

Preface

The preface of this volume collects the prefaces of the post-workshop proceedings of the individual workshops. The actual workshop papers, grouped by event, can be found in the body of this volume.

First International Workshop on Natural Language Processing for Informal Text (NLPIT 2015)

Organizers: Mena B. Habib, University of Twente, The Netherlands; Florian Kunneman, Radboud University, The Netherlands; Maurice van Keulen, University of Twente, The Netherlands

The rapid growth of Internet usage in the last two decades adds new challenges to understanding the informal user generated content (UGC) on the Internet. Textual UGC refers to textual posts on social media, blogs, emails, chat conversations, instant messages, forums, reviews, or advertisements that are created by end-users of an online system. A large portion of language used on textual UGC is informal. Informal text is the style of writing that disregards language grammars and uses a mixture of abbreviations and context dependent terms. The straightforward application of state-of-the-art Natural Language Processing approaches on informal text typically results in a significantly degraded performance due to the following reasons: the lack of sentence structure; the lack of enough context required; the uncommon entities involved; the noisy sparse contents of users' contributions; and the untrusted facts contained.

This was the reason for organizing this workshop on Natural Language Processing for Informal Text (NLPIT) through which we hope to bring the opportunities and challenges involved in informal text processing to the attention of researchers. In particular, we are interested in discussing informal text modelling, normalization, mining, and understanding in addition to various application areas in which UGC is involved. The first NLPIT workshop was held in conjunction with ICWE: the International Conference on Web Engineering held in Rotterdam, The Netherlands, July 23–26, 2015. It was organized by Mena B. Habib and Maurice van Keulen from the University of Twente, and Florian Kunneman from Radboud University, The Netherlands.

The workshop started with a keynote presentation from Nathan Schneider from the University of Edinburgh entitled “Hacking a Way Through the Twitter Language Jungle: Syntactic Annotation, Tagging, and Parsing of English Tweets.” Nathan explained how rich information structures can be extracted from informal text and represented in annotations. Tweets, informal text in general, is in a sense street language, but even street language is almost never entirely ungrammatical. So, even grammatical clues can be extracted, represented in annotations, and used to grasp the meaning of the text. We thank the Centre for Telematics and Information Technology (CTIT) for sponsoring this keynote presentation.

The keynote was followed by 4 research presentations selected from 7 submissions that NLPIT attracted. The common theme of these presentations was Natural Language Processing techniques for a multitude of languages. Among the 4 presentations, we saw Japanese, Tunisian, Kazakh, and Spanish. The first presentation was about extracting ASCII art embedded in English and Japanese texts. The second and fourth presentations were about constructing annotated corpora for use in research for the Tunisian dialect and Spanish, respectively. The third presentation was about word alignment issues in translating between Kazakh and English.

We thank all speakers and the audience for an interesting workshop with fruitful discussions. We furthermore hope that this workshop is the first of a series of NLPIT workshops.

July 2015

Mena Badieh Habib
Florian Kunneman
Maurice van Keulen

Program Committee

Alexandra Balahur	The European Commission's Joint Research Centre (JRC), Italy
Barbara Plank	University of Copenhagen, Denmark
Diana Maynard	University of Sheffield, UK
Djoerd Hiemstra	University of Twente, The Netherlands
Kevin Gimpel	Toyota Technological Institute, USA
Leon Derczynski	University of Sheffield, UK
Marieke van Erp	VU University Amsterdam, The Netherlands
Natalia Konstantinova	University of Wolverhampton, UK
Robert Remus	Universität Leipzig, Germany
Wang Ling	Carnegie Mellon University, USA
Wouter Weerkamp	904Labs, The Netherlands
Zhemini Zhu	University of Twente, The Netherlands

First Workshop on PErvasive WEb Technologies, Trends and Challenges (PEWET 2015)

Organizers: Fernando Ferri, Patrizia Grifoni, Alessia D’Andrea, and Tiziana Guzzo, Istituto di Ricerche sulla Popolazione e le Politiche Sociali (IRPPS), National Research Council, Italy

Pervasive Information Technologies, such as mobile devices, social media, cloud, etc., are increasingly enabling people to easily communicate and to share information and services by means of read-write Web and user generated contents. They influence the way individuals communicate, collaborate, learn, and build relationships. The enormous potential of Pervasive Information Technologies have led scientific communities in different disciplines, from computer science to social science, communication science, and economics, to analyze, study, and provide new theories, models, methods, and case studies. The scientific community is very interested in discussing and developing theories, methods, models, and tools for Pervasive Information Technologies. Challenging activities that have been conducted in Pervasive Information Technologies include social media management tools & platforms, community management strategies, Web applications and services, social structure and community modeling, etc.

To discuss such research topics, the PErvasive WEb Technologies, trends and challenges (PEWET) workshop was organized in conjunction with the 15th International Conference on Web Engineering - ICWE 2015. The workshop, held in Rotterdam, the Netherlands, on June 23–26, 2015, provided a forum for the discussion of Pervasive Web Technologies theories, methods, and experiences. The workshop organizers decided to have an invited talk, and after a review process selected five papers for inclusion in the ICWE workshops proceedings. Each of these submissions was rigorously peer reviewed by at least three experts. The papers were judged according to their originality, significance to theory and practice, readability, and relevance to workshop topics. The invited talk discussed the fractally-organized connectionist networks that according to the speaker may provide a convenient means to achieve what Leibniz calls “an art of complication,” namely an effective way to encapsulate complexity and practically extend the applicability of connectionism to domains such as socio-technical system modeling and design.

The selected papers address two areas: i) Internet technologies, services, and data management and, ii) Web programming, application, and pervasive services.

In the “Internet technologies, services, and data management” area, papers discuss different issues such as retrieval and content management. In the current information retrieval paradigm, the host does not use the query information for content presentation. The retrieval system does not know what happens after the user selects a retrieval result and the host also does not have access to the information which is available to the retrieval system. In the paper titled “Responding to Retrieval: A Proposal to Use Retrieval Information for Better Presentation of Website Content” the author provided a better search experience for the user through better presentation of the content based on the query, and better retrieval results, based on the feedback to the retrieval system from the host server. The retrieval system shares some information with the host server and the host server in turn provides relevant feedback to the retrieval system.

Another issue discussed at the workshop was the modeling and creation of APIs, proposed in the paper titled “Internet-Based Enterprise Innovation through a Community-Based API Builder to Manage APIs” in which an API builder is proposed as a tool for easily creating new APIs connected with existing ones from Cloud-Based Services (CBS).

The Internet of Things (IoT) is addressed in the paper titled “End-User Centered Events Detection and Management in the Internet of Things” where the authors provide the design of a Web environment developed around the concept of event, i.e., simple or complex data streams gathered from physical and social sensors that are encapsulated with contextual information (spacial, temporal, thematic).

In the area “Web programming, application, and pervasive services” papers discuss issues such as the application of asynchronous and modular programming. This issue is complex because asynchronous programming requires uncoupling of a module into two sub-modules, which are non-intuitively connected by a callback method. The separation of the module spurs the birth of another two issues: callback spaghetti and callback hell. Some proposals have been developed, but none of them fully support modular programming and expressiveness without adding a significant complexity. In the paper titled “Proposals for Modular Asynchronous Web Programming: Issues & Challenges” the authors compare and evaluate these proposals, applying them to a non-trivial open source application development.

Another issue is that of “future studies,” referring to studies based on the prediction and analysis of future horizons. The paper titled “Perspectives and Methods in the Development of Technological Tools for Supporting Future Studies in Science and Technology” gives a review of widely adopted approaches in future study activities, with three levels of detail. The first one addresses a wide scale mapping of related disciplines, the second level focuses on traditionally adopted methodologies, and the third one goes into greater detail. The paper also proposes an architecture for an extensible and modular support platform able to offer and integrate tools and functionalities oriented toward the harmonization of aspects related to semantics, document warehousing, and social media aspects. The success of the PEWET workshop would not have been possible without the contribution of the ICWE 2015 organizers and the workshop chairs, Florian Daniel and Oscar Diaz, the PC members, and the authors of the papers, all of whom we would like to sincerely thank.

July 2015

Fernando Ferri
Patrizia Grifoni
Alessia D’Andrea
Tiziana Guzzo

Program Committee

Ahmed Abbasi	University of Virginia, USA
Maria Chiara	CNR, Italy
Caschera	
Maria De Marsico	University of Rome “La Sapienza”, Italy
Arianna D’Ulizia	CNR, Italy

Rajkumar Kannan	Bishop Heber College, India
Marco Padula	National Research Council, Italy
Patrick Paroubek	LIMSI-CNRS, France
Adam	Poznań University of Technology, Poland
Wojciechowski	

First International Workshop in Mining the Social Web (SoWeMine 2015)

Organizers: Spiros Sirmakessis, Technological Institution of Western Greece, Greece; Maria Rigou, University of Patras, Greece; Evanthia Faliagka, Technological Institution of Western Greece, Greece

The rapid development of modern information and communication technologies (ICTs) in the past few years and their introduction into people's daily lives has greatly increased the amount of information available at all levels of their social environment.

People have been steadily turning to the social web for social interaction, news and content consumption, networking, and job seeking. As a result, vast amounts of user information are populating the social Web. In light of these developments the social mining workshop aims to study new and innovative techniques and methodologies on social data mining.

Social mining is a relatively new and fast-growing research area, which includes various tasks such as recommendations, personalization, e-recruitment, opinion mining, sentiment analysis, searching for multimedia data (images, video, etc.).

This workshop aimed to study (and even go beyond) the state of the art on social web mining, a field that merges the topics of social network applications and web mining, which are both major topics of interest for ICWE. The basic scope is to create a forum for professionals and researchers in the fields of personalization, search, text mining, etc. to discuss the application of their techniques and methodologies in this new and very promising research area.

The workshop tried to encourage the discussion on new emergent issues related to current trends derived from the creation and use of modern Web applications.

Six very interesting presentations took place in two sessions

– *Session 1: Information and Knowledge Mining in the Social Web*

- “Sensing Airport Traffic by Mining Location Sharing Social Services” by John Garofalakis, Ioannis Georgoulas, Andreas Komninos, Periklis Ntentopoulos, and Athanasios Plessas, University of Patras, Greece & University of Strathclyde, Glasgow, UK

The paper works with location sharing social services; quite popular among mobile users resulting in a huge social dataset. Authors consider location sharing social services' APIs endpoints as “social sensors” that provide data revealing real-world interactions. They focus on check-ins at airports performing two experiments: one analyzing check-in data collected exclusively from Foursquare and another collecting additional check-in data from Facebook. They compare the two location sharing social platforms' check-ins and show in Foursquare that data can be indicative of the passengers' traffic, while their number is hundreds of times lower than the number of actual traffic observations.

- “An Approach for Mining Social Patterns in the Conceptual Schema of CMS-based Web Applications” by Vassiliki Gkantouna, Athanasios Tsakalidis, Giannis Tzimas, and Emmanouil Viennas, University of Patras, & Technological Educational Institute of Western Greece, Greece

In this work, authors focus on CMS-based web applications that exploit social networking features and propose a model-driven approach to evaluating their hypertext schema in terms of the incorporated design fragments that perform a social network related functionality. Authors have developed a methodology, which, based on the identification and evaluation of design reuse, detects a set of recurrent design solutions denoting either design inconsistencies or effective reusable social design structures that can be used as building blocks for implementing certain social behavior in future designs.

- “An E-recruitment System Exploiting Candidates’ Social Presence” by Evanthia Faliagka, Maria Rigou, and Spiros Sirmakessis, Technological Educational Institution of Western Greece, University of Patras, & Hellenic Open University, Greece

This work aims to help HR Departments in their job. Applicant personality is a crucial criterion in many job positions. Choosing applicants whose personality traits are compatible with job positions is the key issue for HR. The rapid deployment of social web services has made candidates’ social activity much more transparent, giving us the opportunity to infer features of candidate personality with web mining techniques. In this work, a novel approach is proposed and evaluated for automatically extracting candidates’ personality traits based on their social media use.

– *Session 2: Mining the Tweets*

- “#nowplaying on #Spotify: Leveraging Spotify Information on Twitter for Artist Recommendations” by Martin Pichl, Eva Zangerle, and Günther Specht, Institute of Computer Science, University of Innsbruck, Austria

The rise of the Web has opened new distribution channels like online stores and streaming platforms, offering a vast amount of different products. To help customers find products according to their taste on those platforms, recommender systems play an important role. Authors present a music recommendation system exploiting a dataset containing listening histories of users, who posted what they are listening to at the moment on Twitter. As this dataset is updated daily, they propose a genetic algorithm, which allows the recommender system to adopt its input parameters to the extended dataset.

- “Retrieving Relevant and Interesting Tweets during Live Television Broadcasts” by Rianne Kaptein, Yi Zhu, Gijs Koot, Judith Redi, and Omar Niamut, TNO, The Hague & Delft University of Technology, The Netherlands

The use of social TV applications to enhance the experience of live event broadcasts has become an increasingly common practice. An event profile, defined as a set of keywords relevant to an event, can help to track messages related to these events on social networks. Authors propose an event profiler that retrieves relevant and interesting tweets in a continuous stream of event-related tweets as they are posted. For testing the application they have executed a user study. Feedback is collected during a live broadcast by giving the participant the option to like or dislike a tweet, and by judging a selection of tweets on relevancy and interest in a post-experiment questionnaire.

- “Topic Detection in Twitter Using Topology Data Analysis” by Pablo Torres-Tramon, Hugo Hromic, and Bahareh Heravi, Insight Centre for Data Analytics, National University of Ireland, Galway

The authors present automated topic detection in huge datasets in social media. Most of these approaches are based on document clustering and burst detection. These approaches normally represent textual features in standard n -dimensional Euclidean metric spaces. Authors propose a topic detection method based on Topology Data Analysis that transforms the Euclidean feature space into a topological space where the shapes of noisy irrelevant documents are much easier to distinguish from topically-relevant documents.

July 2015

Spiros Sirmakessis
 Maria Rigou
 Evanthia Faliagka

Program Committee

Olfa Nasraoui	University of Louisville, USA
Martin Rajman	EPFL, Switzerland
Evanthia Faliagka	Technological Institution of Western Greece
John Garofalakis	University of Patras, Greece
Maria Rigkou	University of Patras, Greece
Spiros Sioutas	Ionian University, Greece
Spiros Sirmakessis	Technological Educational Institution of Western Greece
John Tsaknakis	Technological Educational Institution of Western Greece
John Tzimas	Technological Educational Institution of Western Greece
Vasilios Verikios	Hellenic Open University, Greece

Current Trends in Web Engineering

15th International Conference, ICWE 2015 Workshops,
NLPIT, PEWET, SoWEMine, Rotterdam, The Netherlands,
June 23-26, 2015. Revised Selected Papers

Daniel, F.; Diaz, O. (Eds.)

2015, XVI, 199 p. 42 illus. in color., Softcover

ISBN: 978-3-319-24799-1