

Web2Touch 2019

Semantic Technologies for Smart Information Sharing and Web Collaboration

Rodrigo Bonacin
CTI Renato Archer and UNIFACCAMP
Campinas, SP Brazil
rodrigo.bonacin@cti.gov.br

Mariagrazia Fugini
DEIB - Politecnico di Milano
Milano, Italy
mariagrazia.fugini@polimi.it

Riccardo Martoglia
FIM - Università di Modena e Reggio
Emilia,
Modena, Italy
riccardo.martoglia@unimo.it

Olga Nabuco
CTI Renato Archer
Campinas, SP Brazil
olga.nabuco@cti.gov.br

Fatiha Sais,
LRI, Paris Sud University
Paris, France
sais@lri.fr

Abstract—This foreword introduces a summary of themes and papers of the *Web2Touch (W2T) 2019* Track at the 28th *IEEE WETICE Conference* held in Capri, June 2019. W2T 2019 includes ten full papers and one short paper. They all address relevant issues in the field of information sharing for collaboration, including, big data analytics, knowledge engineering, linked open data, applications of smart Web technologies, and smart care. The papers are a portfolio of hot issues in research and applications of semantics, smart technologies (e.g., IoT, sensors, remote monitoring devices, and smart contents management) with crucial topics, such as big data analysis, knowledge representation, and smart enterprise management, among the others. This track shows how cooperative technologies based on knowledge representation, intelligent tools, and enhanced Web engineering can enhance collaborative work through smart service design and delivery, so it contributes to change the role of the semantic Web and of applications.

Keywords - *semantic Web; knowledge representation; collaborative Web; linked data; ontology; security; safety; smart care; data analytics.*

I INTRODUCTION

Web2Touch (W2T), since its early editions (2007 onwards), is a workshop aimed at presenting research and developments in the area of collaborative work and smart information sharing using the Web. New advances in Web-based systems are among the drivers for changes in human cooperative activities, which are increasingly performed via the Web. The W2T Track tackles knowledge and data sharing based on Software Engineering, on Artificial Intelligence (AI), and on methods for Knowledge Organization Systems (e.g., ontologies and taxonomies). The focus of the Track is on information and knowledge representation, creation, maintenance, disambiguation, interlinking, trust and security. Reasoning and deep learning techniques, and other technologies related to the world of artificial intelligence, lead to better decisions and/or awareness of events. Therefore, W2T explores the concepts related to “working together” with decision-support

assistance, collective intelligence, semantic-based Web search, smart environments, intentions-based analysis, and other collaborative-based ways for problem solving.

W2T 2019 provides as usual a venue for researchers and practitioners to learn how to handle dynamic Web information, as well as to discuss the applicability and limitations of current and emerging approaches to Web-based knowledge sharing in different domains.

W2T is an open forum for studies in multiple application domains including, for example, Web science, eHealth, eGovernment, eLearning, smart cooperative fruition of evolvable contents in smart cities, digital agriculture, and collaborative systems in general. New advances in Web-based systems are the driver for changes in cooperative activities, risk facing, actions in cybersystems and in smart environments, where human activities are performed preferably via the Web. Web practitioners, users and applications exploit, in rapidly varying ways, the richness of the Web to support the user activities. Based on these considerations, the W2T track tackles knowledge and Web engineering, considering also reasoning and deep learning techniques, machine learning, text and unstructured data analytics, events managements, and so on, since these lead to better decisions or awareness of events. W2T explores themes such as decision support assistance, collective intelligence, semantic search, smart cities and enterprises, and the Internet of Things (IoT). W2T wants to both explore the state-of-the-art on these topics, and report user experience and advanced research topics about future collaborative approaches.

In particular, W2T 2019 collects papers concerning enhanced organization and management of knowledge. It presents models and tools to represent dynamic changes in shared information, smart and context-aware Web applications, as well as new domains of application of semantic techniques, such as Industry 4.0, smart care, Big Data, IoT, enhanced connectivity, and mobile technologies. W2T 2019 is also about practical experiences in both stable and emerging interdisciplinary applications.

Together with big data analytics, currently considered as one of the most important paths to competitive advantage of enterprises, information engineers and application managers are investing a large amount of effort in information extraction, knowledge formatting and data integration phases of information engineering projects. Semantic integration is still a major challenge in Web applications. Over the past three decades, the idea of using semantics for data integration has become increasingly crucial, and has received much attention in the AI, database, Web, and data mining communities. Here, we focus on specific paradigms for semantic data integration, such as ontology-based data access and Web searching, exploration, use, and analysis.

The papers of this edition aim at presenting alternatives to deal with the following issues:

- Developing solutions for smart document management in innovative big data analytics, including the areas of data acquisition, data storage, and data analysis.
- Proposing approaches to unify open data representations, aiming to acquire a single homogenized object by merging the conflicting information of the linked individual objects (i.e., data fusion approaches).
- Exploring semantic technologies in the development of context-aware interactive virtual agents, which are able to detect users' emotional state and express affective behavior aiming to improve their empathy.
- Designing cognitive assistants by exploring complex system modeling, semantic modeling, collective intelligence, and ubiquitous technology.

II ADDRESSED TOPICS

W2T 2019 concerns improvements obtained through enhanced organization and management of knowledge. Examples are models and tools to represent dynamic changes in shared information, smart (context-aware) Web applications, and new domains of application of semantic techniques, such as Industry 4.0, Big Data, social networks, IoT, enhanced connectivity, and mobile technologies. W2T is also about practical experiences in both well established and emerging interdisciplinary applications, including eHealth, smart cities, smart companies, entertainment, eLearning, education, and digital agriculture. Contributions addressing one or more of the following topics were expected:

- Ontologies and Knowledge Graphs.
- Ontology tools, ontology engineering, reuse and integration.
- Knowledge Graph expansion, enrichment, and validation.
- Knowledge and data quality assessment.

- Data profiling and axiom discovery.
- Knowledge and Data provenance.
- Knowledge and Data veracity.
- Semantic sharing and collaborative knowledge management.
- Data integration and interlinking from and across different sources and formats/semantics including big Web data, linked open data, crowdsourced data, social data, knowledge networks data.
- Crowdsourcing techniques for semantic collaboration.
- Use of semantics in mobile Web, wearable and edge computing, cross-device content management and delivery.
- Semantic annotations, semantic of collaborative Web sources and Semantic technologies for information Extraction Transformation and Loading (ETL).
- Collaborative semantic Web techniques and applications.
- Experiences and best practices in semantic Web support for collaborative work and business.
- Semantic techniques for cultural progress and community involvement.
- Collaborative semantic Web in interdisciplinary applications such as Web science, eHealth, education, entertainment, smart environments, smart factories in the Industry 4.0, and digital agriculture.
- Privacy, security and safety in smart environments (cybersecurity and risks in cybersystems).
- Experiences analyzing publicly available datasets, such as: Computer Science Education (e.g., Blackbox, Engage-csedu), MOOC (e.g., Coursera, EdX), Computer Virology (e.g., Genoma, Drebin), Healthcare, Bioinformatics (e.g., Dream Challenges).

III SELECTED PAPERS

A. Full Papers

1. The paper "Semantic Approach as Support in the Analysis of Abstracts in the Literary Review", by Marbilgia Sergio, Talita Sousa Costa, Marcelo S. P. Pessoa and Paulo S. M. Pedro, proposes a novel solution to support researchers in literature review. The paper presents the Sw3T software, which exploits semantic text analysis to classify papers by subject, thus expediting the search and analysis of publications.

2. The paper "Improving Interaction in Integrated Chronic Care Management", by Nabil Badr, Maddalena Sorrentino, Marco De Marco and Mariagrazia Fugini, focuses on health and social care services and examines the key role of information and communication technology in facilitating the integrated care effort. The approach proposes a set of guiding principles helping in the design of resource integration mechanisms in chronic care settings.

3. The paper "A Text Analytics Architecture for Smart Companies", by Mariagrazia Fugini, Jacopo Finocchi, Paolo Locatelli, Filippo Leccardi and Alfredo Lupi, presents an architecture for Big Data Analytics of unstructured content in an enterprise setting. The authors prove the flexibility of their approach through some knowledge extraction experiments in handling heterogeneous documents from different customers.

4. The paper "Mining Developer's Behavior from Web-based IDE Logs", by Pasquale Ardimento, Mario Luca Bernardi, Marta Cimitile and Giuseppe De Ruvo, presents a framework, which allows generating event logs from cloud-based Integrated Development Environments (IDE). A process mining technique is employed to extract the developers' coding processes and compare them in the shared coding environment. Moreover, the paper validates the obtained results by means of a real study involving forty students of an advanced Java programming course.

5. The paper "Content Ontology Design Patterns for Representing Knowledge in the Disaster Risk Domain", by Allan Mazimwe, Imed Hammouda and Anthony Gidudu, tackles the issue of disaster risk management. In order to allow multiple stakeholders to share disaster-related information, the study identifies existing and emerging patterns that can be used to organize disaster knowledge, and then aligns them to the DOLCE1 foundational ontology.

6. The paper "A Deep Learning Framework to Predict Rating using Item Metadata for Cold Start Item", by Fahad Anwaar, Naima Iltaf, Hammad Afzal and Haider Abbas, proposes a content-based method to solve the Cold Start (CS) problem in recommender systems. The method is based on content embedding in deep learning models and predicts the ratings for CS items by exploiting item's textual description including both short item description and user tags.

7. The paper "Empirical Analysis of Semantic Metadata Extraction from Video Lecture Subtitles", by Marcos Vinícius Borges, Julio Cesar dos Reis and Guilherme Gribeler, presents the results of a comparison and an evaluation of existing semantic annotation tools applied to the extraction of semantic metadata from video lectures.

8. The paper "An Ontology-Based Monitoring System in Vineyards of the Burgundy Region", by Amira Mouakher, Rami Belkaroui, Aurélie Bertaux, Ouassila Labbani, Clementine Hugol-Gential and Christophe Nicolle, presents the development of a project of monitoring and recommending a vineyard supervision and control, using ontologies and a sensor system that also feeds the ontology.

9. The paper "An Effective Retrieval Approach for Documents related to Past Civil Engineering Projects", by Christian Esposito and Oscar Tamburis, proposes a pragmatic approach for retrieving appropriate documents from numerous documents generated during civil engineering projects. The approach is based on combining and extending capabilities of the Alfresco document management system and the Lucene/Solr query system.

10. The paper "Soft Ontologies as Fuzzy RDF Statements", by Julio Cesar Dos Reis, Rodrigo Bonacin and Luma Lombello, presents an original approach representing imprecise/uncertain knowledge through soft ontologies expressed in fuzzy RDF. The approach is based on a probabilistic matrix representing the domain concepts associated to the ontological dimensions with a certain degree of belief.

B. Short Paper

11. The paper "Using the LDA method for enriching an ontology from texts", by Amani Mechergui and Anis Tissaoui, presents a top-down approach which deals with semi-automatic ontology construction from texts. It relies on LDA for topic extraction from text and uses a core ontology to restrict the noun phrases to the ones that are specialization of the core concepts.

IV W2T TEAM

Program Committee

For the 2019 edition, we could count upon the precious work done by the members of our Program Committee in reviewing all papers contributing to improve W2T content. We are thankful to:

- * Haider ABBAS, National University of Sciences & Technology, Pakistan.
- * Frederic ANDRES, National Institute of Informatics, Japan.
- * Emna Hachicha BELGHITH, UNamur, Belgium.
- * Ismael BOUASSIDA, ReDCAD, University of Sfax, Tunisia.
- * Paulo Marcos Siqueira BUENO, CFlex MPC, Brazil.
- * Wojciech CELLARY, Poznan University of Economics, Poland.
- * Raja CHIKY, ISEP, France.
- * Marcos DA SILVEIRA, LIST, Luxembourg.
- * Giuseppe DE RUVO, University of Auckland, New Zealand.
- * Juliette DIBIE, AgroParisTech, INRA, France.
- * Julio Cesar DOS REIS, Computer Science Institute, Unicamp, Brazil.
- * Yucong DUAN, CS Department, Hainan University, China.

- * André Ricardo Abed GREGIO, Federal University of Paraná, Brazil.
- * Umair ul HASSAN, The Insight Centre for Data Analytics, Ireland.
- * Sergio ILARRI, University of Zaragoza, Spain.
- * Anum JAVAID, National University of Sciences & Technology, Pakistan.
- * Dilvan MOREIRA, ICMC, USP, Brazil.
- * Eugenio ORLANDI, ACEA and Europol, The Netherlands.
- * Cédric PRUSKI, LIST, Luxembourg.
- * Ivan RICARTE, FT-Unicamp, Brazil.
- * Ramon SALVADOR VALLÉS, UPC-Barcelona Tech, Spain.
- * Lina SOUALMIA, Université de Rouen et CHU de Rouen, France.

Chairs

- Rodrigo BONACIN, CTI, Brazil.
- Mariagrazia FUGINI, Italy.
- Riccardo MARTOGLIA, UniMoRe, Italy.
- Olga NABUCO, CTI, Brazil.
- Fatiha SAÏS, LRI, Paris Sud University, France.

V REFERENCES

Bonacin, R., Fugini, M.G., Martoglia, R. and Nabuco, O., 2018, June. "Web2Touch 2018: Semantic Technologies" in Smart Information Sharing and Web Collaboration. In *2018 IEEE 27th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)* (pp. 251-253). IEEE.