Research Article

Cu Kim Long, Rashmi Agrawal, Ha Quoc Trung, and Hai Van Pham*

# A big data framework for E-Government in Industry 4.0

https://doi.org/10.1515/comp-2020-0191
received May 18, 2020; accepted July 08, 2020

**Abstract:** The next generation of E-Government and healthcare has the potential to increase the more intelligent governance with improvements in transparency, accountability, efficiency, and effectiveness. It enables organizations to use the benefits of information via big data analysis to settle the difficulties effectively. Big Data has emerged which plays a significant role in many sectors around the world. Global trends in taking advantage of the benefits from big data are considered with an overview of the US, European Union, and several developing countries. To deeply understand the utilization of big data in several domains, this study has presented a brief survey of key concepts (such as IoT-enabled data, blockchain-enabled data, and intelligent systems data) to deeply understand the utilization of big data in several domains. Our analysis sets out also the similarities and differences in these concepts. We have also surveyed state-of-the-art technologies including cloud computing, multi-cloud, webservice, and microservice which are used to exploit potential benefits of big data analytics. Furthermore, some typical big data frameworks are surveyed and a big data framework for E-Government is also proposed. Open research questions and challenges are highlighted (for researchers and developers) following our review. Our goal in presenting the novel concepts presented in this article is to promote creative ideas in the research endeavor to perform efficaciously next-generation E-Government in the context of Industry 4.0.

**Keywords:** big data, E-Government, big data analytics, big data framework

*** Corresponding author: Hai Van Pham,** School of Information and Communication Technology, Hanoi University of Science and Technology, Hanoi, Vietnam, e-mail: haipv@soict.hust.edu.vn
**Cu Kim Long:** School of Information and Communication Technology, Hanoi University of Science and Technology, Hanoi, Vietnam, e-mail: longck.2006@gmail.com
**Rashmi Agrawal:** Manav Rachna International Institute of Research and Studies, Faridabad, India, e-mail: drrashmiagrawal78@gmail.com
**Ha Quoc Trung:** Information Communication Center - MOST, Hanoi, Vietnam, e-mail: haquoctrung@gmail.com

## 1 Introduction

The digital transformation can be created by many huge databases. This is one of the major causes that led to the emergence of the "Big Data" concept, which plays a vital role in many sectors around the world to increase the efficiency and effectiveness. There have been rapid socio-technological developments in using the advancements of cloud, IoT, blockchain, and other technologies [1–3]. Moreover, the field of data science has seen exponential growth in data generation and data sharing [4,5]. Big data through digital transformation has contributed to the growth in the volume of data generated [6,7] such as increase in data processing capacity, lower cost of digital storage, affordable and faster communication technologies, proliferation of applications and smart devices, and so on.

Big Data is a significant concept in Industry 4.0 as it contributes to increase the efficiency and effectiveness of governance. Big data analytics (BDA) helps not only companies but also governments through many intelligent systems such as intelligent monitoring systems, intelligent decision-support systems, and intelligent prediction systems. In order to deeply understand big data for E-Government, this work presents a systematic survey of some key subjects such as open government data [8–10], IoT-enabled data [11], blockchain-enabled data [12], intelligent system data [13–15], and social network data. In Industry 4.0, there are some frontier technologies such as blockchain, artificial intelligence (AI), IoT, cloud computing, and BDA used by governments to improve intelligent governance along with transparency, accountability, and improved efficiency and effectiveness [16–23]. Therefore, some countries have studied state-of-the-art technologies to develop the next generation in healthcare community, education, finance of E-Government.

Implementing advanced technologies has resulted in a phenomenon termed Technological Determinism (TD) or Disruptive Innovation (DI) [24]. It relates to the manifestation of a time lag between the development and implementation of new technologies and an understanding of the underlying theory which develops following research into the application of new technologies. The result of the early adoption of new technologies is frequently manifested in socioeconomic and technological problems that are exacerbated in the online world. For example, social networking pushes recommendations based on browsing history and expressed preferences; this can result in subjective reinforcement which reinforces current views held by users without such views being questioned.

In terms of the survey, there are some research questions that need to be addressed: How is the level of the research paper growth per year associated with big data? Which database are works surveyed? Which of the top regions/countries are there research works related to big data as well as interested in investing, researching on big data? In the study presented in this article, we have reviewed 435 published papers (published in the period 2010–2019) sourced using indexing from the ScienceDirect, IEEE Xplore, and Google Scholar research databases. The publications show the key technologies including both proprietary and open-source systems. The search terms used address the key concepts and technologies, and our analysis shows a number of significant and relevant results. Figure 1 shows the number of papers per year from 2010 to 2019, as well as the number of papers by databases, continents, and leading countries.

- From 2010 to 2017, the number of articles grew regularly, from around 5 to 45; the number increased sharply with 80 documents in 2018 and reached a peak of 176 documents in 2019.
- The rate of documents related to Big Data and E-Government in the ScienceDirect, the IEEE Xplore, and Google Scholar databases, at 67, 11, and 22% respectively.
- The top three continents are Europe (187), Asia (149), and America (60). Europe is more around ten times than the remaining continents (Africa and Oceania). The leading countries active in the field are China, India, the US, UK, and Spain.

The literature reviewed has considered big data in Industry 4.0 with concepts, methodologies, supporting techniques, data sources/URLs, domains, and new contributions highlighted in Table 1. Furthermore, some typical big data frameworks (BDFs) are surveyed and a Big Data Framework for E-Government (BDFEG) is also proposed.

The rest of the study is structured as follows: Some key concepts are shown in Section 2. Research background and international effects are set out in Sections 3 and 4, respectively. In Section 5, we introduce a BDF for E-Government and consider open research questions and potential future works. The article closes with Section 6, where we present concluding observations.

# 2 Key concepts

Currently, a number of huge datasets are created at an unprecedented rate from many different sources. The data traffic is driven by a number of technological trends including IoT, cloud computing, blockchain, and the spread of smart devices [17,18]. For instance, there are many powerful systems and distributed applications that are applied using new technologies in various domains such as smart grid systems [19,20], healthcare systems [21], retailing systems [22], and government systems [23].

Big data performs a crucial role in intelligent governance. The new technologies and products have a potentially "game-changing" impact on intelligent governance models and methods. However, as discussed in Section 1, while such technologies may be "game-changing," there is clearly potential for TD [24] which represents a significant issue for all stakeholders in governmental and commercial information systems.

In this section, we present three major factors that generate an amount of huge data in the context of Industry 4.0: IoT-enabled data, blockchain-enabled data, and intelligent system data.

The survey is carried out by six steps. First, the keywords and terms "Big Data" AND "E-Government" are defined. Second, they are searched in the ScienceDirect, IEEE Xplore, and Google Scholar research databases. Third, the contents are classified by the papers according to features such as database, year, region, and country. In summary of big data, the Pivot Table Report function in Microsoft Excel was used to analyze the collected data. Finally, the results are shown in Figure 1.

## 2.1 IoT-enabled data

Networked sensors (also known as smart sensors which may produce raw data or data which have received some pre-processing) accessed using the Internet in cloud-based systems including fog and edge computing (generally known as the IoT) are the source of captured data in smart environment (SE), the data being used for "real-time" analysis and big-data analytics. SE is diverse in both geographical terms (e.g., smart homes and smart

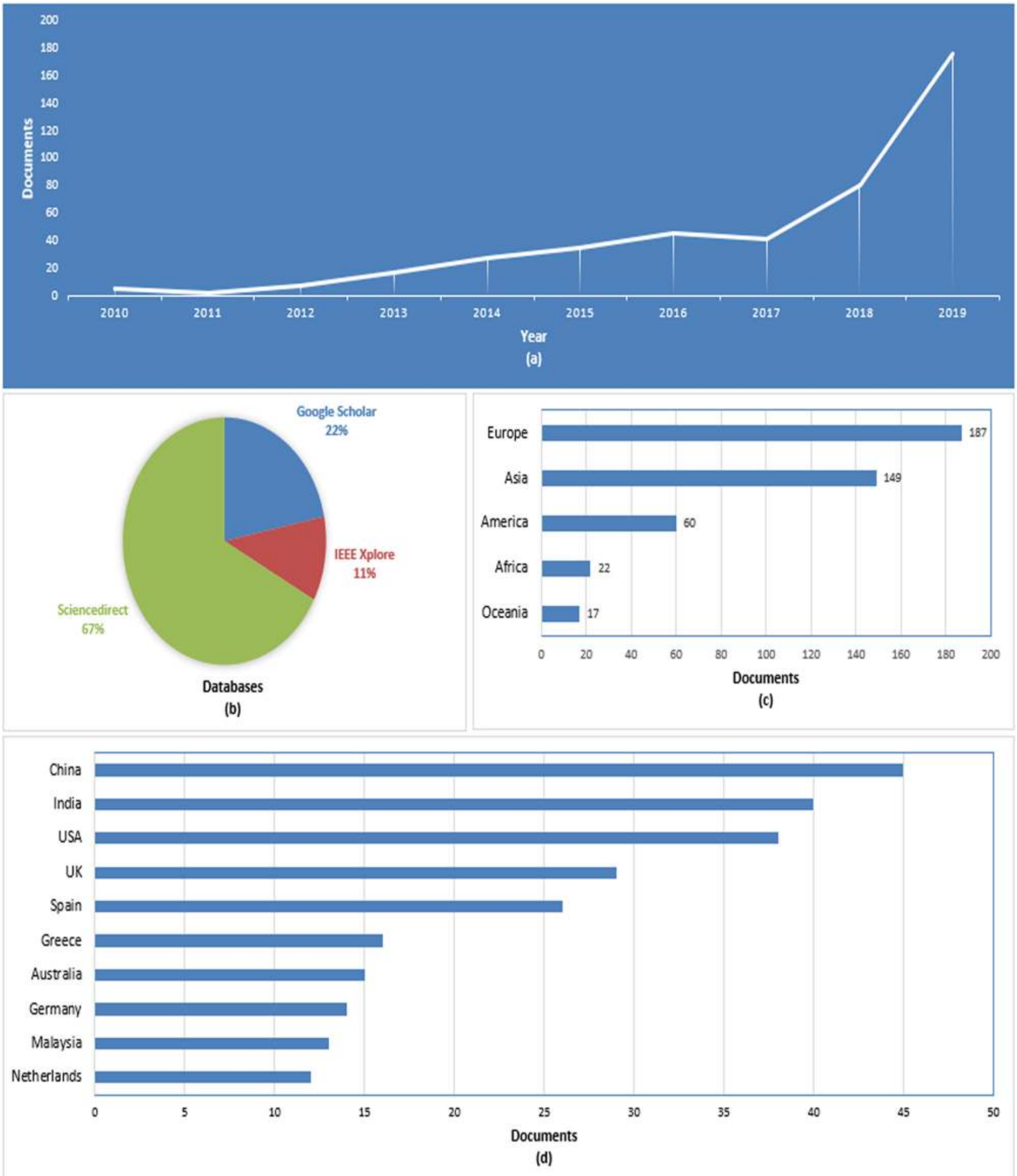


**Figure 1:** Statistics from the ScienceDirect, the IEEE Xplore, and Google Scholar databases (search keywords: "Big Data" AND "E-Government;" Date: 26 April 2020). (a) The number of papers per year; (b) the number of papers by databases; (c) the number of papers by continents; and (d) the number of papers by top countries.

cities [40]) and provides the capability to manage systems in many domains and applications.

The IoT is a primary source of data for big-data analytic solutions [25] along with social networking or

**Table 1:** Comparison and analysis of key concepts

| Concepts | Methodologies | Supporting techniques | Data sources/URL | Domains | New contributions | Refs. |
|---|---|---|---|---|---|---|
| *IoT-enabled data* | • IoT<br>• 5G network<br>• Advanced sensors<br>• BDA<br>• Cloud computing<br>• Security<br>• Storage | • Sensors, wireless adapters, and GPS<br>• Smart device<br>• Gateways and networks technology<br>• Data and signs processing technology<br>• Management service<br>• Cloud computing<br>• Software and algorithms | • Linked Sensor Data<br>• CIC datasets<br>• UCI datasets | • E-Government<br>• Smart cities<br>• Health<br>• Transportation<br>• Agriculture<br>• Environment<br>• Fishery<br>• Manufacturing | • IoT architecture<br>• IoT applications<br>• A coherent IoT ecosystem<br>• The Open IoT platform<br>• Integration AI into IoT system<br>• Data lifecycle<br>• Big data architecture<br>• Big Data technologies for the all IoT domains | [37–43,48–52,54–56,90–100] |
| *Blockchain-enabled data* | • Decentralized techniques<br>• Searching<br>• Cloud computing<br>• Storage | • Peer-to-peer networks<br>• Decentralized Technique<br>• Testing techniques | • Blockchain.info<br>• Bitcoincharts.com<br>• Blockr.io<br>• Bitnodes.earn.com<br>• Blocks.wizb.it<br>• Coinbase.com<br>• Coinmarketcap.com<br>• Coinmap.org<br>• UCI datasets | • Finance<br>• Governments<br>• Healthcare<br>• Real-estate<br>• Voting | • The blockchain's key functional characteristics<br>• Classification blockchain applications<br>• Security of a bitcoin transaction<br>• Smart contract applications<br>• Blockchain-Based Distributed Architecture<br>• Decentralized Security Framework | [3,12,18,57–60,64,67–69,101,102] |

(*Continued*)

**Table 1:** (*continued*)

| Concepts | Methodologies | Supporting techniques | Data sources/URL | Domains | New contributions | Refs. |
|---|---|---|---|---|---|---|
| *Intelligent systems data* | • Statistics<br>• AI<br>• Data mining<br>• Machine learning<br>• Deep learning | • Statistics techniques<br>• Data mining techniques<br>• Machine learning techniques<br>• Deep learning techniques<br>• AI-based techniques | • UCI datasets<br>• Data.gov<br>• Data.gov.uk<br>• Data.gov.in Healthdata.gov<br>• Amazon Web Services public datasets | • Transportation<br>• Smart cities<br>• Agriculture<br>• Manufacturing<br>• Supply chain<br>• Emergency<br>• Healthcare<br>• Environment<br>• Energy<br>• Weather<br>• Education<br>• Finance and stock trading | • The contribution of a data source<br>• The intelligent decision models<br>• The knowledge-based analytics, predictive and consultant systems | [14,21,26,34–39,41,47,50, 58,64,67,76,80–81,103] |

blockchain. Given their diversity, applications using data from the IoT are undergoing continuous development in many fields [26] such as transportation, agriculture, intelligent manufacturing, smart city, and digital government. In fact, there are many IoT-enabled applications used in transportation and logistical delivery companies [27] where vehicle locations and movement can be tracked [28] by methods such as sensors, wireless adapters, and GPS [29–32]. This helps organizations optimize delivery routes as well as manage employees effectively through data-driven systems [33].

In agriculture [34,35], IoT-based systems help farmers become more intelligent. For instance, in IoT-enabled smart farming [36], a system is built for monitoring the light, humidity, temperature, soil moisture, etc. and automating the irrigation system. IoT-based farms prove to be more efficient compared with the traditional ones. The IoT-based applications are applied for data acquisition and sharing, which will significantly impact the agriculture sectors in the near future [37–39].

In an SE, the integration of all the IoT devices in the city can play a vital role in developing smarter city services by digitization [41]. IoT-based applications are also used by governments and public organizations [42–47] to provide public services better for citizens and businesses. In such applications, sensors are deployed at various locations, in objects, and in monitoring systems [48–53]; examples include "real-time:" smart homes, smart parking systems, vehicle networking, surveillance, environmental monitoring (e.g., air quality and weather), and water monitoring.

The impact of the IoT on the Internet is generating existential change with the prediction of 75 billion IoT connections by 2020 [54,55] and 100 billion connected IoT devices by 2025 [56]. It can be seen that the exponential increase in the development and use of IoT applications with the ubiquity of smart devices is driving data generation in a large range of heterogeneous fields and applications.

## 2.2 Blockchain-enabled data

Blockchain is considered a frontier technology and its cryptocurrency applications have provided the basis for bitcoin and Ethereum. In practice, Ethereum (https://www.ethereum.org/) offers distributed and transparent transactions [57]. However, most significantly, blockchain offers novel solutions to help users control actively their transactions.

Blockchain is considered a frontier technology in Industry 4.0 that has attracted attention of stakeholders in many different sectors such as real estate, finance, health, government agencies, and research institutions [58–60]. A huge data volume has been generated from blockchain-based applications in recent years.

Bitcoin is considered a typical case. Blockchain can store the system's transactions [61] in a peer-to-peer and decentralized method. Case studies in the financial domain have shown that blockchain-based applications and platforms in the banking sector reach better results both secure issue and processing speed [58,62].

The benefits of blockchain technology have also attracted the attention of governments, leading to innovation and transformation of government services and processes. Some governments (e.g., US, UK, Singapore, Estonia, Australia, and so forth) have taken advantage of the potential of blockchain to improve public service quality [63–68]. Blockchain is also considered as a potential technology to urge the online services of governments in the future because of enabling secure data sharing and integrity [69].

## 2.3 Intelligent systems data

Intelligent systems are systems that can collect, process, and explore data and exchange them with other ones [70,71]. They can learn themselves during the operation process [72–74] such as chatbots, IBM Watson, and AlphaGo [75–89]. They are used in Information Systems (IS) which are used by both individuals and organizations and are an important component in computerized systems employed in a wide range of fields and applications implemented both on-line and off-line [70,71]. The traditional approach processes data into information without selection, which results in potentially irrelevant information.

Checkland et al. [24] have introduced an intermediate stage termed CAPTA by extending the process. The use of data selection in the CAPTA process is particularly relevant to intelligent systems in SE using the IoT where relevance forms a central role. Moreover, there is the recognized potential issue of latency in "real-time" cloud-based (including fog and edge computing) systems. Intelligent systems must accommodate the principles espoused in CAPTA and accommodate scaled latency capabilities.

## 2.4 Comparative analysis

The aforementioned concepts are very important for the next-generation E-Government as BDA solutions will have a significant impact on the activities of governments in all sectors. To deeply understand the aforementioned ideas, Table 1 shows a comparison of aspects such as methodologies, supporting techniques, data sources, domains, and new contributions.

To recap, as mentioned in Sections 1 and 2, we realized that the huge data sizes are created at an unprecedented rate from sources in many different domains (e.g., transportation, environment, weather, agriculture, energy, tourism, education, healthcare, finance, government, manufacturing, and smart city). To meet the need for storage in the future, there are some open-source products that are used ubiquitously, feature-rich, useful and are applied effectively in the private sector. Besides, BDA provides a basis upon which government agencies can analyze past and current data to identify trends in the data which are then applied to plan and manage services and future projects.

# 3 Research background

## 3.1 Storage

Cloud computing has additional advantages such as cost-efficiency, safety, and so on [104,105]. However, it also has several disadvantages such as data security in the cloud, legal issue, and so forth [106–108]. Despite the various major disadvantages, as mentioned, cloud computing has been pervasively used in some sectors (business, healthcare, manufacturing, education, transportation, etc.) and in nations (US, UK, France, Japan, China, Iran, etc.). It also manifested increasing competitiveness through optimal resource utilization, effective ubiquitous data usage, and so forth [109–111].

Furthermore, multi-cloud has advanced in recent years, and new opportunities and capabilities emerge with relation to web services and microservice. In fact, cloud computing has facilitated to deploy web services. This leads to the advent of new platforms and operating systems. It also sparked the large utilization of applications based on mobile phones or portable devices. Besides, microservice, a variant of the service-oriented architecture (SOA), is a technique to develop applications as a combination of loosely coupled services. Microservice is also known as a specialization of an implementation approach for SOA used to build flexible, independently deployable software systems. Multi-cloud-based service will be one of the hot trends to deploy intelligent systems not only in the private sector but also in the public sector in the near future.

## 3.2 Frontier technologies in Industry 4.0

In the context of Industry 4.0, there are many state-of-the-art technologies that have contributed to increasing huge amounts of data in the past few years, especially IoT, blockchain, and social networks.

First, the IoT has evolved from the convergence of key technologies such as wireless technologies or data analytics technologies [112]. With the advent of smart cities, the IoT is considered a very important technology. It can impact every aspect of our life [113]. However, the security problem is considered the biggest disadvantage of IoT devices because they are attacked easily compared with other devices such as computers or portable devices [114–117].

Second, blockchain promises to be a frontier technology applied in many domains [118]. Currently, there are many cryptocurrency applications in the coin market capitalization, approximately more than 2,200 cryptocurrencies (https://coinmarketcap.com/ as accessed on 26 April 2020). This concept was introduced by Satoshi Nakamoto in 2008, which became an interesting topic nowadays [119]. To deeply understand blockchain, these studies have investigated several main concepts (blockchain, distributed, trusted computing, smart contracts, and proof of work), as well as private, public, and hybrid blockchains [120,121].

Finally, social networks, as well as portable devices, are more and more ubiquitous [122] such as Facebook, Myspace, YouTube, WhatsApp, Google+, Instagram, Twitter, WeChat, and so on. They also create huge datasets and contribute to supporting governments to collect and analyze data. However, government agencies should consider strictly before using social networks their online social network analysis drawbacks and ensure user's information privacy and security [123,124].

## 3.3 BDA

BDA is the core technology for using and developing IoT and AI applications. For instance, smart technology products can recognize handwriting, speech, chat, and so on. Hence, BDA will contribute to making more intelligent systems in the future such as AI systems. They can learn themselves via interaction with humans or other systems by using some intelligent techniques such as machine learning, deep learning, and so forth. However, BDA is still challenging because of the existing major limitations such as the requirements of the capacity of processing computers to be very strong, the novel tools for storing and processing large and unstructured databases must be studied continuously, as well as humans' need for highly accurate prediction systems is more and more increasing [127–129].

BDA is also considered one of the most important technologies in Industry 4.0. It is an integral essential part of creating intelligent systems in Industry 4.0. There are four types of BDA: (i) descriptive data analysis (describes what happened in the past based on graphical reports, images, but does not explain why and predict what will happen); (ii) analyze diagnostic data (explain what happened in the past); (iii) predictive data analysis (predict what can happen); and (iv) prescriptive data analysis (the predictions or recommendations are given based on the large datasets processing results). BDA has a significant impact on many sectors [130], and some major solutions are applied to solve the problems related to BDA including machine learning, deep learning, granular computing, and so on [125,126,131,133].

To sum up, with the development of the cloud computing, especially multi-cloud in recent years, new opportunities and capabilities emerge with relation to webservice and microservice. Moreover, the cutting-edge technologies have contributed to increasing colossal databases, especially IoT, blockchain, and social networks. One of the issues is how to take advantage of the benefits of big data for government agencies to support smart governance? It is understood that it is very necessary to propose a novel methodology about big data to help government agencies to undertake effectively big data projects as well as to understand clearly the steps of data processing for intelligent systems.

# 4 International effects

This section presents a synopsis of the major ongoing big data plans, strategies, as well as projects in some countries or regions in the world.

## 4.1 The United States

Big data research and development initiative was given to solve complicated problems faced by the US government in 2012 [132]. It included 84 programs carried out by six different departments. The biggest data center of the US was completed in 2015 in Utah with a budget of $1.5 billion. The storage capacity was estimated to be approximately a few exabytes [134,135]. Particularly, a federal big data strategic plan was built based on the initiative with seven key areas related to big data research and development [136]. In 2016, once again history confirmed that BDA played an important role in D. J. Trump's victory in the US presidential election campaign (the first one in 2012 in B. Obama's victory).

## 4.2 The European Union (EU)

Industry 4.0 is a new concept that was introduced in 2013 by Germany. Then, the EU also approved the Horizon 2020 program in 2014 [137–139] with funding of approximately €80 billion. Under Horizon 2020, there are many major problems to solve which have big data research and development. The aim is to increase using large data in the public sector in Europe. However, it requires the scalability of data analysis, sample discovery, real-time applications, and the sharing and integration of data from the public sector. Besides, built-in security and privacy mechanisms in big data applications are still limited. Developing predictive analytics as well as modeling and simulation tools for historical data analysis are the main challenges that need to be addressed to use big data effectively.

## 4.3 Other countries

In China, big data is known widely through the China International Big Data Industry Expo (China's Big Data Valley), especially some leading enterprises such as Huawei, Alibaba, Tencent, and Baidu. The strategic plans were given to aim for supporting digital transformation in China like "Made in China 2025" and other similar documents. The results of these strategic plans show that China has gained important achievements in big data collection and analysis to contribute to solving big problems of nation [140,141]. For instance, the Integrated Joint Operations Platform is applied to manage the population in China [142].

In India, many policies and guidelines have been promulgated by NDSAP in recent years such as India's open data policy, government open data license, and open data implementation guidelines [143,179]. Besides, the Government of India takes advantage of the achievements of big data to undertake a lot of big projects such as Aadhaar, Ettal, Ebhasha, DigiLocker, and so on [144–146]. According to Dataquest, the big data industry is currently estimated to be Rs 17.615 crore (equal to ten million) in revenues and is predicted to reach Rs 130.000 crore by 2025 [147].

In Nigeria, ICT is applied to improve the governance process at different levels. Most of these data collected by these organizations are structured in nature and very suitable for current governance services. For example, big data is used in the telecommunication sector to manage customers as well as to provide services efficiently. A critical component of the decision-making process is information derived from the aggregation of data from various sources. Hence, the government effectively explores the huge opportunities in BDA and runs effective E-Government following some guidelines such as a national master plan for E-Government, a national BDA framework, and so forth [148–151].

In Vietnam, the government has had activities to enhance awareness about the benefits of emerging technologies for some years ago such as blockchain, IoT, AI, big data, and so on. Some policies have promulgated in recent years to stimulate research and application of state-of-the-art technologies within IS for not only effective task management in internal government agencies but also serving businesses and citizens better such as The Directive to Enhance the Capacity to Access the Fourth Industrial Revolution, The Artificial Intelligence Development Plan in Vietnam by 2025, and so on. Besides, in the private sector, some of the leading technology enterprises in Vietnam have initially exploited the benefits of cutting-edge technologies to create competitive advantages. For example, while VNPT has the IoT ecosystem and FPT does the online movie and advertising service, VinTech concentrates on researching the AI solutions [152–154].

# 5 A BDFEG

## 5.1 Typical BDFs

In this section, some typical BDFs or architectures will be presented briefly via some attributes including frameworks, time, and components/characteristics. Most of the researchers highlight the necessity of the BDF for reducing the risk when the implementation of big data projects and propose similar BDFs as shown in Table 2.

## 5.2 E-Government in the context of Industry 4.0

In previous studies [165], E-Government is the government's strong application of the ICT's achievements in daily tasks to improve efficiency, while enhancing transparency to serve citizens and businesses better. Particularly, data-driven governance is more and more important for decision-making timely as well as prediction accurately [166,167].

In Industry 4.0, there is a convergence of many frontier technologies, such as blockchain, AI, IoT, supercomputing, machine learning, and deep learning. These technologies promise to bring big opportunities for organizations in various domains such as finance, bank, health, environment, transportation, E-Government, smart cities, and so on. The rapid development of these technologies will help governments not only evolving E-Government effectively but also development sustainably [168–179].

Industry 4.0 is considered a strong impact on the world economy. Technologies will be applied more and more in operation and production such as information technology, smart applications, self-operating systems, and so on. This will lead to some positive results such as increasing productivity, forming new industries, reducing the number of employees, and so on. However, it also requires organizations and businesses to increase initial costs, requires human resources to be more qualified and skills, and so on. Therefore, the governments need also to have the appropriate way to develop the new IS effectively following the trend of Industry 4.0 [180–190].

## 5.3 Proposal for a BDF for E-Government

As analyzed in the aforementioned sections, we realize that it is very necessary to propose a BDF for E-Government in the context of Industry 4.0. To reduce the risk of big data projects as well as to help government agencies understand clearly about the steps of big data implementation to take advantage of its benefits for effective governance, we are based on mostly components referred from reputed frameworks of several countries which did it successfully in the E-Government, and we propose a three-layer framework of Big Data for E-Government, so-called BDFEG. The proposed BDFEG is shown in Figure 2 and its layers and components are recapitulated as follows.

*The platform layer* includes some components such as servers, operating systems, networking systems, storage systems, user interface, and so on.

**Table 2:** Some typical BDFs or architectures

| BDFs | Time | Components/characteristics | Refs. |
|---|---|---|---|
| BDF1 | 2016 and 2017 | ISO 8000 standard: planning, process, data, stakeholders, technology, implementation and evaluation, conformance to the quality | [155,156] |
| BDF2 | 2018 | The framework consists of eight components: organization, stakeholders, scope, policies and standards, optimization, quality, storage, communication and data management | [157] |
| BDF3 | 2018 | The framework has four layers: sources, storage, security and privacy, and applications | [158] |
| BDF4 | 2018 | The Big Data Governance Framework encompasses four main layers: objective, strategy, components, and IT infrastructure | [159] |
| BDF5 | 2019 | This framework includes five main components: planning, organization, operation, implementation, and monitoring | [160] |
| BDF6 | 2019 | The system architecture contains several key components: information consolidation, service registration, storage user profiles, and provide a user interface | [161] |
| BDF7 | 2019 | The architecture framework in real time includes six layers: data, data collection, data storage, data processing, result storage, and visual display | [162] |
| BDF8 | 2019 | The framework consists of three domains: drive domain, support domain, and capability domain | [163] |
| BDF9 | 2019 | The framework comprises three layers: operational platform, security, and data processing | [164] |

*The data layer* has three main components including data collection and governance registration, data preprocessing and classification, and data analytics. This layer provides two methods for processing data, which include online and batch. The key components in the data layer are briefly presented in the following:

- First, intelligent techniques include advanced data processing techniques (data mining, machine learning, deep learning, and so on) as well as platforms or architectures (web services, microservices). They will be used to develop intelligent systems in the future.
- Next, intelligent systems will be built after analyzing data as well as using intelligent techniques. They comprise the decision-making support system, prediction support system, recommendation system, expert systems, and other similar systems.
- In addition, this component also provides two kinds of API (Application Program Interface). The API government open data service supports government agencies to develop intelligent systems, while the API open data service assists in health care, hospitals, businesses, universities, developers, and so on.

*The application and security layer* consists of safety ensuring solutions as well as systems of applications that will be clearer during the physical implementation of the BDFEG. Particularly, the most important component in this layer is intelligent techniques and systems that it refers to significant parts after data analysis.

- First, data collection and governance registration capture data from the data sources (such as internal databases, websites/portal system, online public service system, statistical system, financial system, social networks, and so forth).
- Second, data preprocessing and classification will clean the data gathered by the previous step. This component will be responsible for converting and classifying input data into analysis-ready formats.
- Finally, data analytics includes two types of repositories (data zone and batch data repository). The data zone is used to store data that are gathered from online systems (websites/portals, online public services, social networks, and so on), while the batch data repository is used to store datasets in DBMSs (e.g., Hadoop HDFS, NoSQL, and so forth). The component also comprises three sorts of analytics such as online analytics, real-time analytics, and batch analytics. This component also comprises three sorts of data analytics (such as online analytics, real-time analytics, and batch analytics). Both online analytics and real-time analytics are usually used to handle new data that have just been generated. The latency of online analytics usually less than a defined value (calculated by minutes or seconds), while this of real-time analytics is very small (millisecond). Batch analytics is usually utilized for applications that have high latency (more than 20 minutes) or calculated results base on the huge data volume (TB or even PB).

## 5.4 A case study for validation

In this subsection, we have presented a case study at the Ministry of Science and Technology in Vietnam (MOST) to validate the BDFEG. A diagram is shown in Figure 3 that can be described by the following five main steps:

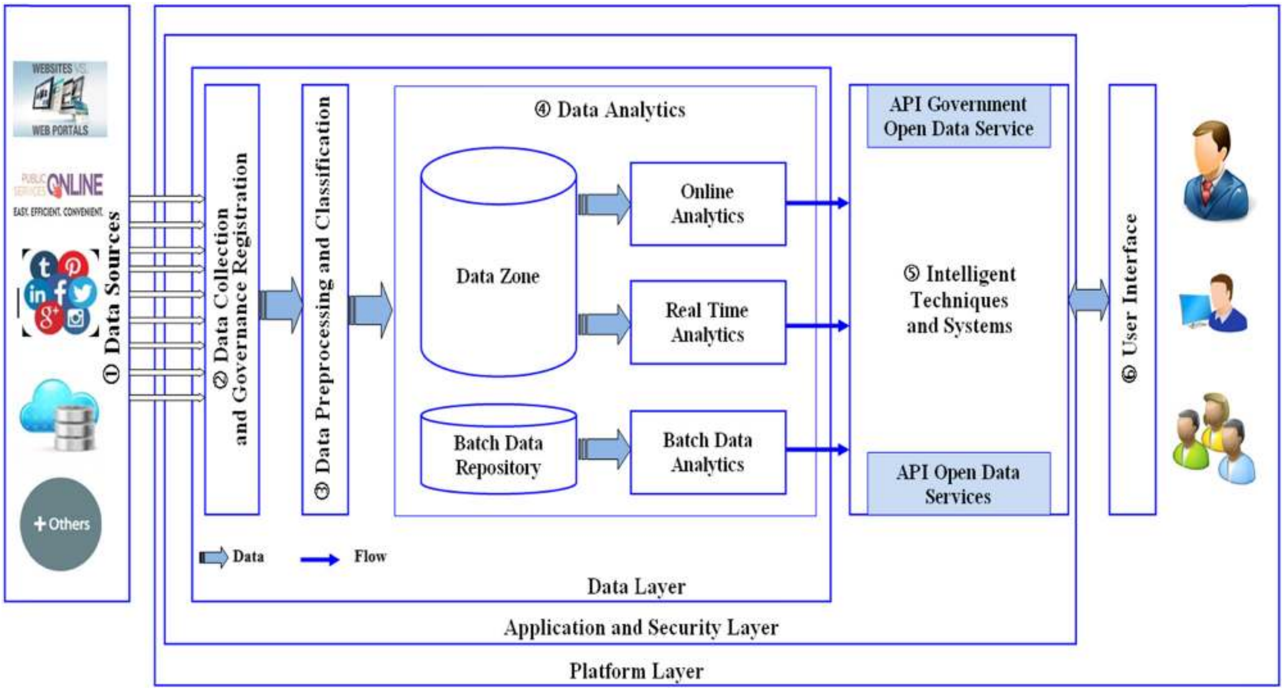


**Figure 2:** A BDFEG.

**Step 1:** The platform is a data center that includes types of devices such as servers, networking devices, operating systems, storage devices, and so on.

**Step 2:** Data sources are gathered from two main sources. First, stream data sources are collected from social networking such as Facebook. Second, historical data sources are aggregated from the existing databases of the units under the MOST like the web or portal databases, online public services databases, human resources management databases, etc.

**Step 3:** The security includes the measures to avoid the unauthorized intrusion from outside. The key measures are used to grant the safety of the system that include the restriction sign-on access to the system (firewalls, intrusion detection system, intrusion prevention system, etc.), the multi-level user authentication (database, application, networking, operations, and monitoring levels), and so on.

**Step 4:** The data processing layer has some key components such as data collection and classification, data storage, and data analytics. This layer is able to gather stream data sources (e.g., Facebook) and batch data from historical databases of the units under the MOST. At this step, Flume and Kafka are used to collect, aggregate, and ingest data to Hadoop. Sqoop is utilized to transfer big data between traditional databases (Microsoft SQL, MySQL) into the Hadoop Distributed File System (HDFS), and inversely. The HDFS, NoSQL databases, and SQL databases are used to store data after collecting and classifying. Then, Storm and Spark are applied to analyze and extract knowledge. Storm can perform both the streaming data and batch data. Spark can process batch data via Spark streaming.

**Step 5:** Intelligent Techniques and Systems include many systems based on the data after collection, classification, and analysis. They are using those used in the case study at the MOST shown in Figure 3 such as Portal MOST, Documents and Tasks Management System, Online Public Services System, Planning and Finance System, R&D Tasks Management System, HRM System, Electronic One-Stop-Shop System, Complaints and Denunciations System, Reporting Statistics System, and many others. Besides, this component also provides two kinds of APIs. The API for government agencies is applied in the systems (Documents and Tasks Management System, Online Public Services System) to exchange most of the documents between the MOST and other ministries as well as local governments. The API for researchers and developers is used in the R&D Tasks Management System to support the academic community and businesses. Until now, this system provides users more than 260,000 papers in local journals, 34,000 R&D projects, and 40 million documents in international R&D databases.

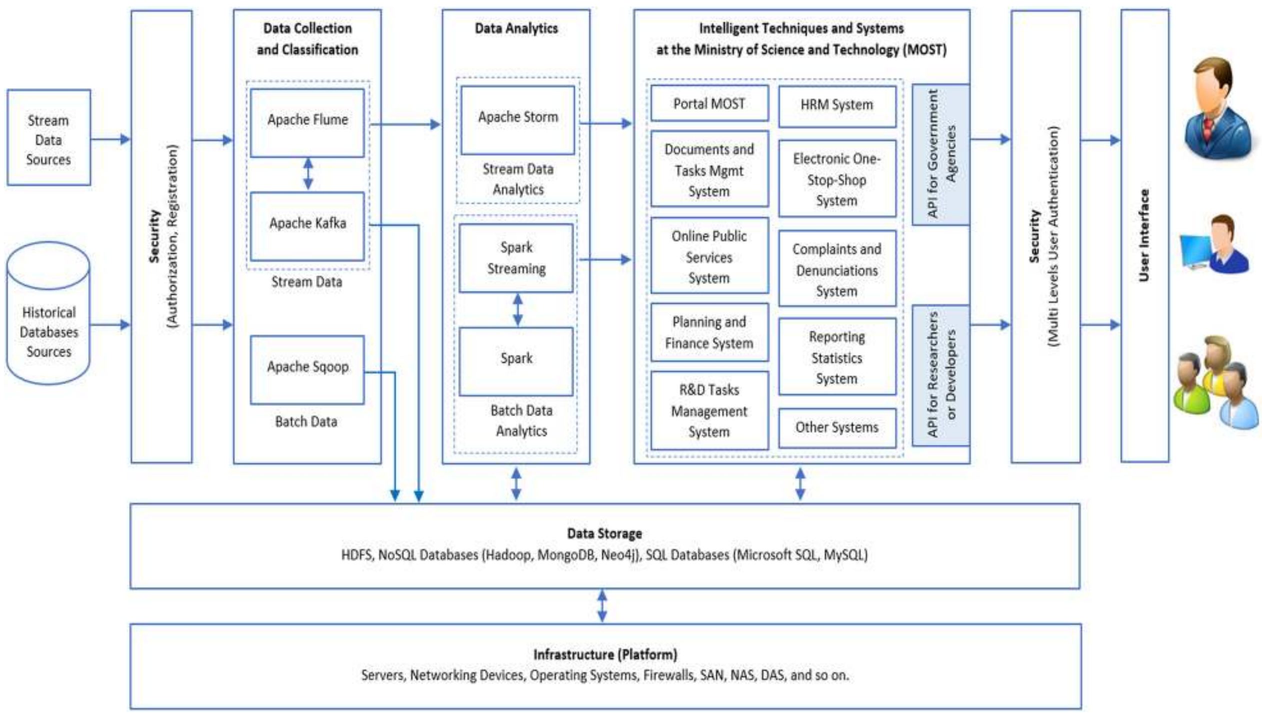


**Figure 3:** A case study at the MOST in Vietnam.

## 6 Conclusions and future works

BDA solutions have been the subject of research and application over many years. Initially, the analytics were carried out using large datasets in traditional relational databases to find emergent properties. The advent of the frontier technologies in Industry 4.0 has driven the concept of big data into a new research area of extremely large and rapidly growing datasets. Using big data effectively is considered a central component in the context of Industry 4.0 in both public and private sectors worldwide and is an important factor for developing the next generation E-Government. Such solutions will provide benefits for governments that can plan policies and national programs with analytics implemented using intelligent techniques and systems.

In this study, we have considered the key concepts, frontier technologies, and BDFs. We have proposed a three-layer framework for Big Data in the context of Industry 4.0. Although the BDFEG gives only at the concept level, it provides a new approach for governments to deploy big data projects. Future works are also emphasized to inspire researchers and developers and are to promote new ideas in the research effort to perform efficaciously next-generation E-Government in the context of Industry 4.0.

As mentioned earlier, many BDFs have been published, but there is no BDF for E-Government. In fact, a BDF for E-Government must be studied to reduce risks for the big data projects of government agencies in the near future. This study proposed a BDF for E-Government, so-called BDFEG. However, the BDFEG is only given at the concept level, and it still needs further research attention to complete in the future works. Some key open research questions and challenges are highlighted (for researchers and developers) following our review as follows:

- The first question is to select the big data platform for the government's projects. Many researchers and developers usually process large datasets on Apache Hadoop. It is considered a wonderful replacement for DBMS that was introduced decades ago. However, one question is given whether the Apache Spark is the alternative of the Apache Hadoop in the future. That is a question that needs to be answered in future research. This is to help the governments select an appropriate platform for big data projects.
- The second question relates to the open government data. The open data issue will help to access easily from businesses, developers, and so on, and will be an excellent opportunity for the academy community to study future works related to big data. Nevertheless, privacy and security are significant issues to consider

strictly. This leads to a big question that is how to harmony between security issues and open data issues.
- The last question relates to the data processing model. In fact, most of the applications of governments still process offline data or batch. To apply online or real-time data processing models, governments need to invest a huge budget like big technology companies. This also leads to a big issue to be solved in the future that is how to reduce the costs of data transmission, storage, and repeated scanning. The results will help governments improve efficiency in using intelligent techniques as well as building smart applications and intelligent systems in the future.

**Acknowledgement:** This work was supported by the Vietnam Ministry of Science and Technology via the technology innovation program until 2020 with the project number: ĐM.40.DA/19. The authors would like to thank greatly Managing Editor (Beata Socha), reviewers, and editors who gave comments and edited this manuscript to improve it for publication.

**Author contributions:** Concept: P. V. Hai and C. K. Long; Methodology: P. V. Hai, C. K. Long, and H. Q. Trung; Works collection: C. K. Long; Analysis and comparison: C. K. Long and P. V. Hai; Writing – original draft paper: C. K. Long; Writing – review and comment internally: R. Agrawal, P. V. Hai, and H. Q. Trung; Checking to plagiarise: R. Agrawal; Refining – the manuscript for publication: C. K. Long and P. V. Hai; All authors have read and agreed to the published version of the manuscript.

**Conflict of interest:** The authors declare no conflict of interest.

**Data availability statement:** As mentioned in the Introduction of the manuscript, all works for surveying are gathered from open databases (the ScienceDirect, IEEE Xplore, and Google Scholar research databases).

## References

[1] M. Alessandro, “AI and big data: A blueprint for a human right, social and ethical impact assessment,” *Comput. Law Secur. Rev.*, vol. 34, no. 4, pp. 754–772, 2018.
[2] J. H. Nord, A. Koohang, and J. Paliszkiewicz, “The Internet of Things: Review and theoretical framework,” *Expert. Syst. Appl.*, vol. 133, pp. 97–108, 2019.
[3] M. Muzammal, Q. Qu, and B. Nasrulin, “Renovating blockchain with distributed databases: An open source system,” *Future Gener. Comput. Syst.*, vol. 90, pp. 105–117, 2019.
[4] H. M. Safhi, B. Frikh, and B. Ouhbi, “Assessing reliability of big data knowledge discovery process,” *Proc. Comput. Sci.*, vol. 148, pp. 30–36, 2019.
[5] R. Kune, P. K. Konugurthi, A. Agarwal, R. R. Chillarige, and R. Buyya, “The anatomy of big data computing,” *Softw. Pract. Exp.*, vol. 46, no. 1, pp. 79–105, 2016.
[6] M. Manzano Surroca, F. Parri, and X. Tarrado, “The new European interoperability framework as a facilitator of digital transformation for citizen empowerment,” *J. Biomed. Inform.*, vol. 82, pp. 94–511, 2019, DOI: 10.1016/j.jbi.2019.103166.
[7] B. Fan, R. Liu, K. Huang, and Y. Zhu, “Defining digital transformation: Results from expert interviews,” *Gov. Inf. Q.*, vol. 36, no. 4, p. 101395, 2019, DOI: 10.1016/j.giq.2019.06.002.
[8] D. Wang, C. Chen, and D. Richards, “A prioritization-based analysis of local open government data portals: A case study of Chinese province-level governments,” *Gov. Inf. Q.*, vol. 35, no. 4, pp. 644–656, 2018.
[9] J. Attard, F. Orlandi, S. Scerri, and S. Auer, “A systematic review of open government data initiatives,” *Gov. Inf. Q.*, vol. 32, no. 4, pp. 399–418, 2015.
[10] Y. Zhao and B. Fan, “Exploring open government data capacity of government agency: Based on the resource-based theory,” *Gov. Inf. Q.*, vol. 35, no. 1, pp. 1–12, 2018.
[11] W. N. Ismail, M. M. Hassan, and H. A. Alsalamah, “Mining of productive periodic-frequent patterns for IoT data analytics,” *Future Gener. Comput. Syst.*, vol. 88, pp. 512–523, 2018.
[12] D. López and B. Farooq, “A multi-layered blockchain framework for smart mobility data-markets,” *Transp. Res. Part. C: Emerg. Technol.*, vol. 111, pp. 588–615, 2020.
[13] E. Bonnevie, J. Goldbarg, A. K. Gallegos-Jeffrey, S. D. Rosenberg, E. Wartella, and J. Smyser, “The internet of things-based decision support system for information processing in intelligent manufacturing using data mining technology,” *Mech. Syst. Signal. Process.*, vol. 110, p. 142, 2020, DOI: 10.1016/j.ymssp.2020.106630.
[14] M. Amoon, T. Altameem, and T. Altameem, “Internet of things sensor assisted security and quality analysis for health care data sets using artificial intelligent based heuristic health management system,” *Measurement*, vol. 161, 2020, DOI: 10.1016/j.measurement.2020.107861.
[15] M. Khovrichev, L. Elkhovskaya, V. Fonin, and M. Balakhontceva, “Intelligent approach for heterogeneous data integration: Information processes analysis engine in clinical remote monitoring systems,” *Proc. Comput. Sci.*, vol. 156, pp. 134–141, 2019.
[16] N. Kawaguchi, “Application of blockchain to supply chain: Flexible Blockchain Technology,” *Proc. Comput. Sci.*, vol. 164, pp. 143–148, 2019.
[17] N. Mani, A. Singh, and S. L. Nimmagadda, “An IoT guided healthcare monitoring system for managing real-time notifications by fog computing services,” *Proc. Comput. Sci.*, vol. 167, pp. 850–859, 2020.
[18] A. Reyna, C. Martín, J. Chen, E. Soler, and M. Díaz, “On blockchain and its integration with IoT. Challenges and

opportunities," *Future Gener. Comput. Syst.*, vol. 88, pp. 173–190, 2018.
[19] P. Niewiadomski, A. Stachowiak, and N. Pawlak, "Knowledge on IT tools based on AI maturity – Industry 4.0," *Persp. Proc. Manuf.*, vol. 39, pp. 574–582, 2019.
[20] G. Aceto, V. Persico, and A. Pescapé, "Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0," *J. Ind. Inf. Integr.*, vol. 18, p. 18, 2020, DOI: 10.1016/j.jii.2020.100129.
[21] A. Kankanhalli, J. Hahn, S. Tan, and G. Gao, "Big data and analytics in healthcare: introduction to the special section," *Inf. Syst. Front.*, vol. 18, no. 2, pp. 233–235, 2016.
[22] M. Elhoseny, A. Abdelaziz, A. S. Salama, A. M. Riad, K. Muhammad, and A. K. Sangaiah, "A hybrid model of Internet of Things and cloud computing to manage big data in health services applications," *Future Gener. Comput. Syst.*, vol. 86, pp. 1383–1394, 2018.
[23] L. Greco, P. Maresca, and J. Caja, "Big data and advanced analytics in Industry 4.0: A comparative analysis across the European Union," *Proc. Manuf.*, vol. 41, pp. 383–390, 2019.
[24] Checkland, et al., *Complex, intelligent, and software intensive systems*, Springer Science and Business Media LLC, 2020.
[25] M. I. Pramanik, R. Y. K. Lau, M. A. K. Azad, M. S. Hossain, M. K. H. Chowdhury, and B. K. Karmaker, "Healthcare informatics and analytics in big data," *Expert. Syst. Appl.*, vol. 152, p. 15215, 2020, DOI: 10.1016/j.eswa.2020.113388.
[26] M. Thibaud, H. Chi, W. Zhou, and S. Piramuthu, "Internet of things (IoT) in high-risk environment, health and safety (EHS) industries: A comprehensive review," *Decis. Support. Syst.*, vol. 108, pp. 79–95, 2018.
[27] D. Zhu, "IoT and big data based cooperative logistical delivery scheduling method and cloud robot system," *Future Gener. Comput. Syst.*, vol. 86, pp. 709–715, 2018.
[28] Z. Zhao, M. Zhang, C. Yang, J. Fang, and G. Q. Huang, "Distributed and collaborative proactive tandem location tracking of vehicle products for warehouse operations," *Comput. Ind. Eng.*, vol. 125, pp. 637–648, 2018.
[29] T. Wang, X. Wang, W. Shi, Z. Zhao, Z. He, and T. Xia, "Target localization and tracking based on improved Bayesian enhanced least-squares algorithm in wireless sensor networks," *Comput. Netw.*, vol. 167, p. 106968, 2020, DOI: 10.1016/j.com-net.2019.106968.
[30] H. Lee, H. Chae, and K. Yi, "A geometric model based 2D LiDAR/radar sensor fusion for tracking surrounding vehicles," *IFAC-PapersOnLine*, vol. 52, no. 8, pp. 130–135, 2019.
[31] M. Simoncini, L. Taccari, F. Sambo, L. Bravi, S. Salti, and A. Lori, "Vehicle classification from low-frequency GPS data with recurrent neural networks," *Transp. Res. Part. C: Emerg. Technol.*, vol. 91, pp. 176–191, 2018.
[32] G. Park, S. B. Choi, D. Hyun, and J. Lee, "Integrated observer approach using in-vehicle sensors and GPS for vehicle state estimation," *Mechatronics*, vol. 50, pp. 134–147, 2018.
[33] A. M. Mota, M. J. Clarkson, P. Almeida, L. Peralta, and N. Matela, "Optimization data on total cost of ownership for conventional and battery electric heavy vehicles driven by humans and by automated driving systems," *Data Brief*, vol. 195, p. 30, 2020, DOI: 10.1016/j.dib.2020.105566.
[34] T. C. Hsu, H. Yang, Y. Chung, and C. Hsu, "A creative IOT agriculture platform for cloud fog computing," *Sustain. Comput.: Inform. Syst.*, vol. 28, 2018, DOI: 10.1016/j.suscom.2018.10.006.
[35] K. Gunasekera, A. N. Borrero, F. Vasuian, and K. P. Bryceson, "Experiences in building an IoT infrastructure for agriculture education," *Proc. Comput. Sci.*, vol. 135, pp. 155–162, 2018.
[36] A. R. Al-Ali, A. Al Nabulsi, S. Mukhopadhyay, M. S. Awal, S. Fernandes, and K. Ailabouni, "IoT-solar energy powered smart farm irrigation system," *J. Electron. Sci. Technol.*, vol. 17, no. 4, pp. 1–14, 2019, DOI: 10.1016/j.jnl-est.20-20.100017.
[37] A. D. Boursianis, M. S. Papadopoulou, P. Diamantoulakis, A. Liopa-Tsakalidi, P. Barouchas, G. Salahas, et al., "Internet of things (IoT) and agricultural unmanned aerial vehicles (UAVs) in smart farming: A comprehensive review," *Internet Things*, p. 100187, 2020, DOI: 10.1016/j.iot.2020.100187.
[38] L. Colizzi, A. Caivano, C. Ardito, G. Desolda, A. Castrignanò, M. Matera, et al., *Chapter 1: Introduction to Agricultural IoT, Agricultural Internet of Things and Decision Support for Precision Smart Farming*, Elsevier Inc, 2020, pp. 1–33, DOI: 10.1016/C2018-0-00051-1.
[39] A. Vij, S. Vijendra, A. Jain, S. Bajaj, A. Bassi, and A. Sharma, "IoT and machine learning approaches for automation of farm irrigation system," *Proc. Comput. Sci.*, vol. 167, pp. 1250–1257, 2020.
[40] S. Alamgir Hossain, M. Anisur Rahman, and M. A. Hossain, "Edge computing framework for enabling situation awareness in IoT based smart city," *J. Parallel Distrib. Comput.*, vol. 122, pp. 226–237, 2018.
[41] M. M. Rathore, A. Paul, W. H. Hong, H. Seo, I. Awan, and S. Saeed, "Exploiting IoT and big data analytics: Defining smart digital city using real-time urban data," *Sustain. Cities Soc.*, vol. 40, pp. 600–610, 2018.
[42] B. W. Wirtz, J. C. Weyerer, and F. T. Schichtel, "An integrative public IoT framework for smart government," *Gov. Inf. Q.*, vol. 36, no. 2, pp. 333–345, 2019.
[43] S. Chatterjee, A. K. Kar, and M. P. Gupta, "Success of IoT in smart cities of India: An empirical analysis," *Gov. Inf. Q.*, vol. 35, no. 3, pp. 349–361, 2018.
[44] N. Al-Nabhan, N. Al-Aboody, and A. B. M. Alim Al Islam, "A hybrid IoT-based approach for emergency evacuation," *Comput. Netw.*, vol. 155, pp. 87–97, 2019.
[45] S. Dey, "Chapter 10: Emerging trends of IoT-based applications in day-to-day life," *Internet Things Biomed. Eng.*, pp. 235–257, 2019.
[46] R. Guirado-Clavijo, J. A. Sanchez-Molina, H. Wang, and F. Bienvenido, "Conceptual data model for IoT in a chain-integrated greenhouse production: Case of the tomato production in Almeria (Spain)," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 102–107, 2018.
[47] A. H. Bagdadee, M. Z. Hoque, and L. Zhang, "IoT based wireless sensor network for power quality control in smart grid," *Proc. Comput. Sci.*, vol. 167, pp. 1148–1160, 2020.
[48] D. Mocrii, Y. Chen, and P. Musilek, "IoT-based smart homes: A review of system architecture, software, communications, privacy and security," *Internet Things*, vol. 1–2, pp. 81–98, 2018.
[49] N. Sharma, H. Parveen Sultana, R. Singh, and S. Patil, "Secure hash authentication in IoT based applications," *Proc. Comput. Sci.*, vol. 165, pp. 328–335, 2019.

[50] R. P. Meenaakshi Sundhari and K. Jaikumar, "IoT assisted hierarchical computation strategic making (HCSM) and dynamic stochastic optimization technique (DSOT) for energy optimization in wireless sensor networks for smart city monitoring," *Comput. Commun.*, vol. 150, pp. 226–234, 2020.
[51] A. R. Hilal, A. Sayedelahl, A. Tabibiazar, M. S. Kamel, and O. A. Basir, "A distributed sensor management for large-scale IoT indoor acoustic surveillance," *Future Gener. Comput. Syst.*, vol. 86, pp. 1170–1184, 2018.
[52] T. P. Fowdur, Y. Beeharry, V. Hurbungs, V. Bassoo, V. Ramnarain-Seetohul, and E. Lun, "Performance analysis and implementation of an adaptive real-time weather forecasting system," *Internet Things*, vol. 3–4, pp. 12–33, 2018.
[53] Y. Chen and D. Han, "Water quality monitoring in smart city: A pilot project," *Autom. Constr.*, vol. 89, pp. 307–316, 2018.
[54] U. Lee, K. Han, H. Cho, K. M. Chung, H. Hong, S. J. Lee, et al., "Intelligent positive computing with mobile, wearable, and IoT devices: Literature review and research directions," *Ad Hoc Netw*, vol. 83, pp. 8–24, 2019.
[55] Y. Bouzembrak, M. Klüche, A. Gavai, and H. J. P. Marvin, "Internet of things in food safety: Literature review and a bibliometric analysis," *Trends Food Sci. Technol.*, vol. 94, pp. 54–64, 2019.
[56] S. Chivarov, P. Kopacek, and N. Chivarov, "Cost oriented humanoid robot communication with IoT devices via MQTT and interaction with a smart home HUB connected devices," *IFAC-PapersOnLine*, vol. 52, no. 25, pp. 104–109, 2019.
[57] P. M. Dhulavvagol, V. H. Bhajantri, and S. G. Totad, "Blockchain ethereum clients performance analysis considering E-voting application," *Proc. Comput. Sci.*, vol. 167, pp. 2506–2515, 2020.
[58] M. Andoni, V. Robu, D. Flynn, S. Abram, D. Geach, D. Jenkins, et al., "Blockchain technology in the energy sector: A systematic review of challenges and opportunities," *Renew. Sustain. Energy Rev.*, vol. 100, pp. 143–174, 2019.
[59] D. Macrinici, C. Cartofeanu, and S. Gao, "Smart contract applications within blockchain technology: A systematic mapping study," *Telemat. Inform.*, vol. 35, no. 8, pp. 2337–2354, 2018.
[60] A. Farouk, A. Alahmadi, S. Ghose, and A. Mashatan, "Blockchain platform for industrial healthcare: Vision and future opportunities," *Comput. Commun.*, vol. 154, pp. 223–235, 2020.
[61] H. Vranken, "Sustainability of bitcoin and blockchains," *Curr. Opin. Environ. Sustain*, vol. 28, pp. 1–9, 2017.
[62] K. Ikeda and M. N. Hamid, "Chapter four: Applications of blockchain in the financial sector and a peer-to-peer global barter web," *Adv. Comput.*, vol. 111, pp. 99–120, 2018.
[63] H. Wang, H. Qin, M. Zhao, X. Wei, and W. Susilo, "Blockchain-based fair payment smart contract for public cloud storage auditing," *Inf. Sci.*, vol. 519, pp. 348–362, 2020.
[64] W. J. Gordon and C. Catalini, "Blockchain technology for healthcare: Facilitating the transition to patient-driven interoperability," *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 224–230, 2018.
[65] G. Kyriakoudes, S. Louca, and B. Behbod, "Cyprus's new national health service and future European health," *Lancet*, vol. 392, no. 10157, p. 1514, 2018.
[66] M. Pawlak, A. Poniszewska-Marańda, and N. Kryvinska, "Towards the intelligent agents for blockchain e-voting system," *Proc. Comput. Sci.*, vol. 141, pp. 239–246, 2018.
[67] P. K. Sharma, S. Rathore, and J. H. Park, "DistArch-SCNet: Blockchain-based distributed architecture with Li-Fi communication for a scalable smart city network," *IEEE Consum. Electron. Mag.*, vol. 7, no. 4, pp. 55–64, 2018.
[68] D. Puthal, N. Malik, S. P. Mohanty, E. Kougianos, and C. Yang, "The Blockchain as a decentralized security framework [future directions]," *IEEE Consum. Electron. Mag.*, vol. 7, no. 2, pp. 18–21, 2018.
[69] C. Liang, Y. Li, and J. Luo, "Blockchain for government services – Use cases, security benefits and challenges," *IEEE Xplore*, vol. 13, pp. 549–56, 2018, DOI: 10.1109/LT.2018.8368494.
[70] K. Li, V. Deolalikar, N. Pradhan, Big data gathering and mining pipelines for CRM using open-source, *IEEE International Conference on Big Data* (*Big Data*), IEEE, USA, 2015. DOI: 10.1109/Big-Data.2015.7364128.
[71] L. Birek, A. Grzywaczewski, R. Iqbal, F. Doctor, and V. Chang, "A novel big data analytics and intelligent technique to predict driver's intent," *Comput. Ind.*, vol. 99, pp. 226–240, 2018.
[72] Y. Huang, Z. Chen, T. Yu, X. Huang, and X. Gu, "Agricultural remote sensing big data: Management and applications," *J. Integr. Agric.*, vol. 17, no. 9, pp. 1915–1931, 2018.
[73] N. Shanmathi and M. Jagannath, "Computerised decision support system for remote health monitoring: A systematic review," *IRBM*, vol. 39, no. 5, pp. 359–367, 2018.
[74] F. Aparicio, M. L. Morales-Botello, M. Rubio, A. Hernando, R. Muñoz, H. López-Fernández, et al., "Perceptions of the use of intelligent information access systems in university level active learning activities among teachers of biomedical subjects," *Int. J. Med. Inform.*, vol. 112, pp. 21–33, 2018.
[75] R. S. Peres, A. Dionisio Rocha, P. Leitao, and J. Barata, "IDARTS – Towards intelligent data analysis and real-time supervision for Industry 4.0," *Comput. Ind.*, vol. 101, pp. 138–146, 2018.
[76] L. Kim, "Intelligent collaborative decision model for simulation of disaster data in cities and urbanlization," *IJAR*, vol. 6, no. 7, pp. 609–616, 2018.
[77] K. A. Pupkov, "Intelligent systems: Development and issues," *Proc. Comput. Sci.*, vol. 103, pp. 581–583, 2017.
[78] A. M. Al-Faifi, B. Song, M. M. Hassan, A. Alamri, and A. Gumaei, "Performance prediction model for cloud service selection from smart data," *Future Gener. Comput. Syst.*, vol. 85, pp. 97–106, 2018.
[79] M. Hiransha, E. A. Gopalakrishnan, M. Vijay Krishna, and K. P. Soman, "NSE stock market prediction using deep-learning models," *Proc. Comput. Sci.*, vol. 132, pp. 1351–1362, 2018.
[80] H. Van Pham, F. Asadi, N. Abut, and I. Kandilli, "Hybrid spiral STC-hedge algebras model in knowledge reasonings for robot coverage path planning and its applications," *Appl. Sci.*, vol. 9, no. 9, p. 1909, 2019, DOI: 10.3390/app9091909.
[81] H. Van Pham and P. Moore, "A proposal for information systems security monitoring based on large datasets," *Int. J. Distrib. Syst. Technol.*, vol. 9, no. 2, pp. 16–26, 2018, DOI: 10.4018/IJDST.2018040102.
[82] H. Van Pham and P. Moore, "Robot coverage path planning under uncertainty using knowledge inference and hedge

algebras," *Machines*, vol. 6, no. 4, p. 46, 2018, DOI: 10.3390/ma-chines6040046.
[83] L. H. Son, P. Van Viet, and P. Van Hai, "Picture inference system: a new fuzzy inference system on picture fuzzy set," *Appl. Intell.*, vol. 46, pp. 652–669, 2017.
[84] T. M. Tuan, N. T. Duc, and P. Van Hai, "Dental diagnosis from X-Ray images using fuzzy rule-based systems," *Int. J. Fuzzy Syst. Appl.*, vol. 16, no. 1, pp. 1–16, 2017.
[85] Y. Kobori, A. Osaka, S. Soh, and H. Okada, "MP15-03 novel application for sexual transmitted infection screening with an AI chatbot," *J. Urol.*, vol. 199, no. 4, pp. e189–e190, 2018.
[86] A. Androutsopoulou, N. Karacapilidis, E. Loukis, and Y. Charalabidis, "Transforming the communication between citizens and government through AI-guided chatbots," *Gov. Inf. Q.*, vol. 36, no. 2, pp. 358–367, 2019.
[87] W. illiamP. Wagner, "Trends in expert system development: A longitudinal content analysis of over thirty years of expert system case studies," *Expert. Syst. Appl.*, vol. 76, pp. 85–96, 2017.
[88] S. Thaker and V. Nagori, "Analysis of fuzzification process in fuzzy expert system," *Proc. Comput. Sci.*, vol. 132, pp. 1308–1316, 2018.
[89] M. L. Mfenjou, A. A. Ari, W. Abdou, and F. Spies, "Methodology and trends for an intelligent transport system in developing countries," *Sustain. Comput.: Inform. Syst.*, vol. 19, pp. 96–111, 2018.
[90] K. K. Patel and S. M. Patel, "Internet of things – IoT: Definition, characteristics, architecture, enabling technologies, application & future challenges," *Int. J. Eng. Sci. Comput.*, vol. 6, no. 5, pp. 6122–6131, 2016.
[91] I. P. Žarko, K. Pripužić, M. Serrano, M. Hauswirth, IoT data management methods and optimisation algorithms for mobile publish/subscribe services in cloud environments, *European Conference on Networks and Communications*, IEEE, Italy, 2014. DOI: 10.1109/EuCNC.2014.6882657.
[92] A. Poniszewska-Maranda, D. Kaczmarek, Selected methods of artificial intelligence for Internet of things conception, *Proceedings of the Federated Conference on Computer Science and Information Systems*, FedCSIS, Poland, 2015, pp. 1343–1348. DOI: 10.15439/2015F161.
[93] P. Yang and L. Xu, "The internet of things (IoT): Informatics methods for IoT-enabled health care," *J. Biomed. Inform.*, vol. 87, pp. 154–156, 2018.
[94] M. H. Salas-Olmedo, B. Moya-Gómez, J. C. García-Palomares, and J. Gutiérrez, "Tourists' digital footprint in cities: Comparing big data sources," *Tour. Manag.*, vol. 66, pp. 13–25, 2018.
[95] D. Blazquez and J. Domenech, "Big data sources and methods for social and economic analyses," *Technol. Forecast. Soc. Change*, vol. 130, pp. 99–113, 2018.
[96] F. Batista e Silva, M. A. Marín Herrera, K. Rosina, R. Ribeiro Barranco, S. Freire, and M. Schiavina, "Analysing spatio-temporal patterns of tourism in Europe at high-resolution with conventional and big data sources," *Tour. Manag.*, vol. 68, pp. 101–115, 2018.
[97] F. E. A. Horita, J. P. de Albuquerque, V. Marchezini, and E. M. Mendiondo, "Bridging the gap between decision-making and emerging big data sources: An application of a model-based framework to disaster management in Brazil," *Decis. Support. Syst.*, vol. 97, pp. 12–22, 2017.
[98] S. Achsas and E. H. Nfaoui, "Improving relational aggregated search from big data sources using stacked autoencoders," *Cognit. Syst. Res.*, vol. 51, pp. 61–71, 2018.
[99] J. G. Enríquez, F. J. Domínguez-Mayo, M. J. Escalona, M. Ross, and G. Staples, "Entity reconciliation in big data sources: A systematic mapping study," *Expert. Syst. Appl.*, vol. 80, pp. 14–27, 2017.
[100] M. Ge, H. Bangui, and B. Buhnova, "Big data for internet of things: A survey," *Future Gener. Comput. Syst.*, vol. 87, pp. 601–614, 2018.
[101] K. Sultan, U. Ruhi, R. Lakhani, Conceptualizing blockchain: Characteristics & applications, *11th IADIS International Conference Information Systems*, IADIS Press, Portugal, 2018, pp. 49–57. ISBN: 978-989-8533-74-6©2018.
[102] C. Ge, Z. Liu, and L. Fang, "A blockchain based decentralized data security mechanism for the internet of things," *J. Parallel Distrib. Comput.*, vol. 141, pp. 1–9, 2020.
[103] Y. Lin, H. Wang, J. Li, and H. Gao, "Data source selection for information integration in big data era," *Inf. Sci.*, vol. 479, pp. 197–213, 2019.
[104] G. L. Stavrinides and H. D. Karatza, "An energy-efficient, QoS-aware and cost-effective scheduling approach for real-time workflow applications in cloud computing systems utilizing DVFS and approximate computations," *Future Gener. Comput. Syst.*, vol. 96, pp. 216–226, 2019.
[105] P. Sun, "Security and privacy protection in cloud computing: Discussions and challenges," *J. Netw. Comput. Appl.*, vol. 160, 2020, DOI: 10.1016/j.jnca.2020.102642.
[106] B. Deepa, S. Srigayathri, and S. Visalakshi, "A review on cloud computing," *Int. J. Trend Res. Dev.*, vol. 1, p. 4, 2017. ISSN: 2394-9333.
[107] S. Parikh, D. Dave, R. Patel, and N. Doshi, "Security and privacy issues in cloud, fog and edge computing," *Proc. Comput. Sci.*, vol. 160, pp. 734–739, 2019.
[108] Z. Zandesh, M. Ghazisaeedi, M. V. Devarakonda, and M. S. Haghighi, "Legal framework for health cloud: A systematic review," *Int. J. Med. Inform.*, vol. 132, p. 103953, 2019, DOI: 10.1016/j.ijmed-inf.2019.103953.
[109] M. S. Mahmoud and Y. Xia, "Chapter 3: Cloud computing," *Netw. Control. Syst.*, pp. 91–125, 2019, DOI: 10.1016/B978-0-12-816119-7.00011-3.
[110] J. Proaño, C. Carrión, and B. Caminero, "Empirical modeling and simulation of a heterogeneous Cloud computing environment," *Parallel Comput*, vol. 83, pp. 118–134, 2019.
[111] F. De la Prieta, S. Rodríguez-González, P. Chamoso, J. M. Corchado, and J. Bajo, "Survey of agent-based cloud computing applications," *Future Gener. Comput. Syst.*, vol. 100, pp. 223–236, 2019.
[112] A. Sehgal, R. Agrawal, R. Bhardwaj, and K. K. Singh, "Reliability analysis of wireless link for IoT applications under shadow-fading conditions," *Proc. Comput. Sci.*, vol. 167, pp. 1515–1523, 2020.
[113] M. A. Khan and K. Salah, "IoT security: Review, blockchain solutions, and open challenges," *Future Gener. Comput. Syst.*, vol. 82, pp. 395–411, 2018.
[114] S. Van Till, "Chapter 10: IoT technology and standards," *Five Technol. Forces Disrupt Secur*, pp. 107–125, 2018, DOI: 10.1016/B978-0-12-805095-8.00010-7.

[115] P. Asghari, A. M. Rahmani, and H. H. S. Javadi, "Service composition approaches in IoT: A systematic review," *J. Netw. Comput. Appl.*, vol. 120, pp. 61–77, 2018.
[116] P. Victer Paul and R. Saraswathi, The internet of things – A comprehensive survey, *Proceedings of 2017 International Conference on Computation of Power, Energy Information and Communication* (*ICCPEIC*) IEEE, India, 2017, DOI: 10.1109/ICCP-EIC.2017.8290405.
[117] P. P. Ray, "A survey on internet of things architectures," *J. King Saud. Univ. – Comput. Inf. Sci.*, vol. 30, no. 3, pp. 291–319, 2018.
[118] G. Drosatos and E. Kaldoudi, "Blockchain applications in the biomedical domain: A scoping review," *Comput. Struct. Biotechnol. J.*, vol. 17, pp. 229–240, 2019.
[119] P. J. Taylor, T. Dargahi, A. Dehghantanha, R. M. Parizi, and K. kChoo, "A systematic literature review of blockchain cyber security," *Digital Commun. Netw.* vol. 6, no. 2, pp. 147–156, 2019, DOI: 10.1016/j.dcan.20-19.01.005.
[120] J. Yli-Huumo, D. Ko, S. Choi, S. Park, and K. Smolander, "Where is current research on blockchain technology? – A systematic review," *PLoS One*, vol. 10, no. 11, p. 0163477, 2016, DOI: 10.1371/journal.pone.0163477.
[121] Y. Lu, "The blockchain: State-of-the-art and research challenges," *J. Ind. Inf. Integr.*, vol. 15, pp. 80–90, 2019.
[122] S. Peng, S. Yu, and P. Mueller, "Social networking big data: Opportunities, solutions, and challenges," *Future Gener. Comput. Syst.*, vol. 86, pp. 1456–1458, 2018.
[123] X. Zou and H. L. Vu, "Academic social networks: Modeling, analysis, mining and applications," *J. Netw. Comput. Appl.*, vol. 132, pp. 86–103, 2019.
[124] C. E. Hendrick, J. N. Cone, J. Cirullo, and J. Maslowsky, "Social networks as an approach to systematic review," *Health Prof. Educ.*, vol. 5, no. 3, pp. 218–224, 2019.
[125] W. A. Günther, M. H. Rezazade Mehrizi, M. Huysman, and F. Feldberg, "Debating big data: A literature review on realizing value from big data," *J. Strategic Inf. Syst.*, vol. 26, no. 3, pp. 191–209, 2017.
[126] Y. N. Malek, A. Kharbouch, H. E. Khoukhi, M. Bakhouya, V. D. Florio, D. E. Ouadghiri, et al., "On the use of IoT and big data technologies for real-time monitoring and data processing," *Proc. Comput. Sci.*, vol. 113, pp. 429–434, 2017.
[127] H. Y. Tran and J. Hu, "Privacy-preserving big data analytics a comprehensive survey," *J. Parallel Distrib. Comput.*, vol. 134, pp. 207–218, 2019.
[128] A. Ajayi, L. Oyedele, O. Akinade, M. Bilal, H. Owolabi, L. Akanbi, et al., "Optimised big data analytics for health and safety hazards prediction in power infrastructure operations," *Saf. Sci.*, vol. 125, 2020, DOI: 10.1016/j.ssci.2020.104656.
[129] R. Iqbal, F. Doctor, B. More, S. Mahmud, and U. Yousuf, "Big data analytics: Computational intelligence techniques and application areas," *Technol. Forecast. Soc. Change*, vol. 153, p. 119253, 2020, DOI: 10.1016/j.techfore.2018.03.024.
[130] P. Galetsi, K. Katsaliaki, and S. Kumar, "Values, challenges and future directions of big data analytics in healthcare: A systematic review," *Soc. Sci. Med.*, vol. 241, p. 112533, 2019, DOI: 10.1016/j.socsci-med.2019.112533.
[131] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for big data," *Inf. Fusion.*, vol. 42, pp. 146–157, 2018.
[132] E. Jardine and A. M. Lindner, "The dark web and cannabis use in the United States: Evidence from a big data research design," *Int. J. Drug. Policy*, vol. 76, p. 102627, 2020, DOI: 10.1016/j.drugpo.2019.102627.
[133] R. C. LaBrie, G. H. Steinke, X. Li, and J. A. Cazier, "Big data analytics sentiment: US-China reaction to data collection by business and government," *Technol. Forecast. Soc. Change*, vol. 130, pp. 45–55, 2018.
[134] S.F.X. Lambert, "Strontium isotope (87Sr/86Sr) data from archaeological sites in Utah, USA," *Data Brief*, vol. 27, pp. 1–10, 2019, DOI: 10.1016/j.dib.2019.104571.
[135] P. Perumalswami, B. Wyatt, A. Harty, A. Mageras, L. Li, M. Miller, et al., "FRI-246-elimination of HCV in a large urban health system in the United States: A big-data approach," *J. Hepatol.*, vol. 70, no. 1, p. e502, 2019, DOI: 10.1016/S0618-8278(19)30991-0.
[136] E. W. Kuiler and C. L. McNeely, "Chapter 10: Federal big data analytics in the health domain: An ontological approach to data interoperability," *Fed. Data Sci. – Transf. Gov. Agric. Policy Using. Artif. Intell.*, pp. 161–176, 2018.
[137] J. Pollex and A. Lenschow, "Surrendering to growth? The European Union's goals for research and technology in the Horizon 2020 framework," *J. Clean. Prod.*, vol. 197, no. 2, pp. 1863–1871, 2018.
[138] L. A. Colombo, M. Pansera, and R. Owen, "The discourse of eco-innovation in the European Union: An analysis of the eco-innovation action plan and horizon 2020," *J. Clean. Prod.*, vol. 214, pp. 653–665, 2020.
[139] J. F. Admiraal, C. Musters, and G. R. de Snoo, "The loss of biodiversity conservation in EU research programmes: Thematic shifts in biodiversity wording in the environment themes of EU research programmes FP7 and Horizon 2020," *J. Nat. Conserv.*, vol. 30, pp. 12–18, 2016.
[140] B. Li, J. Li, Y. Jiang, and X. Lan, "Experience and reflection from China's Xiangya medical big data project," *J. Biomed. Inform.*, vol. 93, pp. 1–6, 2019, DOI: 10.1016/j.jbi.2019.103149.
[141] L. Yadi, S. Yuning, Y. U. Jiayue, X. Yingfa, W. Yiyuan, and Z. Xiaoping, "Big-data-driven model construction and empirical analysis of SMEs credit assessment in China," *Proc. Comput. Sci.*, vol. 147, pp. 613–619, 2019.
[142] W. Zhang, Z. Chong, X. Li, and G. Nie, "Spatial patterns and determinant factors of population flow networks in China: Analysis on tencent location big data," *Cities*, vol. 99, pp. 1–13, 2020, DOI: 10.1016/j.cities.2020.102640.
[143] V. Plutshack, S. Sengupta, A. Sahay, and J. E. Viñuales, "New and renewable energy social enterprises accessing government support: Findings from India," *Energy Policy*, vol. 132, pp. 367–378, 2019.
[144] A. Vats and A. Khan, "India's big data landscape: Challenges and opportunities," *Indian. J. Sci. Technol.*, vol. 10, no. 40, pp. 1–10, 2017, DOI: 10.17485/ijst/2017/v10i40/101542.
[145] A. V. Das, P. R. Donthineni, G. Sai Prashanthi, and S. Basu, "Allergic eye disease in children and adolescents seeking eye care in India: Electronic medical records driven big data analytics report II," *Ocul. Surf.*, vol. 17, no. 4, pp. 683–689, 2019.
[146] P. Navdeep, M. Arora, N. Sharma, Role of big data analytics in analyzing e-governance projects, *10th International Conference on New Trends in Business and Management:*

*An International Perspective, Gian Jyoti E-Journal*, India, vol. 6, no. 2, 2016. ISSN 2250-348X.
[147] S. Mukhopadhyay, H. Bouwman, and M. P. Jaiswal, "An open platform centric approach for scalable government service delivery to the poor: The aadhaar case," *Gov. Inf. Q.*, vol. 36, no. 3, pp. 437–448, 2019.
[148] E. Ifinedo, J. Rikala, and T. Hämäläinen, "Factors affecting Nigerian teacher educators' technology integration: Considering characteristics, knowledge constructs, ICT practices and beliefs," *Comput. Educ.*, vol. 146, p. 103760, 2020, DOI: 10.1016/j.compedu.2019.103760.
[149] O. M. Okunola, J. Rowley, and F. Johnson, "The multi-dimensional digital divide: Perspectives from an e-government portal in Nigeria," *Gov. Inf. Q.*, vol. 34, no. 2, pp. 329–339, 2017.
[150] A. O. Akinola, T. Salau, A. Oluwatayo, O. Babalola, and H. I. Okagbue, "Data on the awareness and adoption of ICT in town planning firms in Lagos state, Nigeria," *Data Brief*, vol. 20, pp. 436–447, 2018.
[151] K. Salisu, *E-Government adoption and framework for big data analytics in Nigeria*, National Information Technology Development Agency (NITDA), 2015. Available from: http://eprints.covenant-university.edu.ng/5284/1/CORRECTED%20PAPER%202-E-GOVERNMENT%20ADOPTION%20IN%20NIG-ERIA%20AND%20FRAMEWORK%20FOR%20BIG%20DATA%20ANALYTICS.-1.pdf.
[152] C. Liang, Y. Li, and J. Luo, "Fast tensor decompositions for big data processing," *Proc. 2016 Int. Conf. Adv. Technol. Commun.*, vol. 13, pp. 549–56, 2016, DOI: 10.1109/ATC.2016.7764776.
[153] D. N. Le, L. Le Tuan, and M. N. Dang Tuan, "Smart-building management system: An Internet-of-Things (IoT) application business model in Vietnam," *Technol. Forecast. Soc. Change*, vol. 141, pp. 22–35, 2019.
[154] Vietnam's Ministry of Industry and Trade and United Nations Development Programme, Industry 4.0 Readiness of Industry Enterprises in Viet Nam, *UNDP - Sustainable Development Goals*, Hanoi, 2019. https://www.vn.undp.org/content/vietnam/en/home/library/I40.html.
[155] B. Rivas, J. Merino, I. Caballero, M. Serrano, and M. Piattini, "Towards a service architecture for master data exchange based on ISO 8000 with support to process large datasets," *Comput. Stand. Interfaces*, vol. 54, no. 2, pp. 94–104, 2017.
[156] K. Timothy, *ISO 8000: An ISO framework for data governance*, British Computer Society, Wolverhampton Branch Meeting, University of Wolverhampton, Babcock Analytic Solutions, UK, 2016.
[157] A. Al-Badi, A. Tarhini, and A. I. Khan, "Exploring big data governance frameworks," *Proc. Comput. Sci.*, vol. 141, pp. 271–277, 2018.
[158] P. Kaur, M. Sharma, and M. Mittal, "Big data and machine learning based secure healthcare framework," *Proc. Comput. Sci.*, vol. 132, pp. 1049–1059, 2018.
[159] H. Yeong Kim and J. Suh Cho, "Data governance framework for big data implementation with NPS case analysis in Korea," *J. Bus. Retail. Manag. Res.*, vol. 12, no. 3, 2018, DOI: 10.24052/jbrmr/v12is03/art-04.
[160] J. Yebenes and M. Zorrilla, "Towards a data governance framework for third generation platforms," *Proc. Comput. Sci.*, vol. 151, pp. 614–621, 2019.
[161] N. N. Teslya, I. A. Ryabchikov, M. V. Petrov, A. A. Taramov, and E. O. Lipkin, "Smart city platform architecture for citizens' mobility support," *Proc. Comput. Sci.*, vol. 150, pp. 646–653, 2019.
[162] Y. Ye, M. Wang, S. Yao, J. N. Jiang, and Q. Liu, "Big data processing framework for manufacturing," *Proc. CIRP*, vol. 83, pp. 661–664, 2019.
[163] Q. Li, L. Lan, N. Zeng, L. You, J. Yin, X. Zhou, et al., "A framework for big data governance to advance RHINs: A case study of China," *IEEE Access*, vol. 7, pp. 50330–50338, 2019.
[164] A. M. S.Osman, "A novel big data analytics framework for smart cities," *Future Gener. Comput. Syst.*, vol. 91, pp. 620–633, 2019.
[165] J. N. Witanto, H. Lim, and M. Atiquzzaman, "Smart government framework with geo-crowdsourcing and social media analysis," *Future Gener. Comput. Syst.*, vol. 89, pp. 1–9, 2018.
[166] D. R. Topor and A. Budson, "A framework for internet of things-enabled smart government: A case of IoT cybersecurity policies and use cases in U.S. federal government," *Gov. Inf. Q.*, vol. 36, no. 2, pp. 346–357, 2019.
[167] H. A. Alaka, L. O. Oyedele, H. A. Owolabi, M. Bilal, S. O. Ajayi, and O. O. Akinade, "A framework for big data analytics approach to failure prediction of construction firms ," *Appl. Comput. Inform.*, vol. 16, pp. 207–222, 2018, DOI: 10.1016/j.aci.2018.04.003.
[168] C. Borrazzo, M. Pacilio, N. Galea, E. Preziosi, M. Carnì, M. Francone, et al., "Big data: Hadoop framework vulnerabilities, security issues and attacks," *Array*, vol. 64, p. 04, 2019, DOI: 10.1016/j.array.2019.100002.
[169] S. Ren, Y. Zhang, Y. Liu, T. Sakao, D. Huisingh, and C. M. V. B. Almeida, "A comprehensive review of big data analytics throughout product lifecycle to support sustainable smart manufacturing: A framework, challenges and future research directions," *J. Clean. Prod.*, vol. 210, pp. 1343–1365, 2019.
[170] A. Oussous, F. Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, "Big data technologies: A survey," *J. King Saud. Univ. – Comput. Inf. Sci.*, vol. 30, no. 4, pp. 431–448, 2018.
[171] S. Karimian-Aliabadi, D. Ardagna, R. Entezari-Maleki, E. Gianniti, and A. Movaghar, "Analytical composite performance models for big data applications," *J. Netw. Comput. Appl.*, vol. 142, pp. 63–75, 2019.
[172] N. A. Ghani, "Social media big data analytics: A survey," *Comput. Hum. Behav.*, vol. 101, pp. 417–428, 2019.
[173] N. Venkatesh, "Comparative analysis of big data, bigdata analytics: Challenges and trends," *Int. Res. J. Eng. Technol.*, vol. 5, no. 5, pp. 1948–1964, 2018.
[174] B. Wang, C. Wu, L. Huang, and L. Kang, "Using data-driven safety decision-making to realize smart safety management in the era of big data: A theoretical perspective on basic questions and their answers," *J. Clean. Prod.*, vol. 210, pp. 1595–1604, 2019.
[175] H. Hu, Y. Wen, T. S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," *IEEE Access*, vol. 2, pp. 652–687, 2014.
[176] G. S. Bhathal and A. Singh, "Big data: Hadoop framework vulnerabilities, security issues and attacks," *Array*, vol. 1–2, pp. 1–8, 2019, DOI: 10.1016/j.ar-ray.2019.100002.

[177] H. M. Safhi, B. Frikh, and B. Ouhbi, “Assessing reliability of big data knowledge discovery process,” *Proc. Comput. Sci.*, vol. 148, pp. 30–36, 2019.

[178] A. Shobanadevi and G. Maragatham, Data mining techniques for IoT and big data – A survey, *Proceedings of 2017 International Conference on Intelligent Sustainable Systems*, IEEE, India, 2018, DOI: 10.1109/ISS1.2017.8389260.

[179] P. Satyam, Big data, smart data, dark data and open data: eGovernment of the future, *Second International Conference on Democracy & eGovernment*, IEEE, Ecuador, 2015, DOI: 10.1109/ICED-EG.2015.7114483.

[180] Y. Yang, H. He, D. Wang and Z. Ding, A framework to data delivery security for big data annotation delivery system, *Proceedings of 2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems*, IEEE, China, 2018, DOI: 10.1109/MASS.2018.00082.

[181] M. Xyntarakis and C. Antoniou, “Chapter 6: Data science and data visualization, mobility patterns,” *Big Data Transp. Anal.*, pp. 107–144, 2019.

[182] W. H. Inmon, D. Linstedt and M. Levins. “Chapter 18.1: An introduction to data visualizations,” *Data Architecture*, Second edition, 2019, pp. 381–395.

[183] Y. Zhang, R. Zhang, Y. Wang, H. Guo, R. Y. Zhong, T. Qu, et al., “Big data driven decision-making for batch-based production systems,” *Proc. CIRP*, vol. 83, pp. 814–818, 2019.

[184] A. Merendino, S. Dibb, M. Meadows, L. Quinn, D. Wilson, L. Simkin, et al., “Big data, big decisions: The impact of big data on board level decision-making,” *J. Bus. Res.*, vol. 93, pp. 67–78, 2018.

[185] L. Huang, C. Wu, B. Wang, and Q. Ouyang, “Big-data-driven safety decision-making: A conceptual framework and its influencing factors,” *Saf. Sci.*, vol. 109, pp. 46–56, 2018.

[186] W. Inoubli, S. Aridhi, H. Mezni, M. Maddouri, and E. Mephu Nguifo, “An experimental survey on big data frameworks,” *Future Gener. Comput. Syst.*, vol. 86, pp. 546–564, 2018.

[187] T. Palonen and R. Viri, “Benchmarking public transport level-of-service using open data,” *Transp. Res. Proc.*, vol. 42, pp. 100–108, 2019.

[188] K. McBride, G. Aavik, M. Toots, T. Kalvet, and R. Krimmer, “How does open government data driven co-creation occur? Six factors and a ‘perfect storm’; insights from Chicago’s food inspection forecasting model,” *Gov. Inf. Q.*, vol. 36, no. 1, pp. 88–97, 2019.

[189] R. K. R. Kummitha, “Cultivating open government data platform ecosystems through governance: Lessons from Buenos Aires, Mexico City and Montevideo,” *Gov. Inf. Q.*, vol. 37, no. 3, p. 101481, 2020, DOI: 10.1016/j.giq.2020.101479.

[190] J. D. Twizeyimana and A. Andersson, “The public value of E-Government – A literature review,” *Gov. Inf. Q.*, vol. 36, no. 2, pp. 167–178, 2019.