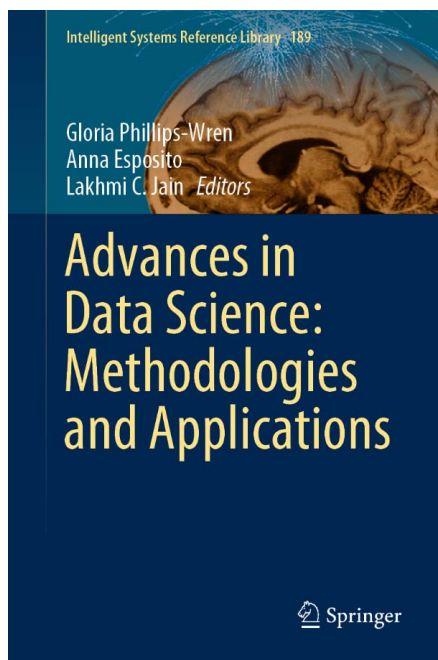


## Book Review



### *Advances in Data Science: Methodologies and Applications*

Editors: Phillips-Wren, Gloria; Esposito, Anna; Jain, Lakhmi C.

Springer 2021, pp. 333, ISBN 978-3-030-51870-7

Vol. 189 in Springer Book Series entitled **Intelligent Systems Reference Library**

<https://www.springer.com/series/8578>

The term 4<sup>th</sup> *Industrial Revolution* (4<sup>th</sup> *IR*) is used to describe the new era into which human civilization is entering [1] and which transforms human societal structure into, so-called, *Society 5.0* [2,3]. Previous IRs in human history have depended on coal and water to drive steam engines and oil to drive internal combustion engines (1<sup>st</sup> IR), the Mass Production Line (2<sup>nd</sup> IR) and advances in Electronics and Information Technology (3<sup>rd</sup> IR). In contrast, the 4<sup>th</sup> IR is characterized by extensive synergy between physical, digital, biological and energy sciences and technologies, with all these scientific disciplines having *data* as their common and uniting thread [4,5].

Indeed, in the recent years, humanity is witnessing a variety of both traditional and novel sensing devices (e.g., wearables) collecting huge amounts of data, a fact that corroborates the eloquent statement that “*Data in the 21st Century is like Oil in the 18th Century*” [6,7]. Good quality data are highly desirable and extensively exploited at all levels of economic activity and affect all aspects of society, offering new opportunities in ways that were unavailable, even unimaginable, only two decades ago and could not be supported by traditional data.

Extensive data collection and information extraction pose dangers to humanity and personal privacy and care must be taken to manage the societal implications of these new technologies [8]. However and despite these risks, nearly all aspects of society are expected to benefit significantly, as significant progress is anticipated in economic development, healthcare and caregiving, famine, poverty and inequality reduction, disaster prevention and efficient energy production and distribution [2,3,5].

Besides traditional economic figures, all sorts of additional and diverse data are currently collected and, to name just few, include data related to mobility, traffic, GPS, outdoors activity, travel, planning, user habits, shopping, customer behavior, entertainment, social interactions, education, medicine, biology, health, forensics, or energy consumption [5,7].

Consequently, the amount of the data is huge. To get a feeling of it and of its rate of increase, it suffices to look at recent statistics and near-future predictions by the market intelligence company IDC [9]. Indeed, the total data created, captured or replicated (but not necessarily stored) in 2018 was estimated to 18 zettabytes, i.e. to  $1.8 \times 10^{13}$  Gbytes [9]. By comparison, a good laptop in the year 2021 carries a disk storage capacity of 128–1.000 Gbytes. Furthermore, it is projected by IDC that by the year 2025, the World’s data will grow to ten times the 2018 figure, i.e. it will rise to about  $1.8 \times 10^{14}$  Gbytes [9]. As eloquently illustrated by IDC, a data of that size would require a CD stack long enough to circle the Earth 222 times and 1.8 billion years to download at the currently available internet speeds [9].

This data needs to be efficiently *managed*. Data management includes storage, transmission, processing and, eventually, conversion of the data into information, knowledge and wisdom [5]. Naturally, a number of scientific and technological issues and questions arise, including:

- What data needs to be collected that carry the desired information?
- How should data be organized and structured to facilitate its management?
- What is the optimal way to process data to extract the desired information?
- What are effective mechanisms to protect data from intruders, especially when dealing with highly sensitive data?
- Which networks and protocols need to be developed to transmit data efficiently?
- What potential legal issues arise when dealing with data, and how can they be legislated?

To address these and other related challenges, new and *interdisciplinary* approaches need to be devised, which have given rise to a multidisciplinary scientific and technological field named *Data Science*. Data Science emerges as a discipline that combines research and expertise from Computer and Information Sciences and Engineering, Artificial Intelligence, Databases and Information Systems, Mathematics, Statistics, Software Engineering, Computer Communication and Networks to answer to challenges with regard to the management of data.

Professors Phillips-Wren, Esposito and Jain have just delivered a high quality book in which they attempt to highlight a broad spectrum of current aspects of both methodologies and application areas of Data Science. The book consists of an introductory chapter and an additional 14 chapters. In the introductory chapter, the editors guide the reader, especially the newcomer to Data Science, to all the concepts, the approaches and some of the open challenges of this fascinating discipline. The remaining 14 chapters have been written by world-recognized researchers who have made significant contributions to their corresponding fields. In these chapters, Data Science applications are presented in an impressive diversity of areas, including *neural and fuzzy processing* (2 chapters), *emotion recognition* (2 chapters), *smart city problems* (2 chapters), *speech and music processing* (2 chapters), *sentiment analysis in tweets* (1 chapter), *abnormal behavior, attention assessment and fall risk assessment* (3 chapters), *data science in robotics* (1 chapter), and *morality linked to neurophysiology* (1 chapter).

Editing a book with the purpose to expose Data Science Methodologies and current Applications to their full extents is not an easy task, as the editors are required to be deeply involved in related research and well-versed in the corresponding state-of-the-art. Professors Phillips-Wren, Esposito and Jain have managed to impress us with the breadth of the topics covered in their book. They are also impressing us with the depth in which each topic is covered. The reader may choose to either concentrate on specific chapters in the book that are of particular interest to him/her or to read through several and diverse chapters in the book and get informed on broader aspects of Data Science Methodologies and Applications. Thus, the book is useful to both the specialized researcher seeking information on specific sub-areas within Data Science or to the newcomer who seeks to get involved in Data Science. I am confident that both types of readers will benefit greatly from the book.

In summary, the editors are warmly congratulated for their superb work. Their book is highly and unreservedly recommended to professors, graduate students, practitioners and other specialists in Data Science and to general readers, all of whom are expected to benefit greatly from it in their researches.

Prof.-Dr. George A. Tsihrintzis  
Department of Informatics  
University of Piraeus  
Greece

## References

- [1] Schwabd, K. The fourth industrial revolution – what it means and how to respond. *Foreign Affairs* (2015) <https://www.foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution>.
- [2] From Industry 4.0 to Society 5.0: the big societal transformation plan of Japan <https://www.i-scoop.eu/industry-4-0/society-5-0/>.
- [3] Society 5.0 [https://www8.cao.go.jp/cstp/english/society5\\_0/index.html](https://www8.cao.go.jp/cstp/english/society5_0/index.html).
- [4] <https://en.wikipedia.org/wiki/Data>.
- [5] Tsihrintzis, G.A., Sotiropoulos, D.N., Jain, L.C. (Eds.). *Machine Learning Paradigms – Advances in Data Analytics*. In: *Intelligent Systems Reference Library Book Series*, vol. 149. Springer (2019).
- [6] Toonders, J. Data is the new oil of the digital economy. *Wired* <https://www.wired.com/insights/2014/07/data-new-oil-digital-economy/>.
- [7] <https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource>.
- [8] Bostrom, N., Čirković, M.M. (eds.). *Global Catastrophic Risks*, Oxford University Press, 2008.
- [9] <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>.