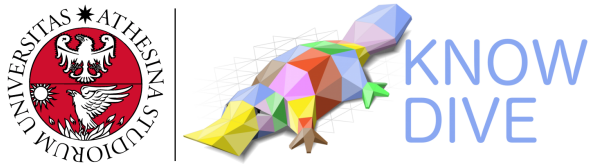
NSURL 2022

Proceedings of the third International Workshop on NLP Solutions for Under Resourced Languages (NSURL 2022) co-located with ICNLSP 2022

December 18, 2022 (virtual)



©2022 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-959429-61-6

http://nsurl.org/

Introduction

Welcome to NSURL2022, the third International Workshop on NLP Solutions for Under Resourced Languages, held online on December 18th 2022, co-located with ICNLSP 2022.
NSURL is a forum for researchers and students to exchange ideas and discuss research and trends in the field of Natural Language Processing and Speech Processing for under resourced languages.
Fourteen papers have been submitted to NSURL 2022. Seven of them have been accepted. All the papers have been presented orally.
We would like to acknowledge the support provided by University of Trento and Data-Scientia. We would like also to express our gratitude to the organizing and the program committees for the hard and valuable contributions.

Abed Alhakim Freihat and Mourad Abbas

Organizers:

Chair: Dr. Abed Alhakim Freihat
Co-chair: Dr. Mourad Abbas

Program Committee:

Abed Alhakim Freihat, University of Trento **(Chair)**
Mourad Abbas, HCLA, Algeria
Ahmed AbuRa'ed, Universitat P. F. Barcelona, Spain
Abdallah Abushmaes, Mawdoo3 Ltd, Jordan
Abdulmohsen Althubaity, The National Center for Artificial Intelligence and Big Data (kacst), KSA
Sunday Ojo, Tshwane University of Technology Pretoria, South Africa
Violetta Cavalli-Sforza, Al Akhawayn University, Morocco
Heshaam Faili, University of Tehran, Iran
Mohammad Gharib, University of Florence, Italy
Osama Hamed, University of Duisburg-Essen, Germany
Linda van Huyssteen Tshwane University of Technology, Pretoria, South Africa
Hadi Khaliliya, Univeristy of Trento, Italy
Mohamed Lichouri, USTHB, Algeria
Nandu C Nair, Univeristy of Trento, Italy
Nasrin Taghizadeh, University of Tehran, Iran

Organizing Committee:

Hadi Khalilia, Univeristy of Trento
Nandu C Nair, Univeristy of Trento

Table of Contents

Syllable Subword Tokens for Open Vocabulary Speech Recognition in Malayalam. 1
Kavya Manohar, A. R. Jayan and Rajeev Rajan

Semi-Supervised and Unsupervised detection of Humour in Code-Mixed Hindi-English Tweets. 8
Chakita Muttaraju, Aakansha Singh, Anusha Kabber and Mamatha H. R.

Toward a Test Set of Dislocations in Persian for Neural Machine Translation. 14
Behnoosh Namdarzadeh, Nicolas Ballier, Lichao Zhu, Guillaume Wisniewski and Jean-Baptiste Yunès

Love Me, Love Me Not: Human-Directed Sentiment Analysis in Arabic. 22
Abraham Israeli, Aviv Naaman, Razan Assi, Yotam Nahum, Kfir Bar and Shai Fine

Eco EvE : Economic Event Extraction. 31
Meriem Mkhinini, Mohamed Ali Ben El Haj Aissa, Aboubacar Sidiki Sidibe, Paul Pike and Aymen Khelifi

An empirical Comparison of Arabic Named Entity Recognition Methods: Application to the ALP Corpus. . . 37
Mohamed Lichouri

ALRT: Cutting Edge Tool for Automatic Generation of Arabic Lexical Recognition Tests. 43
Osama Hamed, Saeed Salah and Abed Alhakim Freihat