

Selected Topics on Systems Modeling and Natural Language Processing: Editorial Introduction to the Issue 7 of CSIMQ

Witold Andrzejewski¹ and Selma Khouri^{2,3}

¹Institute of Computing Science, Poznan University of Technology, Piotrowo 2, 60-965, Poznan, Poland

²LIAS/ISAE-ENSMA, Téléport 2 - 1 avenue Clément Ader, BP 40109, 86961 Futuroscope Chasseneuil Cedex, Poitiers, France

³National High School for Computer Science (ESI), BP 68M OUED SMAR, 16309, El-Harrach, Algiers, Algeria

witold.andrzejewski@cs.put.poznan.pl, selma.khouri@ensma.fr

The seventh issue of Complex Systems Informatics and Modeling Quarterly presents five papers devoted to two distinct research topics: systems modeling and natural language processing (NLP). Both of these subjects are very important in computer science. Through modeling we can simplify the studied problem by concentrating on only one aspect at a time. Moreover, a properly constructed model allows the modeler to work on higher levels of abstraction and not having to concentrate on details. Since the size and complexity of information systems grows rapidly, creating good models of such systems is crucial. The first two papers in this issue cover systems modeling.

“Modeling Enterprise Authorization: A Unified Metamodel and Initial Validation” by Matus Korman et al. deals with the problems of enterprise systems security modeling. This subject has always been very relevant and as such has received a lot of attention in the literature. Most of the proposed approaches to this problem are use case specific and are not able to solve all of the problems that need to be addressed. The authors of the paper start with analyzing the most important authorization engineering approaches, including: Discretionary-, Mandatory-, Role Based-, Attribute Based-, Risk Adaptive- and Token Based Access Control. Based on these solutions, a unifying ontology is designed, which is usable in multiple application contexts in Enterprise Architecture.

Theo P. van der Weide et al. in their paper “Static and Dynamic Aspects of Application Domains: An Inductively Defined Modeling Technique That Allows Decomposition” address a very relevant and rarely researched problem in information development – describing information structure states and how they change in time. The importance of this problem stems from the currently observed tendency to introduce connections between multiple different information systems (e.g., systems of multiple different companies that are part of a single logistics chain). The authors of the paper present a generic modeling technique capable of expressing system data structure, states of relevant concepts and their life cycle in terms of System Dynamics. The authors also provide an inductive definition of their technique which allows easily proving properties of information systems as well as performing decomposition of complex application domains.

An analysis of natural language is slowly becoming a widely used tool in commerce and day to day life. Opinion mining allows recommender systems to provide accurate recommendations based on user-generated reviews. Speech recognition and NLP are the basis for such widely used personal assistants as Apple’s Siri, Microsoft’s Cortana, and Google Now. While a lot of work has already been done on natural language processing, the research usually concerns widely used languages, such as English. Consequently, natural language processing in languages other than

English is a very relevant subject. The paper by Gatis Špats and Ilze Birzniece “Opinion Mining in Latvian Text Using Semantic Polarity Analysis and Machine Learning Approach” is a part of this trend. The authors present main obstacles and challenges for efficient opinion mining of texts written in Latvian, and present results of using two methods for opinion mining: lexicon based method and machine learning method based on Naive Bayes classifier. For the purpose of this paper, authors have prepared a large dataset of annotated Latvian words and made it publicly available.

Following the trend of natural language processing of non-English languages, the paper “REDEN: Named-Entity Linking in Digital Literary Editions Using Linked Data Sets” by Carmen Brando et al. tackles the problem of Named-entity Linking of authors’ names in French literary criticism and essays from the 19th–20th centuries. These names are disambiguated and annotated with links to broad-coverage and domain-specific Linked Data sets, DBpedia and BnF. It is also worth to mention that while the presented algorithm was used in the French literature context, it is much more generic. Moreover, the implementation of the algorithm is available as open source code.

A fifth paper in this issue: “An Approach for Multi-Artifact Testing Through an Ontological Perspective for Behavior-Driven Development” by Thiago Rocha Silva et al., combines both, modeling and analysis of natural language. The paper presents a Behavior-Driven Development based method to improve the iterative development process of interactive systems. The solution allows to automatically verify the consistency of evolving user software requirements with the corresponding development artifacts (including models and tests). The paper uses ontology for specifying tests that can run over multiple artifacts sharing similar concepts. The authors demonstrate the feasibility of the approach via a case study.

We hope that the readers will find this choice of papers interesting. We would like to thank all the authors for submitting their papers to the CSIMQ Journal and CSIMQ Editorial Review Board and external reviewers (Michał Szychowiak, Agnieszka Lawrynowicz, Barbara Catania, Robert Wrembel, Brice Chardin, Jędrzej Potoniec, Elsa Negre, Kathia Oliveira, Mikołaj Morzy, Paweł Boinski, Jarri Porras and Margita Kon-Popovska) who provided multiple important comments and suggestions for the authors.