Preface

In the eighteenth century, the world was at a shifting point with new manufacturing processes being introduced that allowed humans to perform mechanical work much more efficiently than the preceding hand production methods. The steam engine, patented in 1781 by Scottish engineer James Watt, is often considered to be one of the key inventions of that time that triggered the industrial revolution. Due to inventions such as the computer, the Internet, and other digital devices, we are currently witnessing what can be referred to as the digital revolution, i.e., the shift from an analog, mechanical, and electronic world to a digital one.

While the driving force behind the industrial age was to optimize the industrial process using mechanical tools, today we aim to optimize work processes by analyzing data that are created by digital devices. In fact, an increasing number of companies rely on new technologies and services that produce, analyze, and access this data in multiple forms. Consider, for example, an online fashion retailer that assists us in selecting suitable clothing by analyzing our physical appearance and current fashion trends, mobile apps that rely on real-time sensor data to help us to successfully avoid traffic in big cities, or news content providers that assist in comprehending the world around us by providing hierarchically structured news overviews. Truly, the acquisition and analysis of data can be considered to be the new oil that drives our economy.

We usually distinguish between two types of data, namely *big data* and *small data*. A popular definition of big data is provided by IBM, which describes it in terms of four dimensions: volume, variety, velocity, and veracity. Pollock argues in his keynote address at the 2012 European Data Forum in Copenhagen that although there is a big hype around the analysis of big data, the real challenges and opportunities arise from the analysis of small data such as local household energy expenditure or time schedules of local buses. A definition of small data is provided by former McKinsey consultant Allen Bonde who argues at the 2013 Data Pulse Summit in Boston that "small data connects people with timely, meaningful insights (derived from big data and/or 'local' sources), organized and packaged—often visually—to be accessible, understandable, and actionable for everyday

tasks." He also refers to it as "the last mile of big data," i.e., the data users or customers can interact with.

Key techniques for a data-centric optimization of work processes are personalization, data mining, machine learning, knowledge discovery, and information management approaches. In other words, context-aware algorithms are required to understand, interpret, and react upon input data, and adapt their output based on external input parameters. The English physicist Stephen Hawking even argues that the ability to adapt to change is an indicator for intelligence. Following this argument, algorithms that adapt to change can be seen as computationally intelligent—or smart. Therefore, we refer to systems that rely on such computational intelligence as *Smart Information Systems*. As early as 2008, Marissa Mayer, the current CEO of Yahoo! and former vice president of Search Products and User Experience of Google Inc., predicted in an interview held at the LeWeb conference in Paris that "in the future personalized search will be one of the traits of leading search engines." This statement reflects the increasing attention that smart information systems draw from both Academia and Industry. With increasing computational power, smart algorithms enable us to identify patterns, test research hypotheses, or to create data models, hence shedding light on the potential usage of this data.

However, although such techniques have matured over the past few years, there seems to be an increasing gap between current research trends in the analysis of data and the application of data analysis techniques in industry. NASA scientist Kiri Wagstaff even argued during a plenary session of the largest machine learning conference (ICML) in 2012 that "research has lost its connection to the problems of importance for the larger world of science and society." She criticizes that much research is performed on evaluating novel algorithms using limited and artificial datasets, hence breaking adrift from answering the question of what computational analytics techniques can actually be used for.

In this book, we present smart information systems for the private and public sectors. Further, an overview of research questions that can be studied by applying computational intelligence is given, followed by a description of the algorithms, tools, measures, and evaluations used to answer these research questions. Each chapter can be seen as a guideline for transforming raw data into effective smart information services.

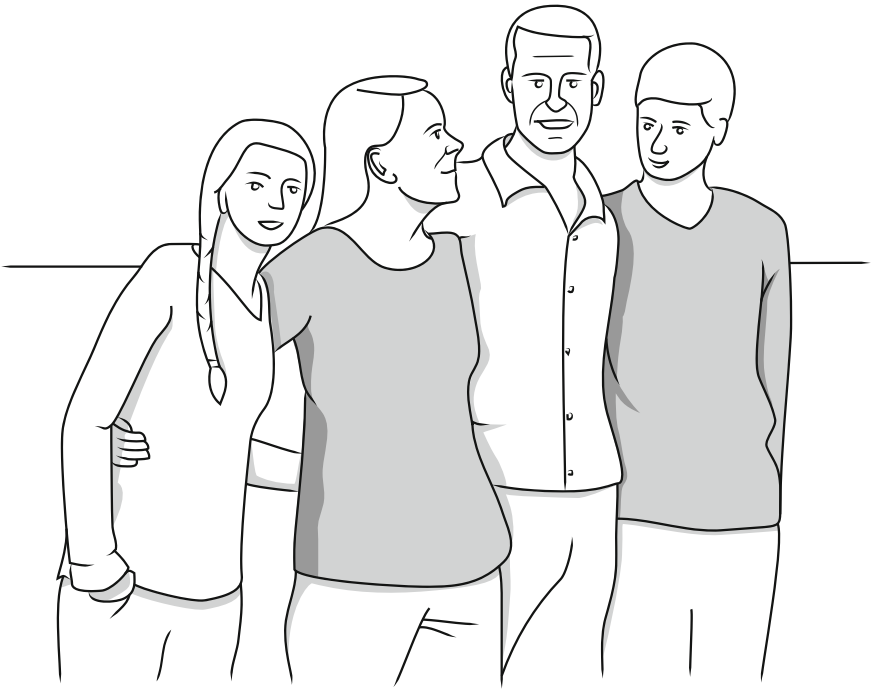
Book Outline

This book illustrates potentials and challenges that arise from analyzing data for the provision of smart information services. In each chapter, we discuss individual use cases. All use cases cover real-world research challenges faced by parties as diverse as leading SMEs, multinational manufacturers, service companies, and the public sector, and are currently being funded (or have very recently been funded) from national and international sources. The book is composed of three sections: In the first

section, we present novel information aggregation services that illustrate how textual data can be employed to generate smart information services. We focus on three different domains, namely information aggregation services for individuals as well as services for the public and private sectors. These use cases showcase how information can be aggregated to provide easier access. In the second section, we outline personalization and recommendation systems that tailor information based on users' individual preferences. Again, we showcase application scenarios under different categories, such as the academic challenges of creating such services and provision of services based on the analysis of data. In the final part, we focus on sensor-based knowledge acquisition services, i.e., we concentrate on the analysis of sensor data that can then be used to provide a clear picture of our world. We present four different scenarios that showcase how computational intelligence allows communities, companies, and individuals to better understand their own environment and products.

Meet the Marks Family

Although the individual use cases can be treated separately from each other, they all have in common that they focus on the analysis of data to provide smart information services that ease our everyday life. In order to illustrate this connection,



Clara, Suzanne, Steven, and Carl Marks. This and other graphical illustrations depicting scenes from the use cases that are presented in this book at courtesy of Sebastian Preuße, Berlin.

each chapter starts with a short episode from the lives of members of the Marks Family.

The Marks—Steven and Suzanne and the two kids Clara and Carl—represent an average family that lives in a suburb of a larger European city. Both parents go to work, their son Carl still goes to high school, their daughter just graduated and is now doing an internship before deciding what to do next with her life. Any resemblance to real persons, living or dead, is purely coincidental. Each episode outlines the need for a smart information system that would help them in typical situations that we all might be facing on a daily basis.

Authors of the Book Chapters

The authors of the book chapters are members, project partners, or associates of the Competence Center Information Retrieval and Machine Learning[1] (CC IRML) of DAI-Labor of Technische Universität Berlin. The lab focuses on the development of intelligent systems and solutions, referred to as "smart services and smart systems," that support us in our everyday life. Key to the success of such systems is the preceding analysis of data which is constantly created by these systems. The authors collaborate closely with industrial partners on a daily basis. Consequently, they have extensive experience in grasping the significant differences between Academia and Industry and are able to communicate this in their chapters, thus bridging the gap between both communities. Their expertise in computational intelligence puts the authors in an ideal situation to provide a state-of-the-art introduction to this field.

Berlin, September 2014 Frank Hopfgartner

[1] http://www.dai-labor.de/en/irml/

http://www.springer.com/978-3-319-14177-0

Smart Information Systems
Computational Intelligence for Real-Life Applications
Hopfgartner, F. (Ed.)
2015, XII, 372 p. 133 illus., 12 illus. in color., Hardcover
ISBN: 978-3-319-14177-0