

# **Studies in Big Data**

Volume 42

## **Series editor**

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland  
e-mail: [kacprzyk@ibspan.waw.pl](mailto:kacprzyk@ibspan.waw.pl)

The series “Studies in Big Data” (SBD) publishes new developments and advances in the various areas of Big Data- quickly and with a high quality. The intent is to cover the theory, research, development, and applications of Big Data, as embedded in the fields of engineering, computer science, physics, economics and life sciences. The books of the series refer to the analysis and understanding of large, complex, and/or distributed data sets generated from recent digital sources coming from sensors or other physical instruments as well as simulations, crowd sourcing, social networks or other internet transactions, such as emails or video click streams and others. The series contains monographs, lecture notes and edited volumes in Big Data spanning the areas of computational intelligence including neural networks, evolutionary computation, soft computing, fuzzy systems, as well as artificial intelligence, data mining, modern statistics and operations research, as well as self-organizing systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

More information about this series at <http://www.springer.com/series/11970>

Ali Emrouznejad · Vincent Charles  
Editors

# Big Data for the Greater Good

 Springer

*Editors*

Ali Emrouznejad  
Operations and Information Management,  
Aston Business School  
Aston University  
Birmingham  
UK

Vincent Charles  
Buckingham Business School  
University of Buckingham  
Buckingham  
UK

ISSN 2197-6503

Studies in Big Data

ISBN 978-3-319-93060-2

<https://doi.org/10.1007/978-3-319-93061-9>

ISSN 2197-6511 (electronic)

ISBN 978-3-319-93061-9 (eBook)

Library of Congress Control Number: 2018942927

© Springer International Publishing AG, part of Springer Nature 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

Today, individuals and organizations are changing the world with Big Data. Data has become a new source of immense economic and social value. Advances in data mining and analytics and the massive increase in computing power and data storage capacity have expanded, by orders of magnitude, the scope of information available to businesses, government, and individuals. Hence, the explosive growth in data and Big Data Analytics has set the next frontier for innovation, competition, productivity, and well-being in almost every sector of our society, from industry to academia and the government. From ROI calculations, revenue impact analyses, and customer insights, to earthquake prediction and environmental scanning, suicide and crime prevention, fraud detection, better medical treatments, and poverty reduction—this is a reminder of all the good that can come from Big Data. But driving knowledge and value from today’s mountains of data also brings policies related to ethics, privacy, security, intellectual property, and liability to the forefront, as the main concern on the agenda of policy makers. To help Big Data deliver its societal promises will require revised principles of monitoring and transparency, thus new types of expertise and institutions.

The present book titled ‘Big Data for the Greater Good’ brings together some of the fascinating uses, thought-provoking changes, and biggest challenges that Big Data can convey to our society. Along theory and applications, the book compiles the authors’ experiences so that these may be aggregated for a better understanding. This book should be of interest to both researchers in the field of Big Data and practitioners from various fields who intend to apply Big Data technologies to improve their strategic and operational decision-making process. Finally, we hope that this book is an invitation to more intensive reflection on Big Data as a source for the common and greater good.

The book is well organized in nine chapters, contributed by authors from all around the globe: Austria, Brazil, Czech Republic, Denmark, France, Germany, Greece, Italy, The Netherlands, UK, and the USA.

Chapter 1 provides an introduction to Big Data, as well as the role it currently plays and could further play in achieving outcomes for the ‘Greater Good’. This chapter discusses the main literature on Big Data for the Greater Good for interested

readers coming from different disciplines. Chapter 2 looks at the means that can be used to extract value from Big Data, and to this end, it explores the intersection between Big Data Analytics and Ethnography. The authors advance that the two approaches to analysing data can complement each other to provide a better sense of the realities of the contexts researched. As such, Big Data Analytics and Ethnography together can assist in the creation of practical solutions that yield a greater societal value.

Chapter 3 complements the first two chapters and provides a global overview of what Big Data is, uncovers its origins, as well as discusses the various definitions of the same, along with the technologies, analysis techniques, issues, challenges, and trends related to Big Data. The chapter further examines the role and profile of the Data Scientist, by means of taking into consideration aspects such as functionality, academic background, and required skills. The authors inspect how Big Data is leading the world towards a new way of social construction, consumption, and processes.

Chapters 4 and 5 deal with the application of Big Data in the field of health care. Chapter 4 focuses on the use of data for in-patient care management in high-acuity spaces, such as operating rooms, intensive care units, and emergency departments. In addition, it discusses a variety of mathematical techniques to assist in managing and mitigating non-actionable alarm signals on monitored patients. Chapter 5 shows how the combination of novel biofeedback-based treatments producing large data sets with Big Data and Cloud-Dew technology can contribute to the greater good of patients with brain disorders. This approach is aimed at optimizing the therapy with regard to the current needs of the patient, improving the efficiency of the therapeutic process, and preventing patient from overstressing during the therapy. The preliminary results are documented using a case study confirming that the approach offers a viable way towards the greater good of the patients.

In the context of increased efforts dedicated to research on Big Data in agricultural and food research, Chap. 6 focuses on the presentation of an innovative and integrated Big Data e-infrastructure solution (AGINFRA+) that aims to enable the sharing of data, algorithms, and results in a scalable and efficient manner across different but interrelated research studies, with an application to the agriculture and food domain. The authors present three use cases for performing agri-food research with the help of the mentioned e-infrastructure. The chapter also analyses the new challenges and directions that will potentially arise for agriculture and food management and policing.

Chapter 7 aims to demonstrate the benefits of data collection and analysis to the maintenance and planning of current and future low-voltage networks. The authors review several agent-based modelling techniques and further present two case studies wherein these techniques are applied to energy modelling on a real low-voltage network in the UK. The chapter shows that Big Data Analytics of supply and demand can contribute to a more efficient usage of renewable sources, which will also result in cutting down carbon emissions.

It is common nowadays for customers to record their experiences in the form of online reviews and blogs. In Chap. 8, the authors examine the case of customer feedback at local, state, and national parks in the New York State Park System. The chapter discusses the design, development, and implementation of software systems that can download, organize, and analyse the voluminous text from the online reviews, analyse them using Natural Language Processing algorithms to perform sentiment analysis and topic modelling, and finally provide facility managers actionable insights to improve visitor experience.

Finally, Chap. 9 discusses the issue of data privacy, which has proven to be a challenge in the Big Data age, but which, nevertheless, can be addressed through modern cryptography. There are two types of solutions to tackle such problematic: one that makes data itself anonymous, but which degrades the value of the data, and the other one that uses Computation on Encrypted Data. This chapter introduces the latter and describes three prototype and pilot applications of the same within privacy-preserving statistics. The applications originate from R&D projects and collaborations between the Danish financial sector and Statistics Denmark.

The chapters contributed to this book should be of considerable interest and provide our readers with informative reading.

Birmingham, UK  
Buckingham, UK  
July 2018

Ali Emrouznejad  
Vincent Charles

# Contents

<b>1</b>	<b>Big Data for the Greater Good: An Introduction</b> . . . . .	<b>1</b>
	Vincent Charles and Ali Emrouznejad	
<b>2</b>	<b>Big Data Analytics and Ethnography: Together for the Greater Good</b> . . . . .	<b>19</b>
	Vincent Charles and Tatiana Gherman	
<b>3</b>	<b>Big Data: A Global Overview</b> . . . . .	<b>35</b>
	Celia Satiko Ishikiriyama and Carlos Francisco Simoes Gomes	
<b>4</b>	<b>Big Data for Predictive Analytics in High Acuity Health Settings</b> . . . . .	<b>51</b>
	John Zaleski	
<b>5</b>	<b>A Novel Big Data-Enabled Approach, Individualizing and Optimizing Brain Disorder Rehabilitation</b> . . . . .	<b>101</b>
	Marketa Janatova, Miroslav Uller, Olga Stepankova, Peter Brezany and Marek Lenart	
<b>6</b>	<b>Big Data in Agricultural and Food Research: Challenges and Opportunities of an Integrated Big Data E-infrastructure</b> . . . . .	<b>129</b>
	Pythagoras Karampiperis, Rob Lokers, Pascal Neveu, Odile Hologne, George Kakalettris, Leonardo Candela, Matthias Filter, Nikos Manouselis, Maritina Stavarakaki and Panagiotis Zervas	
<b>7</b>	<b>Green Neighbourhoods: The Role of Big Data in Low Voltage Networks' Planning</b> . . . . .	<b>151</b>
	Danica Vukadinović Greetham and Laura Hattam	



<b>8</b>	<b>Big Data Improves Visitor Experience at Local, State, and National Parks—Natural Language Processing Applied to Customer Feedback</b> .....	171
	Hari Prasad Udyapuram and Srinagesh Gavirneni	
<b>9</b>	<b>Big Data and Sensitive Data</b> .....	183
	Kurt Nielsen	