LREC 2020 Workshop
Language Resources and Evaluation Conference
11–16 May 2020

1st Joint SLTU and CCURL Workshop
(SLTU-CCURL 2020)

PROCEEDINGS

Editors:
Dorothee Beermann, Laurent Besacier, Sakriani Sakti, and Claudia Soria

Proceedings of the LREC 2020
1st Joint SLTU and CCURL Workshop
(SLTU-CCURL 2020)

Edited by: Dorothee Beermann, Laurent Besacier, Sakriani Sakti, Claudia Soria

ISBN: 979-10-95546-35-1
EAN: 9791095546351

For more information:
European Language Resources Association (ELRA)
9 rue des Cordelières
75013, Paris
France
http://www.elra.info
Email: lrec@elda.org

Introduction

Created in April 2017, SIGUL (http://www.elra.info/en/sig/sigul/) is a joint Special Interest Group of the European Language Resources Association (ELRA) and of the International Speech Communication Association (ISCA). SIGUL intends to bring together a number of professionals involved in the development of language resources and technologies for under-resourced languages. Its main objective is to build a community that not only supports linguistic diversity through technology and ICT but also commits to increase the lesser-resourced languages (regional, minority, or endangered) chances to survive the digital world through language and speech technology.

Before the creation of SIGUL, two workshops addressed language technologies for low resource languages: there have been 6 editions of SLTU (Spoken Language Technologies for Under-resourced languages) which started in 2008; and 3 editions of CCURL (Collaboration and Computing for Under-Resourced Languages) which started in 2014. For 2020, and as a satellite event of LREC, SIGUL board decided to organize the 1st Joint Workshop of SLTU (Spoken Language Technologies for Under-resourced languages) and CCURL (Collaboration and Computing for Under-Resourced Languages) (SLTU-CCURL 2020).

We solicited papers related to all areas of natural language processing, speech and computational linguistics, as well as those at the intersection with digital humanities and documentary linguistics, provided that they address less-resourced languages. One goal of this workshop was to offer a venue where researchers in different disciplines and from varied backgrounds can fruitfully explore new areas of intellectual and practical development while honoring their common interest of sustaining less-resourced languages.

Our programme committee comprised 60 experts in natural language processing and spoken language processing from 19 countries. Each of the 64 submitted papers was reviewed by 3 committee members. We finally accepted 54 papers for the proceedings. We would like to express our sincere thanks to all members of this committee (who worked hard despite the difficult conditions associated with the pandemic) and authors for their great work in making this event a scientifically recognised international Workshop. We would also like to extend our thanks to all our sponsors: Google as platinum sponsor; ELRA, ISCA and ACL/SIGEL for endorsing this event.

Unfortunately, as a consequence of the COVID-19 pandemic, LREC 2020 has been canceled and - as a satellite event of the Conference - SLTU-CCURL 2020 has been canceled as well. We nevertheless hope that you will find these workshop proceedings relevant and stimulating for your own research. We are looking forward to see you soon for future events organised by SIGUL.

SLTU-CCURL-2020 Workshop co-chairs:

Dorothee Beermann (NTNU, Norway)
Laurent Besacier (LIG – Université Grenoble Alpes, France)
Sakriani Sakti (NAIST, Japan)
Claudia Soria (CNR-ILC, Italy)

Organizers

Dorothee Beermann (NTNU, Norway)
Laurent Besacier (LIG – Université Grenoble Alpes, France)
Sakriani Sakti (NAIST, Japan)
Claudia Soria (CNR-ILC, Italy)

Program Committee:

Adrian Doyle (University of Galway, Ireland)
Alexey Karpov (SPIIRAS, Russian Federation)
Alexis Palmer (University of North Texas, USA)
Amir Aharoni (Wikimedia Foundation)
Andras Kornai (Hungarian Academy of Sciences, Hungary)
Angelo Mario Del Grosso (CNR-ILC, Italy)
Antti Arppe (University of Alberta, Canada)
Atticus Harrigan (University of Alberta, Canada)
Charl Van Heerden (Saigen, South Africa)
Daan Van Esch (Google)
Dafydd Gibbon (Bielefeld University, Germany)
Delyth Prys (Bangor University, UK)
Dewi Bryn Jones (Bangor University, UK)
Dorothee Beermann (NTNU, Norway)
Emily Le Chen (University of Illinois, USA)
Federico Boschetti (CNR-ILC, Italy)
Francis Tyers (Indiana University, USA)
Gerard Bailly (GIPSA Lab, CNRS)
Gilles Adda (LIMSI/IMMI CNRS, France)
Heysem Kaya (Utrecht University, The Netherlands)
Hyunji "Hayley" Park (University of Illinois at Urbana-Champaign, USA)
Irina Kipyatkova (SPIIRAS, Russia)
Jeff Good (University at Buffalo, USA)
Jelske Dijkstra (Fryske Akademy, The Netherlands)
John Judge (ADAPT DCU, Ireland)
John Philip McCrae (National University of Ireland Galway, Ireland)
Jonas Fromseier Mortensen (Google)
Jordan Lachler (University of Alberta, Canada)
Joseph Mariani (LIMSI-CNRS, France)
Katherine Schmirler (University of Alberta, Canada)
Kepa Sarasola (University of the Basque Country, Spain)
Kevin Scannell (Saint Louis University, Missouri, USA)
Klara Ceberio (Elhuyar, Spain)
Lane Schwartz (University of Illinois at Urbana-Champaign, USA)
Lars Hellan (NTNU, Norway)
Lars Steinert (University of Bremen, Germany)
Laurent Besacier (LIG-IMAG, France)
Maite Melero (Barcelona Supercomputing Center, Spain)

Marcely Zanon Boito (LIG-IMAG, France)
Mathieu Mangeot-Nagata (LIG-IMAG, France)
Matt Coler (University of Groningen, The Netherlands)
Mohammad A. M. Abushariah (The University of Jordan, Jordan)
Nick Thieberger (University of Melbourne / ARC Centre of Excellence for the Dynamics of Language, Australia)
Omar Farooq (AMU, India)
Pierric Sans (Google)
Pradip K Das (IIT, India)
Richard Littauer (University of Saarland, Germany)
Sahar Ghannay (LIMSI, CNRS, France)
Sakriani Sakti (NAIST, Japan)
Satoshi Nakamura (NAIST, Japan)
Sebastian Stüker (KIT, Germany)
Shyam S Agrawal (KIIT, India)
Sjur Moshagen (UiT The Arctic University of Norway, Norway)
Solomon Teferra Abate (Addis Ababa University, Ethiopia)
Steven Bird (Charles Darwin University, Australia)
Tanja Schultz (Uni-Bremen, Germany)
Thang Vu (Uni-Stuttgart, Germany)
Teresa Lynn (ADAPT Centre, Ireland)
Trond Trosterud (Tromsø University, Norway)
Win Pa Pa (UCS Yangon, Myanmar)

Table of Contents