

Special Issue on Smart Data and Semantics in a Sensor World

Sergio Ilarri ^{1,*} , Laura Po ^{2,*}  and Raquel Trillo-Lado ^{1,*} 

¹ Department of Computer Science and Systems Engineering, I3A, University of Zaragoza, 50018 Zaragoza, Spain

² "Enzo Ferrari" Engineering Department, University of Modena and Reggio Emilia, 41121 Modena, Italy

* Correspondence: silarri@unizar.es (S.I.); laura.po@unimore.it (L.P.); raqueltl@unizar.es (R.T.-L.)

Received: 2 September 2020; Accepted: 8 September 2020; Published: 12 September 2020



1. Introduction

Since its first inception in 2001, the application of the Semantic Web [1,2] has carried out an extensive use of ontologies [3–5], reasoning, and semantics in diverse fields, such as Information Integration, Software Engineering, Bioinformatics, eGovernment, eHealth, and social networks. This widespread use of ontologies has led to an incredible advance in the development of techniques to manipulate, share, reuse, and integrate information across heterogeneous data sources.

In recent years, the growth of the IoT (Internet of Things) required to face the challenges of "Big Data" [6–10]. The cost of sensors is decreasing, while their use is expanding. Moreover, the use of multiple personal smart devices is an emerging trend and all of them can embed sensors to monitor the surrounding environment. Therefore, the number of available sensors is exploding. On the one hand, the flows of sensor data are massive and continuous, and the data could be obtained in real time or with a delay of just a few seconds. Then, the volume of sensor data is increasing continuously every day. On the other hand, the variety of data being generated is also increasing, due to plenty of different devices and different measures to record. There are many kinds of structured and unstructured sensor data in diverse formats. Moreover, data veracity, which is the degree of accuracy or truthfulness of a data set, is an important aspect to consider. In the context of sensor data, it represents the trustworthiness of the data source and the processing of data. The need for more accurate and reliable data was always declared, but often overlooked for the sake of larger and cheaper datasets. Sensor data are uncertain and imprecise by definition; therefore, to increase their quality, allocating enough resources to clean up sensor data properly is needed.

Research areas related to the Semantic Web, like ontology matching/ontology alignment [11,12], are providing efficient methodologies based on the RDF and OWL languages to provide standard ways to convert such datasets into Linked Data sources [13,14]. Due to its interest, in several fields, such as social networks, smart cities, or context-aware mobile applications, it is very relevant to publish the data available as Linked Data. Therefore, the development of techniques to enable users to publish, visualize, and easily manipulate data is in high demand. In addition, the standardization of this linked format has completely revolutionized the way of representing and analyzing data, which requires new graph-based machine learning and data mining techniques to explore such representation.

On the other hand, sensing technologies have become an important field for computer scientists. Sensors are sparsely distributed across the globe, leading to an overwhelming amount of data about our environment. Sensors can range from stationary environmental sensors to drones or autonomous vehicles collecting data, or even to humans acting as sensors using smartphones, and can be used to detect a multitude of observations, from simple phenomena to complex events. Moreover, the Sensor Web [15–17] has realized the idea of a standardized, interoperable platform for everyone to easily share, find, and access sensor data.

However, the various characteristics of sensor data and their corresponding processing requirements, such as their multisource, heterogeneous, real-time, voluminous, streaming, and spatio-temporal features, has led many traditional data processing and integration approaches to show their limitations. Moreover, the lack of integration and communication between sensor networks often isolates important data streams and intensifies the existing problem of having too much data and not enough knowledge.

In this area, Semantic Web technologies [18] have provided particular means to achieve these aims. Specifically, the Semantic Sensor Web (SSW) [15] proposes that sensor data be annotated with semantic metadata, which will increase the interoperability and provide contextual information essential for situational knowledge. Social applications as well as ubiquitous and pervasive computing are examples of areas making use of semantic measured or processed data (e.g., use of semantic techniques for location-based services [19]). Semantization, context awareness, community management, and data visualization are core issues related to this area.

2. Topics of Interest for the Special Issue

The goal of this Special Issue is to provide a venue to show the practical progress made in the area of data management for sensors, particularly regarding the use of semantic techniques to obtain and exploit smart data from the raw sensor data. The terms “semantic web” and “sensor web” were firstly used more than 15 years ago. Despite this, they are mature topics on which the research community is still very active and with significant research challenges to address.

This Special Issue intends to provide insights on recent advances in these topics by soliciting original scientific contributions in the form of theoretical foundations, models, experimental research, and case studies for developing semantic Web-based applications. We aimed to bring together research related to several disciplines, such as Data Management, Knowledge Representation and Engineering, Web of Data, and Sensor Networks, among others. Original research contributions were invited on all aspects of the Semantic Web and Sensor Web, as well as their applications. Particularly, the following themes were within the scope of this Special Issue:

1. **Semantics and Sensor Data:** real-time sensor data streams, data management for sensor data, obtention of smart data from sensors, analytics of sensor data streams, semantic modelling and annotation of sensor data, scaling semantic sensor systems, sensor data representation, acquisition, and cleaning, semantic integration of heterogeneous data sources, semantic data management technologies for sensor data, challenges with managing and integrating real-time and historical sensor data, provenance, and access control and privacy-preserving issues in semantic data and sensor data.
2. **Linked Data:** Linked Data applications and case studies, visualizations, and user interfaces for ontologies, sensor data and linked data, machine learning and data mining for the Web of Data.
3. **Sensor-Based Applications:** semantic modeling of data for smart cities, the mobile web, sensors and semantic streams, smart cities, urban and geospatial data, semantics and sensor data for smart cities, semantics and eGovernment, managing sensor data in transportation applications, collaborative sensing, and spatial crowdsourcing.

We encouraged theoretical, methodological, empirical, and application papers. The submitted papers should describe original work, present significant results, and provide a rigorous, principled, and repeatable evaluation. In addition, the submission of papers incorporating links to data sets and other materials used for evaluation, as well as to live demos and software source code, was appreciated.

3. The Papers

A total of 18 papers were submitted to the Special Issue on Smart Data and Semantics in a Sensor World. After peer review, four of them were accepted and published. Afterwards, an additional paper co-authored by the guest editors was submitted and managed independently by another editor of

Applied Sciences, which after the review process was finally accepted and incorporated also to the special issue.

The paper “Using Adverse Weather Data in Social Media to Assist with City-Level Traffic Situation Awareness and Alerting”, authored by Hao Lu, Yifan Zhu, Kaize Shi, Yisheng Lv, Pengfei Shi, and Zhendong Niu, presents a forecasting and alerting approach for traffic incidents in a city. It exploits temporal, spatial, and meteorological features extracted from social media to estimate the city’s traffic status and suggest the appropriate warning levels. The proposal is evaluated with a traffic incidents dataset (tweets extracted from Sina Weibo, a microblogging service for Chinese-speaking people).

The paper “Ontological Representation of Smart City Data: From Devices to Cities”, co-authored by Paola Espinoza-Arias, María Poveda-Villalón, Raúl García-Castro, and Óscar Corcho, presents a systematic literature review of available smart city ontologies and proposes a set of ontology design patterns. These patterns, which provide core concepts to guide the development of ontologies for smart cities, are defined by considering the ontological requirements for smart city data and the findings of the survey presented.

The paper “MEBN-RM: A Mapping between Multi-Entity Bayesian Network and Relational Model”, co-authored by Cheol Young Park and Kathryn Blackmond Laskey, presents MEBN-RM, a set of mapping rules between key elements of the knowledge representation formalism Multi-Entity Bayesian Network (MEBN), which is the basis of PR-OWL (a probabilistic extension to OWL), and the Relational Model (RM). An algorithm that transforms a relational schema into a MEBN model has been released as an open software tool. The proposal is illustrated with two example use cases for situation awareness (a critical infrastructure defense system and a smart manufacturing system), where sensor data are stored in relational databases and converted to MEBN models, and the performance is evaluated in terms of the mapping speed and the accuracy of the mapping.

The paper “Smart Environmental Data Infrastructures: Bridging the Gap between Earth Sciences and Citizens”, co-authored by José R. R. Viqueira, Sebastián Villarroja, David Mera, and José A. Taboada, focuses on the problem of monitoring and forecasting of environmental conditions and reviews earth science data representation and access standards and technologies. It identifies nine key challenges that need to be tackled for an effective and efficient smart geospatial searching and browsing.

Finally, the paper “Semantic Traffic Sensor Data: The TRAFair Experience”, co-authored by Federico Desimoni, Sergio Ilarri, Laura Po, Federica Rollo, and Raquel Trillo-Lado, tackles the problem of modeling and semantic annotation of traffic data. It presents the tools and techniques used in the TRAFair (Understanding Traffic Flows to Improve Air Quality) project for data modeling and its semantic enrichment in the cities of Modena (Italy) and Zaragoza (Spain). An experimental evaluation shows the performance of the approach proposed to publish Linked Data.

Author Contributions: Paper writing, special issue coordination and management of submitted papers, S.I.; Paper writing and management of submitted papers, L.P.; Paper writing and management of submitted papers, R.T.-L. All authors have read and agreed to the published version of the manuscript.

Funding: We want to thank the support of the projects that fund our work in relation to the topics of this special issue: the TRAFair project 2017-EU-IA-0167, co-financed by the Connecting Europe Facility of the European Union, the project TIN2016-78011-C4-3-R (AEI/FEDER, UE), and the Government of Aragon (Group Reference T64_20R, COSMOS research group).

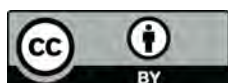
Acknowledgments: This issue would have not been possible without the contributions of the authors who submitted their valuable papers. We would like to thank all the reviewers and the editorial team of Applied Sciences for their work and dedication.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, T.B.; Hendler, J.; Lassila, O. The Semantic Web. *Sci. Am.* **2001**, *284*, 34–43.
2. Shadbolt, N.; Berners-Lee, T.; Hall, W. The Semantic Web Revisited. *IEEE Intell. Syst.* **2006**, *21*, 96–101. [[CrossRef](#)]

3. Gruber, T.R. Toward principles for the design of ontologies used for knowledge sharing? *Int. J. Hum. Comput. Stud.* **1995**, *43*, 907–928. [\[CrossRef\]](#)
4. Wang, T.D.; Parsia, B.; Hendler, J. A Survey of the Web Ontology Landscape. In *International Semantic Web Conference (ISWC)*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2006; Volume 4273, pp. 682–694. [\[CrossRef\]](#)
5. Ding, L.; Kolari, P.; Ding, Z.; Avancha, S. Using Ontologies in the Semantic Web: A Survey. In *Ontologies; Integrated Series in Information Systems*; Springer: Boston, MA, USA, 2007; Volume 14, pp. 79–113. [\[CrossRef\]](#)
6. Ang, L.M.; Seng, K.P. Big Sensor Data Applications in Urban Environments. *Big Data Res.* **2016**, *4*, 1–12. [\[CrossRef\]](#)
7. Katal, A.; Wazid, M.; Goudar, R.H. Big data: Issues, Challenges, Tools and Good practices. In Proceedings of the Sixth International Conference on Contemporary Computing (IC3), Noida, India, 8–10 August 2013; pp. 404–409. [\[CrossRef\]](#)
8. Chen, M.; Mao, S.; Liu, Y. Big Data: A Survey. *Mob. Netw. Appl.* **2014**, *19*, 171–209. [\[CrossRef\]](#)
9. Chen, C.P.; Zhang, C.Y. Data-intensive applications, Challenges, Techniques and Technologies: A survey on Big Data. *Inf. Sci.* **2014**, *275*, 314–347. [\[CrossRef\]](#)
10. Dinakar, J.R.; Vagdevi, S. A study on storage mechanism for heterogeneous sensor data on big data paradigm. In Proceedings of the 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICECCOT), Mysuru, India, 15–16 December 2017; pp. 342–345. [\[CrossRef\]](#)
11. Choi, N.; Song, I.Y.; Han, H. A Survey on Ontology Mapping. *SIGMOD Rec.* **2006**, *35*, 34–41. [\[CrossRef\]](#)
12. Otero-Cerdeira, L.; Rodríguez-Martínez, F.J.; Gómez-Rodríguez, A. Ontology matching: A literature review. *Expert Syst. Appl.* **2015**, *42*, 949–971. [\[CrossRef\]](#)
13. Bizer, C.; Heath, T.; Berners-Lee, T. Linked Data: Principles and State of the Art. Available online: <https://www.w3.org/2008/Talks/WWW2008-W3CTrack-LOD.pdf> (accessed on 12 September 2020).
14. Bizer, C.; Heath, T.; Berners-Lee, T. Linked Data – The Story Thus Far. *Int. J. Semant. Web Inf. Syst.* **2009**, *5*, 1–22. [\[CrossRef\]](#)
15. Sheth, A.; Henson, C.; Sahoo, S.S. Semantic Sensor Web. *IEEE Internet Comput.* **2008**, *12*, 78–83. [\[CrossRef\]](#)
16. Bröring, A.; Echterhoff, J.; Jirka, S.; Simonis, I.; Everding, T.; Stasch, C.; Liang, S.; Lemmens, R. New Generation Sensor Web Enablement. *Sensors* **2011**, *11*, 2652–2699. [\[CrossRef\]](#) [\[PubMed\]](#)
17. van Zyl, T.L.; Simonis, I.; McFerren, G. The Sensor Web: Systems of sensor systems. *Int. J. Digit. Earth* **2009**, *2*, 16–30. [\[CrossRef\]](#)
18. Domingue, J.; Fensel, D.; Hendler, J.A. (Eds.) *Handbook of Semantic Web Technologies*; Springer: Berlin/Heidelberg, Germany, 2011.
19. Ilarri, S.; Illarramendi, A.; Mena, E.; Sheth, A. Semantics in Location-Based Services—Guest Editors' Introduction for Special Issue. *IEEE Internet Comput.* **2011**, *15*, 10–14. [\[CrossRef\]](#)



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).