LREC 2022 Workshop
Language Resources and Evaluation Conference
20-25 June 2022

Towards Digital Language Equality Workshop
(TDLE)

PROCEEDINGS

Editors:
Itziar Aldabe, Begoña Altuna, Aritz Farwell, German Rigau
Preface

This volume documents the Proceedings of the Workshop Towards Digital Language Equality (TDLE), held on 20 June, 2022 as part of the LREC 2022 conference (International Conference on Language Resources and Evaluation).

Language Technology (LT), one of the most important applications of Artificial Intelligence, is revolutionizing many language-related tasks while engendering a rapidly growing and substantial economic impact. Although cross-language communication forms a significant part of this development, LT resources are not equally available to all languages and domains. To make use of Language Technology’s full potential, progress towards a multilingual, efficient, accurate, explainable, ethical, fair and unbiased language understanding is necessary – in short: Digital Language Equality (DLE).

The workshop at LREC brought together researchers and scholars working on policies, initiatives, projects, and research that target DLE at every administrative level. These include models and tools that monitor, measure, catalogue or visualize the evolution and dynamics of DLE through technological factors, (e.g., the available language resources, tools and technologies) and contextual factors (e.g., societal, economic, educational, industrial).

We are thankful to the authors who submitted their work to this workshop. In the end, six papers were accepted. We are also grateful to our Program Committee members and reviewers for their contributions, to Antonios Anastasopoulos who kindly accepted to be our invited speaker and gave a talk on "Measuring Cultural Representativeness and Rethinking LT4All", and to the LREC committee for accepting this workshop as part of LREC 2022.
Organizers
Itziar Aldabe (HiTZ, UPV-EHU)
Begoña Altuna (HiTZ, UPV-EHU)
Aritz Farwell (HiTZ, UPV-EHU)
Federico Gaspari (ADAPT, DCU)
Maria Giagkou (Athena RC/ILSP)
Jan Hajic (Charles University)
Stelios Piperidis (Athena RC/ILSP)
Georg Rehm (DFKI)
German Rigau (HiTZ, UPV-EHU)
Andy Way (ADAPT, DCU)

Program Committee:
Xabier Arregi (HiTZ, UPV/EHU)
Dimitra Anastasiou (Luxembourg Institute of Science and Technology)
Albina Auksoriutė (LIETUVIŲ KALBOS INSTITUTAS)
Jeremy Barnes (HiTZ, UPV/EHU)
Khalid Choukri (ELDA)
Bessie Dendrinos (ECSPM)
Itziar Gonzalez-Dios (HiTZ, UPV/EHU)
Kristine Eide (Language Council of Norway)
Ainara Estarrona (HiTZ, UPV/EHU)
Inma Hernáez (HiTZ, UPV/EHU)
Jaroslava Hlaváčová (CUNI)
Mikel Iruskieta (HiTZ, UPV/EHU)
Sabine Kirchmeier (EFNIL)
Svetla Koeva (INSTITUTE FOR BULGARIAN LANGUAGE)
Krister Linden (UH)
Teresa Lynn (DCU)
Maite Melero (BSC)
Eva Navas (HiTZ, UPV/EHU)
Delyth Prys (BU)
Kepa Sarasola (HiTZ, UPV/EHU)
Claudia Soria (ELEN)
Frieda Steurs (Instituut voor de Nederlandse Taal)
Jana Straková (CUNI)
Tamás Váradi (Nyelvtudományi Kutatóközpont)
Francois Yvon (CNRS)
Table of Contents

_Introducing the Digital Language Equality Metric: Technological Factors_
Federico Gaspari, Owen Gallagher, Georg Rehm, Maria Giagkou, Stelios Piperidis, Jane Dunne and Andy Way ................................................................. 1

_Introducing the Digital Language Equality Metric: Contextual Factors_
Annika Grützner-Zahn and Georg Rehm ................................................................. 13

_Collaborative Metadata Aggregation and Curation in Support of Digital Language Equality Monitoring_
Maria Giagkou, Stelios Piperidis, Penny Labropoulou, Miltos Deligiannis, Athanasia Kolovou and Leon Voukoutis ................................................................. 27

_Measuring HLT Research Equality of European Languages_
Gorka Artola and German Rigau .............................................................................. 36

_National Language Technology Platform for Public Administration_
Marko Tadić, Daša Farkaš, Matea Filko, Artūrs Vasiljevs, Andrejs Vasiljevs, Jānis Ziediņš, Željka Motika, Mark Fishel, Hrafn Loftsson, Jón Guðnason, Claudia Borg, Keith Cortis, Judie Attard and Donatienne Spiteri ................................................................. 46

_The Nós Project: Opening routes for the Galician language in the field of language technologies_
Iria de-Dios-Flores, Carmen Magariños, Adina Ioana Vladu, John E. Ortega, José Ramom Pichel, Marcos García, Pablo Gamallo, Elisa Fernández Rei, Alberto Bugarín-Diz, Manuel González González, Senén Barro and Xosé Luis Regueira ................................................................. 52
Conference Program

14:00–14:45  *Measuring Cultural Representativeness and Rethinking LT4All*
Antonios Anastasopoulos

14:50–15:20  *Developing an agenda and a roadmap for achieving full digital language equality in Europe by 2030*

15:20–15:40  *Introducing the Digital Language Equality Metric: Technological Factors*
Federico Gaspari, Owen Gallagher, Georg Rehm, Maria Giagkou, Stelios Piperidis, Jane Dunne and Andy Way

15:40–16:00  *Introducing the Digital Language Equality Metric: Contextual Factors*
Annika Grützner-Zahn and Georg Rehm

16:30–16:50  *Collaborative Metadata Aggregation and Curation in Support of Digital Language Equality Monitoring*
Maria Giagkou, Stelios Piperidis, Penny Labropoulou, Miltos Deligiannis, Athanasia Kolovou and Leon Voukoutis

16:50–17:10  *Measuring HLT Research Equality of European Languages*
Gorka Artola and German Rigau

17:10–17:30  *National Language Technology Platform for Public Administration*
Marko Tadić, Daša Farkaš, Matea Filip, Artūrs Vasiljevs, Andrejs Vasiljevs, Jānis Ziedinš, Željka Motika, Mark Fishel, Hrafn Loftsson, Jón Guðnason, Claudia Borg, Keith Cortis, Judie Attard and Donatienne Spiteri

17:30–17:50  *The Nós Project: Opening routes for the Galician language in the field of language technologies*
Iria de-Dios-Flores, Carmen Magariños, Adina Ioana Vladu, John E. Ortega, José Ramom Pichel, Marcos García, Pablo Gamallo, Elisa Fernández Rei, Alberto Bugarín-Diz, Manuel González González, Senén Barro and Xosé Luis Regueira