

Vincenzo Morabito

Big Data and Analytics

Strategic and Organizational Impacts

Big Data and Analytics

Vincenzo Morabito

Big Data and Analytics

Strategic and Organizational Impacts

Vincenzo Morabito
Department of Management and Technology
Bocconi University
Milan
Italy

ISBN 978-3-319-10664-9 ISBN 978-3-319-10665-6 (eBook)
DOI 10.1007/978-3-319-10665-6

Library of Congress Control Number: 2014958989

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2015

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.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

Foreword

Few organizations understand how to extract insights and value from the recent explosion of “Big Data.” With a billion plus users on the online social graph doing what they like to do and leaving a digital trail, and with trillions of sensors now being connected in the so-called Internet of Things, organizations need clarity and insights into what lies ahead in deploying these capabilities. While academic scholars are just beginning to appreciate the power of big data analytics and new media to open up a fascinating array of questions from a host of disciplines, the practical applicability of this is still lacking. Big data and analytics touches multiple disciplines ranging from sociology, psychology, and ethics to marketing, statistics, and economics, as well as law and public policy. If harnessed correctly it has the potential to solve a variety of business and societal problems.

This book aims to develop the strategic and organizational impacts of Big Data and analytics for today’s digital business competition and innovation. Written by an academic, the book has nonetheless the main goal to provide a toolbox suitable to be useful to business practice and know-how. To this end Vincenzo as in his former books has structured the content into three parts that guide the reader through how to control and govern the innovation potential of Big Data and Analytics. First, the book focuses on *Strategy* (Part I), analyzing how Big Data and analytics impact on private and public organizations, thus, examining the implications for competitive advantage as well as for government and education. The last chapter provides an overview of Big Data business models, creating a bridge to the content of Part II, which analyzes the managerial challenges of Big Data and analytics governance and evaluation. The conclusive chapter of Part II introduces the reader to the challenges of managing change required by an effective use and absorption of Big Data and analytics, actually trying to complement IT and non-IT managers’ perspective. Finally, Part III discusses through structured and easy to read forms a set of cases of Big Data and analytics initiatives in practice at a global level in 2014.

Use this book as a guide to design your modern analytics-enabled organization. Do not be surprised if it resembles a large-scale real-world laboratory where employees design and conduct experiments and collect the data needed to obtain answers to a variety of questions, from peer influence effects, the influence of

dynamic ties, pricing of digital media, anonymity in online relationships, to designing next-generation recommender systems and enquiries into the changing preference structures of Generation Y and Z consumers. This is a bold new frontier and it is safe to say we ain't seen nothing yet.

Ravi Bapna

Preface

Notwithstanding the interest and the hype that surround Big Data as a key trend as well the claimed business potentiality that it may offer the coupling with a new breed of analytics, the phenomenon has been yet not fully investigated from a strategic and organizational perspective. Indeed, at the moment of writing this book, apart from a series of articles that appeared on the Harvard Business review by McAfee and Brynjolfsson (2012) and on MIT Sloan Management Review by Lavalley et al. (2011) and Davenport et al. (2012), most of the published monographic contributions concern technical, computational, and engineering facets of Big Data and analytics, or oriented toward high-level societal as well as general audience business analyses.

An early joint academics-practitioners effort to provide a unified and comprehensive perspective has been carried out by the White Paper resulting from joint multidisciplinary contributions of more than 130 participants from 26 countries at the World Summit on Big Data and Organization Design held in Paris at the Université Panthéon-Sorbonne during May 16–17, 2013 (Burton et al. 2014). However, it is worth to be mentioned that since 2013 new editorial initiatives have been launched such as, e.g., the Big Data journal (Dumbill 2013). Thus, following up the insights discussed in (Morabito 2014), the present book aims to fill the gap, providing a strategic and organizational perspective on Big Data and analytics, identifying the challenges, ideas, and trends that may represent “food for thought” to practitioners. Accordingly, each topic considered will be analyzed in its technical and managerial aspects, also through the use of case studies and examples. Thus, while relying on academic production as well, the book aims to describe problems from the viewpoints of managers, adopting a clear and easy-to-understand language, in order to capture the interests of top managers and graduate students. Consequently, this book is unique for its intention to synthesize, compare, and comment on major challenges and approaches to Big Data and analytics, being a simple yet ready to consult toolbox for both managers and scholars.

In what follows we provide a brief overview, based on our previous work as well (Morabito 2014), on Big Data drivers and characteristics suitable to introduce their discussion also with regard to analytics in the further chapters of this book, whose outline concludes this introduction.

Big Data Drivers and Characteristics

The spread of social media as a main driver for innovation of products and services and the increasing availability of unstructured data (images, video, audio, etc.) from sensors, cameras, digital devices for monitoring supply chains and stocking in warehouses (i.e., what is actually called *internet of things*), video conferencing systems and voice over IP (VOIP) systems, have contributed to an unmatched availability of information in rapid and constant growth in terms of volume. As for these issues, an interesting definition of “Big Data” has been provided by Edd Dumbill in 2013:

Big data is data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or doesn't fit the structures of your database architectures. To gain value from this data, you must choose an alternative way to process it (Dumbill 2013).

As a consequence of the above scenario and definition, the term “Big Data” is dubbed to indicate the challenges associated with the emergence of data sets whose size and complexity require companies to adopt new tools and models for the management of information. Thus, Big Data require new capabilities (Davenport and Patil 2012) to control external and internal information flows, transforming them into strategic resources to define strategies for products and services that meet customers' needs, increasingly informed and demanding.

However, Big Data computational as well as technical challenges call for a radical change to business models and human resources in terms of information orientation and a unique valorization of a company information asset for investments and support for strategic decisions. At the state of the art the following four dimensions are recognized as characterizing Big Data (IBM; McAfee and Brynjolfsson 2012; Morabito 2014; Pospiech and Felden 2012):

- **Volume:** the first dimension concerns the unmatched quantity of data actually available and storable by businesses (terabytes or even petabytes), through the Internet: for example, 12 terabytes of Tweets are created everyday into improved product sentiment analysis (IBM).
- **Velocity:** the second dimension concerns the dynamics of the volume of data, namely the time-sensitive nature of Big Data, as the speed of their creation and use is often (nearly) real-time.
- **Variety:** the third dimension concerns type of data actually available. Besides, structured data traditionally managed by information systems in organizations, most of the new breed encompasses semi-structured and even unstructured data, ranging from text, log files, audio, video, and images posted, e.g., on social networks to sensor data, click streams, e.g., from Internet of Things.
- **Accessibility:** the fourth dimension concerns the unmatched availability of channels a business may increase and extend its own data and information asset.
- It is worth noting that at the state of the art another dimension is actually considered relevant to Big Data characterization: **Veracity** concerns quality of data and trust of the data actually available at an incomparable degree of volume,

velocity, and variety. Thus, this dimension is relevant to a strategic use of Big Data and analytics by businesses, extending in terms of scale and complexity the issues investigated by information quality scholars (Huang et al. 1999; Madnick et al. 2009; Wang and Strong 1996), for enterprise systems mostly relying on traditional relational database management systems.

As for drivers, (Morabito 2014) identified cloud computing as a relevant one, besides social networks, mobile technologies, and Internet of Things (IoTs). As pointed out by Pospiech and Felden (2012), at the state of the art, cloud computing is considered a key driver of Big Data, for the growing size of available data requires scalable database management systems (DBMS). However, cloud computing faces IT managers and architects the choice of either relying on commercial solutions (mostly expensive) or moving beyond relational database technology, thus, identifying novel data management systems for cloud infrastructures (Agrawal et al. 2010, 2011). Accordingly, at the state of art *NoSQL* (Not Only SQL)¹ data storage systems have been emerging, usually not requiring fixed table schemas and not fully complying nor satisfying the traditional ACID (Atomicity, Consistency, Isolation, and Durability) properties. Among the programming paradigms for processing, generating, and analyzing large data sets, *MapReduce*² and the open source computing framework Hadoop have received a growing interest and adoption in both industry and academia.³

Considering *velocity*, there is a debate in academia about considering Big Data as encompassing both data “stocks” and “flows” (Davenport 2012). For example, at the state of the art Piccoli and Pigni (2013) propose to distinguish the elements of *digital data streams* (DDSs) from “big data”; the latter concerning static data that can be mined for insight. Whereas *digital data streams* (DDSs) are “dynamically evolving sources of data changing over time that have the potential to spur real-time action” (Piccoli and Pigni 2013). Thus, DDSs refer to streams of real-time information by mobile devices and IoTs, that have to be “captured” and analyzed real-time, provided or not they are stored as “Big Data”. The types of use of “big” DDSs may be classified according to those Davenport et al. (2012) have pointed out for Big Data applications to information flows:

¹ Several classifications of the NoSQL databases have been proposed in literature (Han et al. 2011). Here we mention *Key-/Value-Stores* (a map/dictionary allows clients to insert and request values per key) and *Column-Oriented databases* (data are stored and processed by column instead of row). An example of the former is *Amazon’s Dynamo*; whereas *HBase*, *Google’s Bigtable*, and *Cassandra* represent *Column-Oriented databases*. For further details we refer the reader to (Han et al. 2011; Strauch 2010).

² MapReduce exploit, on the one hand, (i) a *map function*, specified by the user to process a key/value pair and to generate a set of intermediate key/value pairs; on the other hand, (ii) a *reduce function* that merges all intermediate values associated with the same intermediate key (Dean and Ghemawat 2008). MapReduce has been used to complete rewrite the production indexing system that produces the data structures used for the Google web search service (Dean and Ghemawat 2008).

³ See for example how IBM has exploited/integrated Hadoop (IBM et al. 2011).

- *Support customer-facing processes:* e.g., to identify fraud or medical patients' health risk.
- *Continuous process monitoring:* e.g., to identify variations in customer sentiments toward a brand or a specific product/service or to exploit sensor data to detect the need for intervention on jet engines, data centers machines, extraction pump, etc.
- *Explore network relationships* on, e.g., LinkedIn, Facebook, and Twitter to identify potential threats or opportunities related to human resources, customers, competitors, etc.

As a consequence, we believe that the distinction between DDSs and Big Data is useful to point out a difference in scope and target of decision making, and analytic activities, depending on the business goals and the type of action required. Indeed, while DDSs may be suitable to be used for marketing and operations issues, such as customer experience management in mobile services, Big Data refer to the information asset an organization is actually able to archive, manage, and exploit for decision making, strategy definition, and business innovation (McAfee and Brynjolfsson 2012).

Having emphasized the specificity of DDS, we now focus on Big Data and analytics applications as also discussed in (Morabito 2014).

As shown in Fig. 1 they cover many industries, spanning from finance (banks and insurance), e.g., improving risk analysis and fraud management, to utility and manufacturing, with a focus on information provided by sensors and IoTs for improved quality control, operations or plants performance, and energy management. Moreover, marketing and service may exploit Big Data for increasing customer experience, through the adoption of social media analytics focused on sentiment analysis, opinion mining, and recommender systems.

As for public sector (further discussed in Chap. 2), Big Data represents an opportunity, on the one hand, e.g., for improving fraud detection as tax evasion control through the integration of a large number of public administration databases; on the other hand, for accountability and transparency of government and administrative activities, due to the increasing relevance and diffusion of *open data* initiatives, making accessible and available for further elaboration by constituencies of large public administration data sets (Cabinet Office 2012; Zuiderwijk et al. 2012), and participation of citizens to the policy making process, thanks to the shift of many government digital initiatives towards an open government perspective (Feller et al. 2011; Lee and Kwak 2012; Di Maio 2010; Nam 2012).

Thus, Big Data seem to have a strategic value for organizations in many industries, confirming the claim by Andrew McAfee and Brynjolfsson (2012) that data-driven decisions are better decisions, relying on evidence of (an unmatched amount of) facts rather than intuition by experts or individuals. Nevertheless, we believe that management challenges and opportunities of Big Data need further discussion and analyses, the state of the art currently privileging their technical facets and characteristics. That is the motivation behind this book, whose outline follows.

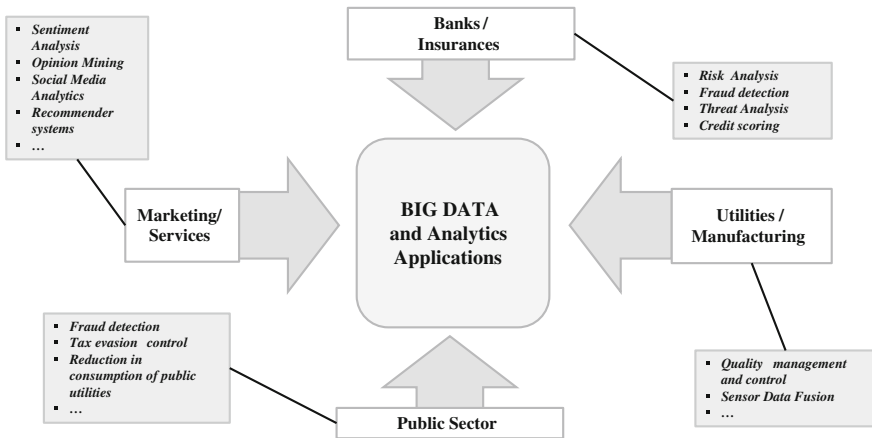


Fig. 1 Big Data Applications. Adapted from (Morabito 2014)

Outline of the Book

The book argument is developed along three main axes, likewise. In particular, we consider first (Part I) *Strategy* issues related to the growing relevance of Big Data and analytics for competitive advantage, also due their empowerment of activities such as, e.g., consumer profiling, market segmentation, and new products or services development. Furthermore, the different chapters will also consider the strategic impact of Big Data and analytics for innovation in domains such as government and education. A discussion of Big Data-driven Business Models conclude this part of the book. Subsequently, (Part II) considers *Organization*, focusing on Big Data and analytics challenges for governance, evaluation, and managing change for Big Data-driven innovation. Finally (Part III), the book will present and review case studies of Big Data *Innovation Practices* at the global level. Thus, Chap. 8 aims to discuss examples of Big Data and analytics applications in practice, providing fact-sheets suitable to build a “map” of 10 interesting digital innovations actually available worldwide. Besides an introduction to the factors considered in the choice of each innovation practice, a specific description of it will be developed. Finally, the conclusion will provide a summary of all arguments of the volume together with general managerial recommendations.

Vincenzo Morabito

References

- Agrawal, D., Das, S., El Abbadi, A.: Big data and cloud computing: new wine or just new bottles? *Proc. VLDB Endow.* 3, 1647–1648 (2010)
- Agrawal, D., Das, S., El Abbadi, A.: Big Data and Cloud Computing: Current State and Future Opportunities. *EDBT, ACM.* pp. 530–533. March 22–24, Sweden (2011)
- Burton, R.M., Mastrangelo, D., Salvador F.(eds.): Big data and organization design. *J. Organ. Des.* 3(1), (2014)
- Cabinet Office UK: Open Data White Paper—Unleashing the Potential. (2012)
- Davenport, T.H., Barth, P., Bean, R.: How “big data” is different. *MIT Sloan Manag. Rev.* 54(1), 43–46 (2012)
- Davenport, T.H., Patil, D.J.: Data scientist: The sexiest job of the 21st century. *Harv. Bus. Rev.* October, (2012)
- Dean, J., Ghemawat, S.: MapReduce: Simplified data processing on large clusters. *Commun. ACM.* 51(1), 1–13 (2008)
- Di Maio, A.: Gartner open government maturity model. Gartner (2010)
- Dumbill, E.: Making sense of big data (editorial). *Big Data.* 1(1), 1–2 (2013)
- Feller, J., Finnegan, P., Nilsson, O.: Open innovation and public administration: Transformational typologies and business model impacts. *Eur. J. Inf. Syst.* 20, 358–374 (2011). doi: [10.1057/Ejis.2010.65](https://doi.org/10.1057/Ejis.2010.65)
- Han, J., Haihong, E., Le, G., Du, J.: Survey on NoSQL database. 6th International Conference on Pervasive Computing and Applications (ICPCA). pp. 363–366 (2011). doi: [10.1109/ICPCA.2011.6106531](https://doi.org/10.1109/ICPCA.2011.6106531)
- Huang, K.T., Lee, Y., Wang, R.Y.: Quality, information and knowledge. Prentice-Hall, Inc (1999)
- IBM, Zikopoulos, P., Eaton, C.: Understanding Big Data: Analytics for Enterprise Class Hadoop and streaming data, 1st edn. McGraw-Hill Osborne Media (2011)
- IBM: What is big data?, <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>. Accessed 7 Jan 2015
- Lavalle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N.: Big Data, Analytics and the Path From Insights to Value. *MIT Sloan Manag. Rev.* 52(2), (2011)
- Lee, G., Kwak, Y.H.: An open government maturity model for social media-based public engagement. *Gov. Inf. Q.* 29(4), 492–503 (2012)
- Madnick, S.E., Wang, R.Y., Lee, Y.W., Zhu, H.: Overview and Framework for Data and Information Quality Research. *J. Data Inf. Qual.* 1, 1–22 (2009). doi: [10.1145/1515693.1516680](https://doi.org/10.1145/1515693.1516680)
- McAfee, A., Brynjolfsson, E.: Big data: The management revolution. *Harv. Bus. Rev.* 61–68 (2012)
- Morabito, V.: Big data. Trends and Challenges in Digital Business Innovation, pp. 3–21 Springer, Cham Heidelberg New York Dordrecht London (2014)
- Morabito, V.: Trends and Challenges in Digital Business Innovation. Springer (2014)
- Nam, T.: Citizens’ attitudes toward open government and government. *Int. Rev. Adm. Sci.* 78(2), 346–368 (2012)
- Piccoli, G., Pigni, F.: Harvesting external data: The potential of digital data streams. *MIS Q. Exec.* 12(1), 143–154 (2013)
- Pospiech, M., Felden, C.: Big data—A State-of-the-Art. *AMCIS 2012.* (2012)
- Strauch, C.: NoSQL databases. *Lect. Notes Stuttgart Media.* 1–8 (2010)
- Wang, R.Y., Strong, D.M.: Beyond accuracy: what data quality means to data consumers. *J. Manag. Inf. Syst.* 12(4), 5–33 (1996)
- Zuiderwijk, A., Janssen, M., Choenni, S.: Open Data Policies: Impediments and Challenges. 12th European Conference on eGovernment (ECEG 2012). pp. 794–802, Barcelona, Spain (2012)

Acknowledgments

This book is the result of the last two years of research, where several people are worth to be acknowledged for their support, useful comments and cooperation. A special mention to Prof. Vincenzo Perrone at Bocconi University, Prof. Vallabh Sambamurthy, Eli Broad Professor at Michigan State University, and Prof. Franco Fontana at LUISS University as main inspiration and mentors.

Moreover, I acknowledge Prof. Giuseppe Soda, Head of the Department of Management and Technology at Bocconi University, and all the other colleagues at the Department, in particular Prof. Arnaldo Camuffo, Prof. Anna Grandori, Prof. Severino Salvemini, and Prof. Giuseppe Airoidi, all formerly at the Institute of Organization and Information Systems at Bocconi University, who have created a rich and rigorous research environment where I am proud to work.

I acknowledge also some colleagues from other universities with whom I've had the pleasure to work, whose conversations, comments, and presentations provided precious insights for this book: among others, Prof. Anindya Ghose at New York University's Leonard N. Stern School of Business, Prof. Vijay Gurbaxani at University of California Irvine, Prof. Saby Mitra at Georgia Institute of Technology, Prof. Ravi Bapna at University of Minnesota Carlson School of Management, George Westerman at MIT Center for Digital Business, Stephanie Woerner at MIT Center for Information Systems Research, Prof. Ritu Agarwal at Robert H. Smith School of Business, Prof. Lynda Applegate at Harvard Business School, Prof. Omar El Sawy at Marshall School of Business, Prof. Marco de Marco at Università Cattolica del Sacro Cuore di Milano, Prof. Tobias Kretschmer, Head of Institute for Strategy, Technology and Organization of Ludwig Maximilians University, Prof. Marinos Themistocleous at the Department of Digital Systems at University of Piraeus, Prof. Chiara Francalanci at Politecnico di Milano, Wolfgang König at Goethe University, Luca Giustiniano at LUISS University, Prof. Zahir Irani at Brunel Business School, Prof. Sinan Aral at NYU Stern School of Business, Prof. Nitham Mohammed Hindi and Prof. Adam Mohamedali Fadlalla of Qatar University, Antonio de Amescua and Román López-Cortijo of Universidad Carlos III de Madrid and Ken and Jane Laudon.

Furthermore, I want to gratefully acknowledge all the companies that have participated to the research interviews, case studies, and surveys.

In particular, for the Financial Institutions: Agos Ducato, Banca Carige, Banca Euromobiliare, Banca Fideuram, Banca d'Italia, Banca Mediolanum, Banco Popolare, Banca Popolare dell'Emilia Romagna, Banca Popolare di Milano, Banca Popolare di Sondrio, Banca Popolare di Vicenza, Banca Popolare di Bari, Banca Sistema, Barclays, BCC Roma, BNL-BNP Paribas, Borsa Italiana, Cariparma Credit Agricole, CACEIS Bank Luxembourg, Carta Si, Cassa Depositi e Prestiti, Cassa di Risparmio di Firenze, Cedacri, Che Banca!, Compass, Corner Bank, Credito Emiliano, Deutsche Bank, Dexia, HypoVereinsbank, Istituto Centrale delle Banche Popolari Italiane, ING Direct, Intesa SanPaolo, Intesa SanPaolo Servitvia, Istituto per le Opere Religiose, Luxemburg Stock Exchange, JP Morgan Chase, Key Client, Mediobanca, Monte Titoli, Banca Monte dei Paschi, Poste Italiane, SEC Servizi, Société Européene de Banque, Standard Chartered, Royal Bank of Scotland, UBI Banca, Unicredit, Unicredit Leasing, Veneto Banca and WeBank.

For the Insurance sector: Allianz, Assimoco, Aspe Re, Cardif, Coface, Ergo Previdenza, Europe Assistance, Assicurazioni Generali, Groupama, Munich RE, Poste Vita, Reale Mutua, Novae, Sara Assicurazioni, UnipolSai, Vittoria Assicurazioni and Zurich.

For the Industrial Sector: ABB, Accenture, Acea, Aci Informatica, Acqua Minerale S. Benedetto, Adidas, Alpitour, Alliance Boots, Amadori, Amazon, Amplifon, Anas, Angelini, ArcelorMittal, Armani, Astaldi, ATAC, ATM, AstraZeneca, Arval, Auchan, Audi, Augusta Westland, Autogrill, Autostrade per l'Italia, Avio, Baglioni Hotels, BMW, BASF, Barilla, Be Consulting, Benetton, Between, Business Integration Partners, Brembo, Bravo Fly, BskyB, BSH, BOSH, Boeing Defence, Cementir, Centrica Energy, Cerved, Chiesi Farmaceutici, CNH Industrial, Coca Cola HBC, Coop Italia, Costa Crociere, D'Amico, Danone, Daimler, De Agostini, Diesel, Dimar, Dolce and Gabbana, General Electric, Ducati, Elettronica, Edipower, Edison, Eni, Enel, ENRC, ERG, Fastweb, Ferservizi, Fincantieri, Ferrari, Ferrovie dello Stato, FCA, Finmeccanica, GlaxosmithKline, GE Capital, GFT Technologies, Grandi Navi Veloci, G4S, Glencore, Gruppo Hera, Gruppo Coin, Gruppo De Agostini, Gtech, Gucci, H3G, Hupac, Infineon, Interroll, Il Sole24Ore, IREN, Istituto Poligrafico e Zecca dello Stato, ITV, Kuwait Petroleum, La Perla, Labelux Group, Lamborghini, Lavazza, Linde, LBBW, Levi's, L'Oréal, Loro Piana, Luxottica, Jaguar Land Rover, Lucite International, MAN, Magneti Marelli, Mapei, Marcegaglia, Mediaset, Menarini, Messaggerie Libri, Miroglio, Mondelez International, Mossi & Ghisolfi, Natuzzi, Novartis, Oerlikon Graziano, OSRAM, Piaggio, Perfetti, Pernod Ricard, Philips, Pirelli, Porsche, ProSiebenSat1, Procter & Gamble, Prysmian, RAI, Rexam, Rolex, Roche, Retonkil Initial, RWE, Saipem, Sandoz, SEA, Seat PG, Selex, Snam, Sorigenia, Sky Italia, Schindler Electroca, Pfizer, RFI, Telecom Italia, Telecom Italia Digital Solution, Telecom Italia Information Technology, Tenaris, Terna, Trenitalia, Tyco, TvevSued, Telefonica, Unilever, Unicoop Firenze, Virgin Atlantic, Volkswagen, Vodafone and Wind.

For the Public Sector: Agenzia per l'Italia Digitale, Comune di Milano, Regione Lombardia and Consip.

I would especially like to acknowledge all the people that have supported me during this years with insights and suggestions. I learned so much from them, and their ideas and competences have inspired my work: Silvio Fraternali, Paolo Cederle, Massimo Milanta, Massimo Schiattarella, Diego Donisi, Marco Sesana, Gianluca Pancaccini, Giovanni Damiani, Gianluigi Castelli, Salvatore Poloni, Milo Gusmeroli, Pierangelo Rigamoti, Danilo Augugliaro, Nazzareno Gregori, Edoardo Romeo, Elvio Sonnino, Pierangelo Mortara, Massimo Messina, Mario Collari, Giuseppe Capponcelli, Massimo Castagnini, Pier Luigi Curcuruto, Giovanni Sordello, Maurizio Montagnese, Umberto Angelucci, Giuseppe Dallona, Gilberto Ceresa, Jesus Marin Rodriguez, Fabio Momola, Rafael Lopez Rueda, Eike Wahl, Marco Cecchella, Maria-Louise Arscott, Antonella Ambriola, Andrea Rigoni, Giovanni Rando Mazzarino, Silvio Sperzani, Samuele Sorato, Alberto Ripepi, Alfredo Montalbano, Gloria Gazzano, Massimo Basso Ricci, Giuseppe De Iaco, Riccardo Amidei, Davide Ferina, Massimo Ferriani, Roberto Burlo, Cristina Bianchini, Dario Scagliotti, Ettore Corsi, Luciano Bartoli, Marco Ternelli, Stewart Alexander, Luca Ghirardi, Francesca Gandini, Vincenzo Tortis, Agostino Ragosa, Sandro Tucci, Vittorio Mondo, Andrea Agosti, Roberto Fonso, Federico Gentili, Nino Lo Banco, Fabio Troiani, Federico Niero, Gianluca Zanutto, Mario Bocca, Marco Zaccanti, Anna Pia Sassano, Fabrizio Lugli, Marco Bertazzoni, Vittorio Boero, Carlo Achermann, Stefano Achermann, Jean-Claude Krieger, Reinhold Grassl, François de Brabant, Maria Cristina Spagnoli, Alessandra Testa, Anna Miseferi, Matteo Attrovio, Nikos Angelopoulos, Igor Bailo, Stefano Levi, Luciano Romeo, Alfio Puglisi, Gennaro Della Valle, Massimo Paltrinieri, Pierantonio Azzalini, Enzo Contento, Marco Fedi, Fiore Della Rosa, Dario Tizzanini, Carlo Capalbo, Simone Battiferri, Vittorio Giusti, Piera Fasoli, Carlo di Lello, Gian Enrico Paglia, George Sifnios, Francesco Varchetta, Gianfranco Casati, Fabio Benasso, Alessandro Marin, Gianluca Guidotti, Fabrizio Virtuani, Luca Verducci, Luca Falco, Francesco Pedrielli, Riccardo Riccobene, Roberto Scolastici, Paola Formaneti, Andrea Mazzucato, Nicoletta Rocca, Mario Breuer, Mario Costantini, Marco Lanza, Marco Poggi, Gianfranco Ardissono, Alex Eugenio Sala, Daniele Bianchi, Giambattista Piacentini, Luigi Zanardi, Valerio Momoni, Daniele Panigati, Maurizio Pescarini, Ermes Franchini, Francesco Mastrandrea, Federico Boni, Mauro Minenna, Massimo Romagnoli, Nicola Grassi, Alessandro Capitani, Mauro Frassetto, Bruno Cocchi, Marco Tempa, Martin Brannigan, Alessandro Guidotti, Gianni Leone, Stefano Signani, Domenico Casalino, Fabrizio Lugoboni, Fabrizio Rocchio, Mauro Bernareggi, Claudio Sorano, Paolo Crovetto, Alberto Ricchiari, Alessandro Musumeci, Luana Barba, Pierluigi Berlucchi, Matthias Schlapp, Ugo Salvi, Danilo Gismondi, Patrick Vandenberghe, Dario Ferri, Claudio Colombatto, Frediano Lorenzin, Paolo Trincianti, Massimiliano Ciferri, Danilo Ughetto, Tiberio Strati, Massimo Nichetti, Stefano Firenze, Vahe Ter Nikogosyan, Giorgio Voltolini, Andrea Maraventano, Thomas Pfitzer, Guido Oppizzi, Alessandro Bruni, Marco Franzi, Guido Albertini, Massimiliano De Gregorio, Vincenzo Russi, Franco Col-lautti, Massimo Dall’Ora, Fabio De Ferrari, Mauro Ferrari, Domenico Solano, Pier Paolo Tamma, Susanna Nardi, Massimo Amato, Alberto Grigoletto, Nunzio Calì, Gianfilippo Pandolfini, Cristiano Cannarsa, Fabio Degli Esposti, Riccardo

Scattaretico, Claudio Basso, Mauro Pianezzola, Marco Zanussi, Davide Carteri, Giulio Tonin, Simonetta Iarlori, Marco Prampolini, Luca Terzaghi, Christian Altomare, Pasquale Tedesco, Michela Quitadamo, Dario Castello, Fabio Boschiero, Aldo Borriore, Paolo Beatini, Maurizio Pellicano, Ottavio Rigodanza, Gianni Fasciotti, Lorenzo Pizzuti, Angelo D'Alessandro, Marcello Guerrini, Michela Quitadamo, Dario Castello, Fabio Boschiero, Aldo Borriore, Paolo Beatini, Pierluigi De Marinis, Fabio Cestola, Roberto Mondonico, Alberto Alberini, Pierluca Ferrari, Umberto Stefani, Elvira Fabrizio, Salvatore Impallomeni, Dario Pagani, Marino Vignati, Giuseppe Rossini, Alfio Puglisi, Renzo Di Antonio, Maurizio Galli, Filippo Vadda, Marco De Paoli, Paolo Cesa, Armando Gervasi, Luigi Di Tria, Marco Gallibariggio, David Alfieri, Mirco Carriglio, Maurizio Castelletti, Roberto Andreoli, Vincenzo Campana, Marco Ravasi, Mauro Viacava, Alessio Pomasan, Salvatore Stefanelli, Roberto Scaramuzza, Marco Zaffaroni, Giuseppe Langer, Francesco Bardelli, Daniele Rizzo, Silvia De Fina, Paulo Morais, Massimiliano Gerli, Andrea Facchini, Massimo Zara, Luca Paleari, Carlo Bozzoli, Luigi Borrelli, Marco Iacomussi, Mario Dio, Giulio Mattiotti, Alessandro Poerio, Fabrizio Frustaci, Roberto Zaccaro, Maurizio Quattrococchi, Gianluca Giovannetti, Pierangelo Colacicco, Silvio Sassatelli, Filippo Passerini, Mario Rech, Claudio Sordi, Tomas Blazquez De La Cruz, Luca Spagnoli, Fabio Oggioni, Luca Severini, Roberto Conte, Alessandro Tintori, Giovanni Ferretti, Alberta Gammicchia, Patrizia Tedesco, Antonio Rainò, Claudio Beveroni, Chiara Manzini, Francesco Del Greco, Lorenzo Tanganelli, Ivano Bosisio, Alessandro Campanini, Giovanni Pietrobelli, Pietro Pacini, Vittorio Padovani, Luciano Dalla Riva, Paolo Pecchiari, Francesco Donatelli, Massimo Palmieri, Alessandro Cucchi, Riccardo Pagnanelli, Raffaella Mastrofilippo, Roberto Coretti, Alessandra Grendele, Davide Casagrande, Lucia Gerini, Filippo Cecchi, Fabio De Maron, Alberto Peralta, Massimo Pernigotti, Massimo Rama, Francisco Souto, Oscar Grignolio, Mario Mella, Massimo Rosso, Filippo Onorato, Stefan Caballo, Ennio Bernardi, Aldo Croci, Giuseppe Genovesi, Maurizio Romanese, Daniele Pagani, Derek Barwise, Guido Vetere, Christophe Pierron, Guenter Lutgen, Andreas Weinberger, Luca Martis, Stefano Levi, Paola Benatti, Massimiliano Baga, Marco Campi, Laura Wegher, Riccardo Sfondrini, Diego Pogliani, Gianluca Pepino, Simona Tonella, José González Osma, Sandeep Sen, Thomas Steinich, Barbara Karuth-Zelle, Ralf Schneider, Rüdiger Schmidt, Wolfgang Gärtner, Alfred Spill, Lissimahos Hatzidimoulas, Marco Damiano Bosco, Mauro Di Pietro Paolo, Paolo Brusegan, Arnold Aschbauer, Robert Wittgen, Peter Kempf, Michael Gorriz, Wilfried Reimann, Abel Archundia Pineda, Jürgen Sturm, Stefan Gaus, Andreas Pfisterer, Peter Rampling, Elke Knobloch, Andrea Weierich, Andreas Lubert, Heinz Laber, Michael Hesse, Markus Lohmann, Andreas König, Herby Marchetti, Rainer Janssen, Frank Rüdiger Poppe, Marcell Assan, Klaus Straub, Robert Blackburn, Wiebe Van der Horst, Martin Stahljans, Mattias Ulbrich, Matthias Schlapp, Jan Brecht, Enzo Contento, Michael Pretz, Gerd Friedrich, Florian Forst, Robert Leindl, Wolfgang Keichel, Stephan Fingerling, Sven Lorenz, Martin Hofmann, Nicolas Burdhardt, Armin Pfoh, Kian Mossanen, Anthony Roberts, John Knowles, Lisa Gibbard, John Hiskett, Richard Wainwright, David Madigan, Matt Hopkins, Gill Lungley, Simon Jobson, Glyn

Hughes, John Herd, Mark Smith, Jeremy Vincent, Guy Lammert, Steve Blackledge, Mark Lichfield, Jacky Lamb, Simon McNamara, Kevin Hanley, Anthony Meadows, Rod Hefford, Stephen Miller, Willem Eelman, Alessandro Ventura, David Bulman, Neil Brown, Alistair Hadfield, Rod Carr and Neil Dyke.

I would especially like to gratefully acknowledge Gianluigi Viscusi at College of Management of Technology (CDM)-École polytechnique fédérale de Lausanne (EPFL), Alan Serrano-Rico at Brunel University, and Nadia Neytcheva Head of Research at the Business Technology Outlook (BTO) Research Program who provided me valuable suggestions and precious support in the coordination of the production process of this book. Furthermore, I acknowledge the support of Business Technology Foundation (Fondazione Business Technology) and all the bright researchers at Business Technology Outlook (BTO) Research Program that have supported me in carrying out interviews, surveys, and data analysis: Florenzo Marra, Giulia Galimberti, Arianna Zago, Alessandro De Pace, Matteo Richiardi, Ezechiele Capitanio, Giovanni Roberto, Alessandro Scannapieco, Massimo Bellini, Tommaso Cenci, Giorgia Cattaneo, Andrada Comanac, Francesco Magro, Marco Castelli, Martino Scanziani, Miguel Miranda, Alice Brocca, Antonio Attinà, Giuseppe Vaccaro, Antonio De Falco, Matteo Pistoletti, Mariya Terzieva and Daniele Durante.

A special acknowledgement goes to the memory of Prof. Antonino Intrieri who provided precious comments and suggestions throughout the years.

Finally I acknowledge my family whose constant support and patience made this book happen.

Vincenzo Morabito

Contents

Part I Strategy

1	Big Data and Analytics for Competitive Advantage	3
1.1	Introduction	3
1.2	Competitive Advantage Definition: Old and New Notions	4
1.2.1	From Sustainable to Dynamic	5
1.2.2	From Company Effects to Network Success.	6
1.3	The Role of Big Data on Gaining Dynamic Competitive Advantage.	6
1.3.1	Big Data Driven Target Marketing	6
1.3.2	Design-Driven Innovation	8
1.3.3	Crowd Innovation.	9
1.4	Big Data Driven Business Models	10
1.5	Organizational Challenges	11
1.5.1	Skill Set Shortages	12
1.5.2	Cultural Barriers.	12
1.5.3	Processes and Structures	13
1.5.4	Technology Maturity Levels	13
1.5.5	Organizational Advantages and Opportunities	13
1.6	Case Studies	14
1.7	Recommendations for Organizations	17
1.7.1	Ask the Right Questions	17
1.7.2	Look Out for Complementary Game Changing Innovations	18
1.7.3	Develop Sound Scenarios	18
1.7.4	Prepare Your Culture	18
1.7.5	Prepare to Change Processes and Structure	19
1.8	Summary.	19
	References	20

2	Big Data and Analytics for Government Innovation	23
2.1	Introduction	23
2.1.1	New Notions of Public Service: Towards a Prosumer Era?	24
2.1.2	Online Direct Democracy	25
2.1.3	Megacities' Global Competition	25
2.2	Public Service Advantages and Opportunities	26
2.2.1	New Sources of Information: Crowdsourcing	26
2.2.2	New Sources of Information: Internet of Things (IoTs)	27
2.2.3	Public Talent in Use	29
2.2.4	Private–Public Partnerships	31
2.2.5	Government Cloud Data	31
2.2.6	Value for Money in Public Service Delivery	32
2.3	Governmental Challenges	33
2.3.1	Data Ownership	33
2.3.2	Data Quality	34
2.3.3	Privacy, Civil Liberties and Equality	34
2.3.4	Talent Recruitment Issues	35
2.4	Case Studies	36
2.5	Recommendations for Organizations	39
2.5.1	Smart City Readiness	39
2.5.2	Learn to Collaborate	40
2.5.3	Civic Education and Online Democracy	41
2.5.4	Legal Framework Development	41
2.6	Summary	42
	References	42
3	Big Data and Education: Massive Digital Education Systems	47
3.1	Introduction	47
3.1.1	From Institutionalized Education to MOOCs	49
3.2	MOOC Educational Model Clusters	51
3.2.1	University-Led MOOCs	51
3.2.2	Peer-to-Peer MOOCs	52
3.3	The Role of Big Data and Analytics	54
3.4	Institutional Advantages and Opportunities from MOOCs	55
3.5	Institutional Challenges from MOOCs	57
3.6	Case Studies	60
3.7	Recommendations for Institutions	62
3.8	Summary	62
	References	63

4	Big Data Driven Business Models	65
4.1	Introduction	65
4.2	Implications of Big Data for Customer Segmentation	69
4.3	Implications of Big Data as a Value Proposition	69
4.4	Implications of Big Data for Channels	70
4.5	The Impact of Big Data on Customer Relationships	71
4.6	The Impact of Big Data on Revenue Stream	72
4.7	The Impact of Big Data on Key Resources and Key Activities	73
4.8	The Impact of Big Data on Key Partnerships	74
4.9	The Impact of Big Data on Cost Structures	75
4.10	Organizational Advantages and Opportunities	76
4.11	Organizational Challenges and Threats	77
4.11.1	Creativity and Innovation Capability Deficit	77
4.11.2	Interrogating Big Data	77
4.11.3	Plug and Play Architectures	78
4.12	Summary	78
	References	79

Part II Organization

5	Big Data Governance	83
5.1	Introduction to Big Data Governance	83
5.1.1	Big Data Types	85
5.1.2	Information Governance Disciplines	87
5.1.3	Industries and Functions	90
5.2	Big Data Maturity Models	91
5.2.1	TDWI Maturity Model	91
5.2.2	Analytics Business Maturity Model	93
5.2.3	DataFlux Data Governance Maturity Model	94
5.2.4	Gartner Maturity Model	95
5.2.5	IBM Data Governance Maturity Model	96
5.3	Organizational Challenges Inherent with Governing Big Data	97
5.4	Organizational Benefits of Governing Big Data	99
5.5	Case Studies	100
5.6	Recommendations for Organizations	101
5.7	Summary	102
	References	103
6	Big Data and Digital Business Evaluation	105
6.1	Introduction	105
6.2	Digital Business Evaluation Using Big Data	106