

Proceedings of the 10th conference
**Human Language Technologies –
the Baltic Perspective**

Baltic HLT 2022

Riga, Latvia
October 6-7, 2022

Edited by
Normunds Grūzītis, Inguna Skadiņa,
Andrejs Vasiljevs

Program Committee

Tanel Alumäe, Tallinn University of Technology, Estonia
Ilze Auziņa, Institute of Mathematics and Computer Science,
University of Latvia
Eduard Barbu, University of Tartu, Estonia
Guntis Bārzdīņš, Institute of Mathematics and Computer Science, University of Latvia
Aleksandrs Berdicevskis, University of Gothenburg, Sweden
Mark Fishel, University of Tartu, Estonia
Filip Ginter, University of Turku, Finland
Normunds Grūzītis, University of Latvia
Heiki-Jaan Kaalep, University of Tartu, Estonia
Kaarel Kaljurand, Nuance Communications, Austria
Jurgita Kapočiūtė-Dzikiėnė, Tilde / Vytautas Magnus University, Lithuania
Kimmo Kettunen, University of Eastern Finland
Margit Langemets, Institute of the Estonian Language, Estonia
Krister Lindén, University of Helsinki, Finland
Einar Meister, Tallinn University of Technology, Estonia
Kadri Muischnek, University of Tartu, Estonia
Costanza Navarretta, University of Copenhagen, Denmark
Maciej Odrodniczuk, Institute of Computer Science,
Polish Academy of Sciences, Poland
Petya Osenova, Bulgarian Academy of Sciences, Bulgaria
Mārcis Pinnis, Tilde, Latvia
Askars Salimbajevs, Tilde, Latvia
Baiba Saulīte, Institute of Mathematics and Computer Science, University of Latvia
Agata Savary, Université Paris-Saclay, France
Inguna Skadiņa, Tilde / University of Latvia
Raivis Skadiņš, Tilde
Michal Škrabal, Institute of the Czech National Corpus, Charles University,
Czech Republic
Andre Tättar, University of Tartu, Estonia
Amalia Todirascu, Université de Strasbourg, France
Andrius Utkas, Vytautas Magnus University, Lithuania
Jurgita Vaičenonienė, Vytautas Magnus University, Lithuania
Andrejs Vasiļjevs, Tilde
Elena Volodina, Språkbanken, University of Gothenburg, Sweden

Organising Committee

Normunds Grūzītis, University of Latvia

Baiba Saulīte, Institute of Mathematics and Computer Science, University of Latvia

Inguna Skadiņa, Tilde / Institute of Mathematics and Computer Science, University of
Latvia

Andrejs Vasiļjevs, Tilde

Organisers

Tilde

Institute of Mathematics and Computer Science, University of Latvia

University of Latvia

Supporters

The conference is supported by the European Regional Development Fund project No. 1.1.1.5/18/I/016 “University of Latvia and Institutes in the European Research Area - Excellence, Activity, Mobility, Capacity” and by the State Research Programme on Digital Resources of the Humanities (project No. VPP-IZM-DH-2020/1-0001).

Preface

For almost twenty years the Baltic HLT conference has provided a forum for sharing new ideas and recent advances in human language technology. This year the conference returns to Riga with the 10th instalment in the series “Human Language Technologies – the Baltic Perspective”.

The first Baltic HLT conference in 2004, was a significant moment for researchers, developers, and industry representatives, as it was the first event to focus on language technology development in the Baltic region. The success of this event inspired the organisation of biennial conferences in Estonia, Latvia and Lithuania.

During the past two decades notable progress has been made in language technology development in all three Baltic countries. Although our languages are often called less resourced or even low resourced significant results have been demonstrated in speech technologies, machine translation, and natural language understanding. These advances have narrowed the gap between high resourced languages and less resourced languages and help people daily.

More than 60 authors submitted papers for this conference. Most of the submissions were peer-reviewed by three reviewers. Twenty-three papers were selected for the conference and, for the first time, published in the Baltic Journal of Modern Computing. Papers in these proceedings represent a wide range of topics, including speech processing, machine translation, application of deep learning techniques to different natural language processing tasks, creation of language resources, and others.

We would like to express our deepest gratitude to all members of the program committee for their hard work. We also thank all the authors for their contributions, which demonstrate the most important recent achievements in the Baltic countries and reveal the latest tendencies in language technology research in our region.

Normunds Grūzītis, Inguna Skadiņa, Andrejs Vasiļjevs