Studies in Computational Intelligence

Volume 941

Series Editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland

The series "Studies in Computational Intelligence" (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent 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.

Indexed by SCOPUS, DBLP, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at http://www.springer.com/series/7092

Rajiv Pandey · Marcin Paprzycki ·
Nidhi Srivastava · Subhash Bhalla ·
Katarzyna Wasielewska-Michniewska
Editors

Semantic IoT: Theory and Applications

Interoperability, Provenance and Beyond

Springer

Editors
Rajiv Pandey
Amity Institute of Information Technology
Amity University Uttar Pradesh Lucknow Campus
Lucknow, Uttar Pradesh, India

Nidhi Srivastava
Amity Institute of Information Technology
Amity University Uttar Pradesh Lucknow Campus
Lucknow, Uttar Pradesh, India

Katarzyna Wasielewska-Michniewska
Systems Research Institute
Polish Academy of Sciences
Warsaw, Poland

Marcin Paprzycki
Systems Research Institute
Polish Academy of Sciences
Warsaw, Poland

Subhash Bhalla
Database System Laboratory
University of Aizu
Aizu-Wakamatsu, Fukushima, Japan

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-3-030-64618-9 ISBN 978-3-030-64619-6 (eBook)
https://doi.org/10.1007/978-3-030-64619-6

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

With an exponentially increasing number of devices connected to the Internet, the Internet of Things (IoT) is encompassing and connecting multiple domains. As a matter of fact, it becomes difficult to point to domains where IoT-based ideas are not explored. This brings about interesting developments related to, what could be considered as, the "second wave of IoT." Here, local and/or domain-specific IoT deployments have to be joined to deliver more complex services to the users. A typical example could be a seaport IoT ecosystem, being combined with a smart city IoT deployment, and with a truck company IoT platform, to optimize logistics and reduce generation of CO_2 (and other forms of pollution). Observe that we assume, here, that the initial deployments were created independently of each other, and they were dedicated to solve specific problems/deliver unique solutions. Moreover, each of them "belongs to" a different stakeholder/group of stakeholders. In addition, it is possible that the newly instantiated ecosystem connects "competitors." For instance, in the above example, multiple (competing) trucking companies are expected to seamlessly work within the developed logistics ecosystem. It should thus be relatively simple to realize that this process brings about (among others): (a) need for interoperability between joining IoT ecosystems and (b) need for provenance and trust management. The first is relatively obvious; systems that were not meant to work together have to communicate and understand each other to deliver required services. Let us now assume that the provenance can be understood as an ability to record the origin and history of pertinent data (elements). In the case of creating multi-stakeholder IoT ecosystems, being able to automatically manage data sharing/exchange/access is crucial, also as one of the key aspects of "building trust."

Let us keep this in mind and reflect on how IoT deployments are realized. The simplest conceptualization involves hardware devices, such as sensors and actuators, connected to gateways, intermediate computing nodes, and the cloud. Depending on the approach, an IoT ecosystem is realized using concepts anchored, among others, in edge, fog, cloud, mist, dew computing, as well as software-defined networking, network virtualization, and stream processing. Obviously, components realizing machine learning/data analytics are also present. All these elements are combined, to create IoT platforms that facilitate publishing, consuming, and analyzing data, within closed and open ecosystems. Taking into account the way that IoT ecosystems are realized, a number of challenges can be identified. Among them of definite importance are (but this list is, obviously, not exhaustive): (i) How to provide common representation and/or shared understanding of data that will enable analysis across (systematically growing) ecosystems? (ii) How to build ecosystems based on data flows? (iii) How to track data provenance? (iv) How to ensure/manage trust? (v) How to search for things/data within ecosystems? (vi) How to store data and assure their quality?

Semantic technologies are often considered among possible ways of addressing these (and other, related) questions. More precisely, in academic research and industrial practice, semantic technologies materialize in the following contexts (this list is, also, not exhaustive, but indicates the breadth of scope of semantic technology usability): (i) representation of artifacts in IoT ecosystems and IoT networks, (ii) providing interoperability between heterogenous IoT artifacts, (ii) representation of provenance information, enabling provenance tracking, trust establishment, and quality assessment, (iv) semantic search, enabling flexible access to data originating in different places across the ecosystem, (v) flexible storage of heterogenous data. Finally, Semantic Web, Web of Things, and Linked Open Data are architectural paradigms, with which the aforementioned solutions are to be integrated, to provide production-ready deployments.

Nevertheless, even though we can observe a systematic uptake of solutions utilizing semantic technologies, in the IoT domain and across the Internet, the total number of actual deployments is not large enough to provide a real-world grounded guidance and best practices to follow. This provides the context for this book, which aims at presenting the current trends in application of semantic technologies in the IoT domain (and closely related areas, such as Semantic Web or Web of Things) and explores enablers that they provide. Moreover, descriptions of real-life use cases, where semantic technologies have been adopted, are included.

The book is divided into four parts and consists of 18 chapters. The first part has five chapters and provides fundamentals that can help understand the remaining parts of the book. The second part deals with IoT data and interoperability and consists of six chapters. The third and fourth parts are devoted to domain-specific and problem-specific applications, which are addressed, in seven chapters. Let us now, briefly, summarize the content of each part.

The first part starts from the chapter by Angelos Chatzimichail, Evangelos Stathopoulos, Dimos Ntioudis, Athina Tsanousa, Maria Rousi, Athanasios Mavropoulos, George Meditskos, Stefanos Vrochidis, and Ioannis Kompatsiaris. They provide an overview of the current trends in application of semantic technologies in the IoT domain, presenting practical applications in multiple domains, such as health care, disaster management, public events, precision agriculture, intelligent transportation, building and infrastructure management. Additionally, research studies on semantic reasoning, aggregation, fusion, and interpretation that aim to intelligently process and ingest sensor data are provided. Moreover, authors elaborate on how semantic technologies can overcome the limitations of device and data heterogeneity. Finally, issues related to the Web of Things and how it augments the IoT concepts are also discussed.

The second chapter was prepared by Jayashree R. Prasad, Priya M. Shelke, and Rajesh S. Prasad. They introduce the basic ideas, concepts, and technologies of the Semantic Web. Authors illustrate these concepts in the context of implementations of Semantic Web desktop and geospatial Semantic Web, as applied to agriculture, health care, and IoT in general. Finally, issues related to data provenance are also considered, specifically within the scope of the PROV data model.

The third chapter has been authored by Reinaldo Padilha França, Ana Carolina Borges Monteiro, Rangel Arthur, and Yuzo Iano. They have focused their attention on semantic searches. Specifically, they provide an overview of the Semantic Web and technology behind the Semantic Web search engines.

In the fourth chapter of this part, Hemanta Kumar Palo considers the fact that there are three different aspects related to the IoT paradigm: (i) things, (ii) Internet, and (iii) semantics. With this in mind, the author reviews and emphasizes the key emerging trends of the semantic technology impacting the IoT. Particularly, the work focuses on different aspects of information modeling, ontology design, semantic interoperability, machine learning, security policy, and processing of semantic data.

Finally, in the last chapter of the first part, Rajiv Pandey and Mrinal Pande discuss the concepts of provenance and trustability, by outlining available models of trust and tools for trust management. Additionally, the chapter introduces an example of trust implementation in an existing ontology, using provenance assertions based on the PROV-DM, proposed by the World Wide Web consortium.

The fact that the second part of this book is devoted to IoT data and its interoperability should not be surprising. It has been argued many times that interoperability is one of the roadblocks of faster uptake of IoT solutions. As a matter of fact, in 2016, the European Commission has funded six independent research projects devoted to IoT interoperability. Keeping this in mind, let us summarize the next six chapters.

The sixth chapter, authored by Arunima Sharma and Ramesh Babu Battula, focuses on complexity and variance in data materializing in IoT deployments, which make it difficult to apply and query available data, to realize user-centered applications. Here, authors propose to use ontologies to resolve issues in data naming. The most popular IoT and selected application domain ontologies are described, and their use is discussed.

The next chapter has been prepared by Amélie Gyrard, Ghislain Atemezing, and Martin Serrano. They recognize heterogeneity of (1) data format, (2) languages to describe sensor metadata, (3) models for structuring sensor datasets, (4) reasoning mechanisms and rule languages to interpret sensor datasets, and (5) applications. In this context, innovative methodologies, to link and associate the data from different domains to improve knowledge discovery, have been discussed. In this context, the chapter is focused on the ontology quality when building sensor-based applications and describes the PerfectO, a Knowledge Directory Services tool. PerfectO assists ontology designers to improve ontologies to be reused in other projects. It selects and classifies a sub-set of tools providing an online interface or a Web service simple to use, which help to enhance ontologies and synthesize a set of practices.

The eighth chapter was authored by Gianfranco E. Modoni, and Marco Sacco and focuses on RDF stores that can be used, among others, in IoT deployments. While multiple RDF stores exist, and may be appropriate and usable for some tasks, they will not fit others. Moreover, a one-size-fits-all solution is not available and, likely, will not be delivered. In this context, a methodological approach to evaluate and rank the relevant functional and non-functional features of the RDF stores is presented. The proposed approach is to help software architects to select which RDF stores best fit their application scenario(s).

The following chapter, written by Vitalina Babenko, Igor Shostak, Mariia Danova, and Olena Feoktystova, deals with creation of ontological knowledge bases in the Semantic Web. Specifically, various tabular structures are considered as sources of knowledge, with the main problem being a contradiction between the wide variety of tabular structures used in knowledge sources and the insufficient efficiency of classical methods for analyzing sources of this type. The implementation of the theoretical results of the study, in the form of algorithmic, mathematical support, as well as experimental studies conducted to determine the upper bound and the nature of the growth of complexity of the method of forming the ontological knowledge bases based on targeted enumeration, has been presented and confirms the validity of the proposed approach.

The tenth chapter, contributed by Beniamino Di Martino and Antonio Esposito, is focused on seamless interoperability of sensor-generated data, which is needed to achieve specific goals. The context of the work is provided by the lack of a universally accepted standard for sensor communications. In the chapter, a prototype tool for the analysis of sensors' API tries to overcome the interoperability issues in a sensor network and provides an instrument to support sensors' orchestration and management. The tool aims not only at automatically analyzing the APIs, but also to derive a semantic representation of them, which can be then used to support the manual annotation with external domain ontologies.

Finally, the last chapter in this part was written by Pratiyush Guleria and Manu Sood. Here, authors focus their attention on interoperability in the healthcare sector. Specifically, they propose to use the foundations of the Semantic Web, in a three-layered framework for IoT interoperability, and have framed a Web ontology structure for semantic interoperability in IoT for the healthcare sector. This is combined with a text analytics model, which performs semantic data classification on a synthetic healthcare dataset, to predict the patient diagnosis using machine learning techniques.

The third part of the book deals with semantic IoT in the context of domain-specific applications. The first chapter was contributed by Gaurav Kant Shankhdhar, Richa Sharma, and Manuj Darbari. Their contribution is focused on the agriculture/farming industry. The considered solution provides a lightweight IoT framework, focused on farming in developing countries like India. In this context, authors have developed a semantically enriched agent-based model

(ABSMSA), which uses SAGRO-Lite lightweight ontology, proposed by the authors, along with the IoT-Lite and the Complex Event Service Ontology (CESO).

The thirteenth chapter was authored by Mahda Noura, Amélie Gyrard, Benjamin Klotz, Raphael Troncy, Soumya Kanti Datta, and Martin Gaedke. It is focused on the automotive industry and attempts at answering the question: How to better understand the knowledge provided by Google results to build future "smart vehicle-centric" applications? Authors present an exhaustive systematic literature survey which is a basis for building a "common sense knowledge" dataset for the automotive sector. The proposed methodology (KEAS: Knowledge Extraction for the Automotive Sector) aims to analyze the most popular "knowledge elements," required to build smart vehicle applications and delivers: (1) keyphrase synonyms, (2) synonyms used to build a corpus of scientific publications, (3) smart car ontologies, and (4) the extraction of the most common terms for the automotive sector.

The next chapter was authored by Regel Gonzalez-Usach, Matilde Julian, Manuel Esteve, and Carlos E. Palau and deals with both interoperability and Ambient Assisted Living (AAL) and Active and Healthy Aging (AHA). Here, results from the European project ACTIVAGE, concerning benefits obtained through the use of IoT in elderly smart homes, by enabling semantic interoperability and cooperation across smart home clusters located in 12 different cities across Europe are presented. The challenges that were solved in the project using real-time semantic translation include: providing interoperability and semantic integration, the management of massive real-time streams of IoT data.

In the fifteenth chapter, authored by Gennady Chuiko, Yaroslav Krainyk, Olga Dvornik, and Yevhen Darnapuk, semantic provenance management for medical data is considered. Specifically, authors consider the presence of semantic data on different levels of a complex health monitoring system. The model of the semantics-based system, for medical data provenance, has been proposed along with the revision of the whole set of available technologies to employ in the semantic engine and analysis of its behavior under different circumstances. Here, inclusion of semantic information into the general dataflow should not only allow evaluating data quality but also discover behavioral patterns to identify problems inside a specific part of the system. Authors consider several scenarios that involve different device types that measure a patient's state.

The last part of this volume is devoted to problem-specific applications. In the sixteenth chapter, Matthew Weber and Edward A. Lee consider semantic localization for the IoT. They base their work on an interesting observation that Euclidean geometry and Newtonian time, with floating-point numbers, may not be the best choice for IoT ecosystems. Hence, they survey location models from robotics, the Internet, cyber-physical systems, and philosophy. As a result, a logical framework, wherein a spatial ontology is defined as a model-theoretic structure, is proposed. It is then argued that space-aware IoT services gain advantages for privacy and interoperability when they are designed for the most abstract spatial-ontologies possible.

Next, in the seventeenth chapter, Arun Kumar and Sharad Sharma consider simplifying trigger-action programming, for a control within an IoT scenario. Authors propose an IoT-based semantic interoperability model, with EUPont Semantic Web ontology, to deliver semantic interoperability among heterogeneous IoT devices for control of end-user applications. They believe that the EUPont could serve as the core information layer for the future IoT end-user programming solutions.

Finally, in the last chapter, written by Jayashree R. Prasad, Shailesh P. Bendale, and Rajesh S. Prasad, authors consider interrelation between IoT ecosystems and software-defined network (SDN), network function virtualization (NFV), cloud technologies, and semantic technologies. Here, the key underlying problem is, again, interoperability between devices. The context is provided by the observation that deployment of 5G (and later, 6G) networks will result in an almost unbound growth in the number of connected (heterogeneous) devices.

We would like to express our gratitude to all reviewers who have evaluated all submissions and helped improve the ones that have been selected for the inclusion in this volume. Thank you for your hard work.

Lucknow, India — Rajiv Pandey
Warsaw, Poland — Marcin Paprzycki
Lucknow, India — Nidhi Srivastava
Aizuwakamatsu, Japan — Subhash Bhalla
Warsaw, Poland — Katarzyna Wasielewska-Michniewska

Contents

IoT Data and Interoperability

Domain-Specific Applications

Problem-Specific Applications