CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

BIG DATA, CLOUD COMPUTING AND IOT

Tools and Applications

EDITED BY
Sita Rani, Pankaj Bhambri,
Aman Kataria, Alex Khang
and Arun Kumar Sivaraman

Big Data, Cloud Computing and IoT

Cloud computing, the Internet of Things (IoT), and big data are three significant technological trends affecting the world's largest corporations. This book discusses big data, cloud computing, and the IoT, with a focus on the benefits and implementation problems. In addition, it examines the many structures and applications pertinent to these disciplines. Also, big data, cloud computing, and the IoT are proposed as possible study avenues.

Features:

- Informs about cloud computing, IoT, and big data, including theoretical foundations and the most recent empirical findings
- Provides essential research on the relationship between various technologies and the aggregate influence they have on solving real-world problems
- Ideal for academicians, developers, researchers, computer scientists, practitioners, information technology professionals, students, scholars, and engineers exploring research on the incorporation of technological innovations to address contemporary societal challenges

Big Data, Cloud Computing and IoT

Tools and Applications

Edited by
Sita Rani
Pankaj Bhambri
Aman Kataria
Alex Khang
Arun Kumar Sivaraman

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

Front cover image: metamorworks/Shutterstock

First edition published 2023
by CRC Press
6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

Library of Congress Cataloging-in-Publication Data
Names: Rani, Sita, editor. | Bhambri, Pankaj, editor. | Kataria, Aman,
editor. | Khang, Alex, editor. | Sivaraman, Arun Kumar, editor.
Title: Big data, cloud computing and IoT : tools and applications / edited
by Sita Rani, Pankaj Bhambri, Aman Kataria, Alex Khang, Arun Kumar Sivaraman.
Description: First edition. | Boca Raton : Chapman & Hall/CRC Press, 2023.
| Includes bibliographical references and index. |
Identifiers: LCCN 2022049025 (print) | LCCN 2022049026 (ebook) | ISBN
9781032284200 (hbk) | ISBN 9781032287430 (pbk) | ISBN 9781003298335 (ebk)
Subjects: LCSH: Big data. | Cloud computing. | Internet of things.
Classification: LCC QA76.9.B45 B5537 2023 (print) | LCC QA76.9.B45
(ebook) | DDC 005.7--dc23/eng/20221214
LC record available at https://lccn.loc.gov/2022049025
LC ebook record available at https://lccn.loc.gov/2022049026

ISBN: 978-1-032-28420-0 (hbk)
ISBN: 978-1-032-28743-0 (pbk)
ISBN: 978-1-003-29833-5 (ebk)

DOI: 10.1201/9781003298335

Typeset in Palatino
by SPi Technologies India Pvt Ltd (Straive)

Contents

Preface

The Internet of Things (IoT), big data, and cloud computing are all independent but complementary fields of study. The integration of three technologies provides synergy and a great chance for organizations to reap the enormous benefits of integration. When this combination is properly conceived, built, implemented, and operated, it can unleash a technical force that can propel innovation forward. Big data, the IoT, and the cloud architectures all work together to bring significant economic advantages. In a way, it is a great fit. The IoT captures data in real time. Data management systems benefit from big data's optimization. Rapid data collection, storage, computation, and dissemination are all features of the cloud. Big data solutions linked with IoT andcloud architecture are the key focus of this book, which is based on appealing business propositions. As a result, the book provides a high-level overview of architecture, solution practices, governance, and the underlying technical approach to developing integrated big data, cloud, and IoT solutions.

The security of critical systems and infrastructure is a serious concern for information and communication technology systems and networks. There are a variety of ways to ensure that messages are coming from trustworthy sources. The contributors cover the most recent research and development in authentication systems, including problems and applications for cloud technologies, the IoT, and big data in this edited reference.

Recent advances in micro-electro andmicro-mechanical system innovation, remote intersections, and computerized devices have enabled the creation of low-cost, multifunctional sensor hubs that are simple to operate, waste little power, and send data wirelessly over short distances. Intelligent sensors, when utilized like an IoT segment, convert a predicted reality factor into a digital information stream that may be communicated to a gateway for further processing.

The book illuminates the IoT, the cloud, and big data, as well as other cutting-edge technologies. This book addresses a variety of contemporary scientific and technical issues, including how to transform the IoT concept into a practical, technically feasible, and financially viable product. Big data and cloud computing are presented as important enablers for the sensing and computation backbone of the IoT. We are pleased that CRC Publications offered us the privilege to publish the book. Additionally, we appreciate our families and loved ones for their support. This book is intended for academics, postgraduates, and practitioners interested in cloud computing, the IoT, and big data.

About the Editors

Sita Rani is working in the Deaprtment of Computer Science and Engineering at Guru Nanak Dev Engineering College, Ludhiana. She has completed her B.Tech (CSE) and M.Tech (CSE) in 2002 and 2008 respectively, from Guru Nanak Dev Engineering College, Ludhiana. She earned her PhD in Computer Science and Engineering from I.K.Gujral Punjab Technical University, Kapurthala, Punjab in 2018. She has more than 19 years of teaching experience. She is an active member of ISTE, IEEE, and IAEngg. She is the receiver of ISTE Section Best Teacher Award, 2020, and International Young Scientist Award, 2021. She has contributed to various research activities while publishing articles in renowned journals and conference proceedings. She has published several National and International patents also. She has delivered many expert talks in All India Council of Technical Education (AICTE)–sponsored Faculty Development Programs and organized many international conferences during her 19 years of teaching experience. She is a member of the editorial boards of many international journals of repute. Her research interest includes parallel and distributed computing, machine learning, and the Internet of Things.

Pankaj Bhambri is working in the Department of Information Technology at Ludhiana's Guru Nanak Dev Engineering College. He also serves as the Institute's Coordinator, Skill Enhancement Cell. He has almost 20 years of experience as a teacher. For a long span, he served with the additional duties of an assistant registrar (Academic), member (Academic Council), member (BoS), member (RAC), hostel warden, APIO, and NSS coordinator for his institute. His research has appeared in a variety of prestigious international/national journals and conference proceedings indexed in SCIE, Scopus and UGC-CARE. Dr. Bhambri has contributed numerous textbooks as an editor/author and has filed several patents, too. He is the life member of ISTE, GENCO Alumni Association, I.E. (India), IIIE, IETE, IETA, I2OR and CSI, among others. As a result of his outstanding social and academic/research accomplishments over the past two decades, Dr. Bhambri has been awarded the ISTE Best Teacher Award-2022; I2OR National Award-2020; Green ThinkerZ Top 100 International Distinguished Educators-2020; I2OR Outstanding Educator Award-2019; SAA Distinguished Alumni Award-2012; CIPS Rashtriya Rattan Award-2008; LCHC Best Teacher Award-2007; and countless other accolades from various government and nonprofit organizations. He has supervised many undergraduate/postgraduate research projects/dissertations and is now supervising multiple Ph.D. research works as well. He organized numerous courses while receiving funding from the

AICTE, TEQIP, and others. Machine learning, bioinformatics, wireless sensor networks, and network security are his areas of interest.

Aman Kataria completed his B.Tech in electronics and communication engineering from Malout Institute of Management and Information Technology, Malout (established by state government), in 2010. He did his master's in electronics and instrumentation control engineering and his doctoral degree in 2013 and 2020, respectively, from Thapar Institute of Engineering and Technology, Patiala. Currently, he is working in Council of Scientific and Industrial Research-Central Scientific Instruments Organization as a project associate. He has also served as a lecturer at the Indian Institute of Information Technology, Una, Himachal Pradesh (under mentorship of National Institute of Technology, Hamirpur) in the Electronics and Communication Department. He has contributed to various research activities while publishing papers in the various Science Citation Index Expanded-and Scopus-indexed journals and conference proceedings. He has published three international patents also. His research interest includes are machine learning, artificial intelligence, image processing, cyber-physical systems, the Internet of Things, and soft computing.

Alex Khang, PHDSC, is a professor in information technology at the Universities of Science and Technology in Vietnam and United States, a software industry expert, an artificial intelligence (AI) and data scientist, a workforce development solutions consultant, and the chief of technology (AI and Data Science Research Center) at the Global Research Institute of Technology and Engineering, North Carolina, United States. He has 28-plusyears of nonstop teaching and research experience in information technology (software development, database technology, AI engineering, data engineering, data science, data analytics, Internet of Things (IoT)–based technologies, and cloud computing) at the Universities of Technology and Science in Vietnam, the European Union, India, and United States.

He has beenthe chair of session for 20-plusinternational conferences, an international keynote speaker for more than 25 international conclaves, an expert tech speaker for over 100 seminars and webinars, an international technical board member for 10-plusinternational organizations,an international editorial board member for more than five journals; a reviewer and evaluator for over 100 journal papers; and an international examiner and evaluator for more than 15 PhD theses in the computer science field.

He is the recipient of the Best Professor of the Year 2021, Researcher of the Year 2021, the Global Teacher Award 2021 (AKS), the Life Time Achievement Award 2021, the Leadership Award 2022 (Educacio World), and many other repute awards. He has contributed to various research activities in fields of AI and data science while publishing many international articles in renowned journals and conference proceedings.

He has published 52 authored books (in computer science between 2000–2010), authored 2 books (software development), edited 4 books, contributed 10 book chapters, and edited 2 books (calling for book chapters) in the fields of AI, data science, big data, IoT, smart city ecosystem, healthcare ecosystem, fintech, and blockchain technology (since 2020).

He has over 28 years of nonstop work as a CEO, a CTO, an engineering director, and a senior consultant in the field of software production and specialized in data engineering for foreign corporations from Germany, Sweden, the United States, Singapore, and multinationals.

Arun Kumar Sivaraman obtained his bachelor's degree in computer science and engineering from Anna University, Chennai, India, and his master's in computer science and engineering in 2010 from the College of Engineering Guindy (CEG) Chennai Campus. He has awarded a PhD, 2017, in computer science and engineering from Manonmaniam Sundaranar University (Govt.), Trinelveli, India. He received a Master of Business Administration, 2020, in education management from Alagappa University, Karaikudi, India. He has more than a decade of professional experience in the industrial, research and development, and academic sectors. He has worked as a lead data engineer for top multinational corporations like Cognizant, Standard Chartered, and Gilead Life Sciences. He worked as project consultant for a healthcare research (R&D) project in "The Research Council, Sultanate of Oman" funded by the Ministry of Health, Oman. He published a book in machine learning titled *Image Processing for Machine Learning* (ISBN: 978-93-5445-509-4). He published many research papers in the Scopus-indexed reputed journal, holds two Indian patents, and got one grant in international patent. For his merit, he got an offer as Lead Data Engineer in Tata Consultancy Services (TCS) 2011, Employment Pass Eligibility Certificate from the government of Singapore in 2012, and the Young Scientist award from the government of Sultanate of Oman 2018. He is currently working as an assistant professor (Sr. Grade) at VIT University, Chennai Campus, India. He is an active coeditor of a couple of special journal issues by TechScience Press (*Computer, Materials and Continua* – IF 4.89). His academic and research expertise covers a wide range of subject areas, including data engineering, data analytics, data science, and machine learning.

List of Contributors

Adalarasu Kanagasabai
SASTRA Deemed to be University
Tamil Nadu, India

Aman Kataria
CSIR-CSIO
Chandigarh, India

Ashish Verma
Military College of Telecommunication Engineering (MCTE-MHOW)
Madhya Pradesh, India

B. S. Shylaja
Ambedkar Institute of Technology
Karnataka, India

Bhavesh Borisaniya
Shantilal Shah Engineering College
Gujarat, India

C. Chethana
BMS Institute of Technology and Management
Karnataka, India

Charanjeet Singh
Gujranwala Guru Nanak Institute of Management and Technology
Punjab, India

Debabrata Sarddar
University of Kalyani
West Bengal, India

Devesh Pratap Singh
Graphic Era deemed to be University
Uttarakhand, India

Enakshmi Nandi
University of Kalyani
West Bengal, India

G. Boopathi Raja
Velalar College of Engineering and Technology
Tamil Nadu, India

Ganesh Reddy Karri
VIT-AP University
Andhra Pradesh, India

Gurjot Kaur Walia
Guru Nanak Dev Engineering College
Punjab, India

Gurwinder Singh
Guru Nanak Dev Engineering College
Punjab, India

Jagannath Mohan
Vellore Institute of Technology
Tamil Nadu

K. P. Maheswari
Fatima College
Tamil Nadu, India

Kailash Kumar Sahu
Pandit Sundarlal Sharma (Open) University
Chhattisgarh, India

Kanagaraj Venusamy
University of Technology and Applied Sciences-AI Mussanah
AI Muladdha, Sultanate of Oman

Keerthik Dhivya Rajakumar
Vellore Institute of Technology
Tamil Nadu

Munish Rattan
Guru Nanak Dev Engineering College
Punjab, India

N. Anuradha
Subbalakshmi Lakshmipathy College of Science
Tamil Nadu, India

Neelam Sharma
KIET Group of Institutions
Uttar Pradesh, India

Neelam Singh
Graphic Era deemed to be University
Uttarakhand, India

Pankaj Bhambri
Guru Nanak Dev Engineering College
Punjab, India

Parul Dubey
G.H. Raisoni College of Engineering
Maharashtra, India

Payel Ray
University of Kalyani
West Bengal, India

Piyush Kumar Pareek
Nitte Meenakshi Institute of Technology, Karnataka, India

Pushkar Dubey
Pandit Sundarlal Sharma (Open) University
Chhattisgarh, India

R. Bhaskar
Ambedkar Institute of Technology
Karnataka, India

R. Nagarajan
Gnanamani College of Technology
Tamil Nadu, India

Rajesh Bodade
Military College of Telecommunication Engineering (MCTE-MHOW)
Madhya Pradesh, India

Ramandeep Kaur
Gujranwala Guru Nanak Institute of Management and Technology
Punjab, India

Ranjan Kumar Mondal
University of Kalyani
West Bengal, India

S. Kannadhasan
Study World College of Engineering
Tamil Nadu, India

S. Nirmala Devi
Subbalakshmi Lakshmipathy College of Science
Tamil Nadu, India

S. R. Deepu
Ambedkar Institute of Technology
Karnataka, India

Sudheer Mangalampalli
VIT-AP University
Andhra Pradesh, India

Saurabh Kumar
The LNM Institute of Information Technology
Rajasthan, India

Shefali Kanwar
Amity University
Uttar Pradesh, India

Shruti Negi
Graphic Era deemed to be University
Uttarakhand, India

Sita Rani
Guru Nanak Dev Engineering College
Punjab, India

Surbhi Gupta
Amity University
Uttar Pradesh, India

Vandana Rawat
Graphic Era deemed to be University
Uttarakhand, India

Vanieka
Vellore Institute of Technology
Madhya Pradesh, India

1

Integration of IoT, Big Data, and Cloud Computing Technologies: Trend of the Era

Sita Rani, Pankaj Bhambri and Aman Kataria

CONTENTS

1.1 Introduction

Today, technology has captured every sphere of day-to-day life. It contributes significantly to providing better services and quality lives to users (Rani et al., 2021b). It has benefited a number of domains, like healthcare, industry, education, agriculture, transportation, banking and finance, among others (Rani et al., 2022c). This technological revolution comprises artificial

DOI: 10.1201/9781003298335-1

intelligence (AI), machine learning (ML), the Internet of Things (IoT), big data analytics, cloud computing, 4G, and 5G. The fundamental aim is to provide all-time connectivity and availability of data (Singh et al., 2021b). Individually as well as together, the IoT, big data, and cloud computing are playing notable roles in many application areas. The development of a strong charging structure network is broadly known as a vital necessity for a huge-scale evolution to electromobility (Bhambri et al., 2022). Kshirsagar et al. (2022) covered research on various methodologies for using neural networks to detect plant leaf diseases. Today, almost every field is making use of neural networks in one or another way (Bhambri et al., 2022).

1.1.1 IoT

The Internet is observed as a communication network that connects people and information (Hussein, 2019). On the other side, the IoT is a network of distributed objects in which each object has the potential to sense data/signals and process them with varying capacities (Soumyalatha, 2016; Rani et al., 2021). Each object/thing has a unique address so it can be located in cyberspace, which helps with the operations and communications among the objects (Singh et al., 2017; Jabraeil Jamali et al., 2020; Rani et al., 2022b). So, the fundamental aim of the IoT is to connect various things and individuals to facilitate all-time connectivity everywhere they are by using various communication networks and services, as shown in Figure 1.1. The IoT is considered one of the phases of the Internet's evolution. The IoT has made it convenient to connect many ordinary devices/gadgets to the Internet and helps achieve many diverse objectives (Kumar et al., 2022; Rani et al., 2022b). Whereas in 2015, only 0.06% of objects were connected to the IoT system, in 2020, approximately 50 billion smart devices have become part of the system and are connected through the Internet, which is a very high number compared to the world's population, shown in Figure 1.2.

With the evolution of the Internet, now it is not only a network of machines but has also emerged as a network of a variety of devices (Rani et al., 2021c; Banerjee et al., 2022). The IoT has become a network of different interconnected objects/things, also called a network of networks (Arya et al., 2022), as depicted in Figure 1.3.

In current times, various devices like smart appliances, vehicles, cyber-physical systems, smartphones, industrial systems, healthcare devices, and many others are using the

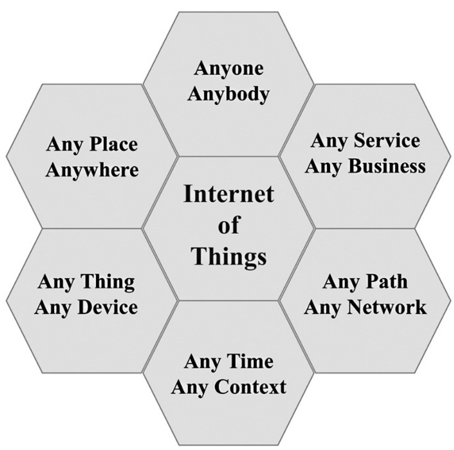

FIGURE 1.1
IoT: anytime, anywhere connectivity (Soumyalatha, 2016).

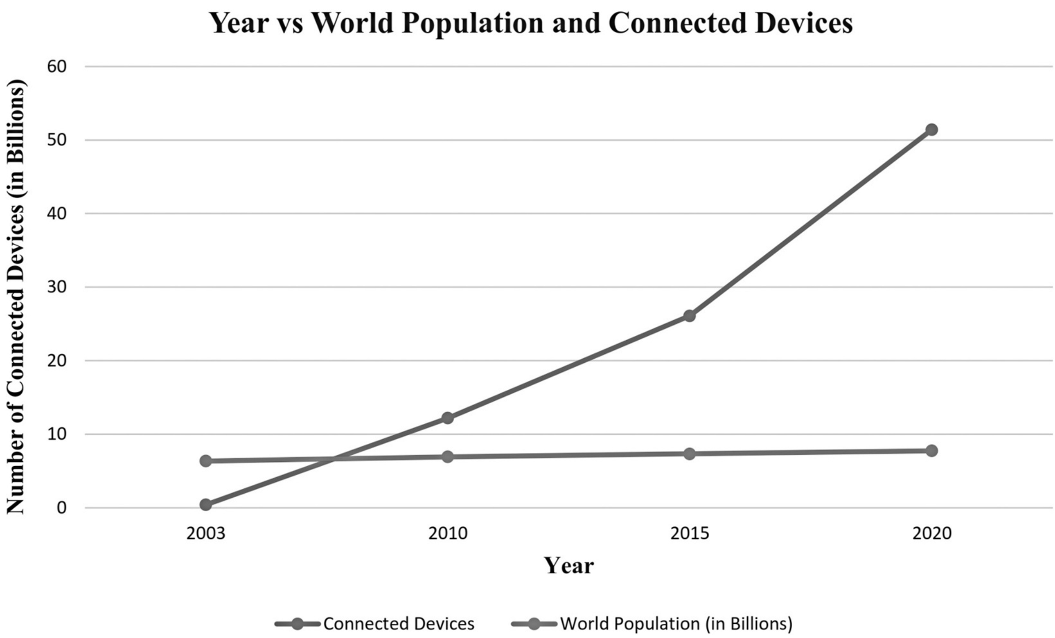

FIGURE 1.2
IoT-connected devices: year/population (Soumyalatha, 2016).

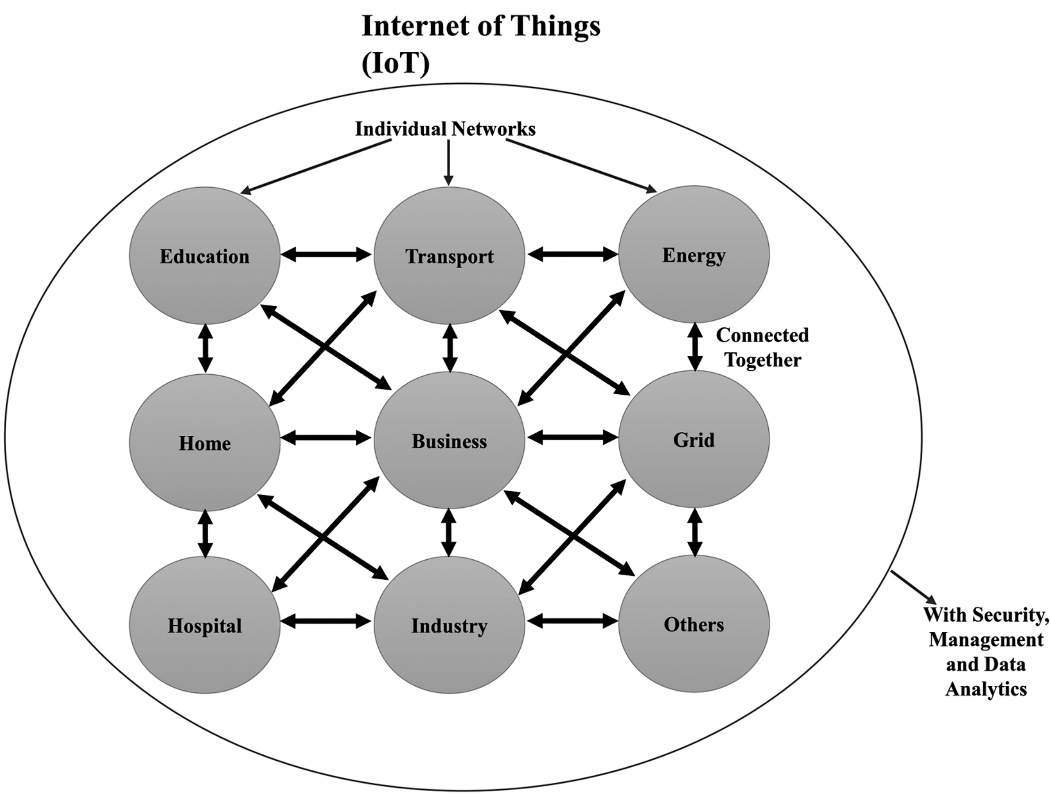

FIGURE 1.3
IoT: network mesh (Hussein, 2019).

Internet for information exchange (Rani et al., 2022a, 2022). Irrespective of the size, these smart devices have the potential of real-time data collection, monitoring, positioning, tracking, and process control (Bhambri et al., 2022).

Keeping all this in view, the IoT platform is supposed to flourish in terms of the number of connected devices and smart objects in a variety of commercial application domains (Rani et al., 2021a; Kothandaraman et al., 2022). The IoT will offer numerous opportunities in the domain of research due to the widespread usage of smart devices in various spheres of day-to-day life (Gupta and Kaur, 2016).

1.1.1.1 Architecture

As the number of smart devices connected through the IoT is increasing rapidly, it needs a flexible and expandable architecture (Aswale et al., 2019). Communication channels should be capable enough to manage voluminous data to avoid congestion (Tandon and Bhambri, 2017). Different IoT models are developed using either of the two architectures, that is, a three-layer architecture or a five-layer architecture, as discussed in the following:

Three-Layer Architecture: It is also called conventional architecture, as shown in Figure 1.4.

- The bottom layer is the Perception Layer, which keeps environmental data sensors. The basic tasks of this layer are to sense the environmental data using sensors and recognize the intelligent devices.
- The second layer, that is, Network Layer interconnects the intelligent objects/ devices, servers, and networked devices. The data gathered in the perception layer are transmitted and processed in this layer.
- The Application Layer provides the application-centered services to the users. It caters to a variety of applications, like smart healthcare, intelligent buildings, and smart homes and cities, among others.

The three-layer IoT architecture only gives an overview but is not ample for all application domains, especially those that require the sharper facets of the IoT (Rani et al., 2022d). To administer this challenge, the five-layer architecture is proposed, depicted in Figure 1.5.

The Perception Layer and the Application Layer do the same functionality as discussed earlier in three-layer architecture. The other layers work as follows:

- The Transport Layer transmits the data gathered by the lower Perception Layer to the upper Processing layer and vice versa. The communication media used by this layer are radiofrequency identification, 3G, near-field communication (NFC), local area network (LAN), and Bluetooth.

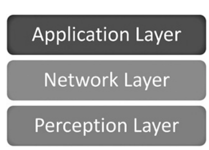

FIGURE 1.4
IoT: three-layer architecture.

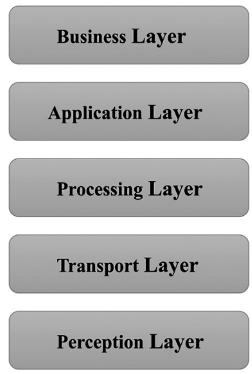

FIGURE 1.5
IoT: five-layer architecture.

- The next layer, that is, the Processing Layer, also known as middleware, stores, analyzes, and processes huge volumes of data received from the Transportation Layer. This layer uses a variety of technologies like databases, big data modules, and cloud computing applications to accomplish its tasks.
- The integrated IoT system is administered by the topmost layer, that is, the Business Layer. It manages the privacy of the users, various applications, and business models.

1.1.1.2 IoT Application Domains

In collaboration with ML and AI, anomaly detection systems are vastly used in behavioral analysis to help identify and predict the prevalence of anomalies (Bhambri et al., 2020). The IoT plays a very significant role in numerous and diverse application domains. Smart devices and gadgets have entered almost all the spheres of day-to-day life for individuals, organizations, and institutions (Rani et al., 2021). Broadly, the IoT has covered many important areas, including healthcare, industry, smart cities, agriculture, and others (Anand and Bhambri, 2018).

- **Smart Cities**

The IoT is contributing significantly to improving the quality of life of residents by improving cities' infrastructure. Some of the prominent applications in the smart city ecosystem consist of intelligent transportation and traffic management, smart buildings, smart parking, smart grids and metering systems, waste management, and smart lighting.

- **Healthcare**

IoT-enabled medical equipment and smart devices are offering many benefits in the healthcare sector, such as monitoring patients remotely, gathering medical data, tracking patient flow, observing the movement of medical staff, and more. Sensor-based devices and smart gadgets significantly enhance the experience of measuring various health parameters, like blood pressure and glucose levels, temperature, cholesterol, and heart rate, among others.

- **Smart Agriculture**

The IoT has the equal potential to improve various processes and services in the agriculture sector. This sensor-based technology is playing a very important role in analyzing various soil parameters and environmental conditions much needed in farming. The IoT has applications in analyzing soil nutrients, sensing microclimatic conditions, analyzing the quantity of vitamins and minerals in various agricultural products, monitoring grazing animals, and so on.

- **Retail and Logistics**

The IoT has made its own place in supply chain management and logistics too. It is used in the domains of product tracking, administering the rotation of articles/products in the warehouses and shelves of the outlets, and more. It also plays an important role in analyzing consignment conditions, studying storage conditions, and even keeping a regular check on gas leakage in industrial environments to ensure safety. Most of these applications are deployed using IoT-facilitated wireless sensor networks.

- **Smart Living**

In this domain, a variety of IoT-equipped smart applications are deployed to provide a better quality of life to residents. Most appliances and equipment can be managed remotely to save time, energy, and other resources. IoT gadgets/devices are also playing a very vital role in the security of individuals and the infrastructure.

- **Smart Environment**

The environment plays a significant role in providing a healthy life to all living beings. There are many factors, like industry waste, vehicle pollution, increasing population, and others, that are continuously damaging the environment. Smart equipment and devices are providing many innovative ways to monitor and manage industrial waste, keep a check on the pollution level in the air, and generate data to regulate government policies (Mohanta et al., 2022).

1.1.2 Big Data

Modern industry revolves around the most vital key term *data*. The rapid development of smart devices, gadgets, sensors, communication networks, computing devices, and storage systems has added a new dimension to data gathering and processing (Yaqoob et al., 2016). The data generated through the various applications in different domains are multiplying each year. *Big data* is a contemporary term initiated from the need of managing voluminous data by big ventures like Facebook, Twitter, Google, Yahoo, and others (Jain et al., 2021). The best way to achieve the balance between energy usage and quality of service is workload-aware energy-efficient virtual machine consolidation (Kaur et al., 2020).

Its key role is to define huge, diversified data, which are usually unstructured and not manageable using traditional techniques and tools. Today, governments, industry, and academia all have become highly attentive toward the unexplored capability of big data (Rodríguez-Mazahua et al., 2016; Rani and Gupta, 2017). The biggest challenge faced by information technology (IT) engineers and data scientists is that the increase in data is very rapid, which makes the administration and analysis of data difficult (Azeem et al., 2021).

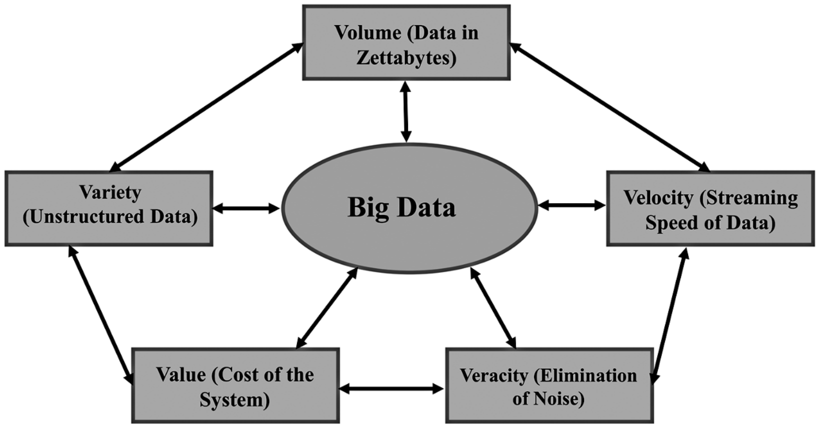

FIGURE 1.6
The big data: 5V model.

Consequently, it becomes important to use standardized tools and techniques to manage, store, and analyze such a voluminous amount of data. Big data models traversed through 3V, 4V, and 5V evolutions, in which each new model is an extension of the previous one (Singh et al., 2021a). The latest development, that is, 5V model (shown in Figure 1.6), is described as follows:

- **Volume**. Due to the rapid development in a variety of domains, the volume of data is increasing tremendously. There is an approximately 40% yearly increase in data and in the size of zettabytes.
- **Velocity**. It describes the timeliness of data. The collection and processing of data must be done rapidly to enhance its commercial utility.
- **Variety**. When the volume of data is huge, it usually comprises semistructured or unstructured data like text, audio, video, and regular structured data.
- **Value**. Today, data are some of the most precious resources. They are administered like commodities that are sold and bought among the parties. Consequently, understanding its value/cost can aid to manage other resources and help in budget planning.
- **Veracity**. As data are gathered from a number of sources, the accuracy of the data is always questionable. Collected data are preprocessed using different techniques and methods before its actual use. It ensures better and more accurate decision-making.

1.1.2.1 Big Data Applications

Big data is an infrastructure that aids in fetching, organizing, and managing huge volumes of data (Chhabra and Bhambri, 2021). It uses various specialized software, methods, and tools to store, sort, access, discover, and process voluminous data (Bhambri and Gupta, 2014). The data, gathered over a long period, are analyzed visually for their value to be administered in various scenarios (Khan and Javaid, 2021). Different software like Hadoop and Apache play a very significant role in deriving inferences from data that are of huge significance to various kinds of industries, such as the following (Shrivastava et al., 2021):

- **Education**. In the education sector, big data plays an important role in grading, restructuring syllabi, predicting future requirements, and in career analysis of the candidates.
- **Healthcare**. The volume of data/samples gathered helps in the detection/prediction of impending pandemic outbreaks, preventive care, and evidence-dependent medicine.
- **Government**. Big data storage techniques and analysis tools are playing a very vital role in the areas of banking, insurance, finance, security from cyber threats, and many more.
- **Media**. Big data analytical techniques are used for customer behavior, audience study and prediction, and more.
- **Transportation**. Big data is used in traffic analysis, rescheduling of traffic routes, ensuring public safety, and so on.
- **Banking**. The study of business statistics, financial security, risk analysis and calculation, forecasting investment funds, and more use big data analytics.

Consequently, big data analytics is aiding many domains for the purpose of analysis and prediction.

1.1.3 Cloud Computing

The concept of cloud computing emerged for the optimal utilization of various computing resources (Nazir et al., 2020). These resources should be paid for only when required/used like other resources, for example, water, electricity, gas, and so on (Bhambri and Kaur, 2014). The various services provided under this concept are catered to and managed by the various data centers that are located at different geographical locations (Kaur and Bhambri, 2019). A cloud computing service was a long-held fantasy in the domain of information and communication technology (ICT), and it will be realized due to the emergence of economic data centers (Bhambri and Chhabra (2022). Another major challenge arising due to the evolution of computing as a service is the security of data/information (Sinha et al., 2020). From the architectural aspect of cloud computing, the most vital entity is the data center (Sarga, 2012). Basically, data centers provide different categories of cloud services to customers/users (Kaur et al., 2015). A number of companies like Google, Amazon, Yahoo, Facebook, and many more have already entered the market and provide cloud services to the customers using pay-on-demand/use model (Singh et al., 2021a). As discussed earlier, they have deployed their data centers at different locations to provide economical services to users.

The security of data and the quality of services are managed by cloud service providers. Cloud users have no control over service parameters, like availability of information, quality, and others (Bhambri and Gupta, 2012a). To obtain and ensure committed quality services, consumers usually go for service-level agreements (SLAs). An SLA comprises all the terms and conditions settled between the consumer and cloud provider. The professional image of a cloud service provider is evaluated by their adherence to the SLA (Bhambri and Gupta, 2012b). When large organizations acquire multiple applications to provide the most reliable and efficient services to their clients, the procedure is termed a Hazy Cloud Process (Jasmine and Gupta, 2012).

1.1.3.1 Cloud Service Models

Cloud computing is a framework for fetching customized access to various ICT resources like storage, communication networks, processing units, applications, software, and online services (Paika and Bhambri, 2013). The fundamental aim of the cloud models is to enhance the availability of the resources at affordable cost. The cloud framework comprises five imperative characteristics, three service delivery models, and four deployment models. The delivery models are further classified as follows:

- **Software as a Service (SaaS)**

In this model, a variety of software/services are mounted on the cloud. These software and services are made available to the consumer on a rental basis and are accessed through Internet/email. SaaS provides instant on-demand services to the customer for different kinds of needs like business applications, software, enterprise resource planning (ERP), supply-chain management (SCM), and others. As SaaS and platform as a service (PaaS), both operate on top of infrastructure as a service (IaaS); any infringement in IaaS will surely impact the security mechanism of both delivery models. It depicts the high degree of dependence in the layered architecture, where, on one side, the layered framework strengthens the security and reliability of the model but comprises technical risks too.

- **PaaS**

This model enables the consumer to access a platform to deploy cloud framework and infrastructure using specific applications without acquiring a personal machine. PaaS basically describes the sharing of platforms and software like operating systems and customized application-specific frameworks. PaaS deploys cloud services considering the constraints of the resources. PaaS is usually opted for by consumers who are also using IaaS.

- **IaaS**

This service model provides an infrastructure for consumers on a pay-as-you-go basis. The various resources which become accessible to the consumers are computational resources, servers, and memory storage in cloud. Under IaaS, the service may be provided by an individual physical or virtual vendor or as a conflux of both. In a cloud-based system, the consumers can pay as per usage where all IaaS, PaaS, and SaaS provide layer-specific services in a row one after another.

IaaS facilitates the consumers with a platform where they can have control over physical resources like communication networks, storage devices, and other computer devices (Bhambri and Gupta, 2013). The second cloud layer, that is, PaaS, provides the access to the cloud resources by handing over control at the system administration level, like operating systems and various runtime services. SaaS is the topmost service dependent on end-user applications. Under this model, most of the control is kept by the cloud owner (Kaur and Bhambri, 2015). IBM has depicted the control of the cloud services with a diagram shown in Figure 1.7.

1.1.3.2 Cloud Deployment Models

The cloud deployment models are based on the infrastructure of the communication networks on which it is hosted. They are characterized by the location of the cloud infrastructure (Chopra et al., 2011). The cloud deployment models are categorized as public cloud, private cloud, community cloud, and hybrid cloud:

- **Public Cloud**. The most frequently used cloud is the public cloud. In this model, both individuals and companies can manage their accounts when subscribing to cloud services.

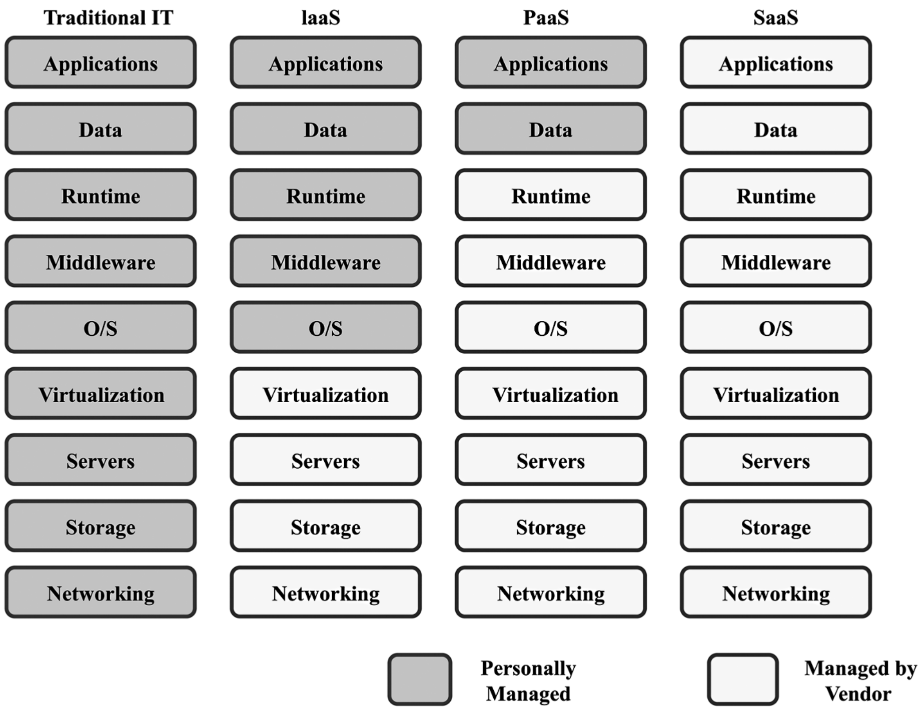

FIGURE 1.7
IBM cloud services: administered by owner and consumer.

- **Private Cloud**. In this model, the cloud services are deployed and administered using a private network for an individual. This type of model can cater to an individual organization or be used in collaboration with some other service provider to cater to the needs of an entity.
- **Community Cloud**. This type of cloud is usually shared by entities with similar requirements. It is used to spread the common policies, employment, promoting the optimal structure and analogy of appliances. Different educational institutions, business organizations, scientific research bodies, health organizations, and many other associations that usually share common interests mostly use community clouds for the exchange of information.
- **Hybrid Cloud**. This model is the confluence of various previously discussed models. It may be the amalgamation of different private clouds, private and public clouds, and so on. This type of cloud model caters the needs of almost all types of consumers. The major challenge faced in this framework is security.

1.1.3.3 Cloud Applications Areas

Today, most organizations, irrespective of their size, benefit from a cloud framework for deploying their data, applications, and software over the cloud. The major objective behind this adoption is to reduce the cost associated with the purchase and maintenance of ICT infrastructure (Harleen, 2016). Cloud computing offers plenty of services to consumers fundamentally through remote access to distant resources. The most prominent cloud computing features are

- on-demand services,
- pooling of available resources,
- better quality services,
- expanded network visibility and availability, and
- quick scalability.

With the preceding features, cloud computing plays a very significant role in different applications domains:

- E-commerce
- Business applications
- Data storage, management, and access services
- Multimedia applications
- Privacy and data security
- Geographical maps and location-dependent applications, among others

Along with this, cloud computing model is used in a number of other domains too.

1.2 Integrated Framework

The three dominating technologies of the current time, that is, big data, the IoT, and cloud computing, are playing significant roles in many real-life applications. The prominent areas that benefit from these technologies are smart cities, Industry 4.0, agriculture, healthcare, transportation, weather forecasting, banking and insurance, and many more (Arunachalam et al., 2021). Applying these technologies in various applications is not done in isolation; the functions integrate with each other, as shown in Figure 1.8. Today,

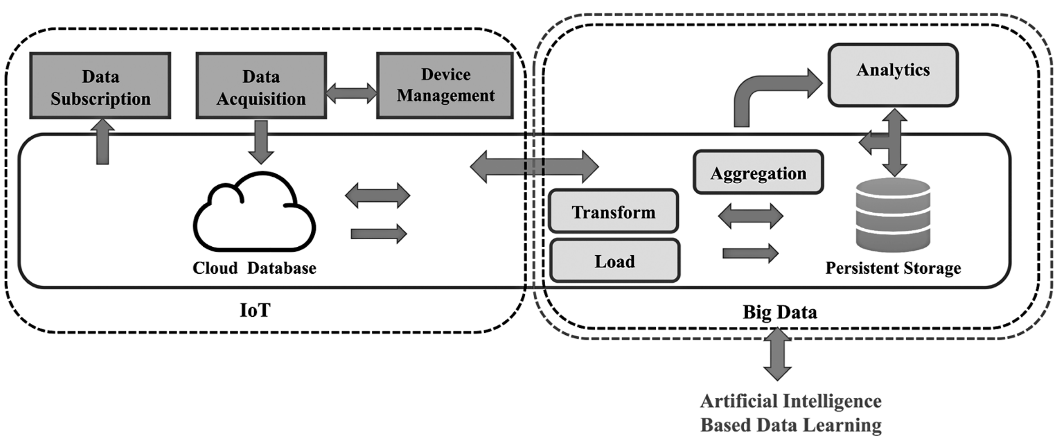

FIGURE 1.8
The integration of the IoT, big data, and cloud computing to realize real-life applications.