Web2Touch 2020-21

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
*F*IM - 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 Saïs
LRI, CNRS and Paris Saclay University
Orsay, France
fatiha.sais@lri.fr

Abstract—This foreword introduces a summary of themes and papers of the *Web2Touch (W2T) 2020-21* Track at the 29th *IEEE WETICE Conference* held as a virtual Conference, in October 2020. W2T 2020-21 includes six full papers and four short papers. 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 address a portfolio of hot issues in research and applications of semantics, smart technologies (e.g., IoT, sensors, devices for tele-monitoring, and smart contents management) with crucial topics, such as big data analysis, knowledge representation, 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 radically change the role of the semantic Web and applications.

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

I INTRODUCTION

Web2Touch (W2T), since its early editions (2007 onwards), is 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 the driver 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 2020-21 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 management, 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 the 2020-21 edition, W2T track includes 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 2020-21 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 it 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:
- Proposing approaches to dealing with heterogeneous and complex data, including the generation of services, the detection of duplicate records and the navigation on multidimensional data;
- Exploring new advanced techniques and semantic web technologies to make personalized recommendations, applied to big scientific datasets, collaborative work environments and health data. As well as pointing to the need to consider semantics in cyber-physical systems solutions.
- Proposing approaches to deal with the quality of linked data, including aspects such as integration, invalidate identity links as well as link maintenance. Such approaches were applied to different scenarios and databases such as e-Government, DBpdia and life sciences databases.

II ADDRESSED TOPICS

W2T 2020-21 concerns improvements obtained through enhanced organization and management of knowledge. Examples are models and tools to represent dynamic changes in shared information, smart (and 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.
    - o Ontology tools, ontology engineering, reuse and integration.
    - o Knowledge Graph expansion, enrichment, and validation.

- o Knowledge and data quality assessment.
- o Data profiling and axiom discovery.
- o Knowledge and Data provenance.
- o Knowledge and Data veracity.
- Semantic sharing and collaborative knowledge management.
    - o 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.
    - o Crowdsourcing techniques for semantic collaboration.
    - o Use of semantics in mobile Web, wearable and edge computing, and cross-device content management and delivery.
    - o Semantic annotations, semantic of collaborative Web sources and Semantic technologies for information Extraction Transformation and Loading (ETL).
- Collaborative semantic Web techniques and applications.
    - o Experiences and best practices in semantic Web support for collaborative work and business.
    - o Semantic techniques for cultural progress and community involvement.
    - o 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.
    - o 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 "A Model-Driven Approach for Semantic Data-as-a-Service Generation", by Hela Taktak, Khouloud Boukadi, Michael Mrissa, Chirine Ghedira-Guégan and Faiez Gargouri, exploits Model-Driven Engineering (MDE) and semantic techniques to provide a new solution for data-as-a-service generation. The approach is applied to the detection of natural disasters and is evaluated on a set of real data sources.

2. The paper "Capturing and Exploiting Citation Knowledge for Recommending Recently Published Papers", by Anita Khadka, Iván Cantador and Miriam Fernandez, proposes a novel citation recommendation approach that exploits citation knowledge to make personalized suggestions. The paper includes experimental analyses on a newly built dataset containing detailed paper citation data.

3. The paper "Duplicate record detection approach based on sentence embeddings", by Lattar Hafsa, Ben Salem Aicha and Hajjami Ben Ghezala Henda, presents a new approach for duplicate record detection where a sentence encoder model is used to embed the values of records' attributes into embeddings vectors. By means of support vector machine algorithms, the effectiveness of the approach is improved and shown to outperform the baselines.

4. The paper "An End-to-End Framework for Integrating and Publishing Linked Open Government Data, by *Rabeb Abida, Emna Hachicha Belghith and Anthony Cleve*, presents an end-to-end framework for converting and publishing linked government statistical data. It is based on linked open data technologies. The aim is to assist users to produce and publish structured linked data in an e-Government context. The issue is to determine the quality of the published data. The proposed framework integrates data refinement and semantic visualization of data.

5. The paper "Dissimilarity-based approach for Identity Link Invalidation", by *Anderson Carlos Ferreira Da Silva, Fatiha Saïs, Emmanuel Waller and Frederic Andres*, proposes a framework to invalidate identity links by dissimilarity and outlier detection in equivalence classes of identity links. The proposed identity link invalidation method is supported by a data fusion method, which couples equivalence classes for a set of owl:sameAs links and assigns a quality degree for each property value. Equivalence classes and their quality assessment are used to detect outliers. Experiments on real datasets like DBpdia have shown the relevance of the approach.

6. The paper "LODMF: A Linked Open Data Maintenance Framework", by *André Regino, Julio Cesar dos Reis and Rodrigo Bonacin*, presents a Linked Open Data Maintenance Framework (LODMF) that aims at supporting the maintenance of links during the evolution of RDF datasets. It allows (i) detecting RDF changes due to broken links, (ii) classifying affected and non-affected links, and (iii) applying maintenance actions in order to update links.

.

B. Short Papers

1. The paper "Managing and recommending resources in web-based collaborative working environments", by *Siying Li, Marie-Helene Abel and Elsa Negre*, presents an approach that considers a web-based Collaborative Work Environment (CWE) as an ontology-based information system, aiming to generate more accurate resource recommendations for users.

2. The paper "Gesture Recognition for Navigation in Multi-Dimensional Data", by *Mariagrazia Fugini, Jacopo Finocchi and Giulia Trasa*, presentes practical experimentations using machine learning techniques for enhancing the gesture based navigation on data collected by IoT sensors.

3. The paper "A Review on Cyber-Physical Systems: Models and Architectures", by *Mohamed Anis Aguida, Samir Ouchani and Mourad Benmalek*, presents a review of initiatives in modeling and designing Cyber-Physical Systems, including solutions for smart grid, smart healthcare, smart manufacturing, and smart city.

4. The paper "Tracing patient PLOD by mobile phones: Mitigation of epidemic risks based on patient locational open data", by *Ikki Ohmukai, Yasunori Yamamoto, Maori Ito and Takashi Okumura*, proposes a framework that applies RDF and open data technology on patient location data to estimate potential risks of patient contacts and mitigate of epidemic risks.

IV W2T TEAM

Program Committee

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:

- Frederic ANDRES, National Institute of Informatics, Japan.
- Ismael BOUASSIDA, ReDCAD, University of Sfax, Tunisia.
- Wojciech CELLARY, Poznan University of Economics, Poland.
- Marcos DA SILVEIRA, LIST, Luxembourg .
- Giuseppe DE RUVO, University of Auckland, New Zealand.
- 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.
- Emna HACHICHA BELGHITH, Université de Namur ASBL, Belgium

- Umair ul HASSAN, The Insight Centre for Data Analytics, Ireland.
- Sergio ILARRI, University of Zaragoza, Spain.
- Anum JAVAID, National University of Sciences & Technology, Pakistan.
- Ana Maria MONTEIRO, Unifaccamp, Brazil
- Dilvan MOREIRA, ICMC, USP, Brazil.
- Cédric PRUSKI, LIST, Luxembourg.
- Ivan RICARTE, FT-Unicamp, Brazil.
- Ferrucio De Franco ROSA, CTI, Brazil.
- Ramon SALVADOR VALLÉS, UPC-Barcelona Tech, Spain.
- Lina SOUALMIA, Université de Rouen et CHU de Rouen, France.
- Mahsa TEIMOURIKIA, CadMakers, Vancouver, Canada.
- Virginie THION, Université Rennes 1, France.

Chairs
- Rodrigo BONACIN, CTI, Brazil
- Mariagrazia FUGINI, Politecnico di Milano, Italy
- Riccardo MARTOGLIA, UniMoRe, Italy
- Olga NABUCO, CTI, Brazil
- Fatiha SAÏS, LRI, Paris Saclay University, France

V REFERENCES

Bonacin, R., Fugini, M., Martoglia, R. Nabuco, O., Sais, F. 2019 "Web2Touch 2019: Semantic Technologies for Smart Information Sharing and Web Collaboration". In 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE) (pp. 255-258). IEEE.