

***Third and final call for papers for the  
Second International Workshop Towards Digital Language Equality (TDLE):  
Focusing on Sustainability  
co-located with LREC-COLING 2024, Saturday 25<sup>th</sup> May 2024, Turin (Italy)***

<https://european-language-equality.eu/tdle-2024/>

***Submission deadline extended to 29<sup>th</sup> February 2024 – we welcome submissions rejected from the main LREC-COLING 2024 conference only if these are relevant and suitable***

## **1 Description and Aims of the Workshop**

The key aim of this half-day workshop co-located with LREC-COLING 2024 (<https://lrec-coling-2024.org/>), to be held in Turin (Italy) on Saturday 25<sup>th</sup> May 2024, is to discuss and promote the importance of sustainability in the design, development, creation, use, distribution and sharing of language data, resources, platforms, infrastructures, tools and technologies, with the intention of achieving Digital Language Equality (DLE). While some important work has recently addressed these crucial areas (e.g. Fort and Couillault, 2016; Hessenthaler et al., 2022; Ramesh et al., 2023; Moorkens et al., 2024), the relevant contributions seem to be as yet unsystematic and relatively isolated. The workshop intends to provide an inclusive forum to encourage in-depth debate and facilitate collaborations to promote the sustainability of resources and technologies in any (combination of) languages, in support of multilingualism and of the overarching goal of DLE.

*The sustainability of language resources and technologies is key to enabling multilingualism and digital language equality in the age of Artificial Intelligence.*

## **2 Topics of Interest**

The *Second International Workshop Towards Digital Language Equality (TDLE)* focuses on sustainability in relation to the design, development, creation, use, distribution and sharing of language data, resources, platforms, infrastructures, tools and technologies, with a view to promoting the broader goal of Digital Language Equality (DLE). The concept of DLE has been firmly established in relation to all languages of Europe (Rehm and Way, 2023), and has the potential to also benefit other languages throughout the world, to support the prosperity of the respective communities at a time of impressive – but as yet very unevenly distributed and severely imbalanced – progress in language-centric Artificial Intelligence (AI), e.g. through large language models (LLMs). The workshop places particular emphasis on multilingualism and on leveling up digital support for languages, domains and applications that have so far been underserved, and wishes to explore ways to develop policies and funding streams to work towards sustainability in connection with DLE, especially in support of regional, minority and territorial languages.

To this end, recognizing that the sustainability of Language Resources and Technologies (LRTs) is key to enabling multilingualism and DLE in the age of AI, topics of particular interest for the workshop on which we invite original contributions covering any (combination of) languages include, but are not limited to, the following:

- research on the factors affecting DLE and the sustainability of LRTs;
- best practices, case studies and validated guidelines related to the design, implementation and improvement of sustainability of written, oral/spoken, signed and/or multimodal LRTs (including LLMs), particularly in support of DLE;
- how multilingual LLM technology can support DLE;
- retrospectively assessing the sustainability of legacy LRTs, and future-proofing new LRTs in the interest of DLE;
- analyzing the costs and benefits of foregrounding sustainability for LRTs;
- the role of metadata, accompanying documentation and licenses in showing and improving the sustainability of LRTs;

- sustainability, fairness and accessibility (e.g. for users with physical or cognitive disabilities, limited computing resources and connectivity) of platforms and infrastructures hosting, distributing and sharing LRTs in the interest of DLE;
- how current data and computing access inequality is affecting DLE (in particular regarding LLMs);
- ecological sustainability and environmental fairness of developing and deploying state-of-the-art LRTs, e.g. LLMs with regard to energy consumption, global warming and climate change;
- developing data and parameter efficient methods to train or adapt language models to new languages;
- how to evaluate, measure, compare and improve the sustainability of LRTs;
- establishing benchmarks and protocols to ensure the sustainability of LRTs;
- how to avoid the potential dangers of developing and using *unfair* and *unsustainable* LRTs, e.g. for malicious, ill-intentioned or harmful purposes;
- ethical, legal, cultural and/or socio-economic implications of (ignoring) fairness and sustainability of LRTs;
- developing and implementing forward-looking policies to promote fairness and long-term sustainability of LRTs to achieve DLE;
- education and training needs and experiences in relation to promoting fairness and sustainability of LRTs and ways to raise broad awareness of DLE and related topics, e.g. among the general public, policy- and decision-makers.

Given this wide-ranging and inclusive remit, the workshop intends to bring together developers, creators, vendors, distributors, brokers, users, evaluators and researchers of written, oral/spoken, signed and/or multimodal LRTs in any (combination of) languages.

### 3 Background and First TDLE Workshop Held in 2022

The second 2024 edition of the workshop builds on the success of the first *Towards Digital Language Equality (TDLE) workshop*,<sup>1</sup> that was held at LREC 2022 in Marseille (France) on 20 June 2022, and whose accepted papers were published in a dedicated volume of proceedings, Aldabe et al. (2022).<sup>2</sup>

Following this well-received inaugural workshop held in June 2022, the second event in the series will be co-located with LREC-COLING 2024 in Turin (Italy) on Saturday 25<sup>th</sup> May 2024, and will focus specifically on the highly relevant topic of the sustainability of LRTs in connection with multilingualism and DLE.

### 4 Submissions

Up-to-date information on the workshop, including materials for authors, guidelines, templates, stylesheet and key dates can be found at the dedicated website <https://european-language-equality.eu/tdle-2024/>. To contact the organizing committee of the workshop directly, you can email [tdle2024.hitz@ehu.eu](mailto:tdle2024.hitz@ehu.eu).

Papers submitted to the workshop should be completely anonymous for double-blind peer review, written in English, and prepared using the official LREC-COLING 2024 author's kit and submission stylesheet/template available at <https://lrec-coling-2024.org/authors-kit/>. The submissions to the workshop should not exceed 8 pages, excluding references, and be saved in unprotected PDF format.

***The submission deadline for the workshop has been extended to 29<sup>th</sup> February 2024. Suitable submissions rejected from the main LREC-COLING 2024 conference programme are welcome, but they should be clearly relevant to the focus of the workshop and address explicitly its topics of interest.*** Papers should be submitted through the START submission management system available at <https://softconf.com/lrec-coling2024/tdle2024/>.

The workshop seeks original papers, i.e. it does *not* accept submissions that have been, or will be, published elsewhere. The workshop allows simultaneous submissions, and in these cases the authors should clearly indicate in the manuscript to which other conference, workshop or venue they have submitted the paper for review. Each paper submitted to the workshop will receive three double-blind peer reviews. Papers accepted for presentation will be included in the proceedings of the workshop.

<sup>1</sup> <https://european-language-equality.eu/tdle-2022/>

<sup>2</sup> [www.lrec-conf.org/proceedings/lrec2022/workshops/TDLE/2022.tdle-1.0.pdf](http://www.lrec-conf.org/proceedings/lrec2022/workshops/TDLE/2022.tdle-1.0.pdf)

In light of the LREC-COLING 2024 Map and the “Share your LRs!” initiative, when submitting their papers through the START system authors are asked to provide essential information about resources (in a broad sense, i.e. also technologies, standards, evaluation kits, etc.) that have been used for the work described in the paper or are a new result of their research. Moreover, ELRA encourages all LREC-COLING authors to share the described LRs (data, tools, services, etc.) to enable their reuse and replicability of experiments (including evaluation ones).

## 5 Key Dates

Paper submission deadline: *extended to Thursday, 29<sup>th</sup> February 2024*

Notification of acceptance: *Thursday, 14<sup>th</sup> March 2024*

Camera-ready papers due: *Friday, 29<sup>th</sup> March 2024*

Half-day workshop date: *Saturday, 25<sup>th</sup> May 2024*

## 6 Workshop Organizers

- Itziar Aldabe (HiTZ Basque Center for Language Technology - Ixa, University of the Basque Country, Spain)
- Begoña Altuna (HiTZ Basque Center for Language Technology - Ixa, University of the Basque Country, Spain)
- Aritz Farwell (HiTZ Basque Center for Language Technology - Ixa, University of the Basque Country, Spain)
- Federico Gaspari (University of Naples “Federico II”, Italy & ADAPT Centre, Dublin City University, Ireland – co-chair)
- Joss Moorkens (School of Applied Language & Intercultural Studies/ADAPT Centre, Dublin City University, Ireland – co-chair)
- Stelios Piperidis (Institute of Language and Speech Processing, Athena Research and Innovation Center in Information, Communication and Knowledge Technologies, Greece)
- Georg Rehm (Speech and Language Technology Lab, Deutsches Forschungszentrum für Künstliche Intelligenz, Germany)
- German Rigau (HiTZ Basque Center for Language Technology - Ixa, University of the Basque Country, Spain)

## 7 Program Committee

- Antonios Anastasopoulos (GMU, USA)
- Anya Belz (ADAPT, DCU, Ireland)
- Steven Bird (CDU, Australia)
- Fred Blain (Uni. Tilburg, Netherlands)
- Franco Cutugno (Uni. Naples “Federico II”, Italy)
- Bessie Dendrinou (NKUA, Greece & ECSPM, Denmark)
- Félix do Carmo (Uni. Surrey, UK)
- Annika Grützner-Zahn (DFKI, Germany)
- Ana Guerberof-Arenas (Uni. Groningen, Netherlands)
- Davyth Hicks (ELEN, Belgium)
- Monja Jannet (ADAPT, DCU, Ireland)
- John Judge (ADAPT, DCU, Ireland)
- Dorothy Kenny (SALIS/CTTS/ADAPT, DCU, Ireland)
- Sabine Kirchmeier (EFNIL, Luxembourg)
- Teresa Lynn (MBZUAI, United Arab Emirates)
- Maite Melero (BSC, Spain)
- Helena Moniz (Uni. Lisbon, Portugal & EAMT)
- Johanna Monti (UniOR, Italy)
- Rachele Raus (UniBO, Italy)

- Wessel Reijers (Uni. Paderborn, Germany)
- Celia Rico Pérez (Universidad Complutense de Madrid, Spain)
- Dimitar Shterionov (TU, Netherlands)
- Carlos S. C. Teixeira (IOTA Localisation Services & Uni. Rovira i Virgili, Spain)
- Antonio Toral (Uni. Groningen, Netherlands)
- Vincent Vandeghinste (Instituut voor de Nederlandse Taal, Netherlands & KU Leuven, Belgium)

## References

- Itziar Aldabe, Begoña Altuna, Aritz Farwell and German Rigau, editors. 2022. *Proceedings of the Workshop Towards Digital Language Equality (TDLE)*. European Language Resources Association, Marseille, France.
- Karèn Fort and Alain Couillault, 2016. “Yes, We Care! Results of the Ethics and Natural Language Processing Surveys”. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*. European Language Resources Association, Portorož, Slovenia. 1593–1600.
- Marius Hessenthaler, Emma Strubell, Dirk Hovy and Anne Lauscher, 2022. “Bridging Fairness and Environmental Sustainability in Natural Language Processing”. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, Abu Dhabi, United Arab Emirates. 7817–7836.
- András Kornai, 2013. “Digital Language Death”. *PLoS ONE*, 8(10):e77056.
- Joss Moorkens, Sheila Castilho, Federico Gaspari, Antonio Toral and Maja Popović, editors. 2024. Special Issue on “Translation Automation and Sustainability”. *Journal of Specialised Translation*. 41 (2024). <https://www.jostrans.org/issue/view/492> (last accessed 14/02/2024).
- Krithika Ramesh, Sunayana Sitaram and Monojit Choudhury, 2023. “Fairness in Language Models Beyond English: Gaps and Challenges”. *Findings of the Association for Computational Linguistics: EACL 2023*. Association for Computational Linguistics, Dubrovnik, Croatia. 2106–2119.
- Georg Rehm and Andy Way, editors. 2023. *European Language Equality: A Strategic Agenda for Digital Language Equality*. Berlin: Springer.