1st Workshop on Natural Language Processing for Requirements Engineering (NLP4RE'18)

Chapter · March 2018

4 authors:

Fabiano Dalpiaz
Utrecht University
150 PUBLICATIONS   **2,581** CITATIONS

SEE PROFILE

Alessio Ferrari
Italian National Research Council
104 PUBLICATIONS   **1,129** CITATIONS

SEE PROFILE

Xavier Franch
Universitat Politècnica de Catalunya
426 PUBLICATIONS   **5,363** CITATIONS

SEE PROFILE

Cristina Palomares
Universitat Politècnica de Catalunya
38 PUBLICATIONS   **364** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Norsk COSI View project

Crowd-Based Requirements Engineering View project

1$^{\text{st}}$ Workshop on Natural Language Processing for Requirements Engineering (NLP4RE'18)

Fabiano Dalpiaz
University of Utrecht
f.dalpiaz@uu.nl

Alessio Ferrari
CNR-ISTI
alessio.ferrari@isti.cnr.it

Xavier Franch
UPC-BarcelonaTech
franch@essi.upc.edu

Cristina Palomares
UPC-BarcelonaTech
cpalomares@essi.upc.edu

1   Preface

Natural language processing (NLP) has played an important role in several computer science areas, and requirements engineering (RE) is not an exception. For over twenty years, several works were published on the application of NLP techniques to address RE specific problems, such as traceability, categorisation, defect detection, model generation, and more. In recent years, the advent of massive and heterogeneous natural language (NL) RE-relevant sources, like tweets and app reviews, has sparked the interest of the RE community in NLP. Furthermore, we witness the novel golden age of NLP technologies, enabled by deep and shallow learning approaches that have improved the accuracy of most NLP tasks, including parsing and machine translation. It is therefore an appropriate moment to create a venue in which researchers on applications of NLP to RE problems can meet, share ideas and create synergies, assisted by experts from the NLP community.

The current document is a preface to the proceedings of the 1$^{\text{st}}$ Workshop on Natural Language Processing for Requirements Engineering (NLP4RE'18, `http://fmt.isti.cnr.it/nlp4re/`), co-located with the 24$^{\text{th}}$ International Working Conference on Requirements Engineering: Foundation for Software Quality (REFSQ 2018) held in Utrecht, the Netherlands. The goal of NLP4RE'18 is to establish a meeting point for the researchers in the field to meet and discuss advances, challenges and barriers, and also to foster collaborations. The workshop features one keynote from Dan M. Berry, University of Waterloo on *Evaluation of NLP Tools for Hairy RE Tasks*. The keynote will illustrate, from an RE perspective, how to soundly assess the effectiveness of NLP tools in RE. Furthermore, NLP4RE'18 includes also an invited talk on *Domain-adaptation of Natural Language Processing Tools for RE*, from Tejaswini Deoskar, Institute for Logic, Language and Computation (ILLC), University of Amsterdam. This talk is given by a computational linguist who is an informed outsider about RE, and discusses solutions for one of the fundamental hurdles for NLP in RE: domain adaptation.

The workshop received 19 submissions. The papers were independently reviewed by three program committee members, and 11 papers were accepted by the co-organizers for presentation at the workshop. The authors of the remaining 8 papers were given the possibility to present a poster. This opportunity was accepted by the authors of 3 papers. Among these papers, 2 are presented also in the current proceedings as poster abstracts.

The papers and the posters can be grouped into four main groups: (1) technical papers discussing RE needs and associated NLP solutions; (2) report papers presenting past, ongoing and future work of research groups interested in NLP for RE; (3) vision papers discussing potential and intriguing research directions for the application of NLP technologies to RE problems; and (4) posters illustrating preliminary works. Below, we summarise the contributions within the different groups.

1.1 Technical Papers on RE Needs and NLP Solutions

Groen *et al.* [GSK+] focus on RE needs, and present a paper aimed at establishing whether NLP is actually necessary in RE. They investigate the problem of extracting requirements-relevant information from the large amounts of available online user feedback about software products such as apps, and they compare the amount of time required for conducting a manual versus automated analysis. They conclude that automated analysis is significantly faster, confirming the need for NLP in RE. Schlutter and Vogelsang [SV] focus on RE solutions instead, and they propose a method to ease requirements comprehension. The method collects disjoint requirements belonging to different components of a larger system, and automatically combines them into a single knowledge representation graph that relates the requirements. The coherent view on requirements concepts and relations provided by the graph can be used to support requirements comprehension and analysis.

1.2 Reports from Research Groups

Report papers from research groups, particularly encouraged by the call for papers of NLP4RE'18, form the large majority of the accepted contributions. Three reports come from within the RE research community [GF, BCFQ, FSM+], three reports belong to the artificial intelligence and NLP community [TH, BG, Tsa], and one report presents experiences from a company [Fem].

Gnesi and Ferrari [GF] present the past and current work of the FMT lab at CNR-ISTI on the usage of NLP for defect detection in requirements, support for product lines' synthesis, and analysis of requirements elicitation interviews. Borull *et al.* [BCFQ] report on the ongoing work of the GESSI group at UPC within the context of the OpenReq EU project, focussing on the identification of similar requirements, to support inter-dependency detection, and reuse. Fucci *et al.* [FSM+] present the work of the MAST group at the University of Hamburg, mostly dedicated to the automated analysis of online user feedback and app reviews for the extraction of requirements-relevant information.

Tsarfaty [Tsa], from the NLP lab at the Open University of Israel, presents the work of the group performed in the context of the ERC-StG research grant named *Natural Language Programming (NLPRO): Turning Texts into Executable Code.* The project aims at developing a novel natural language compiler that takes a natural language description of a system as input, and returns as output the respective executable. Töws and Heuss [TH] report on the research done at Fraunhofer FKIE concerning requirements grouping by means of different similarity evaluation techniques. Bäumer and Geierhos [BG], from the department of Digital Humanities at the University of Paderborn, present the ongoing work of their group concerning the development of the CORDULA platform, oriented to recognise and automatically compensate language inaccuracies (e.g., ambiguity, vagueness and incompleteness) in requirements.

Femmer [Fem] presents the only company report at the workshop, which introduces the Qualicen Requirements Scout, a commercial tool for defect detection in requirements. The author illustrates the initial research performed at TUM that led to the development of the tool, and outlines research outcomes and industrial challenges from his perspective of tool provider.

1.3 Vision Papers

Garigliano *et al.* [GPM] present a vision paper that argues that NLP tools for general RE should be focused on a deep internal semantic representation of the text, which attempts to describe the text meaning in a form that can be different from the original one. This is expected to enable the extraction of *implied* information that is not explicitly given in the text. Friesen *at al.* [FBG], instead, present the idea of improving requirements quality by using chatbot technologies. Chatbots are expected to support the automatic compensation of some deficits in natural language requirement descriptions, by means of direct and guided interaction with the user.

1.4 Posters

Kifetew *et al.* [KPS] present a poster discussing the lessons learned in the SUPERSEDE EU project, aimed at processing user feedback to support software evolution. Specifically, the authors discuss issues related to the analysis of feedback that was made available in German, rather than English, and the consequent need for adaptation of available NLP approaches, mostly tailored for English. Caron *et al.* [CBG] present the idea for a system aimed at extracting information from requirements with a syntactic approach, and they outline the architecture for such system. Gori *et al.* [GOP+] presents an experiment that uses machine learning to detect

requirements defects, and they illustrate the encountered problems and challenges. This last poster does not appear in the current proceedings.

2 Program Committee

We warmly thank all the reviewers of our Program Committee (PC), who helped in the selection of the papers by providing timely and accurate reviews. The PC members of NLP4RE are:

- Daniel M. Berry, University of Waterloo, Canada

- Jörg Dörr, Fraunhofer IESE, Germany

- Henning Femmer, Technical University of Munich, Germany

- Davide Fucci, University of Hamburg, Germany

- Vincenzo Gervasi, University of Pisa, Italy

- Eduard Groen, Fraunhofer IESE, Germany

- Emitzá Guzmán, University of Zurich, Switzerland

- Garm Lucassen, Utrecht University, the Netherlands

- Daniel Méndez, Technical University of Munich, Germany

- Luisa Mich, University of Trento, Italy

- Barbara Paech, University of Heidelberg, Germany

- Mehrdad Sabetzadeh, University of Luxembourg, Luxembourg

- Nicolas Sannier, University of Luxembourg, Luxembourg

- Pete Sawyer, Aston University, UK

- Norbert Seyff, University of Zurich and University of Applied Sciences and Arts Northwestern, Switzerland

- Michael Unterkalmsteiner, Blekinge Institute of Technology, Sweden

- Andreas Vogelsang, TU Berlin, Germany

References

[BCFQ] Ricard Borrull, Dolors Costal, Xavier Franch, and Carme Quer. Research on NLP for RE at UPC: a Report.

[BG]   Frederik Simon Bäumer and Michaela Geierhos. How to Deal with Inaccurate Service Requirements? Insights in Our Current Approach and New Ideas.

[CBG]   Matthew Caron, Frederik Simon Bäumer, and Michaela Geierhos. Back to Basics: Extracting Software Requirements with a Syntactic Approach.

[FBG]   Edwin Friesen, Frederik Simon Bäumer, and Michaela Geierhos. CORDULA: Software Requirements Extraction Utilizing Chatbot as Communication Interface.

[Fem]   Henning Femmer. Requirements Quality Defect Detection with the Qualicen Requirements Scout.

[FSM+] Davide Fucci, Christoph Stanik, Lloyd Montgomery, Zijad Kurtanovic, Timo Johann, and Walid Maalej. Research on NLP for RE at the University of Hamburg: a Report.

[GF]   Stefania Gnesi and Alessio Ferrari. Research on NLP for RE at CNR-ISTI: a Report.

[GOP+] Gloria Gori, Francesco Orsini, Marco Papini, Alessandro Fantechi, and Paolo Frasconi. Defect Detection and Machine Learning for Requirement Engineering: new Roadmaps.

[GPM]   Roberto Garigliano, Dominic Perini, and Luisa Mich. Which Semantics for Requirements Engineering: from Shallow to Deep.

[GSK+] Eduard Groen, Jacqueline Schowalter, Sylwia Kopczyska, Svenja Polst, and Sadaf Alvani. Is there Really a Need for NLP in RE? A Benchmarking Study to Assess Scalability of Manual Analysis.

[KPS]   Fitsum Meshesha Kifetew, Anna Perini, and Angelo Susi. Managing Multi-Lingual User Feedback: the SUPERSEDE project experience.

[SV]   Aaron Schlutter and Andreas Vogelsang. Knowledge Representation of Requirements Documents Using Natural Language Processing.

[TH]   Daniel Töws and Timm Heuss. Research on NLP for RE at Fraunhofer FKIE: a Report on Grouping Requirements.

[Tsa]   Reut Tsarfaty. Natural Language Programming (NLPRO): Turning Texts into Code.