity bill, and to minimize the operational cost of the power grid and energy losses, and to divert the demand load from peak times. IoT devices gather the energy consumption requirements of various home appliances and transmit them to the home control units. Thereafter, the SG control unit schedules the energy consumption of home appliances depending on the users' defined preferences such that each consumer's electricity bill is reduced. Demand-side energy management can be conducted at different levels of the SG. For instance, at a home-level to guard consumers' privacy. It can also be performed at higher levels to benefit the utility companies by producing a more effective scheduling plan.

B. NAN Applications

1) *Smart Distribution*: Smart distribution stems from advanced automated IoT technology and is one of the key components of SG. It is the component that is directly linked to users in the smart grid. Smart distribution grid comprises of the communication system, power distribution remote unit, master unit, and station unit. With use IoT technology, the smart distribution grid can immediately detect the faults in case of any disorder and can overcome the fault instantaneously. IoT technology aids smart distribution grid by allowing different types of sensors for collecting data about temperature, humidity, noise etc. which ensures monitoring and secure functioning of the distribution grid. The first demonstration of smart distribution by using IoT technology in the distribution network of smart grid was conducted in the Henan Hebi IoT demonstration project in Hebi, China. It utilizes noise, temperature and tower tilt sensors, as well as ZigBee, GPRS, 3G, and power fiber. It identifies online monitoring, online inspection, and lifecycle management. This project was executed using 10 kV underground and overhead laying mixed line which covers 45 utility distribution transformers, 68 units of surge arresters, circuit breakers on 20 pillars, a ring counter, and four cable branch boxes. Readers are referred to for detailed description and results of this project.

2) *Smart Patrol*: The patrolling of power generation, transmission and distribution was chiefly a manual task, carried out regularly at specific time intervals. However, the quality and quantity of patrolling are not always as expected as it is influenced by climate conditions and both human and environmental factors. Additionally, it is generally tough for power workers to patrol unattended substation equipment. The IoT technology provides a

promising solution to this issue by introducing Smart patrol. The smart patrol consists of WSN and RFID tags which are connected to the power substation through IoT technology and are used to locate the power equipment so as to enhance patrolling quality and improve the reliability, efficiency, and stability of a power system and its supply. The smart patrol will be beneficial in a number of applications, for instance, patrol staff positioning, equipment status reports, environment monitoring, state maintenance, and standard operations guidance.

C. WAN Applications

1) *Transmission Tower Protection*: An integral part of power transmission, the transmission tower protection is a WAN application of IoT-aided SG systems, created to increase the safety of transmission towers from physical damage by the robbery of components, natural disasters, unsafe construction and growing trees under the foundations. Burglary and intentional damage by people are the major causes of transmission tower damage. Natural calamities like typhoons, strong thunderstorms, and global warming effect also can cause transmission towers to collapse. Moreover, important infrastructure projects, such as highways and high-speed railways, are commonly built near the transmission towers and must sometimes cross high voltage transmission lines. Often the construction companies are not entirely aware of the risks associated in operating near high voltage transmission towers. A number of large construction machines are used by these companies, which pose serious dangers to their workers and also damage transmission lines and towers. Due to lack of communication and information passing between such companies and relevant power transmission departments, there is huge difficulty in inspection and monitoring of all the power transmission facilities, which could cause risks to transmission towers.

Presently, the primary method of protecting transmission towers is manual patrol by the staff. However, a consistent staff-based manual patrol of high voltage transmission lines and towers by is very tough because of manpower realities, the level of knowledge of the staff and divisions of responsibility as well as physical positioning of some towers. This affects the patrolling quality.

The patrolling period differs from 1-10 weeks, indicative of inadequate monitoring and greater security risks. Albeit having some cameras and infrared alarms installed on transmission towers to detect burglary and other potential damage, the stability and accuracy of such equipment are still unsatisfactory.

Assisted by WSNs, IoT technology can allow remote monitoring in overcoming these security problems. The IoT-aided transmission tower protection system contains different sensors

which produce early warnings of threats to high voltage transmission towers, enabling prompt responses. The sensors include vibration sensors, anti-theft bolts, a leaning sensor, and a video camera. These sensors and the sink node form a WSN. The sensors identify any threat and transmit the corresponding signals to the sink node. The sink node receives these signals from the sensors, processes them into data and transmits the data to the monitoring center through the Internet or any other public/private communication network.

Anti-theft bolts are deployed on the lower part of the tower. One vibration sensor is installed underground, in the base of the transmission tower and the other is installed on the tower, roughly 3-5 meters above the ground. The leaning sensor is located close to the vibration sensor on the tower. The camera, directed towards the transmission line, is deployed on the tower about 6-8 meters high. Finally, the sink node is installed in the middle of the transmission tower. The vibration sensors monitor the vibration signals of the ground and tower. When they detect signals that indicate excavation construction close to the transmission tower, they transmit those signals to the sink node.

Upon burglary or occurring vandalism, the vibration sensor on the tower, the leaning sensor, and the anti-theft bolts detect these signals and consequently transmit those signals instantly to the sink node. In both cases, the sink node integrates and processes the signals, and initiates real-time images to be sent from the video camera to the sink if the threat is determined. Similarly, the video cameras directed towards the transmission lines detect big construction machinery and trees that get too close to the transmission lines and transmit real-time images to the sink. In all of these threats, the sink sounds an alarm and via the Internet, it forwards the images to the monitoring center. The real-time data of a series of transmission towers is handled by the monitoring center. The alarm signal and images alert the monitoring center staff to take necessary actions to manage such threats.

2) *Online Monitoring of Power Transmission Lines:* The online monitoring of power transmission lines is also a crucial application of the IoT in the SG, particularly for disaster prevention and mitigation. Recently, the challenges of security, reliability, and stability inherent to high voltage power transmission lines have been highlighted by natural disasters. Typically, there has been manual monitoring of high voltage transmission line monitoring. Sensors measuring conductor temperature, galloping, micrometeorology, wind vibration, and icing can now be used to obtain real-time online monitoring of power transmission lines. This new online power transmission line monitoring system is separated into two parts. Firstly, the sensors are deployed on the power transmission lines between transmission towers to monitor the states of power transmission lines. Secondly, sensors are deployed on the transmission towers to monitor their states and their environmental parameters. The

communication between the power transmission line sensors and the transmission tower sensors is promoted by IoT.

1.5 Pros and Cons

There are massive potential benefits of the IoT in the energy efficiency market, specifically for end-users and integrators. IoT will enable end users to measure, check, and promptly adapt their installation corresponding to the market expectations, the cost of the raw materials or the energy prices. For energy efficiency system integrators, IoT will enhance their capability to tackle customer needs, rapidly decreasing the duration for integration and commissioning. Integrators will use remote diagnostic and fix capability of products to improve the maintainability of their systems.

However, large players (e.g. Schneider Electric, Siemens, ABB, etc.) who have historically safeguarded this highly advantageous market by sustaining an extremely high level of required investment to develop complete and complex solutions and administering specific standardization (e.g. through domain-specific protocols such as IEC61850), dominate this market.

The IoT pledges to demolish these two entry barriers by providing low-cost flexible solutions and employing new communication solutions coming from the world of the Internet. Consequentially, new players are being attracted to this market. SMEs and start-ups may offer low-cost solutions, at least for a full set of niche applications, while large players from the Internet world may also enter as direct competitors of the traditional large players (e.g. using their big data analysis capabilities.). The large players thus are subjected to the main risks - new players approaching the traditional market counter-balances the platform provided by a big, new potential market. Nonetheless, this risk may be alleviated when these two groups of players partner with each other, paving way to a novel, open ecosystem.

1.6 Challenges

According to IEEE Communications Magazine, the open challenges that are posed to this technology are discussed as follows.

Interference Management

Due to the unexpected expanse of wireless devices, the coexistence of devices is very rapidly increasing, causing an interference problem and ensuing frequent data communication errors. Interference can prevent the effective deployment of sensors, WLANs, and other equipment in smart cities. Interference management is a vital challenge to ensure that these devices function without any interference. Off-the-shelf interference management models

and mechanisms for wireless networks can surveil the amount of interference and offer specific solutions to deal with it. However, due to the volume of connected devices, the increased complexity, these models and mechanisms will be unable to entirely and optimally resolve the interference issue in smart cities. Therefore, there is a necessity for robust interference management over smart cities' networks.

Scalable Wireless Solutions

The need for scalable wireless solutions has increased, due to the possibly large number of devices being connected to the Internet in view of creating smart cities. Albeit state-of-the-art wireless technologies, such as RFID, ZigBee, Bluetooth, LoRaWAN, Z-Wave, and other WPAN already supporting low-power device communication, their capabilities are restricted with regards to the number of devices, throughput, and transmission range. Since modern wireless technologies like 802.11 (Wi-Fi) were initially designed to provide high throughput to a limited number of stations situated indoors at a short distance from each other, these technologies will be unsuitable for communication of smart cities. Thus, smart communication is required and needs to be focused on, in the Smart City environment.

Interoperability Support among Heterogeneous Wireless Networks

Smart City networks are usually employed using different wireless network technologies, such as mobile ad hoc networks, Wi-Fi, WiMAX, and wireless mesh networks. However, a critical concern about the interoperability among these heterogeneous wireless networks has arisen. Communication among different wireless networks can be enabled when the problems pertaining to interoperability are addressed. The interplay of the diverse wireless technologies for efficient delivery of value-added services and applications results in several challenging problems: mainly concerning resource allocation, quality of service (QoS) provisioning, architecture, mobility management, and security. As such, it is important to tackle these challenges in the future.

Mobility Management

Numerous applications for mobile users, such as logistics, e-health, and intelligent transportation systems are offered by smart cities. These services depend chiefly on heterogeneous mobile technologies and so, demand varying services spanning from non-real-time (low data rate) applications to real-time (high-speed) multimedia applications provided by different access networks. Thus, designing intelligent mobility management techniques, which exploit various wireless access technologies to attain global roaming, is one of the huge research challenges for the upcoming mobile systems. Additionally,

integration and interoperation of contemporary mobility management techniques in the heterogeneous access networks are needed for the amalgamation of upcoming wireless technologies in smart cities.

High Energy Consumption

Communication and networking from an energy perspective have gained major attention, since the deployment of resource-restricted devices in smart cities. Despite advanced communication technologies, WiMAX and LTEA, facilitating users for massive downloading and uploading speed, the energy consumption rate is notably high and can restrict the realization of these technologies. Although devices in the future may have high specifications for battery life, the energy consumption rate of these modern communication technologies will still be deemed as higher. The reasons for such modernistic high energy consumption include support of multiple parallel transmission, enhancement of the radio network to attain good quality signals, and powerful data transmission. Thus, future smart cities must ensure control and optimization of renewable energy sources and demand-side management programs by delivering real-time information.

2. Big Data

2.1 Overview

The IoT offers an opportunity for sensors and actuator devices to communicate seamlessly within the Smart City environment and allows an increasingly convenient information sharing across platforms (Gubbi et al., 2013). By benefitting from all the opportunities provided by the Internet, the recent adaptation of various wireless technologies makes IoT as the next revolutionary technology. The IoT has recently been employed in the Smart City to generate intelligent systems such as smart grids, smart homes, smart retail, smart water, smart healthcare, smart transportation, and smart energy (Gubbi et al., 2013). However as per Morabito, "a universally agreed definition of a Smart City is yet to be conceived, and recognizing common global trends is challenging" (Morabito, 2015). But a Smart City can be denoted by a simple formula:

Smart City = Digital City + Internet of Things + Cloud Computing

The development of a Smart City can be divided into three categories: an information stage, a digital stage, and an intelligent stage. This involves the construction of a city's information infrastructure into digital city infrastructure. This starts the real-time sharing of resources and information and providing coordination between government, municipality, and citizens. The Smart City focuses on implementing the next-generation information technology to all aspects of life, embedding sensors and equipment in hospitals, power grids, railways, bridges, tunnels, roads, buildings, water systems, dams, oil and gas pipelines as well as other objects globally, thereby constituting the IoT. The Internet revolution resulted in the interconnection between people at an unexpected pace and

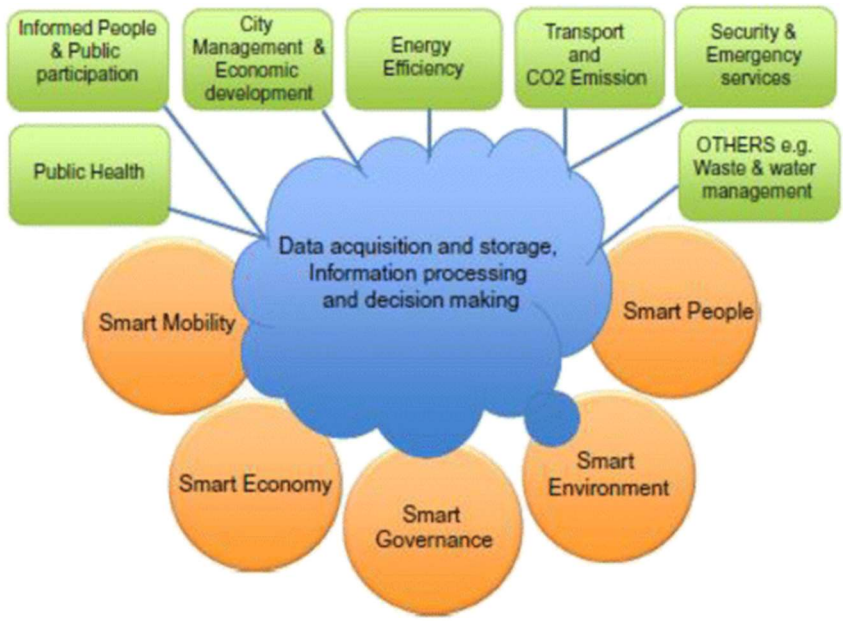


Figure 14 - Using the Cloud to store data generated from different components of a smart city
Source: Applications of big data to smart cities

scale. The subsequent revolution will be the interconnection between objects to form a Smart City, which stresses the interconnection of sensing and actuating devices, thus enabling the sharing of across platforms through a unified framework. Such sharing and the unifying framework are obtained by seamless ubiquitous sensing, data analytics, and information representation with cloud computing. The post-PC era exists presently, where handheld devices such as smartphones are altering our environment by enhancing interactivity and informativeness.

Big data systems are stored, processed, and mined in smart cities efficiently to produce information to boost various Smart City services. Moreover, big data can aid decision makers to plan for any development/expansion in Smart City resources, areas or services. The different characteristics of big data portray its great potential for benefits and advancements. Though possibilities are unending, they are confined by the availability of advanced tools and technologies. The appropriate methods and tools for efficient and effective data analysis can be utilized for big data to attain its goals and further develop the services in smart cities. This will foster collaboration and communication between entities and facilitate the production of additional applications and services, which can further develop the Smart City. Big data and its services can be applied in many sectors in a Smart City, hence offering better customer experiences and services, in turn helping businesses obtain upgraded performance (e.g., higher profits or increased market shares). For instance, healthcare can be made better by improving patient care, diagnosis and treatment tools, preventive care services and healthcare records management. Similarly, transportation systems can optimize routes and schedules, accommodate different demands, and increase environmental cleanliness.

Cloud computing denotes various types of computing models involving many computers or clusters connected through a real-time communication network. Complex large-scale computing tasks such as mining big social network data produced through smartphone applications can be performed using cloud computing. Cloud computing services, such as platform as a service (PaaS), software as a service, and infrastructure as a service, can be joined with IoT. Such a combination can transform every business – great amounts of data can be readily processed with the use of big data technology. Moreover, as Armbrust discusses, cloud computing can offer "the virtual infrastructure for utility computing that integrates monitoring devices, storage devices, analytics tools, visualization platforms, and client delivery". An end-to-end service provisioning for users and businesses to access applications on demand, from any location, can be provided by cloud computing-derived cost-based model with a business framework. Cloud computing also provides the underlying engine via the big data technology such as Hadoop. Hadoop was created to offer an enabling

platform and programming models for the distributed processing of large data sets across various clusters. Two primary components constitute Hadoop: Hadoop Distributed File System and MapReduce that are closely associated with each other. Although the real-time requirements of data storage and processing in the Smart City are considered, the adoption of streaming architecture will guarantee efficient and seamless communication between sensing devices within the Smart City network. Such technology has been recently employed with the implementation of several stream processing platforms, such as Apache S4, Storm, and Spark streaming, which can allow data storage and processing across different interconnected nodes.

2.2 Characteristics of Big Data

Big Data is vital because it allows organizations to gather, store, manage, and manipulate large amounts of data at the right time, at the right speed, to gain the right insights. Moreover, Big Data generators must produce scalable data (Volume) of various types (Variety) under controllable generation rates (Velocity), while maintaining the key characteristics of the raw data (Veracity), the collected data can bring to the intended process, activity or predictive analysis/hypothesis. As no clear definition exists for 'Big Data', it is generally described based on some of its characteristics. Therefore, these five characteristics have been used to define Big Data, also known as 4V's (Volume, Variety, Velocity, and Veracity).

Volume: denotes the quantity of data gathered by a company. This data must be utilized further to gain important knowledge. Enterprises are oversupplied with continually-growing data of all types, easily building up terabytes and even petabytes of information (e.g. turning 12 terabytes of Tweets per day into enhanced product sentiment analysis; or translating 350 billion annual meter readings to better estimate power consumption). Moreover, Demchenko, Grosso, de Laat, and Membrey stated that volume is the most distinctive and significant feature of Big Data, imposing specific requirements to all traditional technologies and tools presently used.

Velocity: denotes the duration in which Big Data can be processed. Some activities are very important and need prompt responses, which is why fast processing maximizes efficiency. For time-sensitive processes such as fraud detection, Big Data flows must be scrutinized and utilized as they stream into the organizations so as to maximize the value of the information (e.g. examine 5 million trade events created each day to identify potential fraud; analyze 500 million daily call detail records in real-time to predict customer churn faster).

Variety: denotes the type of data that Big Data can comprise. This data may be structured or unstructured. Big data consists of any type of data, including structured and unstructured data such as text, sensor data, audio, video, click streams, log files and so on. The analysis of combined data types brings new problems, situations, and so on, such as monitoring hundreds of live video feeds from surveillance cameras to target points of interest, exploiting the 80% data growth in images, video, and documents to improve customer satisfaction.

Value: denotes the key feature of the data which is defined by the added value that the collected data can introduce to the intended process, activity or predictive analysis/hypothesis. Data value will be based on the events or processes they signify such as stochastic, probabilistic, regular or random. Accordingly, the requirements may be executed to collect all data, store for a longer period (for some possible event of interest), etc. In this aspect, the data value is closely related to the data variety and volume.

Veracity: denotes the extent to which a leader trusts information to make a decision. Thus, finding the right correlations in Big Data is highly imperative for the future of business. However, placing trust in Big Data deems to be a major problem since the number and type of source grows and 1 in 3 business leaders fail to trust the information utilized to make decisions.

2.3 Value creation

2.3.1 Data Stages

To achieve smartness, each city must experience the data, information, knowledge and wisdom pyramid. Raw data such as climatic changes, movement of people, consumption of water and energy from all sectors of the city regathered from multiple data sources which are woven into the operational infrastructure of the city through which they are sensed and measured. Once the raw data is collected, it gets encapsulated with control information through network protocols and is then transmitted to the nearest gateway in the city's network which then gets transmitted to the Big Data cloud. The intelligent layer of the Smart City would be able to look,

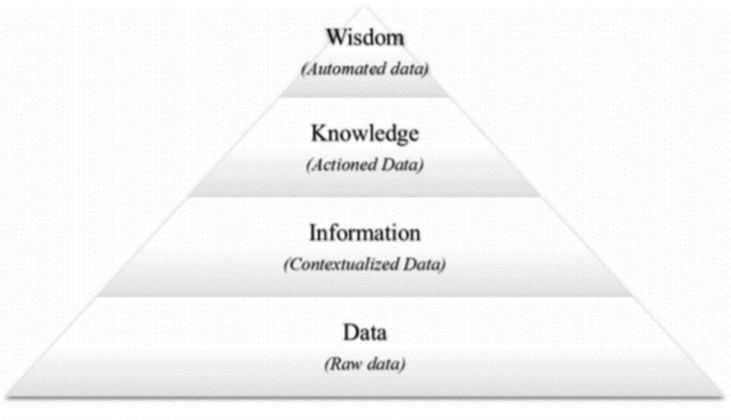


Figure 15 - DIKW Pyramid
Source: A Knowledge Value Chain for Knowledge Management

listen, connect, predict and correct the Big Data that it processes to unlock the different stages of the pyramid thereby achieving and exhibiting intelligence.

2.3.2 Big Data Pipeline

The following are brief descriptions of each stage:

1) Look (Data): The purpose of this stage is to collect and index the data in order to enable fast access during searching, sorting and so on. A Smart City is generally made up of a huge network of sensors which constantly collects data along with the data that are manually collected by humans, e.g., data of sensors, smart meters, satellites, webcams, RFID tags, smart cards, smartphones, social media, citizens' records, location information and much more. The collected data will have to be indexed inside a column-based data store which uses basic structures such as the hash table, tree-based index, multidimensional index, or bitmap index.

2) Listen (Information): This stage is where the incoming data are understood, its attributes are identified, and the unwanted data are filtered out to obtain the most valuable ones.

3) Learn (Knowledge): This stage deals with the extraction of features out of the information using techniques such as grouping and clustering. Upon applying certain criteria and correlating them together a resolution to the entity can be obtained.

4) Connect (Wisdom): This stage deals with the aggregation of different pieces of information and the integration of past and present data to derive experience

5) Predict (Intelligence): This stage deals with predictive analytics which helps to gain insights and make a decision. The rules and relationships are mined from past data in order to be applied to the future sets of data.

6) Correct (Visualization): This stage deals with the presenting (visualizes) of the results obtained from the previous stages in order to help with decision-making. Parameters based on which the outcome was derived and based on which inputs they were predicted is also provided.

2.3.3 Big Data Computing

For a Smart City to function as a real-time ecosystem, a real-time view of the city must be provided at all times which would require the detection and action to be fast and proactive and hence should the data be analyzed. However, the data has to be matched and analyzed together with the historical data and the one obtained from social media in order to understand better and also to increase the accuracy in understanding the rhythm and

everyday behavior of the city. This is made possible as the Big Data Computing Platform uses two main computing technologies namely the Stream computing and Batch computing.

High-speed data environments which is required to provide real-time analytics for the purpose of scoring and detecting insights, all of which has to be done in an instant, make use of Stream Computing. The system usually is made up of a clustered hardware and a massive data parallel processing which analyzed the data on the go without any interruption or blockage. To help with the identification and extraction useful information from real-time data feeds, there are usually models which are embedded in them.

Whereas a Batch Computing system makes use of Hadoop, which is an opensource Apache project and is predominantly used in Big Data environments. It works considerably fast as it parallelizes the data processing across the different computer nodes thereby dealing with large volumes of data in the minimum time possible. Hadoop comprises two main components:

• Hadoop Distributed File System (HDFS): is a data service which helps with the management of files in Big Data environments. In HDFS, the data files that are to be processed are usually fragmented into smaller blocks which are usually between 64-128 MB and is distributed across the computing nodes. A node in the HDFS cluster serves two functions: i) The Name Node function deals with the distribution of the block, tracking of their locations and the operations carried out by them and also stores a detailed metadata of the blocks. The Name node generally makes use of a server with high availability and backs up in a secondary Name Node in order to avoid the single point of failure issues. There is a single Name Node present for each cluster. ii) The Data node deals with the processing of the data blocks. Each HDFS cluster contains a number of data nodes.

• Hadoop MapReduce: can be best described as an engine that can process a vast amount of input data to produce an efficient output. Its operation can be classified into two phases:

1) The Map phase: in which the data blocks from the inputs are received and processed to generate an output which takes the form of a sorted list of key-and-value pairs.

2) The Reduce phase; in which the final output is produced after receiving, analyzing and merging the results from the mappers. This phase makes use of two trackers: a JobTracker (for job submission) and a TaskTracker (for Map and Reduce process execution).

The entire process of MapReduce can be summarized in nine steps, as follows:

i) The data files are loaded/streamed into the HDFS

ii) The JobTracker interacts with the Name Node to obtain the list of the holding nodes of the input files for scheduling map tasks.

iii) The unstructured data is transformed into the required format for processing using the “InputFormat” function after which the files are broken into block through the “InputSplit” function.

iv) The data blocks are transformed into key-value pairs form for the mappers using the "RecordReader" function.

v) Each mapper processes its input data blocks. The outputs which would also be in Key-Value pairs will have to be combined and sorted to produce a result which would appear like (Key, List of values {Value1, Value2,….etc.}). The "List of values"contains the different values of the same key which are gathered from different mappers and hence holds different data.

vi) The intermediate data is partitioned into smaller units, usually through a hash function. The outputs of the mappers are communicated to the HDFS and the Reduce process is not scheduled by the JobTracker until the TaskTracker reports the completion.

vii) The JobTracker schedules the reduce tasks on the nodes. And the mappers’ outputs are transferred to the reducers using the shuffling process so that the outputs that carry the same key are transferred to the same reducer.

viii) The inputs are merged and processed by the reducers and the final output is produced.

ix) The output is taken and organized into a suitable format for the application requirements using a function called “OutputFormat”. After which the “RecordWriter” function writes the final output into the HDFS.

For interaction with the platforms, Hadoop is generally integrated with other open-source technologies. Some of them are as follows:

• HBase (for data storage): it is a columnar, nonrelational, distributed database used for storage of data tables in Big Data environment allowing for random and fast read/writes access.

• Sqoop (for data integration: refers to SQL-to-Hadoop, it executes the Extract Transform Load (ETL) processes for importing data from non-Hadoop environments (e.g., RDBMS) to Hadoop and vice versa.

• Hive (for data mining): is an open-source data warehousing layer built on top of Hadoop’s main components (HDFS and MapReduce), it also provides SQL-like queries using HiveQL.

• Pig (for scripting): uses a high-level language called "Pig Latin" to write programs on top of Hadoop.

• Zookeeper (for coordination): is used for coordinating distributed application in Hadoop clusters.

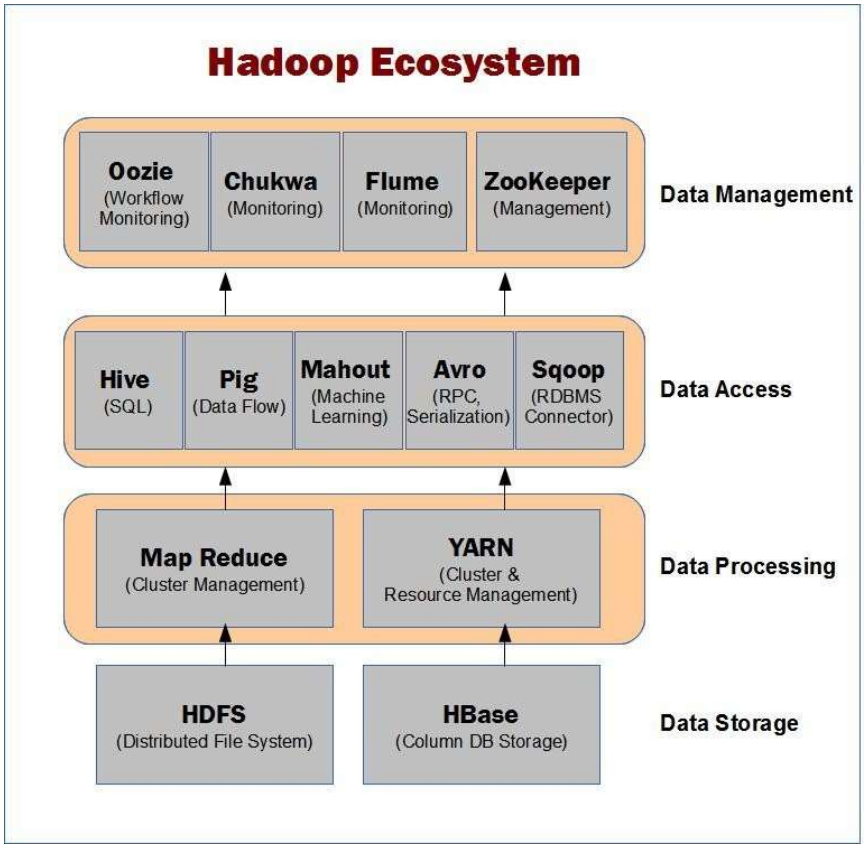


Figure 16 - Hadoop Ecosystem
Source: Udemy

• Mahout (for advanced analytics): is an open-source software which runs primarily on top of Hadoop, it holds implementations of data mining and machine learning algorithms.

Hadoop used to search through the whole data volume, filtering the noise, classifying the data, extracting the information, deriving the features, connecting the information and representing the output in a structured form.

2.3.4 Big Data Analytics and Visualization

The advanced analytics of data is applied in the intelligent layer which influences the operation processes of the city by the actions and predictions it provides. The following are the analysis techniques used:

• Machine learning: refers to the use of complex algorithms which enable computers to predict behaviors based on trained datasets. Machine learning can be used in a number of stages as follows; pattern recognition can be used to identify attributes in the data during

the listening stage, unsupervised and supervised learning to identify and classify the data during learning stage, unsupervised learning to identify and classify data that cannot even be identified by the trained datasets during the learning stage and to integrate the past and present data during connecting stage thereby providing relationships between them.

• A/B testing: where the changes or modifications that can be made to improve a situation is identified using a set of testing data. It is used for giving predictions at the predicting stage and also to increase the quality of data from poor datasets.

• Association rules learning: where unknown relationships between datasets are found.

• Regression: where the changes in the dependent variable are studied when one or more independent variables are modified. These are used in predictive analytics.

• Slicing and Dicing: where data from different dimensions are explored by deriving graphs, scoreboards, and dashboards

• Data mining: where techniques from machine learning and statistics are used to extract knowledge from data in the learning stage. It is used in pattern extraction.

• Semantics analytics: where Natural Language Processing is used to extract information from text data. It can be used in the listening and learning stages.

• Time series analysis: where signal processing and statistics are used to learn and extract value from a sequence of datasets.

• Spatial analysis: where data that comes from Geographical Information Systems (GISs) are analyzed.

• Anomaly identification: where unexpected events are detected by observing changes in the data. For instance, in a Smart Waste system, where trucks collect garbage from containers, this technique can be used to understand the behavioral change that would happen when the sensor detects the container to be full and yet the garbage is not collected.

There are three different types of analytics that are used in Big Data environment, they are as follows,

Descriptive analytics: is used to analyze the collected data and describe what had/is happened/happening in the city usually through reports, dashboards, real-time detections as well as identify why it had/is happened/happening through the use of forensics and data mining.

Predictive analytics: is used during the Predict stage to give insights on how things could turn out to be, it makes use of techniques such as a/b testing, regression, classifications and time series analysis. Based on the predictive analytics predictive models can be produced and they generally comprise of three parts:

• The attributes that are evaluated in the listening and learning stage which usually provides knowledge about a certain element, for instance, the energy consumption from smart meters which would help in understanding the energy consumed by the entity.

• Using the information already available information certain attributes are predicted. For instance, all the available characteristics of the energy consumed could provide detailed information such as the amount of consumption during the night, consumption peak, and so on.

• Using the already derived attributes and its data, the behavior of the element can be predicted. For instance, predicting the behavior of energy consumption to identify when the peak load might happen.

Perspective analytics: A model which is built is usually evaluated and tested using the past data where they are split into testing and training datasets. After the verification of the process, the model will be ready for deployment into the system. Since the data feeds would be continuously flowing in, the model will have to be able to satisfy the changes that are introduced due to it. For this process, the past and the present models are integrated at the connect stage. Under certain conditions, this analysis provides suitable decisions to be made in the given layer, for instance, reduce the cost of the peak load hours.

Furthermore, the intelligent visualization and representation of the data are critical for understanding the outcome of the analytics and for the extracted value. Thereby demanding the intelligent layer to be a well-suited user-friendly interface that helps in the monitoring management of the city based on the data that is collected. Reports, dashboards, real-time scoring, tracking and monitoring interface, and alerts are examples of analytics' visualizations.

2.4 Data Software Platform

A city which actively has multiple technologies embedded in its framework for effective functioning can only be called a Digital City. Whereas to make a city truly a Smart City the multiple systems which are integrated should work together to improve the city’s overall infrastructure in an intelligent manner.

The main technologies that are commonly used by the software platforms are Cyber-Physical Systems (CPS), Internet of Things (IoT), Big Data and Cloud Computing. Based on the requirements of the platform they work together in combination. According to Lee et al. (2013) a Smart City must be able to satisfy five requirements: Sensing and Networks requirements which IoT and CPS can deal with, Processing of Data which can be taken care of Big Data and Cloud Computing, the interfaces between the Smart City services which can be handled by the Cloud Computing Environment, the Security is also a crucial aspect. The International Organization of Standards (ISO) and the International Telecommunication Union (ITU) take into account only of Big Data, IoT and Cloud Computing to be the vital technologies for a Smart City while the National Institute of Standards and Technology (NIST) also considers CPS to be an important part of the Smart Cities. Other technologies which could be useful such as Ubiquitous and Mobile Computing, Machine-to-Machine Communications and Service Oriented Architecture (SOA) are also used in some platforms.

2.4.1 Microservices vs Monolithic Architecture

A Monolithic architecture based Traditional Service-Oriented Architecture (SOA) comprises multiple software features in a single application and database which is interconnected and interdependent. Though the tightly coupled dependence between the components and the functions allow for a single package, the monolithic architecture which the package is built upon lacks the flexibility to be able to support continuous development and delivery which is crucial in the rapidly changing and highly heterogeneous environment that is common today. Whereas in the case of microservices, each of it is dedicated to a specific function of the application and the architecture is generally composed of many such services. The microservices make use of HyperText Transfer Protocol (HTTP), Representational State Transfer (REST) or a message bus asynchronously to communicate amongst themselves. Due to the flexibility that the microservices, continuous, efficient and independent deployment of application function units is made possible. One of the key features that microservices

provide is fine granularity which is the ability of each of the micro-services to have a different framework, language or resource and loose coupling which is the ability of the components to stay independent of its deployment and not be influenced by the others'.

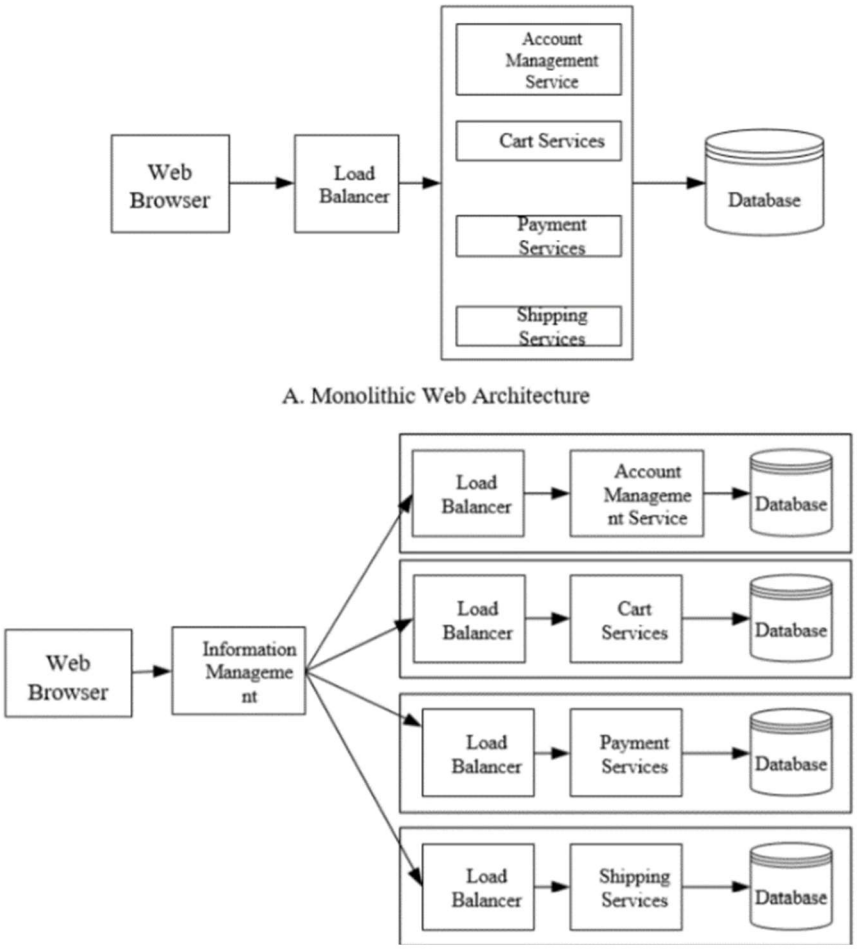


Figure 17 - Monolithic and Microservices architecture
Source: A Microservice-enabled Architecture for Smart Surveillance using Blockchain Technology

The monolithic architecture is being replaced with the microservices architecture including in the commercial web application platform. Services including individual account management, cart services, payment services, and shipping services were all taken care of by the monolithic web design version and all the information is stored and accessible in a single database. The downside, however, was that in the case of any of the services malfunctioning, the information flow for the rest of the system is halted and the system as a whole will have to be restarted. Whereas with the monolithic system, since each service is a layer and collectively accesses a single database even upon malfunctioning of a service the rest of the system can function without being affected. This makes the services independent and more reliable.

To enhance the security and scalability of applications an increased number of smart solutions are investigating the implementation of the microservices architecture. In one of the cases, it was implemented in accordance with IoT to help the planning of the bus rapid system in an intelligent transportation system. This architecture, when applied to a Smart City IoT platform, regards each microservice as a separate department of engineering which

makes it independent and offers more flexibility for selecting the development platform, and even without the use of a middleware communication protocols are simplified.

2.4.2 Platform Architecture

As these data platforms can have multiple configurations and architectures, they are divided into five categories for better understanding, according to the enabling technologies that each platform uses. Almost all of them use Cloud Computing as well as at least one more enabling technology, most commonly IoT and Big Data. Projects which already exist and those that are currently underway are discussed here to understand the various possibilities.

Internet of Things and Cloud Computing.

The platforms that use both IoT and Cloud Computing as enabling technologies are discussed herewith. SmartSantander is an experimental infrastructure to support the development and deployment of Smart City applications and services (Sanchez et al. 2014). The project is centered in Santander, Spain, with smaller facilities in other European cities. The platform processes a large variety of information, including data about traffic conditions, temperature, CO2 emissions, humidity, and luminosity. Currently, the project has placed more than 20,000 sensors in the city. Padova Smart City uses IoT to create a sensor network in the city of Padova, Italy. Using more than 300 sensors, the platform collects environmental data, such as CO2 emissions and air temperature and monitors street lights. A feature highlighted in this platform is the use of common protocols and data formats to allow interoperability among multiple city systems. The European Platform for Intelligent Cities (EPIC) project proposes a complete IoT Middleware to facilitate the use and management of Wireless Sensor Networks (WSN). This middleware aims to deal with the heterogeneity, interoperability, scalability, extensibility, and reconfigurability problems in a WSN.

ClouT proposes a two-layer architecture to collect data from the WSN and manage the sensors and actuators in the city network. The first layer is the Sensors and Actuators Layer, which handles data from the WSN. The second layer, the IoT Kernel Layer, manages and monitors the sensors and actuators network. Open Machine Type Communications (OpenMTC) is a Machine-To-Machine (M2M) based communication platform for Smart Cities. Its goal is to enable efficient communication among many devices, associating them with multiple services. To achieve this, the platform supports standard interfaces to various types of devices, data/event processing methods to achieve real-time performance, and easy application development, providing a software development kit. The analysis of the

aforementioned platforms led to identifying four major functional requirements: WSN management, management of the data collected from the city, services and applications management, and an infrastructure to make the data from the platform available to city applications. The analysis also led to identifying five non-functional requirements: adaptation, interoperability, scalability, extensibility, and reconfigurability. Two weak points of these platforms were identified: (1) the lack of pre-processing components to verify the integrity of the data collected from the city and make small transformations of the data, such as aggregations, and (2) most of the platforms do not include a discussion of security concerns.

Internet of Things, Cloud Computing, and Big Data.

The platforms that use IoT, Cloud Computing and Big Data as enabling technologies are discussed herewith. OpenIoT6 is an open source middleware for the development of IoT-based applications. It has an API to manage the WSN and a directory service to dynamically discover the sensors deployed in the city; it also has a layer for service definition and access. Big Data tools are used to store and analyze the data from the platform. A Smart City project called Vital is built on this platform and uses the term "Cloud of Things" to refer to the use of Cloud Computing and IoT.

The Concinnity project provides a platform for managing data and applications following the PaaS model, with which its authors built Big Sensor Data Applications. However, this platform focuses on multiple data sources, such as the WSN, social networks, and data from platform users. It also includes a service directory where developers can find and publish services facilitating its reuse. OpenIoT and Concinnity offer developers tools to implement applications directly on the platform. OpenIoT allows the mash-up of the services defined in the platform and automatically creates a visual interface for end-users. Concinnity provides a set of development tools, such as a Workflow Editor and Engine, a Service Publisher, and an Application Editor.

Sentilo is a platform that deals with the management of sensors and actuators, designed for Smart Cities that desire openness and interoperability. Sentilo uses IoT concepts to control the WSN and Cloud Computing to share data with the applications. Big Data tools are mainly used to collect and store data from sensors, ensuring platform scalability. Originally designed for Barcelona, after its deployment, the City released the code for the Sentilo project under the LGPL and EUPL open source licenses. The main functional requirements identified for this group of platforms were the management of a WSN, management of data lifecycle (collect, store, process), making the data from the platform publicly available, a service

directory for application developers, and tools for application development. As non-functional requirements, interoperability and scalability were identified.

A weak point of these platforms is the lack of stream processing tools to analyze real-time data from the city, an important requirement for many Smart City applications. Another problem is that most of the platforms do not support the customization of services with citizen data. Despite the privacy problems, offering context-aware, customized services to citizens is highly desirable.

Cloud Computing and Big Data.

A platform for Smart Cities based on Cloud Computing and Big Data, whose main components are data management and service hosting were presented by Vilajosana et al. (2013). It includes an Open Data API that allows third-party applications to access the data stored on the platform. Big Data tools are used to collect data streams and analyze data, such as prediction and inference. SCALable LOgging Platform for Smart City (Scallop4SC). Big Data is also used to process a large volume of data gathered from smart buildings. The platform uses information about the building, such as water and energy consumption, temperature, air humidity, and the amount of garbage generated. Periodically, the buildings send data to the platform for processing. It uses the MapReduce algorithm to achieve the objective of analyzing smart building data.

CiDAP is a Big Data analytics platform deployed into the SmartSantander testbed. The platform uses data collected from SmartSantander and analyzes it to understand the city's behavior. The main components of this platform are the agents, which collect data from the SmartSantander platform; the Big Data repository for storing the data; the Big Data processing for intensive data processing and analytics; and a CityModel server, responsible for interfacing with external applications. This platform uses Apache Spark to process the data. A Smart City architecture based on Big Data to achieve the necessary availability and scalability required for a Smart Cities platform was recently proposed by Khan et al, 2013. The architecture has three layers: one to collect, analyze, and filter data; another to map and aggregate data to make it semantically relevant; and a third layer where users can browse and recover the data processed from the other two layers. The implementation of the architecture uses only open source projects, and the authors have presented tools for all layers.

SMARTY is a project aimed at providing tools and services for mobility and flexible city transport systems. Its software platform collects data from multiple sources, such as traffic

flow, user location, transport service delays, and parking availability. A network of low-cost sensors collects data from the city, and social networks are continuously monitored to retrieve data from citizens. The platform processes the massive amount of data generated by the city with data-mining techniques, such as classification, regression, and clustering.

The main functional requirements identified for this group of platforms were data management, such as collecting, analyzing, and visualizing data; large-scale data processing, such as batch and real-time processing; and the use of semantic techniques combined with Big Data. As nonfunctional requirements, scalability and adaptation were identified. Most of the platforms in this section lack an IoT layer and do not indicate how the data is collected from the city; the exception is CiDAP, which uses the SmartSantander testbed as an IoT middleware. Most also lack a discussion about security concerns.

Cloud Computing.

A two-layered service platform for creating Smart City applications was presented by Piro et al, 2014. The first is a low-level layer that controls the communication among the city WSN devices. The second layer collects the data from the devices and provides services for the development of applications that use the city data. U-City is a platform for creating smart ubiquitous cities that offers several service management features, such as autonomic service discovery, service deployment, and context-aware service execution. It also offers pre-defined services, such as an inference engine, a context-aware data service, and a portal for the platform's management.

Gambas, a middleware for the development of Smart City applications, supports data acquisition, distribution, and integration. The platform also provides an application runtime to facilitate the development and deployment of services using city data and a service registry. The middleware supports context-awareness so Smart City services can adapt to the citizen situation, behavior, and intent. All communication in the platform is encrypted to ensure citizen's privacy and security. Civitas is a middleware to support the development of Smart City services by facilitating the development and deployment of Smart City applications and to avoid the emergence of "information islands", that is, disconnected applications that do not share relevant information. Citizens connect to the middleware via a special device called the Civitas Plug, which ensures privacy and security. The middleware has two main design principles to facilitate application integration: Everything is a Software Object, which promotes the consistency of the software design and reusability of the middleware; and Independence of the City Layout, meaning that city services should work with more than one city layout.

The main functional requirements identified for this group of platforms were service management and data management. As non-functional requirements, security, privacy, and context awareness were identified. A drawback of the platforms presented in this section is that none of them use known frameworks to implement components, such as the inference engine and processing tools, which might make platform maintenance difficult. Another problem is that the platforms do not describe a mechanism to allow external access to the platform data.

Cloud Computing and Cyber-Physical Systems.

Gurgen et al, 2013 presented a middleware for Smart Cities autonomic services, which includes many self- properties, such as self-organization, self- optimization, self-configuration, self-protection, self-healing, self-discovery, and self-description. They justify using cloud computing to provide scalability, reliability, and elasticity to the platform, which provides application developers with the contexts of both individual users and the city. A CPS-based platform, whose main characteristic is self-configuration and self-adaptation capabilities in smart environments (including Smart Cities) was proposed by Privat et al, 2014. This platform provides a shared distributed software infrastructure that collects data and reacts to changes in the environment.

An event-based CPS platform, which uses an event manager to manage and generate cooperation among M2M components was proposed by Wan et al, 2012. This platform provides data and services to third-party applications through a publish/subscribe module. The platform also enables the design of event processing flows to manage mission-critical wireless messages.

The main functional requirements identified for this group of platforms were the autonomic reaction to changes in the city environment, communication among city devices, and a publish/ subscribe mechanism for applications to communicate with the platform. While configurability, adaptation, and context awareness were identified as non-functional requirements. The platforms in this section focus on the deployment, configuration, and execution of CPS devices in the city, but they lack important requirements, such as the monitoring and publication of the data from the devices. They also do not describe any mechanism to verify the data collected from the city, discarding inconsistencies.

2.4.3 Commercial Smart City Initiatives

Big ICT companies, such as IBM, Microsoft, and Oracle are working on Smart City initiatives. Oracle has a Smart City solution based on three platforms. The Smart Innovations Platform enables the communication of the city government with the population via chat, phone, and email. The Smart Process Platform for continuous monitoring and improvement of city services, helping to identify which services to prioritize, extend, consolidate, or discontinue. The Smart Infrastructure Platform enables the integration and interoperability of legacy IT infrastructure and new city services. Many cities around the world, such as New York, Madrid, and Hong Kong, already use Oracle Smart City solutions, according to Oracle.

Cisco is working with cities, such as Amsterdam and Nice, to deploy an IoT infrastructure. In Nice, the city government implemented, with Cisco's support, a four-layered Smart City platform to collect city data. Layer 1 comprises sensors and the networking infrastructure to gather and transmit the data. Layer 2 is responsible for processing, storing, and analyzing data, which occurs in distributed points across the city and thereby boosts the platform's scalability. Layer 3 is a central computation infrastructure enabling the integration, storage, and sharing of city data. Layer 4 comprises Smart City applications developed using the platform services and data. Cisco also presents a list of non-functional requirements that are necessary to handle a Smart City platform, such as security, extensibility, scalability, flexibility, and interoperability. The Hitachi vision of Smart Cities presents three main phases in the ITC infrastructure. First, the data are collected from households, buildings, and other end-user devices. Second, the data are analyzed by information systems. Finally, real-time data are provided for Smart City services and applications. Hitachi is already working on many Smart City projects in Japan, mainly in energy and water distribution.

Microsoft's CityNext project presents ideas for developing Smart Cities initiatives, such as the use of Cloud Computing, Big Data, Internet of Things, and Social Networks. They also introduce four main objectives: engaging the city population, empowering city employees to increase productivity and efficiency, transforming the city with new digital services, and optimizing city operations and infrastructure. Some cities that already use Microsoft services are Buenos Aires, which developed a city dashboard, Tacoma, with an education analytics and research system, and Helsinki, which developed a solution to collect and analyze data from bus sensors to reduce fuel consumption. A detailed system or platform software architecture from Microsoft could not be found. IBM has many different Smart Cities projects, among the most cited of which is the Intelligent Operations Center that helps cities to monitor and manage resources, incidents, and events in real time. One important requirement of this system is the integration of systems and data from the various city

departments and legacy systems. However, relevant technical information about IBM's solutions is not available.

2.4.4 Requirements for Smart City Software Platforms

The functional and non-functional requirements extracted from the analyzed platforms are analyzed to understand what are the key requirements that a Smart City should meet.

Functional Requirements.

The main goal of a Smart Cities platform is to facilitate the development of Smart City applications. Towards this aim, most of the analyzed platforms implement requirements for collecting, managing, and sharing city data and for providing tools to facilitate the development of Smart City applications.

—Data Management: Most Smart Cities platforms implement this requirement, which includes the collection, storage, analysis, and visualization of city data. The analyzed platforms use different techniques for this requirement, such as relational databases, Big Data tools, and customized tools implemented by the platform development team.

—Applications Runtime: Some platforms focus on managing the execution of their applications, aiming to facilitate the applications' deployment and integration. Some platforms provide a complete environment for developers to deploy their applications; while others offer an execution runtime service for applications developed with tools the platform provides.

—WSN Management: Many of the analyzed platforms have a WSN management layer to control and monitor the devices deployed in the city. Most of these platforms use IoT concepts to organize and manage the WSN. Other platforms do not explicitly mention this but indeed include a software layer to manage the city network devices. Some platforms include features to manage all the device activities, such as adding, removing, and monitoring the sensors and actuators. Two platforms describe a WSN deployed in a city: Padova Smart City, with 3,000 sensors, and SmartSantander, with more than 20,000 sensors.

—Data Processing: Some platforms use specific processing components, such as inference engines, workflow processing, and Big Data processing tools. These components process large datasets, and their main purpose is to analyze, verify, aggregate, and filter the data from the city. In addition, some platforms analyze data streams in real time.

—External Data Access: Almost all platforms describe an interface for external applications to access the platform data, most commonly an API. Some platforms use REST, while others

use cloud computing concepts to provide the city data as a service, and one proposes an open data platform. Also, one platform uses the publish/subscribe paradigm to make the data and services available to applications.

—Service Management: Most of the analyzed platforms adopt a Service-Oriented Architecture, in which the platform functionalities are offered by services. Some of them use services to provide features to applications, such as access to raw sensors data and analyzed data, and workflow engines. Others enable developers to deploy services on the platform and make them available to other applications. Some platforms also use service composition and choreographies to create new services or applications.

—Software Engineering Tools: Some platforms provide a set of tools for the development and maintenance of services and applications. For describing and implementing applications, some platforms create visual interfaces. Other platforms provide workflow design tools to define data or service flows and create Smart City applications. Moreover, some platforms use analytics and reporting tools to facilitate the development of data visualization and reports, and two platforms describe the use of a Smart City application SDK.

—Definition of a City Model: Some platforms provide a city model to facilitate the manipulation and understanding of the platform data and to facilitate the integration of the collected data. For example, the city model in Cheng et al. (2015) is used to allow queries in the data from the city sensor network. Privat et al. (2014) use a finite-state model to represent the possible city data flows. Based on the aforementioned functional requirements, it can be observed that the main platform's activities aim to control the city data lifecycle: (1) collecting the data with a WSN, (2) managing the data in the platform, (3) processing the data using city models, and (4) sharing the raw and processed data allowing external access. These activities are highly related to the enabling technologies, such as IoT with the WSN management, Data Management and Processing with Big Data, and Service Management with Cloud Computing.

Non-Functional Requirements.

Most of the non-functional requirements of Smart City platforms relate to large, heterogeneous distributed systems, such as scalability, adaptation, and interoperability. Other non-functional requirements relate to the manipulation of critical and personal data from citizens, such as security and privacy.

—Interoperability: Different devices, systems, applications, and platforms comprise a Smart City environment, all of which must operate in an integrated fashion and may include sensors from multiple vendors, systems implemented in different languages, platforms that

share data and users, and legacy systems should communicate with the new platforms. Previous work in the field adopted several techniques to handle this need, including interoperable objects, adopting generic and standard interfaces, applying Semantic Web to integrate all platform components, and using a naming mechanism to recognize different devices or data sources.

—Scalability: A Smart City platform's number of users and services and amount of data will be massive and increase over time. For example, in the SmartSantander testbed, there were more than 20,000 sensors in a city of 178,000 inhabitants collecting a large amount of city data; and CiDAP collected more than 50GBs of data in three months. This non-functional requirement is relevant to many functional requirements, such as WSN management, data management, and service management.

—Security: Malicious users can make fraudulent use of services and data provided by the platform. Many platforms have a component or describe mechanisms to handle security, avoiding attacks to the city infrastructure and information theft.

—Privacy: A Smart City platform collects and manipulates several citizen-sensitive data, such as medical records, user localization, and consumption habits. The challenge is to use these data while hiding or to avoid saving identifiable information. Some of the strategies used to achieve this requirement are cryptography, tokens to control the access to the data that users can manipulate, and anonymization.

—Context Awareness: As the city and user situation can change over time, many applications and services can provide better results using contextual information such as user information like location, activity, and language, or city information, such as traffic conditions, climate, and air quality. Examples of context use are displaying a different language in an application to a tourist or changing the route of a user to avoid polluted areas.

—Adaptation: Related to context awareness, many platforms adapt their behavior to context to achieve fault-tolerance, choose a closer server to improve efficiency, decide for batch or real-time processing or adapt data from multiple data sources. Adaptation is most used in platforms that use CPS as enabling technology, but other concepts are used to meet this requirement as well, such as semantic technologies.

—Extensibility: The capability to add services, components, and applications to the platform is important for assuring that it meets evolving system requirements and user needs. Hernández-Muñoz et al. (2011) state that easy extensibility is valuable because one cannot know in advance what services a city will need in the future. Scallop4SC uses materialized

views that developers extend to implement their applications. Some platforms employ only open source tools, facilitating the platform's extensibility. CiDAP offers extensibility to enable the use of the platform in cities of different scales.

—Configurability: A Smart City platform has many configuration options and parameters that define its behavior at execution time, such as defining pollution and congestion thresholds and the priority of services. Thus, it is important to allow (re)configuration of the platform's many variables. Two of the platforms highlighted the importance of self-configurability capacities given the massive number of configurations a Smart City platform needs. Other platforms provide a portal to centralize the configurations.

Based on the above non-functional requirements, it can be observed that some of them are very important to many functional requirements, such as scalability, which is valuable to the WSN and data management; security and privacy, which are important for all data requirements; extensibility, which is required for service management; and configurability, which is also important for all functional requirements.

2.5 Artificial Intelligence

It has been almost 60 years since Artificial Intelligence (AI) was defined as an academic discipline of computing science. After 60 years of development, AI is still a hot research field, but the contents of AI has been changed a lot. Leaders of the world's most influential technology firms, Google, Microsoft, and Facebook are demonstrating their interests in AI. AI together with other technologies, such as sensors, IOT, cloud computing, big data, data analysis, and smart decision-making are driving a new industrial revolution, which is called as Industrial 4.0 or the fourth industrial revolution. The use of smart and advanced technologies could automate near half of all existing work activities in the world and it saves some $16 trillion in wages by

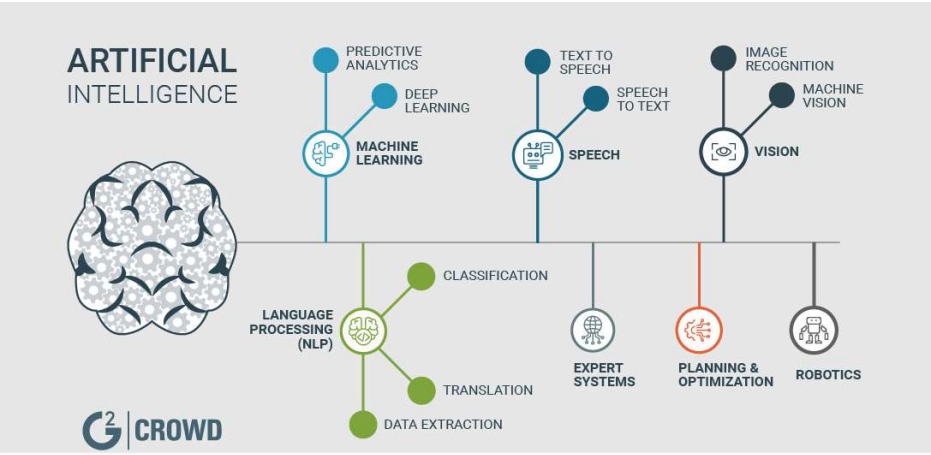


Figure 18 - Artificial Intelligence Overview
Source: G2 Crowd

replacing human operators. It was predicted that revenues from AI by 2020 will be increased by more than $47 billion. The industry will invest in many promised fields, such as a smart factory, smart logistics, medical diagnosis system, big data and decision management and fraud analysis in business.

Artificial Intelligence plays an important role in making safer cities, transport smarter and urban planning optimal. The Smart City offers a rich avenue for applying AI algorithms like machine learning, planning, scheduling, semantics, knowledge representation, data integration, logic, trust, and agents. AI can track citizen's habits, activities, and behavioral characteristics. Data and products can be personalized to meet and anticipate each user's unique and changing needs. Each citizen will have one's own digital personal assistant. Artificial intelligence can help governments handle their regulations monitoring by creating a natural language processing system to read through the legalities of regulations and reassemble the words into a set of computer-understandable rules. IoT, AI, VR, AR, and bots technologies are changing the way data is created, collected, interpreted, and communicated.

As governments begin to rely more on data-driven Artificial Intelligence applications, the new applications lead to new issues, security, and privacy concerns. Each government department also needs to have a transparent system for total audit ability so one can see who did what, and when. Deep Learning platform users should identify erroneous or incomplete data to avoid misleading decisions. The new AI applications introduce a number of business, security and privacy issues which will have to be addressed. It will be important to ensure that these intelligent applications are developed in a way that they will provide the desired benefit and that the user can trust the advice and services provided. It will be important to be able to detect and isolate infected or malicious AI programs immediately and develop the effective policy and laws for governing their development and use so that personal information is safeguarded and not misused.

Artificial Intelligence has been part of our imaginations and issues in research labs since computers became available and popular in real life. To be more specific, it's in 1956, when a handful of computer scientists rallied around the term at the Dartmouth Conferences and birthed the field of AI. The idea is to make machines be able to carry out tasks in a way that it could be considered "smart" or "intelligent". The concept is very broad from today's perspective. AI can be divided into two types: Computational Intelligence (Numerical AI or soft computing). Computational Intelligence can be further classified by several groups: 1. Artificial Neural Networks (ANN); 2. Fuzzy Logic Systems (FLS); 3. Evolutional Computing (EC); and 4. Swarm Intelligence (SI).

2.5.1 Computational Intelligence

Computational Intelligence (CI) or Soft Computing (SC) can be called Numerical AI (verse Symbolic AI) has become a rapid developing field in computer science and been an advanced information processing technology classified CI into several groups, such as Artificial Neural Networks (ANN), Fuzzy Logic Systems (FLS) and Evolutionary Computing (Genetic Algorithms). The human brain is able to process data and information quickly and accurately because of its biological network structure and approximate reasoning ability. Both the functions of Artificial Neural Networks and Fuzzy Logic Systems are based on the mechanisms of the human biological brain. The ANN simulates physiological features of the human brain and has been applied for non-linear mapping by a numerical approach. The FLS simulates psychological features of the human brain and has been applied for linguistic translating by the use of membership functions.

2.5.2 Artificial Neural Networks (ANN)

Because the functions of Artificial Neural Networks are based on the biological neural network structure of the human's brain, they are able to analyze complex/complicated problems and to make correct decisions. The powerful and paralleled computing architecture of ANN demonstrates their learning abilities. ANN can learn from historical data, previous examples, and experience to do modeling, classification, clustering, and predication. If the environment changes, ANN is able to modify their behavior. There are many ANNs application areas such as dynamic system modeling, pattern recognition, data classification, business predication, and multi-dimension mapping and novelty detection.

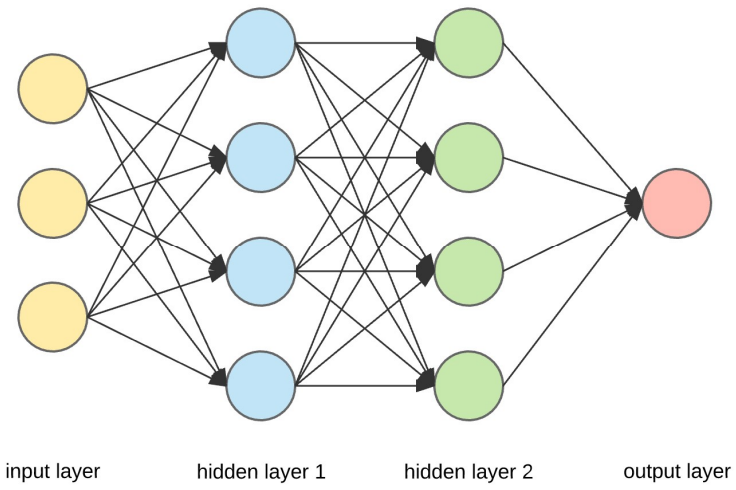


Figure 19 - Artificial Neural Network
Source: Towards Data Science

ANN uses artificial, software-based computing technologies that approximate the function of neurons in a brain are connected together. They form a 'neural network' which receives an input; analyses it; makes a determination about it and is informed if its determination is correct. If the output is wrong, the connections between the neurons are adjusted by the

algorithm, which will change future predictions. Initially, the network will be wrong many times. But as millions of examples are fed in, the connections between neurons will be tuned so the neural network makes correct determinations on almost all occasions. Practice makes (nearly) perfect.

2.5.3 Deep Learning (DL)

From the view of the structure, DL is on kind of ANNs or it can be said that the application of Artificial Neural Networks (ANNs) to learning tasks that contain more than one hidden layer. The types of DL approaches have been growing increasingly richer, encompassing variety of Artificial Neural Networks with multiple processing layers, unsupervised or supervised learning algorithms. There exist about more than 10–15 types of DL algorithms. Four types of DL are presented, which are mostly used in predictive maintenance field: 1. Deep Feedforward Networks (DFN); 2. Long Short-Term Memory (LSTM); 3. Convolutional Networks (CN); 4. Deep Belief Networks

Deep learning is useful because it undertakes both the tasks of feature specification and optimization. Deep learning uses the same principles of ANNs, which approximate the function of neurons in a brain, are connected together. Using this approach, the following tasks can be done:

- Speech and image recognition;
- Nature language translation in real-time;
- Voice control;
- Prediction of genetic variation;
- Customer relationship management;
- Medical diagnosis;
- Prediction of a machine health condition; and more.

The main difference between ordinary ANNs and Deep Learning Algorithms is that in DL the features are not designed by developers but learned/extracted from data itself through a generalized self-learning procedure. Deep learning is able to offer a more suitable and convenient way of treating feature extraction problem.

However, DL is not well suited to every problem. It typically requires large data sets for training. It has a 'black box' problem—it can be difficult to know how a neural network developed its predictions.

There are many commercial applications driven by DL technology. The best example is that Google has developed their AlphaGo system based on the Deep Learning Algorithm by learning the game of Go well enough to beat a professional Go player. It makes people focus more on AI and DL technologies.

2.6 Applications of the Big Data Analytics and AI

2.6.1 Predictive Maintenance

Deep learning has proven to show superior performance in certain domains such as object recognition and image classification. It has also gained popularity in domains such as finance where time-series data plays an important role. Predictive Maintenance is also a domain where data is collected over time to monitor the state of an asset with the goal of finding patterns to predict failures which can also benefit from certain deep learning algorithms.

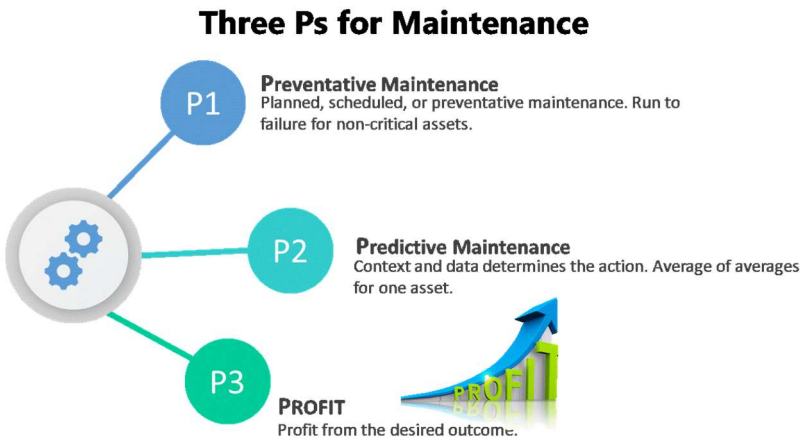


Figure 20 - Planned, Preventive, and Predictive Maintenance
Source: ARC Advisory Group

More recently, DL, as the latest research area of ANN, has accelerated its application in maintenance and service. In the research area of predictive maintenance, conventional ANN methods have proved their values in fault diagnosis, e.g., BP supervised ANN and SOM unsupervised ANN perform well in the fault detection and classification, however, when the target condition is beyond history data, they don't have the ability to predict potential failures. This problem is common and challenging since, in many situations, one may not have the data under faults, some machines may run several years without any failures). However, it may exist potential faults that may occur one day, and when it happens, it may cause terrible disasters in both economy and personal safety.

An important part of modeling a failure predictor is selecting or constructing the right features, i.e. selecting existing features from the data set, or constructing derivative features, which are most suitable for solving the learning task.

Traditionally, the features are selected manually, relying on the experience of process engineers who understand the physical and mechanical processes in the analyzed system. Unfortunately, manual feature selection suffers from different kinds of bias and is very labor intensive. Moreover, the selected features are specific to a particular learning task and cannot be easily reused in a different task (e.g. the features which are effective for predicting failures in one production line will not necessarily be effective in a different line). A general framework has been developed for predictive maintenance based on DL. It is expected that it may overcome the disadvantage of ANN in predictive maintenance mentioned above.

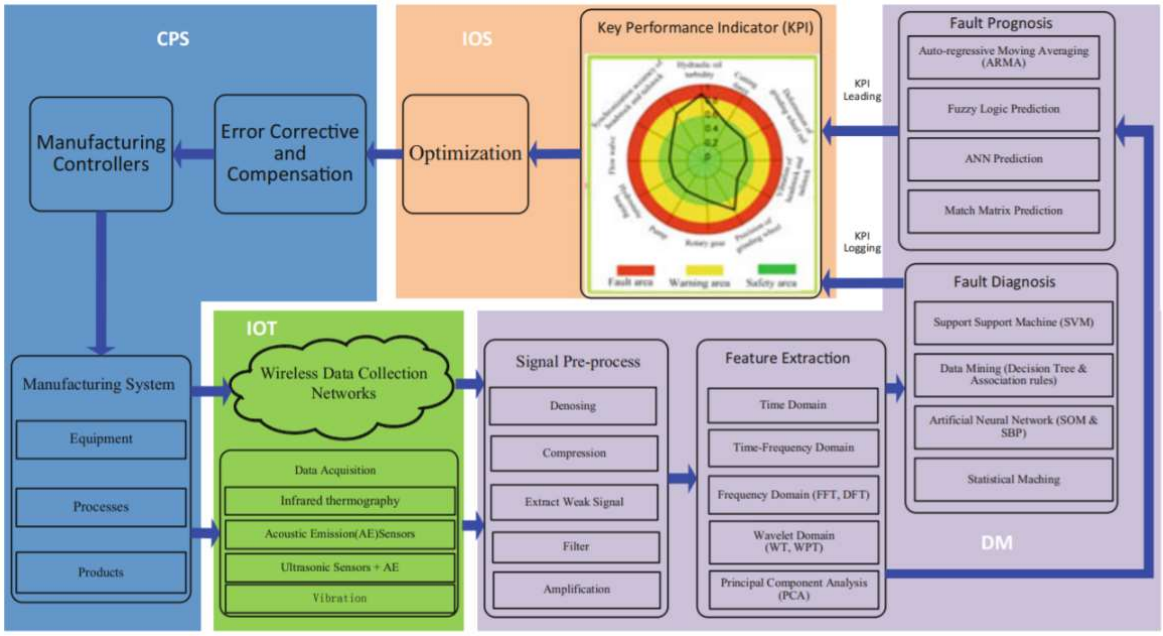


Figure 21 - Intelligent predictive maintenance for fault diagnosis
Source: Industry 4.0 scenario

The framework has been tested in several projects: Green Monitoring, Monitor X and rotational machinery rig in KDL lab at NTNU. Deep Learning techniques investigated in the Green Monitor project offer an alternative to manual feature selection. It refers to automatically extract features from the raw data that are most suitable for solving a particular learning task. Predictive Maintenance can benefit from such automatic feature extraction to reduce effort, cost and delay that are associated with extracting good features.

One interesting commercial application of DL for predictive maintenance is that 3DSignals (ultrasonic sensors), relies on the DL algorithm to monitor sawing machines, learn the audio patterns from troubled machines. Then the system can predict failures of machines to

realize predictive maintenance strategy. 3DSignals has already begun talking to use the deep learning service to detect cavitation problems in a hydropower plant.

Chevron has launched an effort to predict maintenance problems in its oil fields and refineries, a capability that many companies have been working for years to cultivate and is just now gathering momentum. In 2017, Chevron signed a seven-year deal to make Microsoft's Azure its primary cloud supplier. The partnership will give Chevron's engineers access to data in one cloud repository instead of in different silos within the organization.

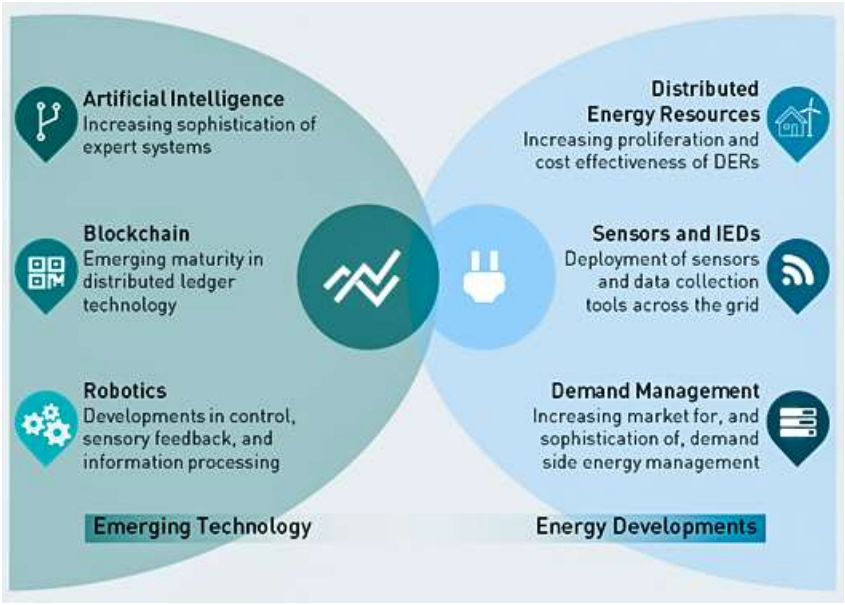


Figure 22 - Management and operation of the future power system and its components
Source: Watt Logic

In recent years, advancements have been made in the quality and affordability of sensors, as well as the cloud-based platforms required to gather and analyze data being streamed from devices in the field. It's also become easier and quicker to outfit equipment with wireless sensors, whereas in the past, sensors typically required weeks-worth of wiring and installation. Chevron expects to outfit oil equipment with sensors for predictive maintenance by 2019 in a wide-scale pilot program, with full adoption for many of the machines expected by 2024.

Mitsubishi Electric is investing in AI research as the company believes this will help and support the manufacturing industry. AI can be used to adjust parameters, perform voice/face recognition, search for problems, and assist in both preventive and predictive maintenance. Mitsubishi Electric has introduced Maisart (Mitsubishi Electric's AI State-of-the-ART) - a compact AI system that offers scalability and can be adopted in manufacturing industries. Mitsubishi Electric has already implemented its AI solution-on-a-chip in their air conditioning systems and now have plans to introduce speech recognition function into motion control and other factory automation products. Successful case studies on a driver and car health monitoring; how AI separates speech from background noises;

preventive/predictive maintenance; and maintenance with augmented reality were discussed.

2.6.2 Demand Side Energy Management

Moghram et. al. in his paper evaluated five load forecasting techniques - multiple linear regression, stochastic time series, general exponential smoothing, state space method, and AI or artificial intelligence-based approach. The authors implemented the techniques to generate an hourly, load forecast for the next day using the data from a southeastern utility in the US. The authors briefly described the implementation of each technique and compared and analyzed the results. Authors K. Liu et.al. proposed and evaluated three techniques for load forecasting, which were ANN, FL, and a time-series auto-regressive (AR) model. Although the conclusion that AR-based model is less efficient than the other two models wasn't clearly explained, the load series data was also considered as a stationary data, which contradicts the fact that load profiles are dynamic sources.

Statistical modeling technique was discussed in more recent papers for load forecasting. For example, a regression-based load forecasting approach using the PG&E dataset was proposed. Other regression-based approaches proposed deals with the use of a weighted least squares technique, temperature modeling (implementing various heating and cooling functions), weekday and weekend modeling etc. As a modification to the basic regression technique, Haida et. al proposed a transformation technique to model the nonlinear relationship between load and weather variables. The transformation technique was used with the Tokyo Power Utility Corporation dataset to forecast short-term load. A unique approach of forecasting a daily cumulative energy consumption forecast before an hourly load forecast was given by Ruzic et. al., wherein a two-step multiple linear regression (MLR) models were used for prediction. The first step of the MLR was used to predict the cumulative energy consumption of the day, and the next step predicted the hourly load profile. Works in proposing a probability density-based estimation of load forecasting. The load forecast was the conditional expectation of the load given the explanatory variables including time, weather conditions, etc.

Time series analysis of load data is another way to forecast load profile. Autoregressive (AR) and Autoregressive Moving Average (ARMA) techniques have been used for load forecasting in recent years. A combination of auto-regressive moving average model and regression techniques has been presented. The regression part is used to predict the peak and valley loads and ARIMA models are applied to the data to make the hourly load forecast. A 3rd

order polynomial for the temperature attribute was proposed to reflect the nonlinear relationship between the load and temperature. A supervised time-series model, that takes a pre-defined manual input as the primary forecast and then formulates a regression model using the available data has been proposed.

Artificial Neural Networks (ANN) are also highly used in performing load forecasting. Although models based on statistical methods generally perform well, in case of an abrupt change in the model attributes or the presence of statistical glitch in data, deficiencies arise and the prediction accuracy dips. This greatly affects the load patterns and load profile. AI-based techniques like ANN and fuzzy logic can cope with this kind of problem and perform predictions without any loss of accuracy. An ANN-based load prediction model using the back-propagation model and the radial basis function model respectively, with a performance review of both the models. Has also been presented as well as the use of ANN to perform real-time load forecasting was proposed. Real-time data from a local utility is used to forecast the load on an hourly basis. To further improve the prediction efficiency, hybrid schemes employing Support Vector Machines (SVM) and ANN has been proposed. The proposed model consists of two modules, the first one is used to predict the peak load and the second module is used to predict the hourly load.

In addition to ANN, Fuzzy Networks or Fuzzy Neural Networks (FNN) also forms one of the most used load forecasting techniques. Fuzzy logic is an approach to make partial decisions, not a complete 0 or a complete 1, but more of a fraction between 0 and 1. The idea resonates with the idea of the likelihood of an event to be true or false, rather than completely true or false. A long-term load forecasting technique using a fuzzy logic approach was proposed and according to the authors, fuzzy logic outperforms ANN in long-term forecasting, due to the increased gap between the weather conditions and load profile. Authors of "A novel approach to short-term load forecasting using fuzzy neural networks" proposed a unique fuzzy network for load prediction for each individual day of a week, which leads to a load forecast model that forecasts the peak and the valley load and calculates the hourly load profile using the available load data.

Despite the advancement in load forecasting techniques, none of the techniques guarantees a 100% prediction accuracy. Also, there is no single benchmark technique that can be used in any case of load forecasting. The desirable model varies with the data availability and the forecast objectives. Although advanced models like ANN and fuzzy logic offer a high degree of accuracy, they also increase the prediction complexity of the whole system.

In the UK, Upside Energy uses machine learning to manage a portfolio of storage assets to support the grid and help with its flexibility. It has developed a cloud service that aggregates

the energy stored in systems people and businesses already own such as uninterruptible power supplies (UPS), solar PV systems (solar panels), electric vehicles (EV), domestic heating systems, etc. This creates a Virtual Energy Store that could be sold to the grid to help it balance supply and demand. The revenue generated from these services are shared with device owners and manufacturers.

Aggregator **Open Energi** is developing such an AI model. It collects between 10,000 and 25,000 messages per second relating to 30 different data points and performs tens of millions of switches per year. This data is being used to train a deep learning model which combines asset-level constraints from a bottom-up approach with portfolio-level modeling that can tweak the outputs to improve the aggregated solution.

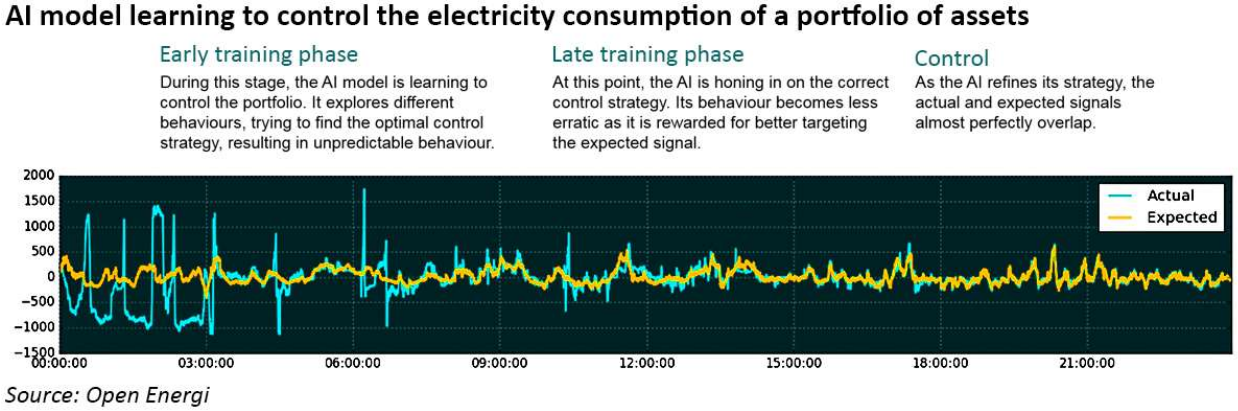


Figure 23 - Demand Side Energy Management

The model can look at a sequence of actions leading to the rescheduling of power consumption and make grid-scale predictions identifying the costs of various actions.

An early entrant in this field is Swiss-German firm **Alpiq**, which launched an intelligent system called GridSense in 2014, which aims to make imperceptible changes to the energy consumption of domestic or commercial equipment into which it is integrated (such as boilers, heat pumps, charging stations for electric vehicles) based on measurements such as grid load, weather forecasts and electricity charges.

By recording information on past usages of the accumulator at a point of connection to a grid it is possible to locally estimate the future usage of the accumulator; and programming the energy flow between the grid and the accumulator, which may be in either direction, on the basis of the estimated information. This allows the charging process to be optimized, using machine learning and data mining techniques in order to determine the optimal level of charging when the vehicle is connected to the grid, avoiding the need for full charging if only a small amount of energy is needed for the next trip.

The complexity of the electricity system is increasing, new technologies such as renewable generation, storage, and electric vehicles are changing the fundamental dynamics of the grid, and the growth in demand-side response, changing consumer behaviors, and the emergence of connected devices are creating a major shift in the interactions affecting the system.

2.6.3 Load Forecasting

Forecasting is an integral part of any sphere of human activity and energetics is not an exception. With the recent integration of smart grid systems to today's power systems and the increasing penetration of renewable energy sources, the process for meticulous electric load forecasting becomes more complex, calling for more effective techniques for the electric load forecasting in order to accomplish excellent planning, management, and operation of electric power systems. Moreover, due to nonlinear and nonstationary features of electric loads, which are affected by seasonal effects, weather conditions, socioeconomic dynamics, and random effects, electric load signals are characterized by high unpredictability, making the electric load forecasting a very arduous challenge.

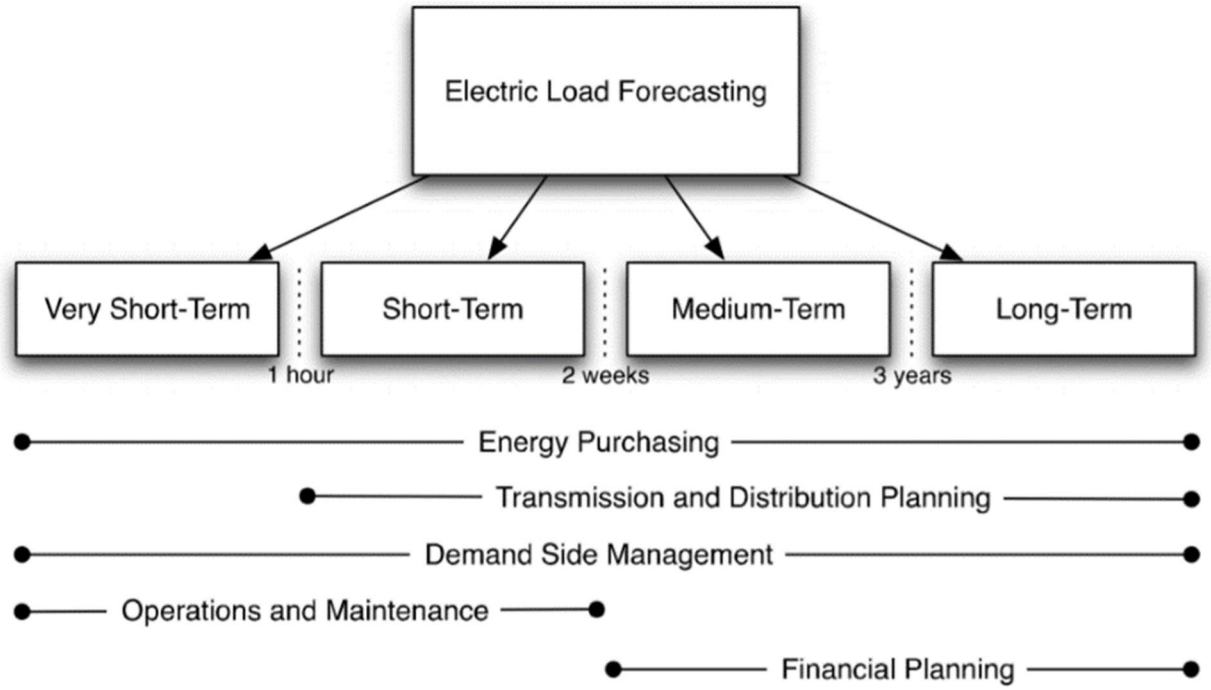


Figure 24 - Electric load forecasting applications and classification
Source: A Review of Artificial Intelligence Techniques

With respect to the time period, there are four categories of electric load forecasting: (1) long-term, among 3-year and 50- year electric load is predicted, (2) if the forecast ranges from 2 weeks to 3-year, then it is considered as medium-term electric load forecasting, (3) short-term electric load forecasting (or short-term load forecasting, generally abbreviated

as STLF in the literature) refers to hour, day or week ahead predictions, and (4) very short-term electric load forecasting which includes few minutes to an hour ahead forecasting of electric loads.

For strategic planning of the development of the electric power systems, both long-term and medium-term forecasts have a great significance which includes scheduling of construction of new generation and transmission capacity, maintenance scheduling, as well as long-term demand side measurement and management planning. However, an accurate STLF technique can alleviate operating costs, keep energy markets efficient, and provide a better understanding of the dynamics of the monitored system. On the contrary, an erroneous prediction might cause either a load overestimation, which leads to the excess of supply and reserve and consequently more costs and contracts curtailments for market participants or a load underestimation resulting in failures in gathering adequate provisions, hence more expensive complementary services.

The techniques developed for load forecasting are commonly analyzed in two categories named as analytical techniques and artificial intelligence (AI) techniques. Widely used analytical methods are linear regression method, Box-Jenkins method, and nonparametric regression method. For STLF, the electric load is highly connected with meteorological factors such as temperature, humidity, wind speed, and etc. The change in holidays, weekdays, weekends, the day before and after holidays also has impacts on the load forecast. The analytical methods work well under normal daily circumstances, but they can't give contenting results while dealing with meteorological, sociological or economic changes, hence they are not updated depending on time.

Currently, conventional models using analytical techniques are inadequate owing to the presence of nonlinear phenomena in the stochastic process of electric load forecasting. Therefore, AI techniques have indicated the capability of learning complex nonlinear relationships, which are difficult to model, and accordingly making them popular. In the STLF literature, there are various applications of artificial intelligent techniques which can be categorized as artificial neural network (ANN), support vector machine (SVM), adaptive neuro-fuzzy inference system (ANFIS), genetic algorithm (GA), fuzzy logic (FL), self-organizing map (SOM), and extreme learning machines (ELM).

In the UK, Google's DeepMind has teamed up with National Grid to predict supply and demand peaks and hopes to reduce national energy usage by 10%. DeepMind's algorithms have enabled Google to cut the amount of energy used by the cooling systems in its data centers by 40%, cutting their overall electricity consumption by 15%.

Meanwhile in the US, the Department of Energy's SLAC National Accelerator Laboratory is sponsoring a project known as GRIP, for Grid Resilience and Intelligence Project, to combine artificial intelligence with large amounts of data to identify places where the electricity grid is vulnerable to disruption, so these areas can be reinforced in advance and recover faster when failures do occur. The eventual goal is an autonomous grid that seamlessly absorbs routine power fluctuations from clean energy sources like solar and wind and quickly responds to disruptive events – from major storms to eclipse-induced dips in solar power – with minimal intervention from humans.

The project will use both machine learning, where computers ingest large amounts of data and teach themselves how a system behaves, and artificial intelligence, which uses the knowledge the machines have acquired to solve problems. For example, a grid can be divided into "islands," or microgrids, that can be isolated to prevent a power disruption from spreading and taking the whole system down.

Siemens is coordinating a major research project in Germany, designed to determine the extent to which existing control center technology can be exploited to better understand grid fluctuations and to identify when and where entirely new structures and architectures will be needed.

At the moment, control center operators have low visibility of grid fluctuations – they can only see the amount of electricity being transported at each location, and whether a line is overloaded. The "DynaGridCenter" project will transmit measurement data to the control center, to be analyzed in real time. The technologies for monitoring the grid more closely are already being implemented. So-called phasor measurement units (PMUs) transmit the height and the phase angle of current and voltage every 20 milliseconds. In this way, they supplement the measured data that formerly had been transmitted in the second's range by adding a highly dynamic component. The PMUs in the grid are synchronized to a common time source, which allows them to be directly compared with each other. This shows up unwanted vibrations and very fast transient effects in the network. This data is captured using existing measurement technology which has not historically been fed to the control center despite being installed at various points on the grid.

2.6.4 Smart transport

Because of the demands on the city's physical infrastructure are going to continue to increase, it is going to be an unavoidable choice for using Intelligent transportation systems. By using smart technologies such as the Internet of Everything (IoE) that visualize all the transportation network items and use Big Data platforms to understand the behavior of the

movements, smart cities will extract values that help decision making, to achieve more vehicles but less traffic.

According to the Smart City's three layers: At instrumented layer, data is collected from multiple sources that identify and measure traffic speed and volume on city's roads, such sources are GPS, cameras, radar, and sensors embedded in roads and vehicles. This makes every single highway, lane, intersection, and vehicle representing a data point. More data points are collected from data of people's flow from the telecommunication networks through mobile phone devices. At Interconnected layer, all the data from sources with the map of the city's transportation network are combined through the city's network with the Big Data platform. At Intelligent layer, traffic analysis is applied to represent flow conditions and prediction capabilities for traffic management to identify the problem, enabling officials to respond to dangerous road conditions, accidents or growing traffic density in near real-time, by implementing system-wide changes to keep traffic moving and alerting drivers what to expect, through smartphones and built-in navigation devices, allowing them to find alternative routes, also providing the drivers with real-time information about available parking for use or book. Additionally, the insights that executed from analyzing that data in Big Data platform is used for helping cities with long-term planning such as where to plan a public transportation and parking places.

An example of using automated traffic surveillance and control system is found in Los Angeles. The system could adjust the time delay between light changes whenever issues arise. So, if there’s an accident that causes one or more lanes to be closed on any highway in the city, it can adjust the lights and give more time to let cars caught up in it all pass through. Alternatively, it can also be used to help keep public transport running on time. If the buses are late, the system can help them to pass through the lights faster and get back on schedule. The LA system synchronizes all 4,500 traffic lights in the metropolis. The project is totally unique with regards to its size and its scope and is reputed to have cost more than $400 million to implement. According to officials, the average time to drive 5 miles in the city before was 20 minutes. With the new system, this has been reduced to just 17.2 minutes.

Another example is the London Public Transit which has recorded a data where around 8 million trips are made via rail, buses, and tubes which is taken from a smart card "Oyster" used by 85 percent of total passengers. This data yields 45 million journeys a week, a billion or more every half year. This data set is endless until there is a change in technology. They have a record of 1 billion people those who tapped "in" and "out" of the public transport systems in greater London during a period of 6 months. The dataset remarks the entry and

exit location of the people, the timing at which the traveler enters and exits the system. This data is then related to the detailed data on buses and timetables from Transport for London (TfL). This combined information is helpful in assigning the close stations and lines and figuring out where the passenger might divert to.

2.7 Challenges for big data and data analytics

Big data raises a number of issues for public policymakers. These include questions about how the value of data should be determined to establish economic statistics or calculate tax, or whether the lack of standardization in big-data formats will create anti-competitive situations and reduce consumer value and much more. However, sources generally agree that the main policy issues related to big data analytics are privacy and personal data protection, data ownership, barriers to the free flow of data, skills gaps in labor markets and an emerging new digital divide.

—Privacy and personal data protection: The 2016 General Data Protection Regulation (GDPR), which will apply from 2018, has reinforced EU data protection standards, considered by the European Commission to be the highest in the world. Data subjects must agree to data collection and consent to the purpose for which the data is used. Whilst restrictions on the flow of personal data within the EU are banned, there must be adequate protection measures in place before personal data can move outside Europe's borders. A large proportion of big data is not personal (e.g. weather information, satellite imaging, operational machine data); however, some big data may include elements that link directly to a person (e.g. name, address, card or phone numbers) and hence be considered personal data. Even when this data has been aggregated and 'pseudonymized' to remove explicit identifiers, analytical techniques applied to very large datasets make it technically possible to 're-identify' a person a large percentage of the time. The danger is that the use of this personal big data can lead to surveillance, unwanted disclosure of private information and discriminatory profiling. Maintaining privacy depends on appropriate security measures to protect the big data sources and the commitment of the data curators (including third parties to whom data may be transferred) to ensure misuse does not happen. It should be noted, however, that some have argued that legal measures to ensure privacy cannot be relied upon to prevent post- collection misuse. The European Commission argues that the high standards set in the GDPR act as a competitive advantage for European businesses, in as much as they foster trust on the part of citizens who are consequently more willing to share data. Other observers, EPRS Big data and data analytics, however, believe that European companies are losing out in the application of big data to American firms because

of stricter regulation, and that strong privacy rights for citizens in terms of personal data may, in fact, prevent the realization of the potential benefits from big data, as costs will outweigh efficiency gains. Restrictions on the use of personal data – to a specific purpose, context or for a limited period – could be problematic if they are interpreted strictly, since the benefits of big data may come from unanticipated uses long after the original purpose has been served, and 'public interest' exceptions are limited. Other observers have noted that notions of privacy change from one generation to another: achieving the right trade-offs between privacy/security and innovation/ convenience may require a better understanding of privacy in the digital world.

—Data ownership: 'Ownership' is considered by a number of observers to be an overly simple or impractical concept in relation to big data. Big data does not have one owner but typically comes with a complex set of rights and privileges associated with different stakeholders. For example, large quantities of technical data are generated in a 'smart car' but what rights to that data accrue to the owner of the car, the driver, the dealer that sold the car or the automobile manufacturer? Providing guidelines and model arrangements could increase legal certainty for businesses, simplify the conclusion of contracts, and facilitate the use of big data to improve outcomes. On the other hand, complex regulations that hinder companies from buying, selling or exchanging data could have a negative effect on applying big data to solve real-world problems.

Ownership and rights also need to be seen in the framework of innovation and competition policies (e.g. abuse of dominant position and risks of consumer lock-ins) and against the background of data flow in the single market. The Belgian secretary of state for privacy recently argued that governments should be free to sell anonymized patient data to pharmaceutical companies, as long as patients can see what has been done with their data and they end up benefiting from the resulting reduced costs and innovative medicines.

—Data localization and barriers to the free flow of data: Data localization refers to requirements to store and process data within a set of geographic boundaries. Governments may adopt data localization measures to safeguard data privacy or to promote the local digital industry. The General Data Protection Regulation prohibits data localization measures for personal data inside the borders of the EU but bans the export of personal data outside the EU unless it is protected by comparable measures abroad. However, data localization can apply to non-personal data, including both inherently non-personal data such as climate information, and data that has been aggregated or pseudonymized. For example, plans for a consolidated German federal IT infrastructure (the Bundescloud) indicate that any protected information, such as business secrets, must be exclusively processed in Germany.

Data localization can also result from regulations on the handling of data that act as barriers to the free flow of data, such as requirements to obtain the consent of data subjects, the rights of subjects to review data for accuracy, and legal obligations to notify security breaches. Indeed, some stakeholders have seen data localization not as a policy yielding concrete results but rather as a way of increasing pressure on other nations to increase their own data protection measures.

Nevertheless, data localization policies are generally considered to run counter to the design of the internet, to decrease information security by concentrating storage in one place, and to increase costs for local businesses due to loss of scale, while only offering a limited incentive to local development. They also run counter to liberal trade measures, since data constitutes a significant part of international trade in goods and services. The total effect of regulatory barriers to the free flow of data in the EU is considered by some economists to have a negative impact on the growth of between 0.4% and 1.1% of GDP.

—Data Management: Many authors also cite data management as a challenge, because the platform must store and process a large amount of data and use efficient and scalable data storage and processing algorithms. Data Analysis is also a challenge because it is difficult to extract useful knowledge. Another challenge is data trustworthiness it is claimed that a high number of data sources makes it difficult to ensure that all data are correct.

—Heterogeneity: This is a challenge because the different devices in a Smart City generate diverse data. Managing data across all city systems because of these variations in data is an active problem. Researchers state that a Smart City platform should define standards across heterogeneous devices, systems, and domains.

—Energy Management: Some authors cite Energy Consumption as a challenge all platform components must face, such as sensors, actuators, and servers. Moreover, energy management in a Smart City health care application is vital, because applications or services in domains like this cannot fail during power outages.

—Communication: Since the Smart Cities of the future will incorporate a massive number of devices, enabling communication among these devices will be a challenge. The domains in a Smart City that depend on mission-critical communication and has to ensure reliability, such as health care and public safety. In addition, good communication mechanisms are required to share platform data with applications.

—Scalability: In the coming decades, the number of connected devices in a Smart City will continually increase, requiring a strong level of scalability in the associated software platform. Moreover, the number of users, services, and data stored will increase with

population growth and during special events in the city. A Smart City platform must support large-scale, efficient services. As an example, Sinaeepourfard et al. (2016) estimated that the city of Barcelona will need more than 1 million sensors to cover the entire city, generating more than 8GB of data every day.

—Security: Unauthorized users accessing city services without permission can cause serious harm. City networks will have to be safe from cyber-terrorism and cyber-vandalism. Gurgen et al, 2013 highlighted the importance of security in CPS platforms, as such systems control aspects of the city infrastructure, which a malicious user can corrupt, for example, by tampering with traffic lights and light posts.

—Lack of Testbed: The lack of testbeds is a challenge to the development of platforms for Smart Cities. Without testbeds, it is hard to perform tests and experimentation to discover the real challenges that deploying a Smart City platform will present. Smart City Simulators could be a much lower-cost alternative for experimentation.

—City Models: Some authors also argue that it is hard to fully understand a city and describe an effective and efficient model for it. It is also necessary to create a useful model of the city to make intelligent decisions and modeling is required to observe and understand city activity, as well as to avoid generating unnecessary and empty models. A unified model of the city is required so the massive amount of heterogeneous data generated can be shared among applications and services.

—Platform Maintenance: Three works state that deploying and maintaining the platform is a challenge. The difficulty of maintaining a middleware to manage millions or billions of devices connected to the platform. Similarly, the administration of the platform can be a challenge, due to its size and many devices spread across the city and addressing coordination issues in the sensor nodes can be a problem, again because of the city sensor network size.

3. Blockchain Technology

3.1 Overview

Blockchains being undisputable digital ledger systems implemented in a distributed fashion (i.e., without a central repository) and usually without a central authority. Fundamentally, blockchains allow a community of users to record transactions in a ledger that is public to that community and no transaction can be modified once published. In 2008, the blockchain concept was innovatively incorporated with several other technologies and computing concepts to allow the creation of modern cryptocurrencies: electronic money protected through cryptographic mechanisms instead of a central repository. Bitcoin was the pioneer for such blockchain based approaches. These currency blockchain systems are novel since they store value too, apart from the information storage. The value is connected to a digital wallet—an electronic device (or software) that allows an individual to make electronic transactions. The wallets, as described by NISTIR (2018), "*are used to sign transactions sent from one wallet to another, recording the transferred value publicly, allowing all participants of the network to independently verify the validity of the transactions. Each participant can keep a full record of all transactions, making the network resilient to attempts to alter that record (or forge transactions) later*."

Numerous media and technological attention have been placing blockchain in the spotlight for few years now. However, this limelight is associated to the "magic" of blockchain rather than its underlying functioning, which is yet to be well-understood. This gap in knowledge will be covered in the following section, along with the consideration of using blockchain in every sector being discussed.
Blockchain technology is the foundation of modern cryptocurrencies, so named owing to blockchain's heavy usage of cryptographic functions. Users utilize public and private keys to digitally sign and transact securely within the system. Users of the blockchain may solve puzzles using cryptographic hashing with aims of being rewarded with a fixed amount of the cryptocurrency. Blockchain has broader applications, not just confined to cryptocurrencies. The broader applicability is demonstrated here while still focusing to a large extent on the cryptocurrency use case (since that is the primary use case today). Understanding critical aspects of blockchain is vital for organizations aiming to use the technology. Great difficulty exists in changing any data that is already on the blockchain, and that alterations to the blockchain software may cause forking of the blockchain. Blockchain does not work like a database whereby changes can be made through a simple query. Another pivotal aspect of blockchain technology is the manner of agreement or "attaining consensus" among participants about valid transactions. Many such models are available, each with pros and cons for a specific business case – some store wealth, some are platforms for smart contracts (software which is deployed on the blockchain itself and executed by the

computers running that blockchain). There is constant development of new blockchain technologies to augment efficiency of current systems and facilitate new use cases. It is important to differentiate between permissionless blockchain implementations and limited implementations - former has no reading/writing restrictions while latter restricts usage to certain people or authorities, permitting intricate controls. Despite the above diversity of existing and new blockchain systems, most employ few common core concepts. Each transaction involves one or more addresses and a recording of what happened, and it is digitally signed. Blockchains are comprised of blocks, each block being a group of transactions. All the transactions in a block are grouped together, along with a cryptographic hash of the previous block. Finally, a new hash is created for the current block's header to be recorded within the block data itself as well as within the next block. Over time, each block is then chained to the previous block in the chain by adding the hash of the previous block to the header of the current block.

Each technology in a blockchain system tries to rectify previous difficulties by combining existing, proven concepts in specific manners. The utilization of the functional blockchain technology has its problems too, such as dealing with malicious users, how controls are applied, and the drawbacks of any blockchain implementation. Nevertheless, blockchain certainly will pave way for novel solutions.

3.2 Blockchain Categorization

Blockchain networks can be categorized based on their permission model, which determines who can maintain them (e.g., publish blocks). If anyone can publish a new block, it is permissionless. If only particular users can publish blocks, it is permissioned. If simplified, a permissioned blockchain network is like a corporate intranet that is controlled, while a permissionless blockchain network is like the public internet, where anyone can participate. Permissioned blockchain networks are often used for a group of organizations and individuals, typically referred to as a consortium. This distinction is necessary to understand as it impacts some of the blockchain components discussed later in this document.

3.2.1 Permissionless

Permissionless blockchain networks are decentralized ledger platforms open to anyone publishing blocks, without needing permission from any authority. Permissionless blockchain platforms are frequently open source software, freely available to anyone who wishes to download them. Since anyone has the right to publish blocks, this results in the property that anyone can read the blockchain as well as issue transactions on the blockchain (through including those transactions within published blocks). Any blockchain network user within a permissionless blockchain network can read and write to the ledger. However,

permissionless blockchain networks are open to all to participate, so malicious users may attempt to publish blocks in a way that subverts the system (discussed in detail later). To prevent this, permissionless blockchain networks often utilize a multiparty agreement or 'consensus' system that necessitates users to expend or maintain resources when attempting to publish blocks. This prevents malicious users from easily subverting the system. Examples of such consensus models include proof of work and proof of stake methods. The consensus systems in permissionless blockchain networks usually encourage non-malicious behavior through rewarding the publishers of protocol-conforming blocks with a native cryptocurrency.

3.2.2 Permissioned

Permissioned blockchain networks are ones where users publishing blocks must be authorized by some authority (be it centralized or decentralized). Since only authorized users are maintaining the blockchain, it is possible to restrict read access and to restrict who can issue transactions. Hence, permissioned blockchain networks may allow anyone to read the blockchain or they may restrict read access to authorized individuals. They also may allow anyone to submit transactions to be incorporated in the blockchain or, again, they may restrict this access only to authorized individuals. Permissioned blockchain networks may be instantiated and maintained using open source or closed source software.

Permissioned blockchain networks can have the same traceability of digital assets as they pass through the blockchain, as well as the same distributed, resilient, and redundant data storage system as a permissionless blockchain networks. They also use consensus models for publishing blocks, but these methods often do not need the expense or maintenance of resources (as is the case with current permissionless blockchain networks). This is possible through the need for establishment of one's identity, to participate as a member of the permissioned blockchain network. As such, those maintaining the blockchain have a level of trust with each other, since they were all authorized to publish blocks and since their authorization can be revoked if they misbehave. Consensus models in permissioned blockchain networks are then usually faster and less computationally expensive.

3.3 Blockchain Architecture

Though they appear to be complex, understanding Blockchain systems can easily be done by thorough learning of each component of the technology. Blockchain utilizes of familiar computer science mechanisms along with cryptographic techniques to provide a solid base for the financial services.

3.3.1 Hashes

A key component of blockchain technology is the use of cryptographic hash functions for

many operations. Hashing is a method of applying a cryptographic hash function to data, which calculates a relatively unique output (called a message digest, or just digest) for an input of nearly any size (e.g., a file, text, or image). It enables individuals to independently take input data, hash that data, and derive the same result – proving that there was no change in the data. Even the smallest change to the input (e.g., changing a single bit) will result in a entirely different output digest.

3.3.2 Transactions

A transaction represents an interaction between parties. With cryptocurrencies, for example, a transaction represents a transfer of the cryptocurrency between blockchain network users. For business-to-business scenarios, a transaction could be a manner of recording activities occurring on digital or physical assets. Each block in a blockchain can contain zero or more transactions. For some blockchain implementations, a constant supply of new blocks (even with zero transactions) is crucial to maintain the security of the blockchain network; by having a constant supply of new blocks being published, it prohibits malicious users from ever "catching up" and manufacturing a longer, altered blockchain.

A single cryptocurrency transaction typically requires at least the following information, but can contain more:

• Inputs – – The inputs are usually a list of the digital assets to be transferred. A transaction will reference the source of the digital asset (providing provenance) – either the prior transaction where it was given to the sender, or for the case of new digital assets, the origin event. Since the input to the transaction is a reference to past events, the digital assets do not change. In the case of cryptocurrencies, this means that value cannot be removed or added from existing digital assets. Instead, a single digital asset can be split into multiple new digital assets (each with lesser value) or multiple digital assets can be joined to form fewer new digital assets (with a correspondingly greater value). The splitting or joining of assets will be specified within the transaction output. The sender must also offer proof that they have access to the referenced inputs, generally by digitally signing the transaction – proving access to the private key.

• Outputs – The outputs are usually the accounts, which will be the recipients of the digital assets along with how much digital asset they will receive. Each output specifies the number of digital assets to be transferred to the new owner(s), the identifier of the new owner(s), and a set of conditions the new owners must meet to spend that value. If the digital assets provided are more than required, the extra funds must be explicitly sent back to the sender (denoting a mechanism to "make change").

While primarily used to transfer digital assets, transactions can be more broadly used to transfer data. In a simple case, someone may simply want to permanently and publicly post

data on the blockchain. In the case of smart contract systems, transactions can be utilized to send data, process that data, and store some result on the blockchain. For instance, a transaction can be used to change an attribute of a digitized asset like the location of a shipment within a blockchain technology-based supply chain system.

Regardless of how the data is formed and transacted, determining the validity and authenticity of a transaction is imperative. The validity of a transaction ensures that the transaction meets the protocol requirements and any formalized data formats or smart contract requirements specific to the blockchain implementation. The authenticity of a transaction is also vital, as it determines that the sender of digital assets had access to those digital assets. Transactions are digitally signed by the sender's associated private key, typically and can be verified at any time using the associated public key.

3.3.3 Asymmetric-Key Cryptography

Blockchain technology uses asymmetric-key cryptography (also termed, public key cryptography). Asymmetric-key cryptography uses a pair of keys: a public key and a private key that are mathematically related to each other. The public key is made public without decreasing the security of the process, but the private key must remain secret if the data is to retain its cryptographic protection. Even though there is a relationship between the two keys, the private key cannot efficiently be determined based on knowledge of the public key. One can encrypt with a private key and then decrypt with the public key. In turn, one can encrypt with a public key and then decrypt with a private key.

Asymmetric-key cryptography permits a trust relationship between users who do not know or trust one another, by providing a mechanism to verify the integrity and authenticity of transactions while at the same time allowing transactions to remain public. To do this, the transactions are 'digitally signed'. This means that a private key is used to encrypt a transaction such that anyone with the public key can decrypt it. As the public key is freely available, encrypting the transaction with the private key proves that the signer of the transaction has access to the private key. On the other hand, one can encrypt data with a user's public key such that only users with access to the private key can decrypt it. A disadvantage is that asymmetric-key cryptography is often slow to compute.

This contrasts with symmetric-key cryptography in which a single secret key is used to both encrypt and decrypt. With symmetric-key cryptography users must already have a trust relationship established with one another to exchange the pre-shared key. In a symmetric system, any encrypted data that can be decrypted with the pre-shared key confirms it was sent by another user with access to the pre-shared key; no user without access to the pre-shared key will be able to view the decrypted data. Compared to asymmetric-key cryptography, symmetric-key cryptography is very fast to compute. Because of this, when one claims to be encrypting something using asymmetric-key cryptography, oftentimes the

data is encrypted with symmetric-key cryptography and then the symmetric-key is encrypted using asymmetric-key cryptography. This 'trick' can greatly speed up asymmetric-key cryptography.

3.3.4 Addresses and Address Derivation

Some blockchain networks make use of an address, which is a short, alphanumeric string of characters derived from the blockchain network user's public key using a cryptographic hash function, along with some additional data (e.g., version number, checksums). Most blockchain implementations make use of addresses as the "to" and "from" endpoints in a transaction. Addresses are shorter than the public keys and are not secret. One method to generate an address is to create a public key, applying a cryptographic hash function to it, and converting the hash to text:

Public key ⟶ Cryptographic hash function ⟶ Address

Each blockchain implementation may execute a different method to derive an address. For permissionless blockchain networks, which allow anonymous account creation, a blockchain network user can generate as many asymmetric-key pairs, and thus addresses as desired, enabling a varying degree of pseudo-anonymity. Addresses may act as the public-facing identifier in a blockchain network for a user, and frequently an address will be converted into a QR code for easier use with mobile devices.

Blockchain network users may not be the only source of addresses within blockchain networks. It is important to provide a method of accessing a smart contract once it has been deployed within a blockchain network. For Ethereum, smart contracts are accessible via a special address called a contract account. This account address is created when a smart contract is deployed (the address for a contract account is deterministically computed from the smart contract creator's address). This contract account allows for the contract to be executed whenever it receives a transaction, and thereafter create additional smart contracts.

3.3.5 Ledgers

A ledger is a collection of transactions. Throughout history, pen and paper ledgers have been used to keep track of the exchange of goods and services. In modern times, ledgers have been stored digitally, often in large databases owned and operated by a centralized trusted third party (i.e., the owner of the ledger) on behalf of a community of users. These ledgers with centralized ownership can be employed in a centralized or distributed fashion (i.e., just one server or a coordinating cluster of servers).
There is growing interest in exploring the distributed ownership of the ledger. Blockchain technology enables such an approach using both distributed ownership as well as a

distributed physical architecture. The distributed physical architecture of blockchain networks often involve a much larger set of computers than is typical for centrally managed distributed physical architecture. This interest in distributed ownership of ledgers is due to possible trust, security, and reliability concerns related to ledgers with centralized ownership.

3.3.6 Blocks

Blockchain network users submit candidate transactions to the blockchain network via software (desktop applications, smartphone applications, digital wallets, web services, etc.). The software sends these transactions to a node or nodes within the blockchain network. The chosen nodes may be non-publishing full nodes as well as publishing nodes. The submitted transactions are then propagated to the other nodes in the network. However, this by itself does not place the transaction in the blockchain. For many blockchain implementations, once a pending transaction has been distributed to nodes, it must then wait in a queue till it is added to the blockchain by a publishing node.

Transactions are added to the blockchain when a publishing node publishes a block. A block contains a block header and block data. The block header contains metadata for this block. The block data comprises a list of validated and authentic transactions which have been submitted to the blockchain network. Validity and authenticity are guaranteed by checking that the transaction is correctly formatted and that the providers of digital assets in each transaction (listed in the transaction's 'input' values) have each cryptographically signed the transaction. This verifies that the providers of digital assets for a transaction had access to the private key which could sign over the available digital assets. The other full nodes will check the validity and authenticity of all transactions in a published block and will not accept a block if it contains invalid transactions.

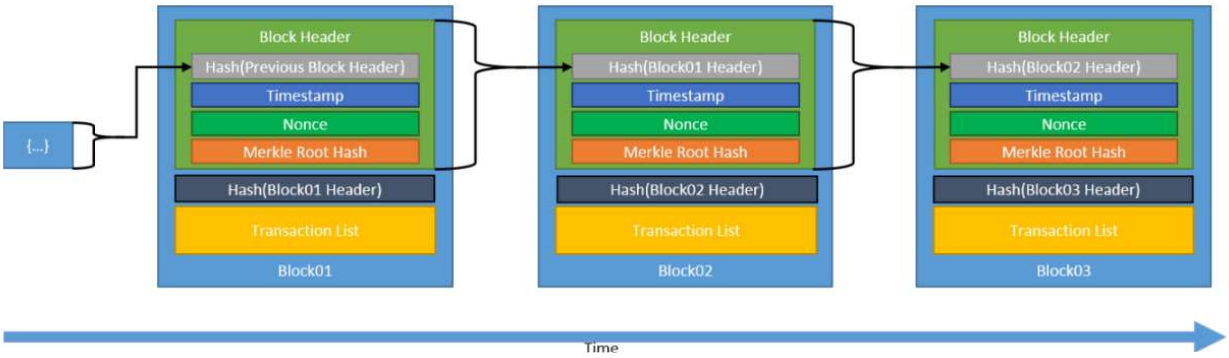


Figure 25 - Generic Chain of Blocks
Source: Blockchain Technology Overview, NIST

3.3.6 Chaining Blocks

Blocks are chained together through each block containing the hash digest of the previous block's header, thus forming the blockchain. If a previously published block were modified,

it would have a different hash. This, consequently, would cause all subsequent blocks to also have different hashes since they include the hash of the previous block. This makes it possible to easily detect and reject altered blocks.

3.4 Consensus use in chaining

A vital aspect of blockchain technology is determining which user publishes the next block. This is solved through implementing one of many possible consensus models. For permissionless blockchain networks there are typically many publishing nodes competing at the same time to publish the next block. They usually do this to win cryptocurrency and/or transaction fees. They are commonly, mutually distrusting users, who may only

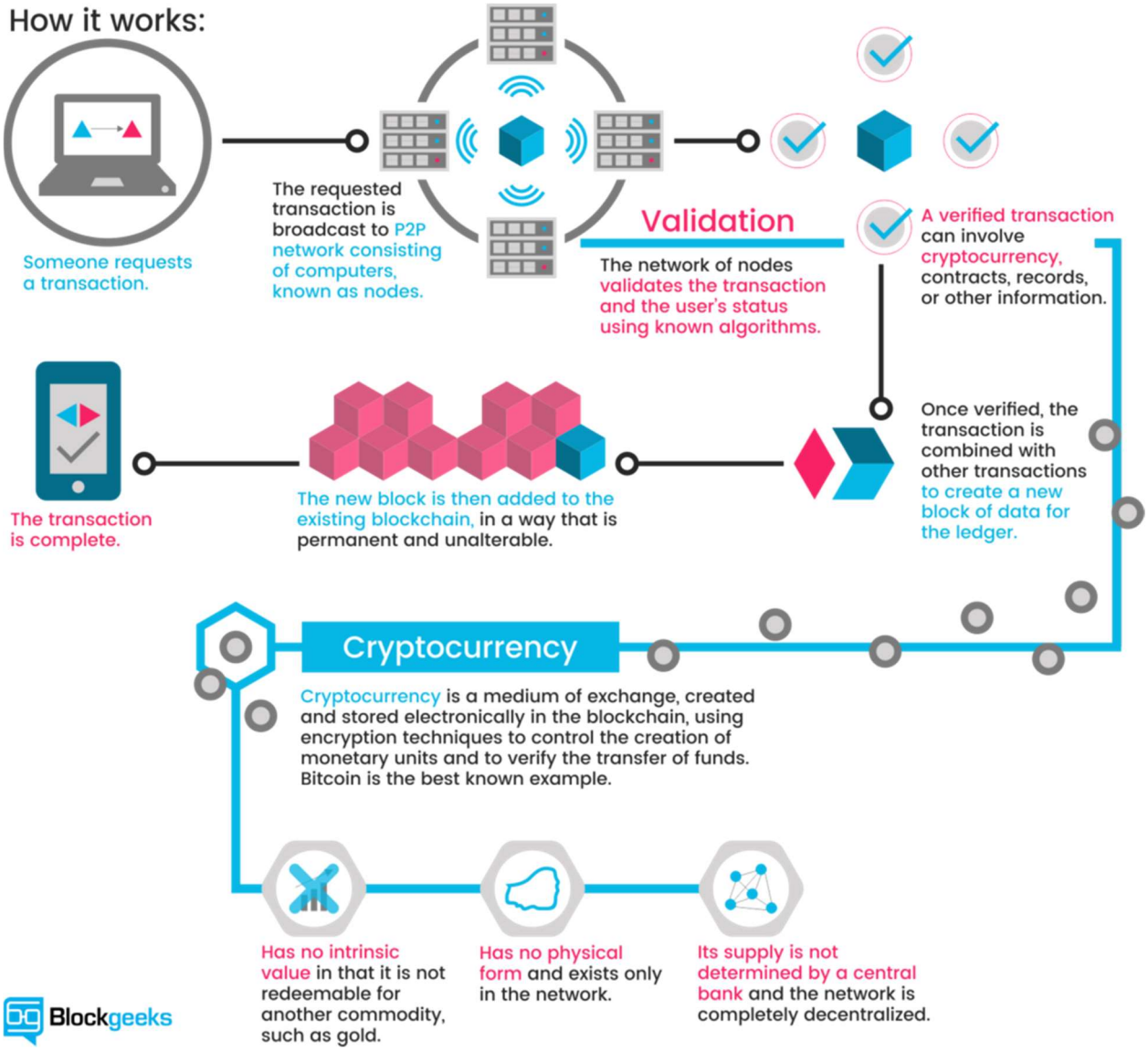


Figure 26 - Blockchain Validation
Source: Blockgeeks

know each other by their public addresses. Each publishing node is probably motivated by a desire for financial gain, not the well-being of the other publishing nodes or even the

network itself. In such cases, the questions of *'Why would a user propagate a block that another user is attempting to publish?'* and *'Who resolves conflicts when multiple nodes publish a block at approximately the same time?'* need to be addressed. To make this work, blockchain technologies use consensus models to enable a group of mutually distrusting users to work together.

When a user joins a blockchain network, they agree to the initial state of the system. This is recorded in the only pre-configured block, the genesis block. Every blockchain network has a published genesis block and every block must be added to the blockchain after it, depending on the agreed-upon consensus model. Regardless of the model, however, each block must be valid and therefore can be validated independently by each blockchain network user. By merging the initial state and the ability to verify every block since then, users can independently agree on the current state of the blockchain. It is crucial to note that if there were ever two valid chains presented to a full node, the default mechanism in most blockchain networks is that the 'longer' chain is viewed as the correct one and will be adopted. This is because it has had the most amount of work put into it. Nevertheless, in practice, software handles everything and the users do not need to be aware of these details.

A central feature of blockchain technology is that there is no need to have a trusted third party provide the state of the system—every user within the system can verify the system's integrity. To add a new block to the blockchain, all nodes must come to a common agreement over time. Yet, some temporary disagreement is permitted. For permissionless blockchain networks, the consensus model must work even in the presence of possibly malicious users since these users might attempt to disrupt or take over the blockchain. It is important to note that for permissioned blockchain networks, legal remedies may be used if a user act maliciously.

There are several such consensus model of which the noteworthy ones are discussed below:

The Proof of Work model, being the oldest of the lot, is the model followed by majority of the cryptocurrencies including Bitcoin and Ethereum. However, the disadvantage with this model stems from its energy intensive computational method, the annual estimate of which is 50TWh and it has been strongly criticized in the past few years. As a result of which more energy efficient Consensus models such as Proof of Stake, Round Robin, Proof of Authority, Proof of Elapsed Time and several such Consensus Models have been developed.

3.4.1 Proof of Work Consensus Model

In the proof of work (PoW) model, a user publishes the next block by being the first to solve a computationally intensive puzzle. The solution to this puzzle is the "proof" they have performed work. The puzzle is designed such that solving the puzzle is difficult but checking

that a solution is valid is easy. This enables all other full nodes to easily validate any proposed next blocks, and any proposed block that did not fulfill the puzzle would be rejected.

A common puzzle method is to require that the hash digest of a block header be less than a target value. Publishing nodes make many small changes to their block header (e.g., changing the nonce) trying to find a hash digest that meets the requirement. For each attempt, the publishing node must compute the hash for the entire block header. Hashing the block header many times becomes a computationally intensive process. The target value may be altered over time to adjust the difficulty (up or down) to influence how often blocks are being published.

For example, Bitcoin, which uses the proof of work model, adjusts the puzzle difficulty every 2016 blocks to influence the block publication rate to be around once every ten minutes. The adjustment is made to the difficulty level of the puzzle, and essentially either increases or decreases the number of leading zeros required. By increasing the number of leading zeros, it increases the difficulty of the puzzle, because any solution must be less than the difficulty level – meaning there are fewer possible solutions. By decreasing the number of leading zeros, it decreases the difficulty level, because there are more possible solutions. This adjustment is to maintain the computational difficulty of the puzzle, and therefore uphold the core security mechanism of the Bitcoin network. Available computing power increases over time, as does the number of publishing nodes, so the puzzle difficulty is typically increasing.

Adjustments to the difficulty target intends to ensure that no entity can take over block production, but as a result the puzzle solving computations require significant resource consumption. Due to the significant resource consumption of some proof of work blockchain networks, there is a move to add publishing nodes to areas where there is a excess supply of cheap electricity.

An important aspect of this model is that the work put into a puzzle does not influence one's likelihood of solving the current or future puzzles as the puzzles are independent. So, when a user receives a completed and valid block from another user, they are incentivized to stop and remove their current work and to start building off the newly received block instead because they know the other publishing nodes will be building off it.

There is currently no known shortcut to this process; publishing nodes must expend computation effort, time, and resources to find the correct nonce value for the target. Often the publishing nodes attempt to crack this computationally difficult puzzle to claim a reward of some kind (usually in the form of a cryptocurrency offered by the blockchain network). The prospect of being rewarded for extending and sustaining the blockchain is referred to as a reward system or incentive model.

Once a publishing node has performed this work, they send their block with a valid nonce to full nodes in the blockchain network. The recipient full nodes verify that the new block satisfies the puzzle requirement, then add the block to their copy of the blockchain and resend the block to their peer nodes. The new block gets quickly distributed throughout the network of participating nodes, in this way. Verification of the nonce is easy, as only a single hash needs to be done to check if it solves the puzzle.

3.4.2 Proof of Stake Consensus Model

The proof of stake (PoS) model is grounded on the idea that the more stake a user has invested into the system, the more likely they will want the system to succeed, and the less likely they will want to disrupt it. Stake is often an amount of cryptocurrency that the blockchain network user has invested into the system (through differing means, such as by locking it via a special transaction type, or holding it within special wallet software, or by sending it to a specific address,). Once staked, the cryptocurrency is generally unable to be spent. Proof of stake blockchain networks utilizes the amount of stake a user has as a determining factor for publishing new blocks. Thus, the likelihood of a blockchain network user publishing a new block is tied to the ratio of their stake to the overall blockchain network amount of staked cryptocurrency.

With this consensus model, there is no need to conduct resource intensive computations (involving time, electricity, and processing power) as found in proof of work. Since this consensus model utilizes fewer resources, some blockchain networks have decided to forego a block creation reward; these systems are designed so that all the cryptocurrency is already distributed among users rather than new cryptocurrency being generated at a constant pace. In such systems, the reward for block publication is then typically the earning of user provided transaction fees.

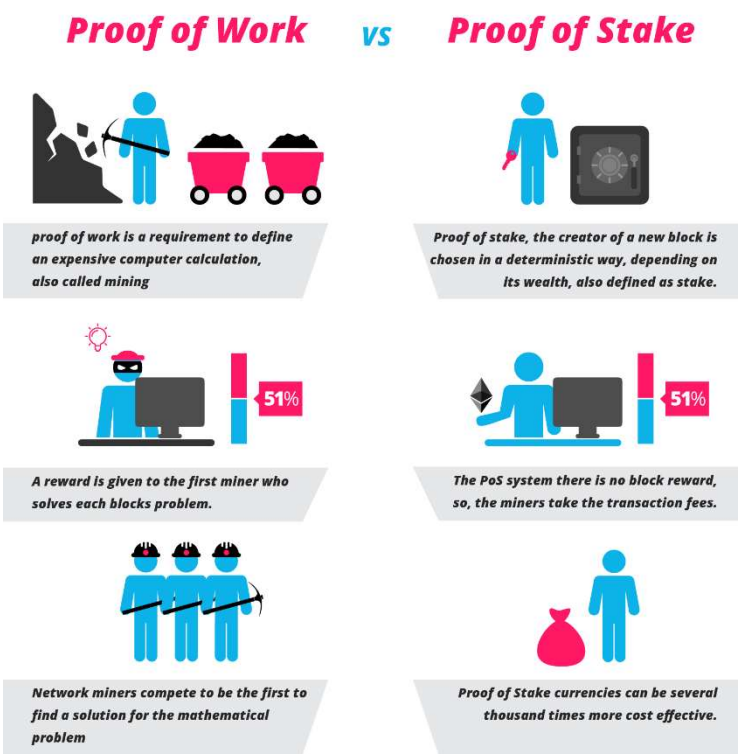


Figure 27 – Proof of Work vs Proof of Stake
Source: Blockgeeks

The methods for how the blockchain network use the stake can differ. Here four approaches are discussed: random selection of staked users, multi-round voting, coin aging systems and delegate systems. Regardless of the exact approach, users with more stake are more likely to publish new blocks.

3.4.3 Round Robin Consensus Model

Round Robin is a consensus model that is used by some permissioned blockchain networks. Within this model of consensus, nodes take turns in generating blocks. Round Robin Consensus has a long history based in distributed system architecture. To handle situations where a publishing node is not available to publish a block on its turn, these systems may include a time limit to allow available nodes to publish blocks so that unavailable nodes will not cause a stop in block publication. This model guarantees no one node creates the majority of the blocks. It benefits from a straightforward approach, lacks cryptographic puzzles, and has low power requirements.

Since there is a need for trust amongst nodes, Round Robin does not work well in the permissionless blockchain networks used by most cryptocurrencies. This is due to the fact that malicious nodes could endlessly add additional nodes to increase their odds of publishing new blocks. In the worst case, they could use this to subvert the right functioning of the blockchain network.

3.4.4 Proof of Authority/Proof of Identity Consensus Model

The proof of authority (also referred to as proof of identity) consensus model relies on the partial trust of publishing nodes through their known link to real world identities. Publishing nodes must have their identities proven and verifiable within the blockchain network (e.g., identifying documents which have been confirmed and notarized and included on the blockchain). The concept is that the publishing node is staking its identity/reputation to publish new blocks. Blockchain network users directly affect a publishing node's reputation based on the publishing node's behavior. Similar to gaining gain reputation by acting in a manner that the blockchain network users agree with, publishing nodes can also lose reputation by behaving in a way that the blockchain network users disagree with. The lower the reputation, the less likelihood of being able to publish a block. Therefore, it favorable for a publishing node to maintain a high reputation. This algorithm only applies to permissioned blockchain networks with high levels of trust.

3.4.5 Proof of Elapsed Time Consensus Model

Within the proof of elapsed time (PoET) consensus model, each publishing node requests a wait time from a secure hardware time source within their computer system. The secure hardware time source will create a random wait time and return it to the publishing node software. Publishing nodes take the random time they are given and become idle for that duration. Once a publishing node wakes up from the idle state, it creates and publishes a block to the blockchain network, alerting the other nodes of the new block; any publishing node that is still idle will stop waiting, and the entire process starts over. This model necessitates that a random time was used, since if the time to wait was not selected at random a malicious publishing node would just wait the minimum amount of time by default to dominate the system. This model also requires that the publishing node waited the actual time and did not begin early. These requirements are being fulfilled by executing software in a trusted execution environment found on some computer processors.

Verified and trusted software can run in these secure execution environments and cannot be changed by outside programs. A publishing node would query software running in this secure environment for a random time and then wait for that time to pass. After waiting the assigned time, the publishing node could request a signed certificate that the publishing node waited the randomly assigned time. The publishing node then publishes the certificate along with the block.

3.5 Smart Contracts

The term smart contract dates to 1994, defined by Nick Szabo as “a computerized transaction protocol that executes the terms of a contract. The general objectives of smart contract design are to satisfy common contractual conditions (such as payment terms, liens, confidentiality, and even enforcement), minimize exceptions both malicious and accidental, and minimize the need for trusted intermediaries.”

Smart contracts extend and leverage blockchain technology. A smart contract is a collection of code and data (sometimes indicated as functions and state) that is deployed using cryptographically signed transactions on the blockchain network (e.g., Ethereum’s smart contracts, Hyperledger Fabric’s chaincode). The smart contract is implemented by nodes within the blockchain network; all nodes that execute the smart contract must derive the same results from the execution, and the results of execution are recorded on the blockchain.

Blockchain network users can create transactions which send data to public functions offered by a smart contract. The smart contract executes the suitable method with the user provided data to perform a service. The code, being on the blockchain, is also tamper evident

and tamper resistant and thus can be used (among other purposes) as a trusted third party. A smart contract can perform calculations, store information, expose properties to reflect a publicly exposed state and, if appropriate, automatically send funds to other accounts. It does not essentially even have to perform a financial function. For instance, there exists an Ethereum smart contract that publicly generates trustworthy random numbers. It is vital to understand that not every blockchain can run smart contracts.

The smart contract code can represent a multi-party transaction, usually in the context of a business process. In a multi-party scenario, the benefit is that this can provide attestable data and transparency, which in turn can foster trust, provide insight. This further enables better business decisions, reduces costs from reconciliation that exists in traditional business to business applications, and reduces the time to complete a transaction.

Smart contracts must be deterministic - given an input, they will always produce the same output based on that input. Additionally, all the nodes executing the smart contract must agree on the new state that is attained after the execution. To achieve this, smart contracts cannot operate on data outside of what is directly passed into it (e.g., smart contracts cannot obtain web services data from within the smart contract – it would need to be passed in as a parameter). Any smart contract which uses data from outside the context of its own system is deemed to use an 'Oracle'.

For many blockchain implementations, the publishing nodes execute the smart contract code concurrently when publishing new blocks. There are some blockchain implementations in which there are publishing nodes which do not execute smart contract code, but instead validate the results of the nodes that do. For smart contract enabled permissionless blockchain networks, (such as Ethereum) the user issuing a transaction to a smart contract will have to pay for the cost of the code execution. There is a limit on how much execution time can be consumed by a call to a smart contract, depending on the complexity of the code. If this limit is exceeded, execution stops, and the transaction is discarded. This mechanism not only rewards the publishers for executing the smart contract code, but also prohibits malicious users from deploying and then accessing smart contracts that will conduct a denial of service on the publishing nodes by consuming all resources (e.g., using infinite loops).

For smart contract enabled permissioned blockchain networks, such as those utilizing Hyperledger Fabric's chaincode, there may not be a requirement for users to pay for smart contract code execution. These networks are designed around having known participants, and other methods of preventing bad behavior can be employed (e.g., revoking access).

3.5.1 Smart Contracts vs Traditional Contracts

Smart contracts are based on software languages, or "codes". A code is the set of instructions forming the software which is executed by a computer in an abstract manner to produce real-world output. Coding and legal drafting are two different concepts. Codes are prepared by human language but in a computer-readable manner that takes effect by electronic execution. This is performed by computer drivers in a predetermined pattern and process, which cannot run in case of error or ambiguity. Alternatively, traditional contract obligations are based on human language, which takes effect upon the parties' agreement and is done in line with its interpretation of the human language text, which can itself contain errors and ambiguous text. Therefore, there is a gap between "code" and "contract", which is required to be erased so as to give life to smart contracts.

There is no doubt about growing interest in using and testing blockchain technology in a variety of mediums and purposes, by questioning – and sometimes belittling – the "traditional contract". Conversely, certain sectors including global financial sector take this technology extremely seriously and have begun developing and testing the technology on automated online payment systems.

It is clear that different applications of blockchain technology for smart contracts will be seen. Simple derivative products such as options can be used as an example of a transaction that can be executed through a smart contract platform. For option transactions, the smart contract and blockchain platform would operate as the record-keeper for the title, custodian and intermediary, as well as the clearing and settlement system. Recently, a consortium including BP and Royal Dutch Shell announced that they will develop a blockchain-based digital platform for energy commodities trading by the end of 2018 in order to diminish the administrative operational risks and costs of physical energy trading.

Smart contracts have certain practical advantages over traditional paper contracts:

- Certainty: Since the smart contracts are executed by computer codes, there is no room for any ambiguous natural language maybe used in traditional contracts.

- Speed: Like any types of computer automation, smart contracts are executed and implemented almost promptly without human involvement.

- Cost: Although there will be upfront costs for development and implementation of the smart contract platform, the operating costs connected with the performance of smart contracts will be low, since there will be almost no human involvement at the performance phase.

Despite their advantages, there are no concrete smart contract solutions implemented in practice yet. However, there are projects run by startups and financial institutions, which

are considered as experimental (e.g. Ethereum, Corda and Decentralized Autonomous Organization).

3.6 Smart City Architecture

The significant advancements in IoTs and wireless communications have made it easy to connect a range of devices and enable them to transmit data ubiquitously even from remote locations. However, these systems are more instrumented with open data such as locations, personal and financial information, and therefore, must be capable to defend against security attacks. The Kaspersky Lab shows that smart terminals such as bicycle rental terminals, self-service machines, and information kiosks have a number of security flaws. These devices can be exploited by the cybercriminals for they may get access to personal and financial information of the users. It is also worth noting that implementation of traditional security mechanisms into a city's critical infrastructure to make it smarter has failed. Thus, new solutions based on the nature of the data (private or public) and communication platforms need to be developed to provide data confidentiality, privacy, and integrity. Hence, a security framework based on blockchain technology which allows to communicate the entities in a Smart City without compromising privacy and security is proposed. The main advantage of using blockchain is that it is resilient against many threats. Moreover, it offers a number of unique features such as scalability, better fault tolerance capability, improved reliability, faster and efficient operation. Therefore, amalgamation of blockchain technology with devices in a Smart City will create a common platform where all devices would be able to communicate securely in a distributed environment.

3.6.1 Security Framework

1) Physical Layer: Smart City devices are equipped with sensors and actuators which gather and forward data to the upper layer protocols (IoT). Some of these devices such as Nest thermostat and Acer Fitbit are susceptible to security attacks due to lax encryption and access control mechanisms. Furthermore, there is no single standard for smart devices so that the data produced by them can be shared and combined to provide cross-functionality. Vendors require an agreed-upon implementation and communication standards to overcome these problems in smart devices.

2) Communication Layer: In this layer, Smart City networks use various communication mechanisms such as Bluetooth, 6LoWPAN, Wi-Fi, Ethernet, 3G, and 4G to exchange information among different systems. The blockchain protocols need to be integrated with this layer to allow security and privacy of transmitted data. For example, the transaction records can be converted into blocks using telehash, which can be broadcast in the network. Protocols like BitTorrent can be used for peer to peer communication whereas Ethereum can offer smart contract functionalities. However, integration of existing communication

protocols with blockchain is a major challenge as the requirements vary from application to application. A potential solution can be employing multiple blockchains with the help of a blockchain access layer to provide application specific functionalities.

3) Database Layer: In blockchain, distributed ledger is a type of decentralized database that stores records consecutively. Each record in the ledger includes a time stamp and a unique cryptographic signature. The complete transaction history of the ledger is verifiable and auditable by any legitimate user. There are two different types of distributed ledger in practice: i) permissionless and ii) permissioned. The key benefits of permissionless ledger are that it is censorship resistant and transparent. However, the public ledger has to maintain complex shared records and it consumes more time to reach the consensus compared to the private ledger. Further, public ledgers may also be subjected to anonymous attacks. Therefore, it is recommended to use private ledgers to ensure scalability, performance, and security for real-time applications like traffic systems in a Smart City.

4) Interface Layer: This layer contains numerous smart applications which collaborate with each other to make effective decisions. For instance, a smart phone application can provide location information to the smart home system so that it turns on the air conditioner 5 minutes prior to reach at home. However, the applications should be integrated carefully since vulnerabilities in one application may give intruders access to other dependent processes.

3.7 Use cases

Declining costs of electronic equipment like sensors and processors, and improved wireless broadband connections have accelerated the feasibility of smart cities, allowing municipalities to collect more significant data related to many kinds of transactions. Collecting data from these various sources, collectively, can mean more precise solutions and policies to get at the root of an issue. The availability of increasingly diverse city data, and the capacity of cities to systematically evaluate their goals, creates an opportunity to present solutions that benefit both residents and local government administrations. For example, moving to a new city infrastructure data onto the blockchain ledger could streamline and improve municipal operations for the city and improve service delivery for the resident. Open data policies like blockchain stimulate local innovators, like software developers, to find patterns in the data. These patterns can lead to development of solutions that offer conveniences and better public services for the residents who depend on them.

Blockchain can also increase the speed of public services as well as widen their capabilities and in some cases even decrease costs. From governance to transportation, alternative energy and even healthcare, blockchain offers exciting possibilities.

3.7.1 Transportation on the Blockchain

As self-driving vehicles continue to be tested in cities nationwide, they bring with them large implications for the states they navigate, like Arizona, Nevada, Pennsylvania and others. Fully self-driving cars will have the capability of connecting with others on the road for a seamless experience. Every autonomous vehicle will have a unique operational number linked to a blockchain-enabled virtual wallet preloaded with digital currency in the same way one might prepay a transit card. As computing power increases in cars, the potential benefits range from real-time traffic negotiations between vehicles to allow for better traffic flow to pricing structures that would allow one vehicle to pay a premium to go faster than those around it.

Imagine variable tolling at the microlevel of an individual car. If you are in a rush, your car could negotiate with vehicles around you for right-of-way at a preset premium. Drivers that are in less of a hurry would move out of the way automatically and be compensated for increasing their travel time. Blockchain would allow for car owners to negotiate rates and exchange money in real-time with no mediator or transaction costs.

Moreover, paying for traditional tolls or even fuel — whether gas or electric — could be just as seamless. Once at your destination, the vehicle parks at a metered space or remains in a shared mode where it is continually carrying other passengers throughout the day with payments sent and received through blockchain. If the owner chooses not to share their vehicle, the meter opens a transaction with your digital wallet, and is paid — a fraction of a penny for every second the car is parked in that spot — in real time.

It must be noted, however, that it will be incumbent on policymakers to alleviate potential inequities that could arise due to these types of real-time electronic transactions. At the same time that someone has their car set to automatically negotiating better, more seamless travel by paying a higher rate, lower income drivers will be at an immediate disadvantage, because they cannot afford to pay higher rates. This type of dystopic outcome could be further aggravated by others that seek to drive around clogging the system to be paid to 'get out of the way.' Suffice it to say, the techno-optimist viewpoint may prevail, and optimal situations with mobility working seamlessly as traffic flows smoother than ever is not far from the future. Again, the policy environment will be critical to set the rules of the road.

Once self-driving vehicles become widespread, cities may experience a decrease in parking revenue as the cars multi-task with a variety of passengers instead of sitting idle in a parking space. In this case, when self-driving cars are parked, cities will receive metered parking revenue in real-time, and will no longer have to pay for enforcement. Residents need not worry about fines or feeding the meter. This shift presents a challenge for many cities. Pittsburgh, for example, derives 15 percent of its city budget from parking fines. That would

be a significant loss. Further, if more people turn to the ease of ride sharing, cities stand to lose public transit revenue. Blockchain may provide a way for cities to tax ride sharing transactions and make up the difference.

These examples show how blockchain allows for economies, like that of parking revenue in cities, to be safe, transparent and instantaneous, and to have little or no transaction costs.

3.7.2 Smart Governance

Several nations across the globe have used the secure and unchangeable platform technology to pursue new forms and applications of governance. As more of the administrative aspects of individuals' lives move into the digital world, it is sensible for the governments to experiment in this realm while remaining vigilant in addressing the security and privacy issues that surface during any discussion about digitalization of public infrastructure and processes.

There are more than 100 projects in more than 30 countries which use the Blockchain technology and are being carried out by governments, a few of them are discussed below.

Estonia is one of the countries with very high E-Government Development Index. Specifically, it ranks in the 13thposition globally based on the 2016 UN Global E-Government Index. It also ranks as one of the most innovative countries in the world; ranking at 24thposition out of the 128 countries surveyed in the 2017 edition of the Global Innovation Index report. Since 2014, the topic of Blockchain innovation has gained significant popularity among private and public institutions in Estonia. Several prototypes and concepts involving Blockchain technology have been announced by the government of Estonia. Three notable cases of these innovations involving management of access to health records, provision of notary services to e-residents and authentication of shareholders for e-voting in meetings are briefly described below.

Migration of government data to Blockchain: The initiative aimed at securing access to over 1 million public health records to eliminate unauthorized access to the records without the need of a centralized trust party in or outside government. The initiative relies on the technology developed by Guardtime; a Blockchain start-up. The solution is based on Guardtime's Keyless signatures technology which can establish the integrity of any data without the use and exchange of the traditional private and public keys. The keyless signature infrastructure (KSI) Blockchain will be integrated with the e-Health Authority health (Oracle)database for "real-time visibility" into the state of patient records. This initiative is expected to significantly improve the process used in recording and updating health records in terms of efficiency (including cost) and effectiveness. The use of Blockchain technology will provide the creation of a secured and trusted care records into electronic chains of events while preserving the provenance and integrity of those health records. The

solution will also enable strong identity proofing by preserving immutable records of the declared identities of both patients and healthcare professionals. Equally important, the initiative will empower patients through the recording of consent decisions and patient directives within the secured healthcare record.

E-voting for E-Resident Shareholders: the US stock-market firm Nasdaq in collaboration with the Estonian e-residency program aims to provide e-Residents and Estonian citizens who are shareholders in firms listed on the Tallinn Stock Exchange an opportunity to vote securely online in shareholder meetings. The Estonia's e-residency platform will be used to authenticate e-resident shareholders while the Nasdaq's Blockchain technology will be employed to record votes securely. The agility and size of Estonia coupled with its robust Information Society created the favorable environment for the Nasdaq-Estonian Government collaboration in piloting the e-voting program.

United Kingdom –The country ranks in 1st position in the 2016 E-Government Development Index and the 3rd place in the 2016 Global Innovation Index. UK Government through its Office of Science published a report on Distributed Ledger Technology: Beyond blockchain. The report expressed the transformational potential of distributed ledger and also advanced a number of technologies, governance, security and privacy, and trust and interoperability related recommendations. Furthermore, the UK government believes that it stands in a good position to leverage the benefits and address the challenges related to the use of distributed ledgers in the public service and economy because of the digital capability, innovative financial services, the effective research community and growing private service. Some of the ongoing blockchain based initiatives in the UK include:

Distributed ledger based Gross Settlement System: The Bank of England is currently working on replacing its current real-time gross settlement (RTGS) system to be ready for future demands. Specifically, the future system must address the following strategic RTGS requirements: 1) capability of responding to the changing structure of the financial system; 2) recognizing that payment system users want simpler and more resilient pathways for their payments; 3) capability of interfacing with a range of new technologies being used in the private sector, including distributed ledgers, if/when they achieve critical mass; 4) to remain highly resilient to the increasingly diverse range of threats to continuity of service, and 5) develop capacity to support the future evolution of regulatory and monetary policy tools. From the bank point of view, the new system will change a lot of features between the existing system which was built in 1996 and its successor. Some of these changes will include and enhanced security, which could be provided through the use of distributed ledger/blockchain solutions.

Blockchain for benefit payment: The government is currently test-running a blockchain based social welfare payment mobile app. Claimants in receipt of this payment are advised

to download the app on their phones which will enable them to receive and spend their benefit payments. With their consent, their transactions are being recorded on a distributed ledger to support their financial management. This initiative focuses on adding an additional layer of richer data and identity onto payments so that a deeper and more effective relationship can be established between the government and claimants. The aim of this project is to identify the possibility for welfare payment to citizens to be sent through a secure app and also to see if people reliant on welfare payments would benefit from this approach. This new system consists of a mobile app and a Blockchain system that records payments sent and received by beneficiaries. This initiative is a joint effort of the Department of Work and Pensions, Barclays, Npower, University of London and UK-based blockchain start-up GovCoin.

Paying research grant through Blockchain: Monitoring and controlling the use of grants is incredibly complex. The government considers that a blockchain accessible to all the parties involved might be a better way of solving that problem. The government presently is looking into any sort of Blockchain technique, Bitcoin is one of those. Furthermore, it is open to all ideas because of the fact that there are a number of areas Blockchains can be used, including government grants which can be used to track the money and it gets taxpayers a better deal, potentially. The government is currently exploring future technologies so that new ways of doing old things can be identified to reshape the state through the best use of modern technology.

3.7.3 Smart Energy

Blockchain also has the potential to positively affect a city's energy sector. In 2015, renewable energy made up 64 percent of all new electricity generating capacities created in the United States. As the nation diversifies its energy sector, several grid innovations like microgrids, energy storage technology and smart meters could improve efficiency and democratize the platform.

Mechanism:

A. Distributed Energy Exchange Architecture

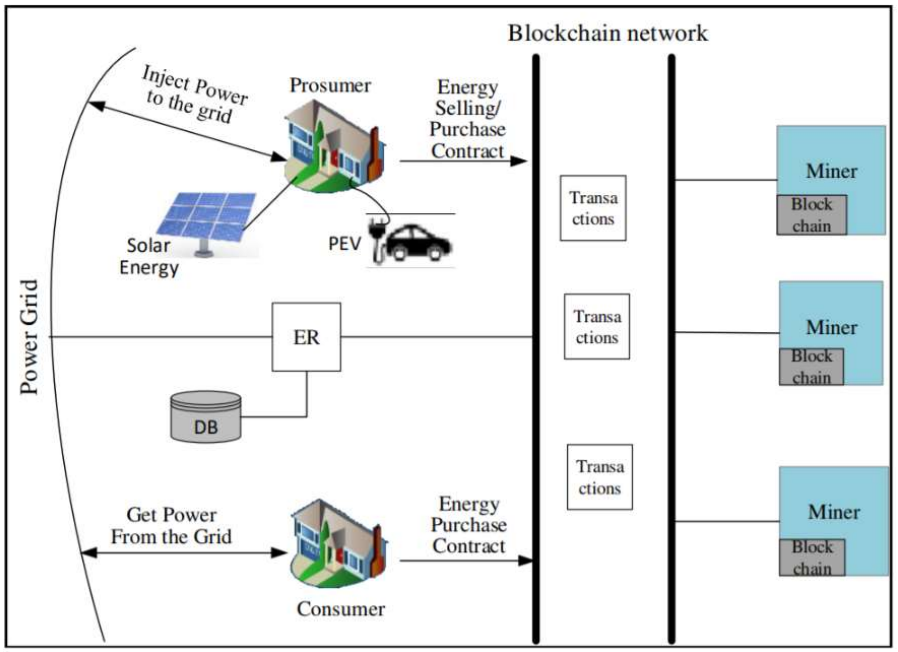


Figure 28 - Distributed Energy Exchange Architecture
Source: Using Blockchains to Secure Distributed Energy Exchange

A decentralized energy system which utilizes Blockchains generally includes of an energy router, prosumers, consumers and miners. *Prosumers* are users who sell and buy energy to and from the system and they usually own their private energy production infrastructure. The energy locally stored by them is exchanged through the smart power grid with other users using the energy router. Consumers buy energy from the system but do not produce it. Miners are principally responsible for the management of operations pertaining to the Blockchain transactions and might or might not take part in the energy transactions. The role of miners can also be played by other users in the systems such as the prosumers, consumers or the power supplier.

B. The energy exchange process

The energy exchange process takes place in the following stages:

- Energy is injected into the smart grid by the prosumers
- A smart contract is created by the prosumers for selling the energy created
- The created smart contract transactions are mined by the miners
- A smart contract for purchase of the energy is created by the consumers

- The purchased energy is delivered to the consumers through the energy router

The energy router registers in the database the amount of energy that prosumers produce and inject into the smart grid network and gives a unique code for the transaction. Thereafter, the prosumer creates a Smart Contract after defining the required information and broadcasts it to the Blockchain network. The miners on the network receive this smart contract and verify if the amount of energy claimed in it has already been injected into the grid and if the transaction is signed by the contract owner. Upon this verification through the energy router, the contract is added to the Blockchain and it would be publicly available to users on the network who can then make offers for purchase.

Consumers can check the list of available energy transactions present on the Blockchain network and submit purchase requests for the one that would fit their requirement. This purchase request, which would include the unique identity of the contract and the payment related to the prosumer, will be received and verified by the miners. This verification would involve checking if the requester has enough credit/money to pay for the amount of energy quoted and if the transaction is signed by the consumer. Upon successful verification the energy router is signaled to release the specified energy amount to the consumer.

To promote a sustainable pathway of such transactions, the energy router and the miners will have to be suitably incentivized. The energy router can either be allowed to utilize a percentage of the energy transacted or a payment compensation from the prosumer or consumer can be made as an incentive. Whereas the miners can benefit by receiving electronic money or energy according to the energy transaction carried out.

The application of Blockchain is primarily using two pathways: i) Transactions and Trading and ii) Smart Charging/Discharging.

3.7.4 Solutions and existing projects

Whole Sale Energy Trading:

In electricity (and gas) trading, trades are begun on an online exchange, or via a broker, after the initiating trader consults an index agency to gather pricing intelligence. After closing the trade, both traders separately enter the transaction details in their respective IT systems (known as “energy trading and risk management” [ETRM] systems). Both parties’ back offices retrieve the transaction details from their ETRM systems and exchange the data with each other, and/ or with the broker, to confirm and reconcile the trade. This step is accomplished either by automated confirmation systems, like EFETnet in Europe, or through traditional communication channels (emails, calls, fax) and spreadsheets. The trade is then settled physically through a TSO (or pipeline or shipment for gas). It is also settled financially through a clearinghouse or bank. Finally, both actors report the transaction details to the relevant auditors and regulators as per their obligations.

This process employs siloed IT systems and sometimes inefficient communications. It can lead to high transaction costs (costly exchange and broker fees, pricing agencies, etc.) and operational costs (time-consuming reconciliation issues, costly back office processes, etc.). Blockchain technology could decrease the transaction costs for trading large volumes by making operational processes more efficient and by connecting the trading desks of all parties. Some foresee blockchain-based trading platforms eliminating the need for brokers and clearinghouses. By reducing transaction costs, blockchain could also enable participants to trade in smaller volumes.

Pilot projects such as Ponton's "Enerchain" and Blockchain Technology Limited (BTL)'s "Interbit" platforms aim to reduce the costs tied to wholesale energy trading. Software and energy market automation company Ponton has developed "Enerchain," a proof of concept blockchain-based clearing platform for wholesale energy trades that does not rely upon a centralized exchange or brokers. Enerchain enables wholesale energy traders to anonymously send orders to a decentralized "orderbook" that can be accessed by other traders. The trading volumes that occur on the Enerchain platform are still very small compared to total volumes on the European Energy Exchange (EEX). Nonetheless, Enerchain has been expanding. It began in 2017 as a consortium of 15 European energy trading firms. As of April 2018, the consortium had grown to 42 firms.

BTL recently conducted a twelve-week pilot project specifically targeting reconciliation issues in the European gas market. In partnership with Wien Energy, BP, Eni Trading & Shipping and other energy companies, the pilot sought to decrease the manual management of post trade communications. Instead of sending trade details via email, trades were logged into a blockchain which counterparties could verify in real time. The pilot relied upon BTL's proprietary blockchain platform, Interbit, whereby one blockchain could possibly exist for every bilateral relation and have all those blockchains connect to one general directory blockchain. In 2018, BTL announced a partnership with Eni Trading & Shipping, Total, Gazprom Marketing & Trading Limited and other companies to use the Interbit blockchain platform to deliver gas trading reconciliation through to settlement and delivery of trades. This enterprise solution is being called One Office and is a revenue generating project for BTL.

Energy Metering

The Underlying Issue: While the new generation of smart meters enables tracking of the energy produced and used by consumers, mechanical meters are still very common. The incompatibility problem among distributors and suppliers can be seen in the United Kingdom, for example, where no clear standards for smart meters exist. This issue causes hardware losing its "smart" functionality when changing distributor or supplier.

Why Blockchain is a Good Solution: As a distributed ledger, the storage of energy production/consumption data and distribution fees in the Blockchain brings trust and transparency to those settlement processes. Furthermore, using an infrastructure of distributed nodes in the network, this solution can help to alleviate the risk of having a central server authority that would act as a single point of failure.

Solution Approach & Benefits: Based on consumer behavior, the production and consumption of energy might fluctuate greatly, and adequate information management is not always put in place. Transparency about amount of energy consumed and information accuracy can be improved with crypto tags. Non-removable, sealed crypto tags verify a meter's integrity (untampered nature), and therefore create trust in the consumption data provided. Via NFC, the owner, previously authorized via biometric authentication, can access the tag data and upload consumption information.

By using this set-up, both utility companies and energy customers can be provided with a secure, cost-effective and scalable solution, enabling them to:

1. Bring together the digital identity of the meter, customers and their smartphones for efficient billing

2. Track data via Blockchain traceability and easily generate awareness of energy consumption

3. Demonstrate the benefits and thus increase the incentive to upgrade legacy IT systems.

For energy metering and consumer services in general, an industry-specific framework establishes fundamental rights among stakeholders, in this case the protection of consumer data. The infrastructure, specifically used distributed ledger protocol, heavily relies on the requirements of certain use cases. For energy metering, a protocol that combines the use of smart contracts with a large network and open adaption policy is required. Using application programming interfaces, convenient monitoring of energy flows and metering statistics, smart contract agreements and payment options could be possible.

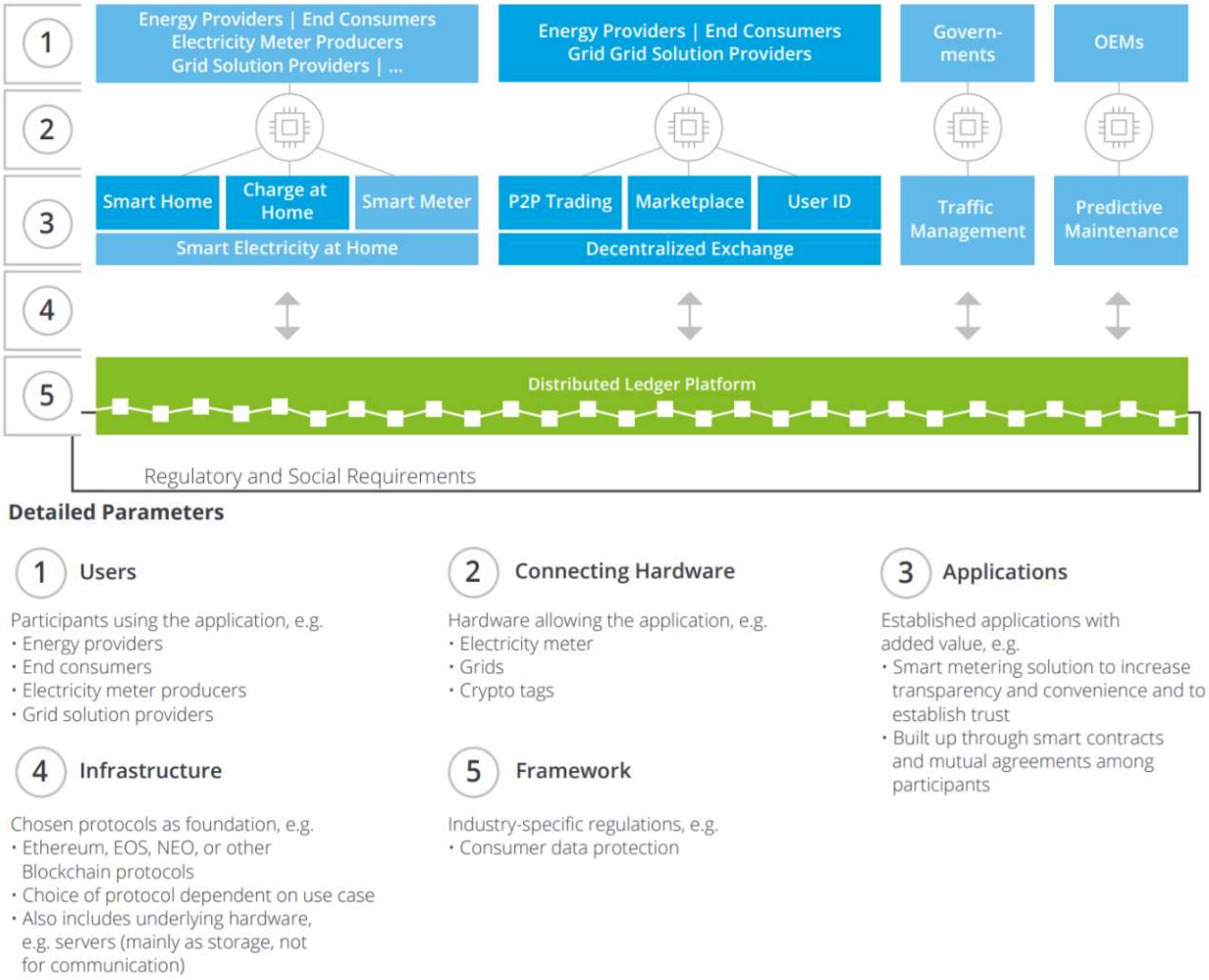


Figure 29 - Distributed Ledger Platform
Source: Deloitte

Local Peer-To-Peer Markets

Owing to the small amounts of electricity produced by domestic suppliers and their likely supply intermittency, due to their renewable-based production, there would have to be significant differences between the current system of energy supply and consumption, and the one envisioned with a greater installed capacity of microgeneration:

• Instead of sourcing their electricity from a single supplier, consumers would buy and sell electricity across an open market, fundamentally, swapping their energy supplier on a minute-by-minute basis.

• Households will be able to buy electricity depending on their own personal preference, whether it be distance from the generator, type of generation, or just to buy at the best price.

To enable this, a P2P trading platform would need to be formed, in which households could market their electricity exports for other households to buy. Once the platform is in place,

optimization methods could be developed to automate the trading process to either increase monetary benefits, support local generators, or even to smooth load curves on the grid as a whole.

Electricity transmission results in energy losses - the further electricity travels, the more of it is lost. If the true cost of losses were reflected in sale price, then generators would be able to offer a better price to nearby customers. This could benefit both the environment and customers. Thus, the key benefits for domestic suppliers using this platform are:

• Ability to exercise independent choice on the purchase and sale of electricity according to their personal preferences and needs;

• Added monetary benefits through optimizing for the most favorable energy transactions at any given time;

• Increased independence from the grid in case of power supply issues from MPPs.

The platform is comprised of a number of components that work together to enable a decentralized, open market. A household supplier is a micro-generator that generates electricity, uses some of its own product, sells excess and purchases electricity on the market when its own generation falls short. Each household supplier has a smart meter installed which records electricity consumption and generation data. This data is subsequently passed to an oracle that controls the movement of content from the real-world on to the blockchain. In this case, the oracle tokenizes their electricity generation and links it to a household supplier's address on the blockchain ledger. This tokenization would allow the generation to be represented as a sub-currency on the network, so it could be traded (e.g., 1 token per kWh generated). For the oracle to accomplish this task, householders would first need to register to the oracle to associate their microgeneration equipment with their wallet address. During this process the oracle would also perform a quality assurance check to ensure their equipment was installed by a registered installer and has been certified. This would be done by checking schemes such as the Microgeneration Certification Scheme.

Each household supplier that trades on this platform must have a wallet on the blockchain which provides an address to identify that supplier on the network. The trade balance of each participant on the platform is maintained by a smart contract known as the export registry, and the electricity trading is enabled through a smart contract called the market.

The start-up LO3 Energy is setting up pilot projects, globally, to demonstrate peer-to-peer electricity trading. Its most high-profile project is the Brooklyn Microgrid. The project's goal is to network thousands of Brooklyn residents in a self-sufficient microgrid, which is a small electricity network with its own sources of supply that can function independently of the main grid. Such a microgrid could improve the resilience of electricity supply to Brooklyn residents in the face of natural disasters that might cause the main grid to shut down.

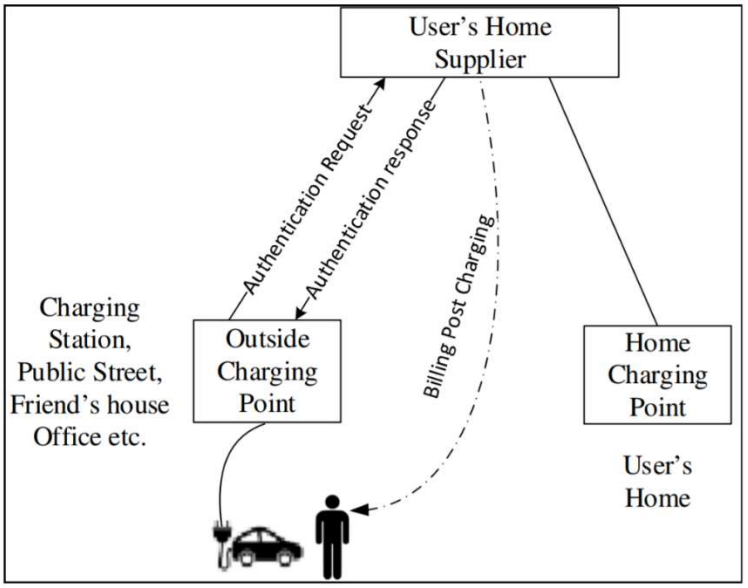
Moreover, by resourcefully harnessing distributed generation sources such as solar panels, the microgrid could theoretically need less expensive infrastructure to produce and deliver energy within Brooklyn, reducing the bills of customers who currently pay high rates to cover their share of the costs of the main grid. To enable customers to efficiently use their distributed energy resources, LO3 is developing a blockchain platform to facilitate peer-to-peer energy trading.

The initial version of the Brooklyn Microgrid is a far cry from LO3's ultimate vision. The pilot project comprises fewer than sixty prosumers. A larger number of participants can virtually trade electricity with one another, but they are not physically connected by a microgrid. Instead, most of the project's participants simply continue to use the main grid. When two participants "trade" electricity and one pay the other, the physical flow of electricity remains unchanged—for example, one participant feeds excess solar power back into the distribution grid, and the other participant consumes electricity from the grid. In fact, the participants cannot even transact electricity, because the utility has a monopoly over electricity sales, and rather, can only trade renewable energy certificates.

As a result, the first iteration of the Brooklyn Microgrid does not provide resilience, cost, or sustainability benefits. (LO3 argues that some customers can decrease their costs by selling surplus distributed energy through the virtual microgrid, and this might encourage the deployment of additional rooftop solar panels. But at a system-wide level, this practice is unlikely to reduce costs or carbon emissions.) Importantly, using a blockchain ledger to facilitate energy trading is only one component of employing a microgrid. Microgrids also require both software and hardware to keep the system in balance and interact with the main grid.

For LO3 to achieve more of its vision it will require the collaboration with utilities and regulators. It hopes to work with the New York utility Con Edison to present Brooklyn Microgrid participants with a single bill that integrates their transactions with other participants as well as their cost of service by the utility. LO3 also hopes to convince regulators to allow it to legally broker sales of electricity among project participants. It will be several years before the company can construct a physical microgrid that can operate independently from the main grid and doing so will almost certainly need the cooperation of the utility and state authorities. LO3 has sought industry partners around the world upon recognizing this. For example, in South Australia, it has partnered with an authorized energy supplier and is working with regulators to set up an energy trading platform like the Brooklyn Microgrid.

In Austria, the companies Verbund and Salzburg AG have developed a blockchain P2P proof of concept that enables tenants to exchange shares of the generated electricity from their roof via a distributed blockchain app on an android tablet. These shares are stored on a proof-of-work blockchain, which is run by the tenant themselves. The grid operator Salzburg Netz GmbH then collects the transaction data via a read-only access and allocates the own consumption to the individual household bills.



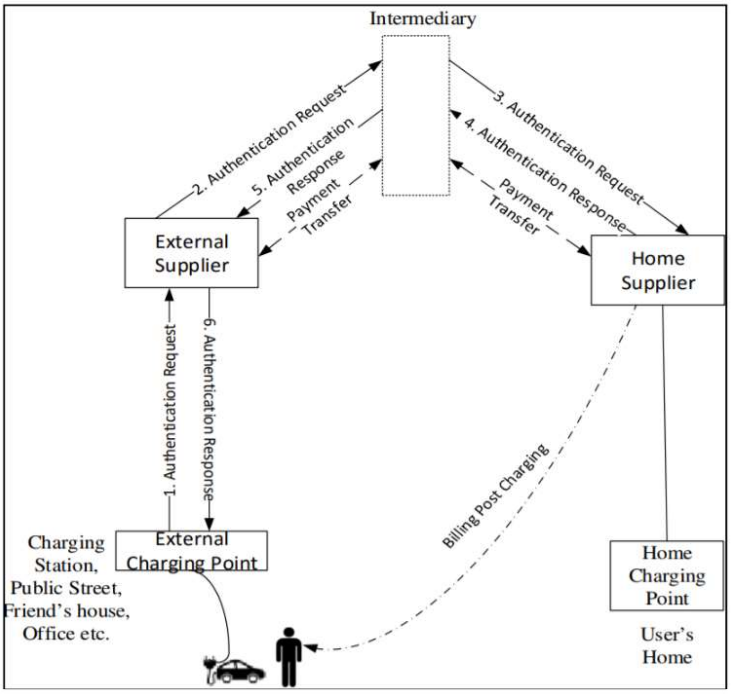


Figure 30 - Internal and External Roaming charging in Smart Grid systems
Source: Using Blockchains to Secure Distributed Energy Exchange

Smart Charging/Discharging

This scenario discusses how peer-to-peer charging and discharging of PEVs could be accomplished using Blockchains. Before presenting the scenario that explains the Blockchain based approach of PEV charging/discharging, the difficulties and proposed approaches to perform PEV charging/discharging based on the existing system are reviewed. This would help the reader appreciate the advantages of using Blockchains for PEV charging/discharging over other approaches. In traditional power system, users need to register with suppliers to get power to their homes or offices and are billed based on the power consumed at that fixed location. This approach is not convenient for PEV charging/discharging. PEV users are mobile and are assumed to be able to charge/discharge anywhere. Charging outside one's home charging point is known as roaming charging. Roaming charging can further be divided into two categories: Internal roaming charging (IRC) and external roaming charging (ERC). If the user charges outside his home but at a charging

location that is within the same supplier as the user's home supplier, it is called IRC. However, if the user charges at a charging location that resides outside the home supplier network, it is referred to as ERC.

Existing power systems require additional capabilities such as remote authentication and payment transaction mechanisms to support roaming charging. Moreover, authentication and payment mechanisms should be equipped with appropriate privacy preserving techniques to protect user's privacy from charging stations and external supplies. As can be seen from the figures, the charging station authenticates the user by contacting the home supplier as the user is unknown to the charging station. The user is billed by the home supplier after the completion of charging. Moreover, during ERC, payment transaction is performed between the home supplier and the external supplier which might involve the usage of a trusted third party as an intermediary. Payment transactions are performed in one of two directions depending on the type of service (charging or discharging). These proposed protocols work well for energy exchange between PEV users and utility companies. However, they do not support peer-to-peer distributed energy exchange. Blockchain allows roaming PEVs to dynamically participate in distributed energy trading without the need to rely on a central supplier for authentication and payment transaction processes. There is no need for pre-registration and contract establishment. Every PEV user can engage in the energy trading dynamically independent of third parties.

With the advancements in Machine Learning and AI, an algorithm can be created which would alter the process of charging/discharging of the EV's to optimize the cost. Another approach would be to alter the time in which the charging/discharging is done so as to supplement the factor of convenience.

Electric Vehicle Charging and Coordination

As electric vehicles (EVs) become more prominent, system operators are faced with the challenges of supplying new EV-related mobile load and, potentially, using surplus stored energy to improve system flexibility. Blockchain technology could improve EV charging coordination by facilitating energy payments at charging stations, and by enabling drivers to make charging decisions based on map and real-time pricing data.

An example of an active project in this space is MotionWerk's "Share&Charge" app. In 2016, Innogy (a subsidiary of German utility RWE) partnered with German blockchain startup Slock.it to create a P2P service allowing EV and charging point owners to rent their charging infrastructure to each other autonomously without the need for an intermediary. By May 2017, Innogy's "Innovation Hub" incubator had spun out a startup, MotionWerk. Its first product, "Share&Charge," allowed EV owners to charge their vehicles by making digital payments using a mobile app. Charging point owner used the application to make their infrastructure available, set tariff structures and to collect fees. Until April 2018, the service

was available to about 1,000 EV owners with 1,250 private and public charging points registered in Germany. The system used an e-wallet and smart contracts on the public Ethereum blockchain as P2P transaction layer, including a Euro-backed "Mobility Token." Share&Charge was the world's first e-mobility transaction platform that used blockchain. Based on end-customer experience and learnings from different pilot initiatives in the EU and the US that MotionWerk conducted (e.g. the Oslo2Rome project), Share&Charge is currently transforming into an open source and decentralized digital protocol for electric vehicle charging. It is envisaged to allow charge point operators and e-mobility service providers to fully decentralize their e-mobility assets to, next to other benefits, simplify processes of controlling, payment and settlement of charging EVs.

Share&Charge is also being tested outside Germany. US based EV charger company eMotorWerks (an Enel group company) has been testing a blockchain-based peer-to-peer charging marketplace in California, allowing drivers to pay each other for use of their home chargers. eMotorWerks is using the Share&Charge platform.

3.8 Risks and Limitations

Blockchain-driven innovation and disruption provides great opportunity for the different economies of the world, but many of the best and most meaningful applications for this technology still lie ahead. The pace of innovation is fast. Perhaps it's even faster than you imagine. At its broadest, blockchain is sparking a new computing paradigm that offers decentralization, coordination and collaboration in a secure and autonomous way. If successful, it will revolutionize diverse sectors of the economies and bring efficiencies of scale that previously required large, centralized operators in order to have an impact. Here, some specific challenges reflected today in other countries' cities are discussed.

3.8.1 Sustainability and Technical Challenges

Cities powered by blockchain can be more resilient and secure, but challenges and barriers to adoption of the technology persist. This is a lesson well learned by the world's first digital nation, Estonia. In April 2007, a wave of cyberattacks likely originating from Russia targeted the country's digital infrastructure. Since this attack, Estonia has become Europe's hub for cybersecurity companies as well as home to NATO's Cooperative Cyber Defense Centre of Excellence. Building on momentum in this space, in Dubai, local government has created a special office charged with transforming Dubai into a Smart City. It's called Smart Dubai Office and was launched as the first city-wide blockchain strategy in October 2016 with the aim of becoming the first blockchain powered city in the world by 2020. While the innovation charges ahead, it has experienced some technical issues. For instance, Bitcoin, the popular blockchain application, has struggled with issues related to storage and sustainability. One

estimate shows that a single bitcoin transaction uses three thousand times more power than a credit card payment.

3.8.2 Regulatory Challenges

The validity of cryptocurrencies (digital assets) as well as blockchain applications are still uncharted territory. As mentioned earlier, a number of U.S. states have taken steps to validate the electronic signatures within blockchains as enforceable under contract law. While cryptocurrency exchanges began cropping up around the world, the Securities and Exchange Commission rejected two Bitcoin ETF proposals in 2017 while the Internal Revenue Service issued a John Doe Summons to Coinbase, Inc., a digital currency exchange, for user and transaction history. Both actions centered on concerns around the anonymity of users and possibilities for tax evasion and money laundering.

New York State Department of Financial Services (DFS) Superintendent Ben Lawsky became the first leader to address this regulatory challenge, requiring and issuing the first BitLicense. The New York DFS now requires anyone who stores, holds, transmits, buys or sells cryptocurrencies for commercial purposes to obtain a BitLicense, which registers them as a commercial agent allowed to hold and use cryptocurrencies for this purpose. The same requirement applies for anyone who performs or exchanges services for commercial purposes, and controls, administers or issues virtual currency. Many in the industry decried this regulation as a step towards protecting incumbent businesses and threatening innovation. Illinois has since taken a light-touch regulatory approach in the form of iterative forums and issuing nonbinding guidance.

Regulatory bodies' slow pace in adapting to blockchain-powered innovation has delayed the advancement of projects such as the Brooklyn Microgrid. Although the project operates on an energy credit system, it has not yet been recognized as a utility that can buy and sell power. This means it could potentially lag behind other innovations. Government can engage with and encourage blockchain innovators, much like the Central Banks of Australia and Canada have, through in-depth studies, beta testing and grant awards, and creating and encouraging dialogue between regulators and the industry. Launching initiatives that coordinate between public agencies and the innovator community has been successful in informing public administrations about the unique concerns and possibilities of blockchain technology. Isolating experimental missteps and moving partnerships into limited trials has become an ideal way to begin slow and gradual implementation.

3.8.3 Financial Speculation

While this is not an issue for private blockchain networks or blockchain applications beyond financial transactions, the explosive market capitalization and valuation of Bitcoin and other cryptocurrencies may not be sustainable. Recently, the market capitalization for cryptocurrencies surged exponentially while at the same time concerns have grown that the technological hype is driving a bubble similar to the dot-com boom of the 1990s. A proliferation of alternative cryptocurrencies that are exchangeable and tied to Ethereum (a common open source coin platform for startups issuing their own coins to build from), known as altcoins, has led to an increasing demand for the cryptocurrency.

Ethereum has swelled dramatically. As many new startups increased demand for Ethereum, the price of the cryptocurrency skyrocketed, leading holders of the currency to invest in more startups with smaller amounts of their total wealth and attracting less well-informed investors to the market. Total market capitalization of cryptocurrencies, including Bitcoin, stood at over $350 billion in March of 2018. This could be a sign that the cryptocurrency ecosystem is maturing. On the other hand, with recent volatility sending cryptocurrencies soaring to more than $800 billion in early 2018, this market capitalization could also be the marker of an unsustainable build in demand.

4. Discussion

A city can be called a "Smart City" when it makes use of advanced technologies to understand the various systems facilitating its function as well as the environment. It primarily analyses the data and makes the required changes to tackle existing issue such as mass urbanization, energy efficiency, as well as to improve the quality of life of its residents. In this thesis, the technologies that were discussed includes IoT, Big Data, Artificial Intelligence, Machine Learning and Blockchain. Apart from the mechanism and application, the benefits, limitations and challenges of these technologies were also evaluated. The pros and cons of these technologies must be carefully taken into consideration whilst building a framework for an effectively functioning Smart City.

One of the major challenges in this respect is the issue of interoperability, which is widely prevalent due to the diverse modes of communication networks and the lack of regulation or achieving unison by the manufacturers to make the products compatible with one another. Every manufacturer would probably adhere to the communication network they built their product on for business reasons and citing design consideration. But for a Smart City to be successful, the various data streams, platforms and services must be compatible with each other, thereby promoting interoperability between each other. This successfully breaks the existing technical barrier and allows access to more ground for exploring the various possibilities. For instance, considering two Smart Meter manufacturers A and B, the data streams generated by both of them must be similar so that they can connect to a common cloud which processes their data.

As the population of such technology-based products and services are growing exponentially, another major concern is the scalability of the technologies as they are often limited by their technological limits and constraints. Further R&D focused on the scalability and interoperability of the technologies could contribute to their state of maturity, e.g. the peer-to-peer energy transactions which are tested on small scales as of now can be implemented on a much larger basis without complications. The more the demonstrations are carried out regarding the maturity of the technology, the greater the positive influence will be on its reliability. The easier availability and acceptability for the active deployments for such solutions accounts for this positive influence. Especially with the use of upcoming technologies such as Machine learning and Artificial Intelligence, reliability has been turning out to be of crucial importance in incorporating them into solutions of Smart Cities. For instance, Machine Learning algorithms can analyze the electricity consumption data, understand the user behavior, predict the possible consumption upon correlating with the

weather conditions and provide an estimated electricity bill. This data provides a complete analysis of the user's behavior which could be misused in multiple ways and this has to be prevented. According to Behavioral Economics, people are generally critical about the reliability of any product or service they consume and are inclined to maximize the benefit from it. This also applies to the technological services and goods that people consume and since it actively influences their everyday lives and deals with their personal and private data, they are more critical of it.

The privacy and security of the data involved is a critical concern as mentioned above. With the data breaches like Ransomware and the famous case of Facebook, privacy has grown to be a pivotal concern for any kind of digital technology that deals with people's data. The solution provided by Blockchain Technology is viable and provides a much secure environment due to its powerful cryptographic techniques and decentralized ledger feature. However, the pace in which it is being incorporated into existing systems have been considerably slow. This could possibly be due to the slowness in its transaction speed. This could affect the processing of bulks of information flowing in from the numerous sensors across a Smart City or due to the lack of supportive regulations which will have to be addressed in the future.

A framework that could be built by incorporating the different technologies that were discussed above would indeed pave way for constituting a highly efficient and functional Smart City. However, there are a number of other crucial factors such as the readiness and maturity of these technologies that should be taken into account with regards to the potential progress and development of these technologies to be applied in a real-life working environment, as most of the projects and applications that were discussed were prototypes and experimental projects. Upon successful outcomes of such experimental projects, their application at a larger scale would require certain changes to be made in the existing policies and regulations of the respective setting, be it a community or a country. Along with these hurdles the willingness and the ability of the cities and countries to invest capital and labor into such a project is crucial. These rudimentary concerns are briefly discussed in this following section.

4.1 Technical Readiness

It can be understood that a Smart City makes use of information and communication technology (ICT) to enhance its sustainability, livability and workability. The entirety of this process can be categorized by three core processes: First would be the collection of the data through sensors, other devices and existing systems. Followed by the communication of the collected information through wired or wireless networks. And lastly the crunching of the

data collected in order to obtain useful and valuable information from it. The data collection and communication would be taken care of by the ***IoT systems*** while the analytics of the data collected will be taken care of by the Big Data Analytics.

The data crunching process can be further classified into i) Presenting, ii) Perfecting and iii) Predicting. The presenting function will be taken care of by the ***Big Data Software Platform*** which would allow users to access, control and make use of the desired data. The Perfecting and Predicting functions will be taken care of by the ***Artificial Intelligence*** and ***Machine Learning*** in accordance with Big Data as they will be able to automate and predict a multitude of tasks. This would in turn immensely reduce the time and energy required for such functions to be carried out manually, thereby improving the operational efficiency. The ***Blockchain technology*** and Smart contracts could be integrated to improve the security and autonomy of processes thereby contributing to the creation of a secure network.

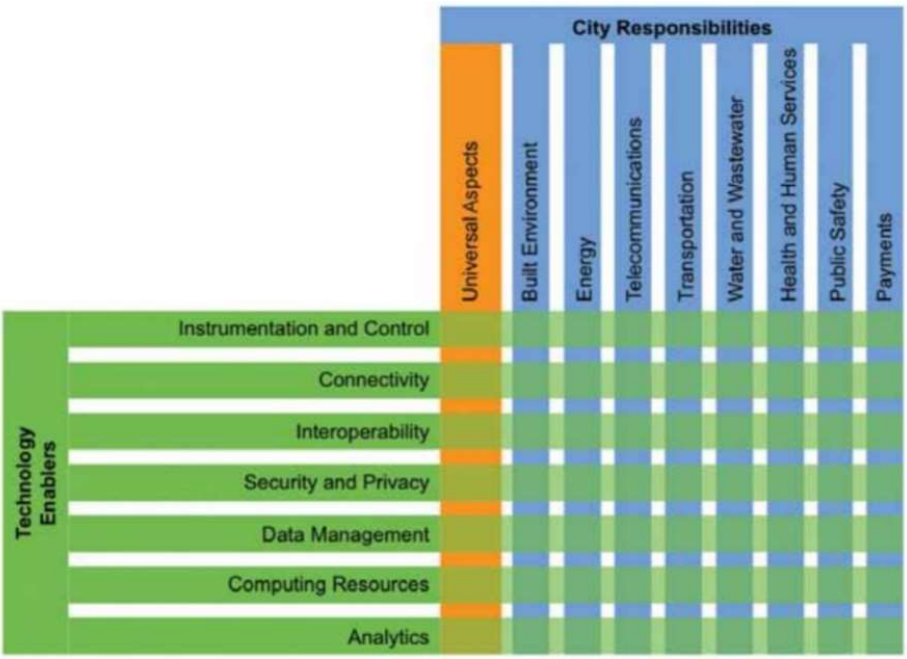


Figure 31 - Aligning responsibilities and enablers.
Source: Smart Cities Readiness Guide

Most cities which proclaim to be smart lack such a unified framework which majorly is due to the financial constraints in the past and the present. This often results in division of the different departments of the city into "siloed" departments which have little to no interaction between them. Thereby creating "islands of automation" built to solve a single problem within a single department yet the sharing of data between them remains difficult. By using an integrated approach and having a system wide view the siloed model can be overcome, which would enable the re-use of data by different departments. However, such

a framework tends to be complex and large scaled hence demanding holistic thinking and hard work which makes the technology readiness viable but, the human ability to coordinate and collaborate between the departments and the technology silos, questionable.

A comprehensive overview of the different factors including the ones discussed above for planning the framework for a Smart City has been published by the Smart Cities Council (The Smart Cities Readiness Guide). The different Technology enablers and the different City responsibilities are matched with each other and a percentage of the current implementation progress is scored. Then the final chart can be used to figure out the areas that require immediate looking into and the rest of the developmental process based on the prioritization.

Conclusively, though the technologies are not at their epitome of maturity, they are certainly mature enough to be implemented into the existing systems of a city. With time the possibilities of integration and the benefits from their applications such as Smart Energy, Smart Mobility and so on can be enhanced and better utilized, helping in achieving the futuristic dream of a Smart City.

4.2 City Readiness

As mentioned in the 2018 report of *Smart Cities: Digital solutions for a more livable future* by the McKinsey Global Institute, 50 cities globally were selected and were analyzed and scored over a combined maximum score of 110 points based on the three layers that collectively reflect their "smartness" which are the technology base, the number of applications they are using in various domains and public adoption.

The strength of the technology base is given a maximum of 37 points and is scored based on the sensors, communication and open data portal. While the leading cities in North America, Europe, China and East Asia take a more sophisticated approach towards building a world-class digital infrastructure the cities in developing countries such as India, Africa and Latin America seem to be lagging due to capital-intensiveness of the process. Whereas the deployment of Smart City applications is divided among the various sectors of an economy they are collectively given a maximum of 55 points. And the citizen experience of the technologies is classified into awareness, usage and satisfaction and is collectively given a maximum of 30 points.

The findings indicate that a political framework of a city alone does not define or direct the city's ability to become smarter. And even the most cutting-edge smart cities have a long way to become a fully functional Smart City. Cities which have more centralized governments compared to the ones with democratic societies and those that are wealthier,

are better equipped to establish comprehensive communication and sensor networks. The vice versa doesn't hold true, where it is not necessarily true that every high-income city is

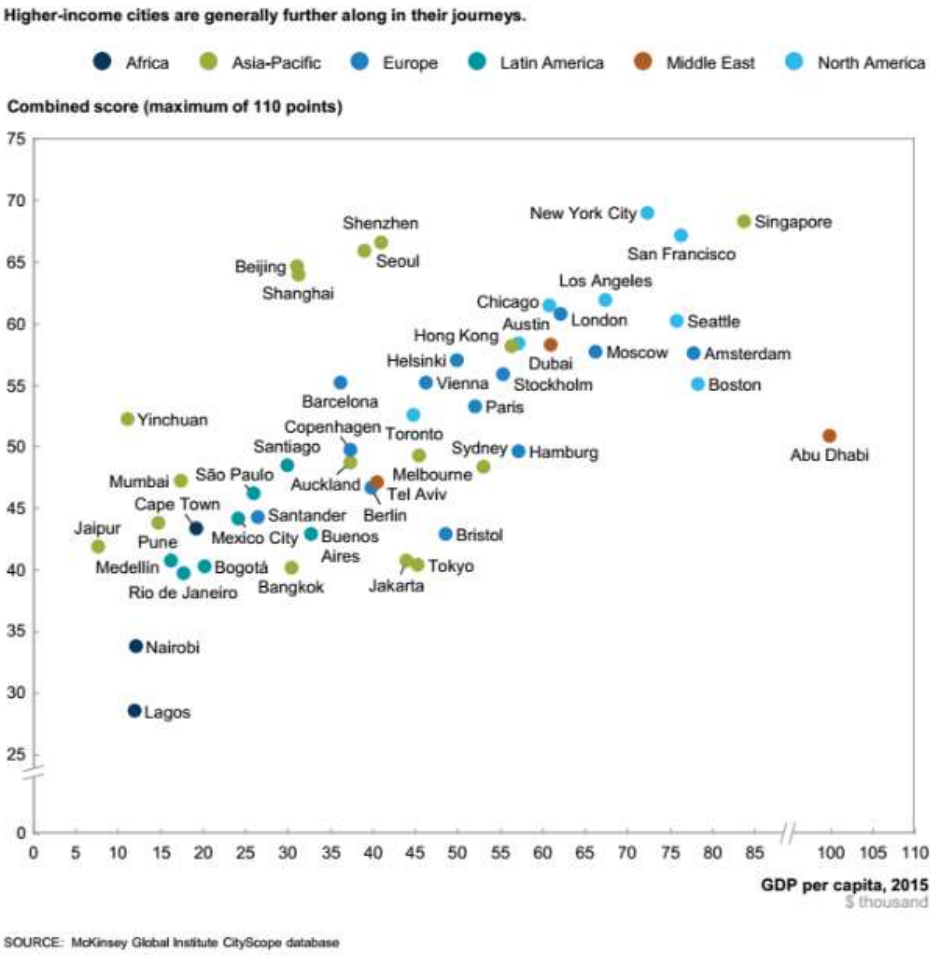


Figure 32 - Where Cities Stand: A Snapshot of Deployment

driving towards implementing all possible applications.

An interesting pattern of low public awareness and usage in some high-income cities with older population can be observed, which could be caused due to a multitude of factors such as resistance to change, being contended with the existing systems, high expectations for technology or the lack of communication and comprehension. However, Chinese cities appear to be adopting such systems at a considerably fast pace. The economic future of developing economies is primarily determined in the cities and the cities of the Asian continent seem to be actively adapting to the positive changes and appear to bear the torch for the Smart City movement in the future.

It can be concluded that a sufficient amount of capital is vital for the research and establishment of the technologies in a city whereas a greater level of familiarization,

acceptance and appreciation from the population is necessary to steer towards a common goal of a better and efficiently functioning Smart City.

4.3 Policy Readiness

Considering the pace of development of these technologies the pace in which the policies are modified to be able to accommodate them have been unarguably slower. Only by making changes in the existing policies can a new system be welcomed into the pre-existing functional setup. Though the older setup could be less efficient than the new one it is often not enough to prove it to the masses only on paper. For instance, with the announcement of COP21 goals several countries have announced their target year for ending the sale of Internal Combustion Engine vehicles and several countries have announced incentive schemes for electric vehicles whereby people will be actively encouraged to switch to Electric Vehicles after a given point in time. This is the active impact policies can have on technology and the people using it.

Policy makers around the world have embraced the smart cities concept and put funding into initiatives and pilot programs. China highlighted Smart City development as one of the major economic priorities in its 12th Five-Year Plan (covering 2011–15), and more than 500 cities across the country have already developed strategies or launched pilot projects. The EU's Horizon 2020 initiative, a €77 billion research and innovation program established in 2013, made smart cities an important part of the agenda which heavily uses IoT. Soon after, India announced plans to create 100 smart cities across the country. The US Department of Transportation launched a Smart City challenge in 2015, encouraging cities across the country to submit innovative plans and compete for grant money. Around the world, governments have partnered with private sector companies to create purpose-built smart cities from the ground up.

City leaders can also decide to push things even further, for example, by encouraging ride sharing, mandating night deliveries, or allocating dedicated lanes for autonomous cars. Individual cities will make their own choices about what kind of future to pursue. Cities on one end of the spectrum may exercise caution or resist change, while others could take bold leaps such as banning private vehicles from the urban core. They have a wide variety of policy tools at their disposal: they can set mandates, incentives, subsidies, and standards; they can convert government fleets; and they can support the build-out of vehicle-charging infrastructure. Cities would do well to engage with the public as they map out their implementation path and address concerns regarding safety, employment, and affordability.

Policies and regulations which ensure the safety, security and privacy of users' data must also be passed for ensuring a safer environment to the users which would thereby give them a sense of security with regards to using the different technologies and platforms. The European Union enforced the General Data Protection Regulation (GDPR) on 25th May 2018, which ensures protection and privacy of data for all individuals within the EU and the EEA. This allows the users to have control over their personal data and to choose if their data can be used for business purposes or not. As stated earlier and further emphasized by the GDPR, privacy and security are of utmost importance with regards to evolving technology and ensuing development of smart cities. The employment of technologies like blockchain ideally caters to this demand and can therefore be key for future policy-making.

Hence, it can be said that the policies and regulations play a crucial role in the shift to the technology-based Smart Cities and also in ensuring the safety, security and privacy of the users upon the shift. Along with guaranteeing the safety and security such policies also promote and drive the masses towards a more efficient and easier to use system which would also be sustainable.

Overall, the different technologies can function independently and constitute the architecture of a smart city or can be combined with one another in an interwoven manner to build a stronger architecture for the smart city. Along with the technologies, the people must be made aware of the functionalities and benefits of these technologies which would allow them to readily accept the novel concept of smart technology and contribute to an efficiently and sustainably functioning Smart City. Implementation of IoT, Big Data and Blockchain Technology for the creation of such a Smart City is certainly possible but each technology presents with its own complexities. These nuances need to be further explored and understood, which could plausibly take several more years of research and development before these technologies could be considered mature.

5. Conclusion and Future Recommendations

In the past, smart technologies were primarily considered only as tools for improvising a city's efficiency. With the rapid and exponential development in the technology sector, data obtained from sensors and high-tech operational centers that process the data are proving to be revolutionary in managing complex operations and automating infrastructure systems. Solutions based on these technologies are being implemented in real-life cases. The benefits of which apply to a plethora of sectors and can be observed from the infographic below.

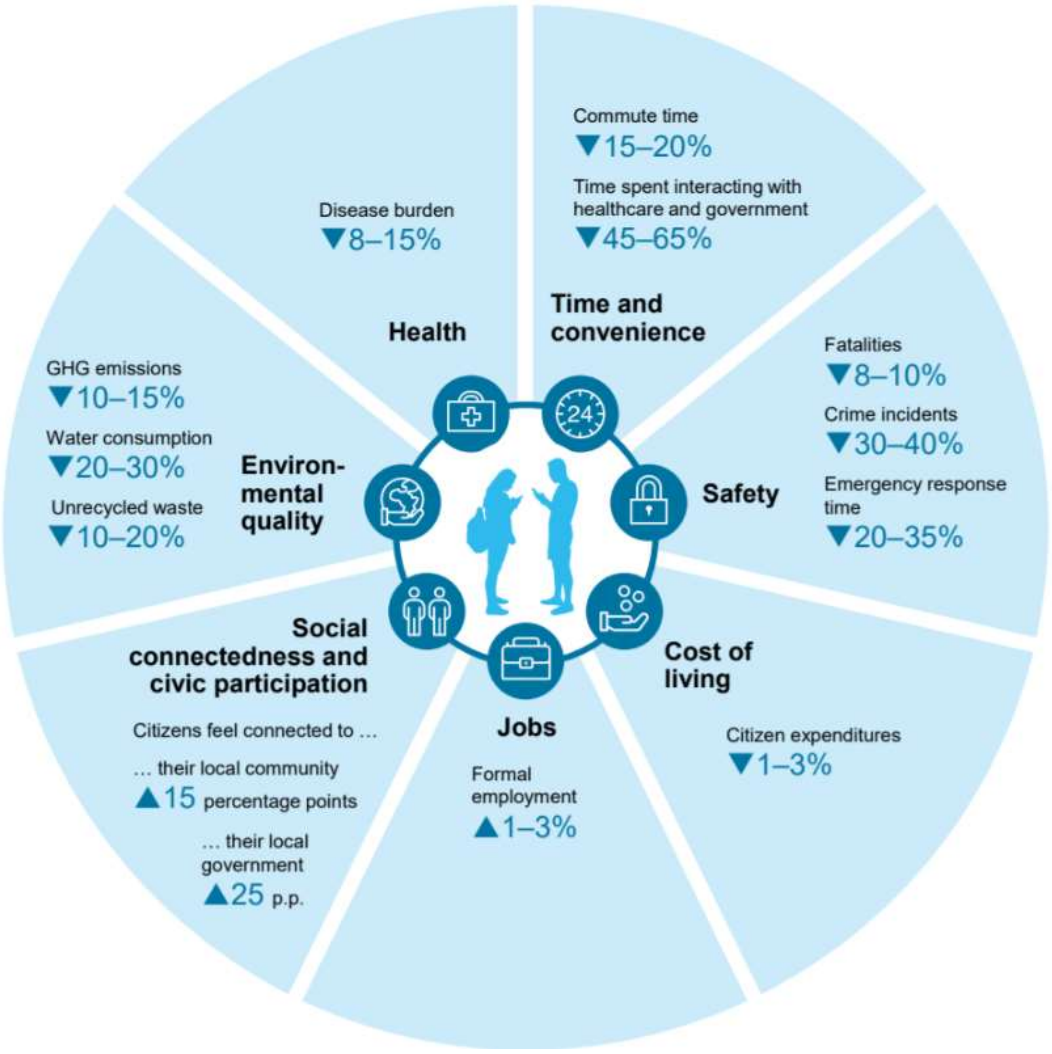


Figure 33 - Potential improvement through current generation of smart city applications
Source: McKinsey Global Institute analysis

An analysis of the various technologies that can be employed to build a fully functional Smart City was carried out and their intricacies were explored. It is well evident that these technologies are applicable and effective for building a strong framework for a Smart City.

The Internet of Things takes care of the information collection using the multiple sensors which are installed for instance in Smart Meters to predict the amount of consumption of energy. The Big Data Analytics helps with the processing of the data collected by the IoT and helps in creating value from it which can ultimately be used for a multitude of applications. The Artificial Intelligence and Machine Learning technologies can be integrated with the Big Data Platform to help with advanced prediction with increased accuracy and to automate processes which would otherwise consume a lot of man hours. These find numerous applications in a Smart City from delivering Travel Transit time estimates to the prediction of the amount of electricity a PV system would generate on a particular day taking into account the changes in the weather and so on. The Blockchain Technology in addition to helping with the secure transactional network also contributes to a Decentralized and Distributed Ledger System which forms the basis for the functioning of the Peer-to-Peer Energy transactions which possess a great potential in the Renewable Energy Sector.

These technologies eliminate difficulties that the conventional systems constantly face such as manual recording of data, complex procedures for analyzing data and potential risks for data breaches and theft and so on. However, they do have certain limitations such as their interoperability, scalability, reliability and so on. These issues will have to be addressed in order to be able to term these technologies remarkable.

Individual components of each tech influence the effectiveness of the Smart City and so does the interplay between each technology. From a broader perspective these technologies are certainly anchoring the creation of a highly functional smart city. However, as with the evolution of any technology, in accordance to supply and demand (energy issues posed within a community), more investigation is needed for an in depth understanding of these technologies and thus, enhancing the development of a Smart City.

References

1. Aaltodoc.aalto.fi. (2018). [online] Available at: https://aaltodoc.aalto.fi/bitstream/handle/123456789/13899/master_Ahmed_Hussnain_2014.pdf [Accessed 10 Oct. 2018].

2. Ahmad, T. and Chen, H. (2018). Utility companies strategy for short-term energy demand forecasting using machine learning based models. *Sustainable Cities and Society*, 39, pp.401-417.

3. Aitzhan, N. and Svetinovic, D. (2018). Security and Privacy in Decentralized Energy Trading Through Multi-Signatures, Blockchain and Anonymous Messaging Streams. *IEEE Transactions on Dependable and Secure Computing*, 15(5), pp.840-852.

4. Al Nuaimi, E., Al Neyadi, H., Mohamed, N. and Al-Jaroodi, J. (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications*, 6(1).

5. Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M. and Ayyash, M. (2015). Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Communications Surveys & Tutorials*, 17(4), pp.2347-2376.

6. Alshawish, R., Alfagih, S. and Musbah, M. (2016). Big data applications in smart cities. *2016 International Conference on Engineering & MIS (ICEMIS)*.

7. Analytics Magazine. (2018). *Smart cities: A world of opportunity in data - Analytics Magazine*. [online] Available at: http://analytics-magazine.org/smart-cities-a-world-of-opportunity-in-data/ [Accessed 22 Nov. 2018].

8. Ara, T., Gajkumar Shah, P. and Prabhakar, M. (2016). Internet of Things Architecture and Applications: A Survey. *Indian Journal of Science and Technology*, 9(45).

9. Assets.kpmg.com. (2018). [online] Available at: https://assets.kpmg.com/content/dam/kpmg/pdf/2016/02/ch-pub-20160122-clarity-on-data-analytics-en.pdf [Accessed 1 Dec. 2018].

10. Atis.org. (2018). *Data Sharing Framework for Smart Cities | ATIS*. [online] Available at: http://atis.org/smart-cities-data-sharing/ [Accessed 26 Nov. 2018].

11. Babu, T. and Swathi, P. (2018). Internet of Things (Iot) & Big Data Analytics for Smart Cities-A Case Study. *SSRN Electronic Journal*.

12. Barns, S. (2018). Smart cities and urban data platforms: Designing interfaces for smart governance. *City, Culture and Society*, 12, pp.5-12.

13. Blockgeeks. (2018). *What is Blockchain Technology? A Step-by-Step Guide For Beginners*. [online] Available at: https://blockgeeks.com/guides/what-is-blockchain-technology/#How_Does_Blockchain_Work [Accessed 8 Nov. 2018].

14. Cheng, B., Longo, S., Cirillo, F., Bauer, M. and Kovacs, E. (2015). Building a Big Data Platform for Smart Cities: Experience and Lessons from Santander. *2015 IEEE International Congress on Big Data*.

15. Choudhary, S., B. Sathe, R. and Kachare, A. (2017). Smart Cities Based on Internet of Things (IoT) -A Review. *International Journal of Engineering Trends and Technology*, 48(8), pp.434-439.

16. Das, S., Dey, A., Pal, A. and Roy, N. (2015). Applications of Artificial Intelligence in Machine Learning: Review and Prospect. *International Journal of Computer Applications*, 115(9), pp.31-41.

17. Datasciencecentral.com. (2018). *Artificial Intelligence vs. Machine Learning vs. Deep Learning*. [online] Available at: https://www.datasciencecentral.com/profiles/blogs/artificial-intelligence-vs-machine-learning-vs-deep-learning [Accessed 27 Jul. 2018].

18. Deloitte India. (2018). *Connecting the next billion | Deloitte India | Technology, Media & Telecommunications*. [online] Available at: https://www2.deloitte.com/in/en/pages/technology-media-and-telecommunications/articles/connecting-the-next-billion.html [Accessed 18 Nov. 2018].

19. Di Silvestre, M., Gallo, P., Ippolito, M., Sanseverino, E. and Zizzo, G. (2018). A Technical Approach to the Energy Blockchain in Microgrids. *IEEE Transactions on Industrial Informatics*, 14(11), pp.4792-4803.

20. DIGITAL SOLUTIONS FOR A MORE LIVABLE FUTURE. (2018). McKinsey & Company.

21. Disruptor Daily. (2018). *Top 10 Smart Energy Startups to Know in 2018 - Disruptor Daily*. [online] Available at: https://www.disruptordaily.com/top-10-smart-energy-startups/ [Accessed 9 Jun. 2018].

22. Ejaz, W., Naeem, M., Shahid, A., Anpalagan, A. and Jo, M. (2017). Efficient Energy Management for the Internet of Things in Smart Cities. *IEEE Communications Magazine*, 55(1), pp.84-91.

23. Engerati - The Smart Energy Network. (2018). *How the internet of things is transforming energy*. [online] Available at: https://www.engerati.com/article/how-internet-things-transforming-energy [Accessed 11 Aug. 2018].

24. Europarl.europa.eu. (2018). [online] Available at: http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/507480/IPOL-ITRE_ET(2014)507480_EN.pdf [Accessed 17 Jul. 2018].

25. European Commission - European Commission. (2018). *Smart cities*. [online] Available at: https://ec.europa.eu/info/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities_en [Accessed 29 Jun. 2018].

26. Forbes.com. (2018). *How Cities Are Getting Smart Using Artificial Intelligence*. [online] Available at: https://www.forbes.com/sites/tomvanderark/2018/06/26/how-cities-are-getting-smart-using-artificial-intelligence/#3c0592938036 [Accessed 16 Oct. 2018].

27. Geetha Pratyusha, M., Misra, Y. and Anil Kumar, M. (2018). IoT based reconfigurable smart city architecture. *International Journal of Engineering & Technology*, 7(2.7), p.175.

28. GlobalLogic Israel. (2018). *The Role of Telecommunications in Smart Cities | GlobalLogic Israel*. [online] Available at: https://www.globallogic.com/il/gl_news/the-role-of-telecommunications-in-smart-cities/ [Accessed 10 Nov. 2018].

29. Governmentciomedia.com. (2018). *4 Examples of How AI Can Make Cities Smarter*. [online] Available at: https://www.governmentciomedia.com/4-examples-how-ai-can-make-cities-smarter [Accessed 30 Aug. 2018].

30. Hassani, H., Huang, X. and Silva, E. (2018). Big-Crypto: Big Data, Blockchain and Cryptocurrency. *Big Data and Cognitive Computing*, 2(4), p.34.

31. Hippold, S. (2018). *Use AI to Make Cities Smarter*. [online] Gartner.com. Available at: https://www.gartner.com/smarterwithgartner/use-ai-to-make-cities-smarter/ [Accessed 20 Oct. 2018].

32. Huawei Enterprise. (2018). *Smart Education — Huawei solutions*. [online] Available at: https://e.huawei.com/en/solutions/industries/education [Accessed 14 Oct. 2018].

33. Iec.ch. (2018). *IEC - White Paper > IoT 2020: Smart and secure IoT platform*. [online] Available at: https://www.iec.ch/whitepaper/iotplatform/ [Accessed 4 Aug. 2018].

34. Internetsociety.org. (2018). [online] Available at: https://www.internetsociety.org/wp-content/uploads/2017/08/ISOC-IoT-Overview-20151221-en.pdf [Accessed 10 Jul. 2018].

35. IoT Agenda. (2018). *What is internet of things (IoT)? - Definition from WhatIs.com*. [online] Available at: https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT [Accessed 13 Jun. 2018].

36. Jha, S., Bilalovic, J., Jha, A., Patel, N. and Zhang, H. (2017). Renewable energy: Present research and future scope of Artificial Intelligence. *Renewable and Sustainable Energy Reviews*, 77, pp.297-317.

37. Jiang, H., Wang, K., Wang, Y., Gao, M. and Zhang, Y. (2016). Energy big data: A survey. *IEEE Access*, 4, pp.3844-3861.

38. Jun, M. (2018). Blockchain government - a next form of infrastructure for the twenty-first century. *Journal of Open Innovation: Technology, Market, and Complexity*, 4(1).

39. Kai, C., Li, H., Xu, L., Li, Y. and Jiang, T. (2018). Energy-Efficient Device-to-Device Communications for Green Smart Cities. *IEEE Transactions on Industrial Informatics*, 14(4), pp.1542-1551.

40. Karim, M., Currie, J. and Lie, T. (2018). A machine learning based optimized energy dispatching scheme for restoring a hybrid microgrid. *Electric Power Systems Research*, 155, pp.206-215.

41. Khorsheed, E. (2018). Long-term energy peak load forecasting models: A hybrid statistical approach. *2018 Advances in Science and Engineering Technology International Conferences (ASET)*.

42. Kinney, S. (2018). *Three smart transportation case studies*. [online] Enterprise IoT Insights. Available at: https://enterpriseiotinsights.com/20180126/transportation/three-smart-transportation-case-studies-tag17-tag99 [Accessed 10 Sep. 2018].

43. KPMG. (2018). *@gov – Data driven government*. [online] Available at: https://home.kpmg.com/au/en/home/insights/2017/04/gov-data-driven-government.html [Accessed 13 Nov. 2018].

44. KPMG. (2018). *Going beyond the data*. [online] Available at: https://home.kpmg.com/xx/en/home/about/leading-insights/going-beyond-the-data.html [Accessed 11 Oct. 2018].

45. Lämmel, P., Tcholtchev, N. and Schieferdecker, I. (2017). Enhancing Cloud based Data Platforms for Smart Cities with Authentication and Authorization Features. *Companion Proceedings of the10th International Conference on Utility and Cloud Computing - UCC '17 Companion*.

46. Li, D., Cao, J. and Yao, Y. (2015). Big data in smart cities. *Science China Information Sciences*, 58(10), pp.1-12.

47. Lorica, B. (2018). *How intelligent data platforms are powering smart cities*. [online] O'Reilly Media. Available at: https://www.oreilly.com/ideas/how-intelligent-data-platforms-are-powering-smart-cities [Accessed 10 Dec. 2018].

48. Mahapatra, C., Moharana, A. and Leung, V. (2017). Energy Management in Smart Cities Based on Internet of Things: Peak Demand Reduction and Energy Savings. *Sensors*, 17(12), p.2812.

49. Mandic-Lukic, J., Milinkovic, B. and Simic, N. (2016). Communication Solutions for Smart Grids, Smart Cities and Smart Buildings. *Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2016)*.

50. Manjili, Y., Vega, R. and Jamshidi, M. (2018). Data-Analytic-Based Adaptive Solar Energy Forecasting Framework. *IEEE Systems Journal*, 12(1), pp.285-296.

51. Mckinsey.com. (2018). [online] Available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/The%20Internet%20of%20Things%20The%20value%20of%20digitizing%20the%20physical%20world/The-Internet-of-things-Mapping-the-value-beyond-the-hype.ashx [Accessed 12 Aug. 2018].

52. Mckinsey.com. (2018). [online] Available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/Analytics%20comes%20of%20age/Analytics-comes-of-age.ashx [Accessed 17 Oct. 2018].

53. Mckinsey.com. (2018). [online] Available at: https://www.mckinsey.com/~/media/McKinsey/Industries/Capital%20Projects%20and%20Infrastructure/Our%20Insights/Voices%20on%20Infrastructure%20Turning%20the%20smart%20city%20opportunity%20into%20reality/Voices-December-2017-WEB.ashx [Accessed 29 Sep. 2018].

54. Mckinsey.com. (2018). [online] Available at: https://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx [Accessed 10 Nov. 2018].

55. New Horizons for a Data-Driven Economy. (2016). .

56. Noor, S., Guo, M., van Dam, K., Shah, N. and Wang, X. (2018). Energy Demand Side Management with supply constraints: Game theoretic Approach. *Energy Procedia*, 145, pp.368-373.

57. Noor, S., Yang, W., Guo, M., van Dam, K. and Wang, X. (2018). Energy Demand Side Management within micro-grid networks enhanced by blockchain. *Applied Energy*, 228, pp.1385-1398.

58. Nvlpubs.nist.gov. (2018). [online] Available at: https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8202.pdf [Accessed 16 Oct. 2018].

59. Ourenergypolicy.org. (2018). [online] Available at: http://www.ourenergypolicy.org/wp-content/uploads/2017/08/SparkCognition_Artificial_Intelligence_and_the_Internet_of_Energy__IoE_.pdf [Accessed 10 Aug. 2018].

60. Pages.nist.gov. (2018). [online] Available at: https://pages.nist.gov/GCTC/uploads/blueprints/20170824-City-Platform-Supercluster-Report-FINAL.pdf [Accessed 16 Nov. 2018].

61. Park, L., Lee, S. and Chang, H. (2018). A Sustainable Home Energy Prosumer-Chain Methodology with Energy Tags over the Blockchain. *Sustainability*, 10(3), p.658.

62. Pop, C., Cioara, T., Antal, M., Anghel, I., Salomie, I. and Bertoncini, M. (2018). Blockchain Based Decentralized Management of Demand Response Programs in Smart Energy Grids. *Sensors*, 18(2), p.162.

63. PwC. (2018). *2018 Artificial Intelligence (AI) Predictions*. [online] Available at: https://www.pwc.com/us/en/services/consulting/library/artificial-intelligence-predictions.html [Accessed 23 Nov. 2018].

64. ReadWrite. (2018). *How the IoT and Related Tech Are Helping to Update the Energy Sector - ReadWrite*. [online] Available at: https://readwrite.com/2018/02/28/iot-related-tech-helping-update-energy-sector/ [Accessed 10 Jun. 2018].

65. Role of Big Data and Analytics in Smart Cities. (2016). *International Journal of Science and Research (IJSR)*, 5(2), pp.12-23.

66. S A, B. and Umamakeswari, A. (2018). Role of Blockchain in the Internet-of-Things (Iot). *International Journal of Engineering & Technology*, 7(2.24), p.109.

67. Saleem, Y., Crespi, N., Rehmani, M. and Copeland, R. (2018). *Internet of Things-aided Smart Grid: Technologies, Architectures, Applications, Prototypes, and Future Research Directions*. [online] Arxiv.org. Available at: https://arxiv.org/abs/1704.08977v1 [Accessed 1 Sep. 2018].

68. Santana, E., Chaves, A., Gerosa, M., Kon, F. and Milojicic, D. (2017). Software Platforms for Smart Cities. *ACM Computing Surveys*, 50(6), pp.1-37.

69. Sethi, P. and Sarangi, S. (2017). Internet of Things: Architectures, Protocols, and Applications. *Journal of Electrical and Computer Engineering*, 2017, pp.1-25.

70. Smarra, F., Jain, A., de Rubeis, T., Ambrosini, D., D'Innocenzo, A. and Mangharam, R. (2018). Data-driven model predictive control using random forests for building energy optimization and climate control. *Applied Energy*, 226, pp.1252-1272.

71. Smart cities readiness guide. (2014). [Redmond (Estats Units d'Amèrica): Smart Cities Council.

72. Stokab.se. (2018). [online] Available at: https://www.stokab.se/Documents/Nyheter%20bilagor/SmartCityInfraEn.pdf [Accessed 1 Nov. 2018].

73. Terroso-Saenz, F., Gonzalez-Vidal, A., Cuenca-Jara, J. and Skarmeta, A. (2017). An open architecture for IoT data analytics in smart cities. *2017 Global Wireless Summit (GWS)*.

74. Things, I. (2018). *Cisco IoT Solutions | Internet of Things Services*. [online] Cisco. Available at: https://www.cisco.com/c/en/us/solutions/internet-of-things/overview.html#~stickynav=1 [Accessed 9 Jun. 2018].

75. Unleashingit.com. (2018). [online] Available at: http://www.unleashingit.com/securepower/assets/wind-river-internet-of-things-smart-buildings-and-homes-use-case.pdf [Accessed 4 Aug. 2018].

76. Watt-Logic. (2018). *Using artificial intelligence and machine learning to manage the electricity grids of the future - Watt-Logic*. [online] Available at: http://watt-logic.com/2017/09/28/ai-in-energy/ [Accessed 23 Aug. 2018].

77. Woyke, E. (2018). *A smarter smart city*. [online] MIT Technology Review. Available at: https://www.technologyreview.com/s/610249/a-smarter-smart-city/ [Accessed 15 Jun. 2018].

78. Wu, J. and Tran, N. (2018). Application of Blockchain Technology in Sustainable Energy Systems: An Overview. *Sustainability*, 10(9), p.3067.

79. Wu, S., Chen, T., Wu, Y. and Lytras, M. (2018). Smart Cities in Taiwan: A Perspective on Big Data Applications. *Sustainability*, 10(2), p.106.

80. Www2.deloitte.com. (2018). [online] Available at: https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/public-sector/deloitte-nl-ps-smart-cities-report.pdf [Accessed 1 Nov. 2018].

81. Yaqoob, I., Hashem, I., Mehmood, Y., Gani, A., Mokhtar, S. and Guizani, S. (2017). Enabling Communication Technologies for Smart Cities. *IEEE Communications Magazine*, 55(1), pp.112-120.

82. Zhao, S., Wang, B., Li, Y. and Li, Y. (2018). Integrated Energy Transaction Mechanisms Based on Blockchain Technology. *Energies*, 11(9), p.2412.

83. Zheng, Z., Xie, S., Dai, H., Chen, X. and Wang, H. (2017). An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. *2017 IEEE International Congress on Big Data (BigData Congress)*.

84. Zhou, Y. and Li, X. (2015). Vehicle to grid technology: A review. *2015 34th Chinese Control Conference (CCC)*.