TSAR 2023

Second Workshop on Text Simplification, Accessibility and Readability

associated with

The 14th International Conference on Recent Advances in Natural Language Processing'2023

Proceedings of the Workshop

September 7th, 2023
Varna, Bulgaria

Workshop on Text Simplification,
Accessibility and Readability
Associated with the International Conference
Recent Advances in Natural Language Processing'2023

PROCEEDINGS

Varna, Bulgaria
7th September 2023

ISBN 978-954-452-086-1

Designed by INCOMA Ltd.
Shoumen, BULGARIA

Preface

Welcome to the proceedings of the 2nd edition of the Workshop on Text Simplification, Accessibility and Readability (TSAR), hosted at the 14th Conference on Recent Advances in Natural Language Processing (RANLP 2023), in Varna, Bulgaria.

This year, we received 24 submissions to the workshop. These submissions covered a variety of current topics of interest to the TSAR community. Three papers examined the area of lexical simplification, considering context-awareness, lexical deletion and the fine-tuning of large language models (LLMs). Five papers considered simplification at the whole-text level, considering cognitive disabilities, the capacity of LLMs to adapt to genre-specific simplification, operations for simplification, coherence in document level simplification and tools for simplification of texts. Additionally, research was presented on accessibility and readability, covering easy-language translation, mediation and evaluation. The research was linguistically diverse, proposing solutions for English, Spanish, Swedish and French.

All submissions were peer-reviewed by the members of the program committee which includes distinguished specialists in text simplification, accessibility, and readability. Out of the 24 submissions to the workshop, 10 were rejected, 11 were accepted and 3 were accepted subject to improvements in line with reviewer feedback. Out of 14 accepted papers, 6 were selected to be presented orally and 8 as posters, which were presented during a lightning-talk session.

The workshop is held in-person, with online attendance for authors who were unable to attend due to constraints beyond the organisers control. The program encompasses: a keynote speech by Dr. Victoria Yaneva, National Board of Medical Examiners, USA; two oral sessions, comprising six presentations; a round of lightning talks to introduce the poster presentations; and a hosted discussion session on current issues and trends in text simplification, accessibility and readability research.

We would like to thank the members of the program committee for their timely help in reviewing the submissions and all the authors for submitting their papers to the workshop. We also thank the organisers of RANLP for hosting the workshop and their kind support in producing these proceedings.

TSAR Organizing Committee

Sanja Štajner, Matthew Shardlow, Fernando Alva-Manchego, Horacio Saggion

Organizing Committee

- Sanja Štajner, Independent Researcher
- Matthew Shardlow, Manchester Metropolitan University
- Fernando Alva-Manchego, Cardiff University
- Horacio Saggion, Universitat Pompeu Fabra

Programme Committee

- Bruce W. Lee (University of Pennsylvania)
- Christina Niklaus (University of St. Gallen)
- Giulia Venturi (Institute of Computational Linguistics "Antonio Zampolli" (ILC-CNR))
- Jaap Kamps (University of Amsterdam)
- Jasper Degraeuwe (Ghent University)
- Jipeng Qiang (Yangzhou University)
- Kim Cheng Sheang (Universitat Pompeu Fabra)
- Laura Vásquez-Rodríguez (University of Manchester)
- Liana Ermakova (HCTI EA-4249, Université de Bretagne Occidentale)
- Maja Popović (ADAPT, Dublin City University)
- Mounica Maddela (Georgia Institute of Technology)
- Natalia Grabar (CNRS STL UMR8163, Université de Lille)
- Oliver Alonzo (Rochester Institute of Technology)
- Philippe Laban (Salesforce Research)
- Piotr Przybyła (Universitat Pompeu Fabra)
- Regina Stodden (Heinrich Heine University Düsseldorf)
- Rémi Cardon (CENTAL, ILC, Université Catholique de Louvain)
- Sarah Ebling (University of Zurich)
- Susana Bautista (Universidad Francisco de Vitoria)
- Sweta Agrawal (University of Maryland)
- Tadashi Nomoto (National Institute of Japanese Literature)
- Tannon Kew (University of Zurich)

Table of Contents

Using ChatGPT as a CAT tool in Easy Language translation
Silvana Deilen, Sergio Hernández Garrido, Ekaterina Lapshinova-Koltunski and Christiane Maaß 1

Context-aware Swedish Lexical Simplification
Emil Graichen and Arne Jonsson 11

TextSimplifier: A Modular, Extensible, and Context Sensitive Simplification Framework for Improved Natural Language Understanding
Sandaru Seneviratne, Eleni Daskalaki and Hanna Suominen 21

Cross-lingual Mediation: Readability Effects
Maria Kunilovskaya, Ruslan Mitkov and Eveline Wandl-Vogt 33

Simplification by Lexical Deletion
Matthew Shardlow and Piotr Przybyła 44

Comparing Generic and Expert Models for Genre-Specific Text Simplification
Zihao LI, Matthew Shardlow and Fernando Alva-Manchego 51

Automatic Text Simplification for People with Cognitive Disabilities: Resource Creation within the Clear-Text Project
Isabel Espinosa-Zaragoza, José Abreu-Salas, Paloma Moreda and Manuel Palomar 68

Towards Sentence-level Text Readability Assessment for French
Duy Van Ngo and Yannick Parmentier 78

Document-level Text Simplification with Coherence Evaluation
Laura Vásquez-Rodríguez, Matthew Shardlow, Piotr Przybyła and Sophia Ananiadou 85

LSLlama: Fine-Tuned LLaMA for Lexical Simplification
Anthony Baez and Horacio Saggion 102

LC-Score: Reference-less estimation of Text Comprehension Difficulty
Paul Tardy, Charlotte Roze and Paul Poupet 109

On Operations in Automatic Text Simplification
Rémi Cardon and Adrien Bibal 116

An automated tool with human supervision to adapt difficult texts into Plain Language
Paul Poupet, Morgane Hauguel, Erwan Boehm, Charlotte Roze and Paul Tardy 131

Beyond Vocabulary: Capturing Readability from Children's Difficulty
Arif Ahmed 134