BIG DATA IN PRACTICE

"A comprehensive compendium of why, how, and to what effects Big Data analytics are used in today's world"
James Kobielus, Big Data Evangelist at IBM

HOW 45 SUCCESSFUL COMPANIES USED BIG DATA ANALYTICS TO DELIVER EXTRAORDINARY RESULTS

BERNARD MARR

WILEY

“Amazing. That was my first word, when I started reading this book. Fascinating was the next. Amazing, because once again, Bernard masterfully takes a complex subject, and translates it into something anyone can understand. Fascinating because the detailed real-life customer examples immediately inspired me to think about my own customers and partners, and how they could emulate the success of these companies. Bernard's book is a must have for all Big Data practitioners and Big Data hopefuls!”

Shawn Ahmed, Senior Director, Business Analytics and IoT at Splunk

“Finally a book that stops talking theory and starts talking facts. Providing real-life and tangible insights for practices, processes, technology and *teams* that support Big Data, across a portfolio of organizations and industries. We often think Big Data is big business and big cost, however some of the most interesting examples show how small businesses can use smart data to make a real difference. The businesses in the book illustrate how Big Data is fundamentally about the customer, and generating a data-driven customer strategy that influences both staff and customers at every touch point of the customer journey.”

Adrian Clowes, Head of Data and Analytics at Center Parcs UK

“*Big Data in Practice* by Bernard Marr is the most complete book on the Big Data and analytics ecosystem. The many real-life examples make it equally relevant for the novice as well as experienced data scientists.”

Fouad Bendris, Business Technologist, Big Data Lead at Hewlett Packard Enterprise

“Bernard Marr is one of the leading authors in the domain of Big Data. Throughout *Big Data in Practice* Marr generously shares some of his keen insights into the practical value delivered to a huge range of different businesses from their Big Data initiatives. This fascinating book provides excellent clues as to the secret sauce required in order to successfully deliver competitive advantage through Big Data analytics. The logical structure of the book means that it is as easy to consume in one sitting as it is to pick up from time to time. This is a must-read for any Big Data sceptics or business leaders looking for inspiration.”

Will Cashman, Head of Customer Analytics at AIB

“The business of business is now data! Bernard Marr's book delivers concrete, valuable, and diverse insights on Big Data use cases, success stories, and lessons learned from numerous business domains. After diving into this book, you will have all the knowledge you need to crush the Big Data hype machine, to soar to new heights of data analytics ROI, and to gain competitive advantage from the data within your organization.”

Kirk Borne, Principal Data Scientist at Booz Allen Hamilton, USA

“Big Data is disrupting every aspect of business. You're holding a book that provides powerful examples of how companies strive to defy outmoded business models and design new ones with Big Data in mind.”

Henrik von Scheel, Google Advisory Board Member

“Bernard Marr provides a comprehensive overview of how far Big Data has come in past years. With inspiring examples he clearly shows how large, and small, organizations can benefit from Big Data. This book is a must-read for any organization that wants to be a data-driven business.”

Mark van Rijmenam, Author *Think Bigger* and Founder of Datafloq

“This is one of those unique business books that is as useful as it is interesting. Bernard has provided us with a unique, inside look at how leading organizations are leveraging new technology to deliver real value out of data and completely transforming the way we think, work, and live.”

Stuart Frankel, CEO at Narrative Science Inc.

“Big Data can be a confusing subject for even sophisticated data analysts. Bernard has done a fantastic job of illustrating the true business benefits of Big Data. In this book you find out succinctly how leading companies are getting real value from Big Data – highly recommended read!'

Arthur Lee, Vice President of Qlik Analytics at Qlik

“If you are searching for the missing link between Big Data technology and achieving business value – look no further! From the world of science to entertainment, Bernard Marr delivers it – and, importantly, shares with us the recipes for success.”

Achim Granzen, Chief Technologist Analytics at Hewlett Packard Enterprise

“A comprehensive compendium of why, how, and to what effects Big Data analytics are used in today's world.”

James Kobielus, Big Data Evangelist at IBM

“A treasure chest of Big Data use cases.”

Stefan Groschupf, CEO at Datameer, Inc.

BIG DATA IN PRACTICE

HOW 45 SUCCESSFUL COMPANIES USED BIG DATA ANALYTICS TO DELIVER EXTRAORDINARY RESULTS

BERNARD MARR

WILEY

This edition first published 2016

Registered office
John Wiley and Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Library of Congress Cataloging-in-Publication Data is available

A catalogue record for this book is available from the British Library.

ISBN 978-1-119-23138-7 (hbk) ISBN 978-1-119-23139-4 (ebk)
ISBN 978-1-119-23141-7 (ebk) ISBN 978-1-119-27882-5 (ebk)

Cover Design: Wiley
Cover Image: © vs148/Shutterstock

This book is dedicated to the people who mean most to me:
My wife
Claire and our three children Sophia, James and Oliver.

CONTENTS

INTRODUCTION

We are witnessing a movement that will completely transform any part of business and society. The word we have given to this movement is Big Data and it will change everything, from the way banks and shops operate to the way we treat cancer and protect our world from terrorism. No matter what job you are in and no matter what industry you work in, Big Data will transform it.

Some people believe that Big Data is just a big fad that will go away if they ignore it for long enough. It won't! The hype around Big Data and the name may disappear (which wouldn't be a great loss), but the phenomenon will stay and only gather momentum. What we call Big Data today will simply become the new normal in a few years' time, when all businesses and government organizations use large volumes of data to improve what they do and how they do it.

I work every day with companies and government organizations on Big Data projects and thought it would be a good idea to share how Big Data is used today, across lots of different industries, among big and small companies, to deliver real value. But first things first, let's just look at what Big Data actually means.

What Is Big Data?

Big Data basically refers to the fact that we can now collect and analyse data in ways that was simply impossible even a few years ago. There are two things that are fuelling this Big Data movement: the fact we have more data on anything and our improved ability to store and analyse any data.

More Data On Everything

Everything we do in our increasingly digitized world leaves a data trail. This means the amount of data available is literally exploding. We have created more data in the past two years than in the entire previous history of mankind. By 2020, it is predicted that about 1.7 megabytes of new data will be created every second, for every human being on the planet. This data is coming not just from the tens of millions of messages and emails we send each other every second via email, WhatsApp, Facebook, Twitter, etc. but also from the one trillion digital photos we take each year and the increasing amounts of video data we generate (every single minute we currently upload about 300 hours of new video to YouTube and we share almost three million videos on Facebook). On top of that, we have data from all the sensors we are now surrounded by. The latest smartphones have sensors to tell where we are (GPS), how fast we are moving (accelerometer), what the weather is like around us (barometer), what force we are using to press the touch screen (touch sensor) and much more. By 2020, we will have over six billion smartphones in the world – all full of sensors that collect data. But not only our phones are getting smart, we now have smart TVs, smart watches, smart meters, smart kettles, fridges, tennis rackets and even smart light bulbs. In fact, by 2020, we will have over 50 billion devices that are connected to the Internet. All this means that the amount of data and the variety of data (from sensor data, to text and video) in the world will grow to unimaginable levels.

Ability To Analyse Everything

All this Big Data is worth very little unless we are able to turn it into insights. In order to do that we need to capture and analyse the data. In the past, there were limitations to the amount of data that could be stored in databases – the

more data there was, the slower the system became. This can now be overcome with new techniques that allow us to store and analyse data across different databases, in distributed locations, connected via networks. So-called distributed computing means huge amounts of data can be stored (in little bits across lots of databases) and analysed by sharing the analysis between different servers (each performing a small part of the analysis).

Google were instrumental in developing distributed computing technology, enabling them to search the Internet. Today, about 1000 computers are involved in answering a single search query, which takes no more than 0.2 seconds to complete. We currently search 3.5 billion times a day on Google alone.

Distributed computing tools such as Hadoop manage the storage and analysis of Big Data across connected databases and servers. What's more, Big Data storage and analysis technology is now available to rent in a software-as-a-service (SAAS) model, which makes Big Data analytics accessible to anyone, even those with low budgets and limited IT support.

Finally, we are seeing amazing advancements in the way we can analyse data. Algorithms can now look at photos, identify who is on them and then search the Internet for other pictures of that person. Algorithms can now understand spoken words, translate them into written text and analyse this text for content, meaning and sentiment (e.g. are we saying nice things or not-so-nice things?). More and more advanced algorithms emerge every day to help us understand our world and predict the future. Couple all this with machine learning and artificial intelligence (the ability of algorithms to learn and make decisions independently) and you can hopefully see that the

developments and opportunities here are very exciting and evolving very quickly.

Big Data Opportunities

With this book I wanted to showcase the current state of the art in Big Data and provide an overview of how companies and organizations across all different industries are using Big Data to deliver value in diverse areas. You will see I have covered areas including how retailers (both traditional bricks ’n’ mortar companies as well as online ones) use Big Data to predict trends and consumer behaviours, how governments are using Big Data to foil terrorist plots, even how a tiny family butcher or a zoo use Big Data to improve performance, as well as the use of Big Data in cities, telecoms, sports, gambling, fashion, manufacturing, research, motor racing, video gaming and everything in between.

Instead of putting their heads in the sand or getting lost in this startling new world of Big Data, the companies I have featured here have figured out smart ways to use data in order to deliver strategic value. In my previous book, *Big Data: Using SMART Big Data, Analytics and Metrics to Make Better Decisions and Improve Performance* (also published by Wiley), I go into more detail on how any company can figure out how to use Big Data to deliver value.

I am convinced that Big Data, unlike any other trend at the moment, will affect everyone and everything we do. You can read this book cover to cover for a complete overview of current Big Data use cases or you can use it as a reference book and dive in and out of the areas you find most interesting or are relevant to you or your clients. I hope you enjoy it!