

Preface

Many collaboration and decision-making settings are nowadays associated with huge, ever-increasing amounts of multiple types of data, obtained from diverse sources, which often have a low signal-to-noise ratio for addressing the problem at hand. These data may also vary in terms of subjectivity and importance, ranging from individual opinions and estimations to broadly accepted practices and trustworthy measurements and scientific results. Additional problems start when we want to consider and exploit accumulated volumes of data, which may have been collected over a few weeks or months, and meaningfully analyze them toward making a decision. Admittedly, when things get complex, we need to identify, understand and exploit data patterns; we need to aggregate appropriate volumes of data from multiple sources, and then mine them for insights that would never emerge from manual inspection or analysis of any single data source. In these settings, “big data” analytics technology currently receives much criticism, in that it does not provide proper insight into what the data means. To make sense of big data and come with discoveries that help improve decision making in practical contexts, human intelligence should be also exploited. We need to provide the appropriate ways to nurture and capture this human intelligence in order to extract the necessary insights and improve the way machines deal with complex situations.

This book reports on cutting-edge research toward efficiently and effectively addressing the above issues. This research has been carried out in the context of an EU-funded FP7 project, namely Dicode (<http://dicode-project.eu>), which aimed at facilitating and augmenting collaboration and decision making in data-intensive and cognitively-complex settings. To do so, whenever appropriate, Dicode built on prominent high-performance computing paradigms and large data processing technologies to meaningfully search, analyze, and aggregate data existing in diverse, extremely large, and rapidly evolving sources. At the same time, particular emphasis was given to the deepening of our insights about the proper exploitation of big data, as well as to collaboration and sense-making support issues. Building on current advancements, the solution proposed by the Dicode project brings together the reasoning capabilities of both the machine and the humans. It can be viewed as an innovative “workbench” incorporating and orchestrating a set of interoperable services that reduce the data-intensiveness and complexity overload at critical decision points to a manageable level, thus permitting stakeholders to be more productive and effective in their work practices.

Chapter 1, “The Dicode Project” by Nikos Karacapilidis, introduces the overall context of the project, and reports on its scientific and technical objectives, the exploitation of its results and its potential impact. Moreover, it sketches the key success indicators of the project, together with the actions taken toward ensuring their accomplishment.

Chapter 2, “Data Intensiveness and Cognitive Complexity in Contemporary Collaboration and Decision Making Settings” by Spyros Christodoulou, Nikos Karacapilidis, Manolis Tzagarakis, Vania Dimitrova and Guillermo de la Calle, reviews the state of the art on collaboration and decision-making support in contemporary settings. Related issues concerning integration technologies are also discussed. The methodologies, tools, and approaches discussed in the chapter are considered with respect to the information overload and cognitive complexity dimensions. The chapter aims to provide useful insights concerning the exploitation and advancement of existing collaboration and decision-making support technologies.

Chapter 3, “Requirements for Big Data Analytics Supporting Decision Making: A Sensemaking Perspective” by Lydia Lau, Fan Yang-Turner and Nikos Karacapilidis, aims to advance our understanding on the synergy between human and machine intelligence in tackling big data analysis. It does so by exploring and using sense-making models to inform the development of a generic conceptual architecture as a means to frame the requirements of such an analysis and to position the role of both technology and human in this synergetic relationship. Two contrasting real-world use case studies were undertaken to test the applicability of the proposed approach.

Chapter 4, “Making Sense of Linked Data: A Semantic Exploration Approach” by Dhavalkumar Thakker, Vania Dimitrova, Lydia Lau, Fan Yang-Turner and Dimoklis Despotakis, presents an experimental study with a uni-focal semantic data browser over several datasets linked via domain ontologies, which is used to inform what intelligent features are needed in order to assist exploratory search through Linked Data. The chapter reports main problems experienced by users while conducting exploratory search tasks, based on which requirements for algorithmic support to address the observed issues are elicited. In addition, a semantic signposting approach for extending a semantic data browser is proposed as a way to address the derived requirements.

Chapter 5, “The Dicode Data Mining Services” by Natalja Friesen, Max Jakob, Jörg Kindermann, Doris Maassen, Axel Poigné, Stefan Rüping and Daniel Trabold, provides an overview of the data mining services developed in the context of the Dicode project. It addresses the usability of the services and indicates which big data technologies are being used to deal with very large data collections. It is shown that these services intend to help in clearly defined steps of the sense-making process, where human capacity is most limited and the impact of automatic solutions is most profound. This includes recommendation services to search and filter information, text mining services to search for new information and unknown relations in data, and subgroup discovery services to find and evaluate hypotheses on data.

Chapter 6, “The Dicode Collaboration and Decision Making Support Services” by Manolis Tzagarakis, Nikos Karacapilidis, Spyros Christodoulou, Fan Yang-Turner and Lydia Lau, presents a series of innovative services developed in the context of the Dicode project to facilitate and augment collaboration, sense-making, and decision making in knowledge intensive environments. The ultimate goals of the proposed solution are to make it easier for users to follow the evolution of an ongoing collaboration, comprehend it in its entirety, and meaningfully aggregate data in order to resolve the issue under consideration. A tool that enables the monitoring and investigation of the collective behavior of teams with respect to sense-making tasks is also presented.

Chapter 7, “Integrating Dicode Services: The Dicode Workbench” by Guillermo de la Calle, Eduardo Alonso-Martínez, Martha Rojas-Vera and Miguel García-Remesal, presents the innovative approach developed in the Dicode project regarding the integration of services and applications. A flexible, scalable, and customizable information and computation infrastructure to exploit the competences of stakeholders and information workers is presented in detail. The proposed approach pays much attention to usability and ease-of-use issues. The chapter reports on two major outcomes of the Dicode project regarding integration issues: the Dicode Workbench and the Dicode Integration Framework.

Chapter 8, “Clinico-Genomic Research Assimilator: A Dicode Use Case” by Georgia Tsiliki and Sophia Kossida, reports on the practical use of the Dicode platform in the biomedical research context. Through a real scenario, it is shown that the platform enables researchers to efficiently and effectively collaborate and make decisions by meaningfully assembling, mining and analyzing available large-scale volumes of complex multifaceted data residing in different sources. Evaluation results are included and thoroughly assessed.

Chapter 9, “Opinion Mining from Unstructured Web 2.0 Data: A Dicode Use Case” by Ralf Löffler, reports on the use of Dicode Workbench and Dicode services in the Social Media Monitoring context. Recognizing that Social Web has given the consumers a voice and Social Media has huge impact on brands and products today, the chapter discusses how the Dicode platform can support a collaborative work environment and offer technical solutions that improve the overall quality in the social media processes. Evaluation results are also included and assessed.

Chapter 10, “Data Mining in Data-Intensive and Cognitively-Complex Settings: Lessons Learned from the Dicode Project” by Natalja Friesen, Jörg Kindermann, Doris Maassen and Stefan Rüping, reports on practical lessons learned while developing the Dicode’s data mining services and using them in data-intensive and cognitively complex settings. Various sources were taken into consideration to establish these lessons, including user feedbacks obtained from evaluation studies, discussion in teams, as well as observation of services’ usage. The lessons are presented in a way that could aid people who engage in various phases of developing similar kind of systems.

Chapter 11, “Collaboration and Decision Making in Data-Intensive and Cognitively-Complex Settings: Lessons Learned from the Dicode Project” by Spyros

Christodoulou, Manolis Tzarakis, Nikos Karacapilidis, Fan Yang-Turner, Lydia Lau and Vania Dimitrova, discusses practical lessons learned during the development of innovative collaboration and decision-making support services in the context of Dicode. These lessons concern: (i) the methodology followed for the development of the abovementioned Dicode services, (ii) the facilitation and enhancement of collaboration and decision making in data-intensive and/or cognitively complex settings, and (iii) related technological and integration issues. Detailed evaluation reports, interviews, and discussions within the development teams, as well as analysis of the use of the developed services by end users through the associated log files, provided valuable feedback for the formulation and compilation of these lessons.

The results of the Dicode project, as reported in this book, are expected to advance the state of the art in approaches on: (i) the proper exploitation of big data and the integrated consideration of data mining and sense-making issues, (ii) recommender systems, with respect to recommendations in heterogeneous, multifaceted data and the identification of hidden links in complex data types, (iii) understanding text to drastically reduce the annotation effort for extracting relations, (iv) Web 2.0 collaboration support tools in terms of interoperability with third-party tools and integration of appropriate reasoning services, and (v) decision-making support applications, by integrating knowledge management and decision making features as well as by building on the synergy of human and machine argumentation-based reasoning.

The advancements reported in the book shape innovative work methodologies for dealing with the problems of information overload and cognitive complexity in diverse collaboration and decision-making contexts. Both individual and collaborative sense-making is augmented through the meaningful exploitation of prominent data processing and data analysis technologies. The proposed solution is user-friendly and built on the synergy of human and machine intelligence. It masks the overall complexity of the underlying issues, thus allowing stakeholders to easily interact with large and complex data, providing them with meaningful recommendations upon which they can base their decisions and actions. Moreover, machine-tractable knowledge concerning the full life cycle of collaboration and decision making is accumulated and maintained.

Nikos Karacapilidis

Mastering Data-Intensive Collaboration and Decision
Making

Research and practical applications in the Dicode
project

Karacapilidis, N. (Ed.)

2014, X, 226 p. 98 illus., Hardcover

ISBN: 978-3-319-02611-4