

TEICAI 2024

**The 1st Workshop Towards Ethical and Inclusive
Conversational AI: Language Attitudes, Linguistic Diversity,
and Language Rights**

Proceedings of the Workshop

March 22, 2024

The TEICAI organisers and presenters gratefully acknowledge the support from the EU COST Action CA19102 "Language in the Human-Machine Era".

Supported by



©2024 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 979-8-89176-074-5

Introduction

Welcome to the 1st workshop Towards Ethical and Inclusive Conversational AI: Language Attitudes, Linguistic Diversity, and Language Rights, co-located with EACL 2024!

Our work on the intersection of language ideologies, language rights, and conversational AI started in 2021 within EU COST Action "Language in the Human-Machine Era"(LITHME). Researchers working on the crossroads of these topics contributed to three editions of international LITHME conferences in 2021-2023, and more focused work on language ideologies and conversational AI has been performed during the CONVERSATIONS 2022 workshop. Finally, we decided to give it the shape of a workshop at EACL 2024 in order to reach a broader academic audience and intensify the exchange within the language technology community.

While we have seen a rapid increase and diversification of language-based technologies over the last two decades, their development is still very much driven by technology designers and technologically relevant concerns. Users' needs and their communicative affordances have played a minor role and are often only addressed when commercial interests are at risk. For example, the dialogue patterns of conversational AI tend to lack adaptability to different user groups and their sociocultural contexts. Overall, technologies are based on structural notions of language and are still not able to do justice to the complexity of human communication.

In addition, the human rights regulation of AI-based conversational technology is scarce in many areas. When it comes to language rights, striking a balance between technological advancement and ensuring that conversational technology contributes positively to linguistic diversity and inclusion is essential. Addressing this requires a holistic approach that involves dialogue and collaboration among technologists, linguists, policymakers, and communities affected by conversational AI systems.

In addition, training data used to build conversational AI of the newest generation is mostly conceptually and medially written language. Furthermore, existing methods prioritize the formal and mechanized aspects of language, neglecting the gestural, social, and emotional elements that are fundamental to communication. Signed languages, due to their inherently multimodal nature and spatial-contextual dependencies are still excluded from the conversational AI space.

While the socio-linguistic community is intensively discussing the issues mentioned above, these topics have not yet completely arrived in the Computational Linguistics and NLP community. The goal of the workshop is to bring together researchers from all communities and intensify the academic exchange in order to cultivate a multidisciplinary approach to the development of conversational AI that can better serve diverse global audiences.

The program of TEICAI 2024 includes two keynote talks, seven paper presentations and one round-table discussion.

The organizing committee would like to express its appreciation to the authors who submitted papers, the reviewers, the panelists, and the invited speakers for their invaluable contributions. We are already looking forward to the workshop's next edition!

Best wishes,

The TEICAI organizing committee

Organizing Committee

Organizing Committee Members

Sviatlana Höhn, LuxAI, Luxembourg

Nina Hosseini-Kivanani, University of Luxembourg, Luxembourg

Dimitra Anastasiou, Luxembourg Institute of Science and Technology, Luxembourg

Bettina Migge, University College Dublin, Ireland

Angela Soltan, State University of Moldova, Moldova

Doris Dippold, University of Surrey, UK

Ekaterina Kamlovskaya, Luxembourg National Data Service, Luxembourg

Fred Philippy, Zortify, Luxembourg

Program Committee

Program Committee

Dimitra Anastasiou, Luxembourg Institute of Science and Technology, Luxembourg
Matt Coler, University of Groningen, Netherlands
Jelske Dijkstra, Fryske Akademy / Mercator Research Centre, Netherlands
Doris Dippold, University of Surrey, UK
Siwen Guo, Zortify, Luxembourg
Leopold Hillah, University of Luxembourg, Luxembourg
Sviatlana Höhn, LuxAI, Luxembourg
Raquel Lázaro Gutiérrez, University of Alcalá, Spain
Daniel Leix Palumbo, University of Groningen, Netherlands
Bettina Migge, University College Dublin, Ireland
Javier Moreno-Rivero, City University of New York, USA
Aria Nourbakhsh, University of Luxembourg, Luxembourg
Fred Philippy, Zortify, Luxembourg
Angela Soltan, State University of Moldova, Moldova

Invited Speakers

Justine Cassell, INRIA Paris, France
Chakravarthi Bharathi Raja, University of Galway, Ireland

Keynote Talk: Culturally-Aware Educational Language Technologies

Justine Cassell

INRIA Paris, France

2024-03-22 09:30:00 – Room: Bastion 1

Abstract: Children seamlessly shift their ways of speaking, adopting and adapting language they hear spoken by adults around them, as well as constructing their own variants. These speaking styles play an important role as children experiment with who they want to be, and how they want to be perceived. They also allow children growing up in situations where different dialects or languages are in contact to mark their affiliation to dominant and minority ethnic, racial, and gender identities. Young people who move between marginalized and mainstream communities often report code-switching as a way to maintain affiliation with their home community, while also making their way in a world where the standard dialect is associated with various kinds of success. However, moving back and forth may lead to what Ogbu (2008) has called “oppositional culture” to describe how school systems may inadvertently set up a situation where the student feels the need to define her identity contra the expectations of the school, and for that reason to refuse the dialect that the school insists on. Unfortunately, language technologies, including and very persistently, educational technologies, may inadvertently carry negative stereotypes about ethnicity and how it is carried in language, leading to stress and lack of a sense of agency as children try to navigate a path that allows them to benefit both from the support of their community and the opportunities offered by mainstream education. My students and I have examined this issue by building “culturally-aware” educational language technologies, and specifically “virtual peers” that either speak only the child’s own dialect, or that model a code-switching strategy between what linguists call low-prestige and high-prestige dialects. Results using a variety of methodological approaches, in both one-shot and longitudinal studies, demonstrate the positive impact of technology such as these that take issues of culture, and of power, into account on children’s school performance. On the other hand, careful assessments of the children’s reactions to the technologies shows that they will need further development to improve the children’s own internalized biases against low-prestige dialect speakers.

Keynote Talk: Language Technology for Equality, Diversity and Inclusion

Bharathi Raja Chakravarthi

University of Galway, Ireland

2024-03-22 14:40:00 – Room: Bastion 1

Abstract: Equality, Diversity and Inclusion (EDI) is an important agenda across every field throughout the world. Language as a major part of communication should be inclusive and treat everyone with equality. Today’s large internet community uses language technology (LT) and has a direct impact on people across the globe. EDI is crucial to ensure everyone is valued and included, so it is necessary to build LT that serves this purpose. Recent results have shown that big data and deep learning are entrenching existing biases and that some algorithms are even naturally biased due to problems such as ‘regression to the mode’. Our focus is on creating LT that will be more inclusive of gender, racial, sexual orientation, persons with disability. Over the past few years, systems have been developed to control online content and eliminate abusive, offensive or hate speech content. However, people in power sometimes misuse this form of censorship to obstruct the democratic right of freedom of speech. Therefore, it is imperative

that research should take a positive reinforcement approach towards online content that is encouraging, positive and supportive. Until now, most studies have focused on solving this problem of negativity in the English language, though the problem is much more than just harmful content. Furthermore, it is multilingual as well.

Roundtable: The magic components of success: multidisciplinary collaboration among language rights, language ideology and conversational AI communities.

2024-03-22 16:00:00 – Room: Bastion 1

Abstract: Discussions about difficulties and key success factors in multidisciplinary projects have been ongoing since multidisciplinary collaborations exist, and the problems mentioned in those discussions frequently echo each other and include problems with communication, differences in terminology, theories and research methods, different approaches to knowledge creation and different community cultures. While these issues are still valid in the TEICAI community, we would like to look into more specific issues related to the specific communes of language ideologies, language rights and conversational AI, and structural issues related to research organisation, funding and academic culture in general. In the round table, we want to highlight personal perspectives of researchers involved in multidisciplinary collaborations on the intersection of conversational AI, language ideologies and language rights. We want to listen to perspectives of researchers in different career stages:

- working mainly on technology and collaborate with researchers in language ideologies and/or language rights;
- working mainly in language rights and/or language ideologies and collaborating with researchers working mainly on technology aspects of conversational AI.

Further, we want to zoom out and understand how these personal experiences are embedded in a bigger picture of academic research.

Following the end of the roundtable, we will be dedicated to translating these discussions into a tangible set of recommendations aimed at fostering an ethical and inclusive future for conversational AI. These guidelines will be published across a variety of channels, aiming to inspire and provide direction to both the existing and next generations of field practitioners.

Panellists

- Nina Hosseini-Kivanani (PhD student, University of Luxembourg, LU)
- Doris Dippold (Senior Lecturer, University of Surrey, UK)
- Valentina Pyatkin (Postdoctoral Researcher, Allen Institute for AI, USA)
- Justine Cassel (Professor, INRIA Paris, FR and CMU, USA)

Moderator: Sviatlana Höhn (Postdoctoral researcher, LuxAI, LU)

Table of Contents

<i>How Do Conversational Agents in Healthcare Impact on Patient Agency?</i> Kerstin Denecke	1
<i>Why academia should cut back general enthusiasm about CAs</i> Alessia Giulimondi	9
<i>Bridging the Language Gap: Integrating Language Variations into Conversational AI Agents for Enhanced User Engagement</i> Marcellus Amadeus, Jose Roberto Homeli da Silva and Joao Victor Pessoa Rocha	16
<i>Socio-cultural adapted chatbots: Harnessing Knowledge Graphs and Large Language Models for enhanced context awareness</i> Jader Camboim de Sá, Dimitra Anastasiou, Marcos Da Silveira and Cédric Pruski	21
<i>How should Conversational Agent systems respond to sexual harassment?</i> Laura De Grazia, Alex Peiró Lilja, Mireia Farrús Cabeceran and Mariona Taulé	28
<i>Non-Referential Functions of Language in Social Agents: The Case of Social Proximity</i> Sviatlana Höhn	36
<i>Making a Long Story Short in Conversation Modeling</i> Yufei Tao, Tiernan Mines and Ameeta Agrawal	42

Program

Friday, March 22, 2024

09:00 - 09:30 *Welcome and Introduction*

09:30 - 10:30 *Keynote Talk by Justine Cassel: "Culturally-Aware Educational Language Technologies"*

10:30 - 11:00 *Coffee Break*

11:00 - 12:00 *Paper Session 1*

Why academia should cut back general enthusiasm about CAs

Alessia Giulimondi

Socio-cultural adapted chatbots: Harnessing Knowledge Graphs and Large Language Models for enhanced context awareness

Jader Camboim de Sá, Dimitra Anastasiou, Marcos Da Silveira and Cédric Pruski

Bridging the Language Gap: Integrating Language Variations into Conversational AI Agents for Enhanced User Engagement

Marcellus Amadeus, Jose Roberto Homeli da Silva and Joao Victor Pessoa Rocha

Non-Referential Functions of Language in Social Agents: The Case of Social Proximity

Sviatlana Höhn

12:00 - 13:45 *Lunch Break*

13:45 - 14:40 *Paper Session 2*

How Do Conversational Agents in Healthcare Impact on Patient Agency?

Kerstin Denecke

How should Conversational Agent systems respond to sexual harassment?

Laura De Grazia, Alex Peiró Lilja, Mireia Farrús Cabeceran and Mariona Taulé

Making a Long Story Short in Conversation Modeling

Yufei Tao, Tiernan Mines and Ameeta Agrawal

Friday, March 22, 2024 (continued)

- 14:40 - 15:30 *Keynote Talk by Bharathi Raja Chakravarthi: "Language Technology for Equality, Diversity and Inclusion"*
- 15:30 - 16:00 *Coffee Break*
- 16:00 - 16:40 *Round-Table: "The magic components of success: multidisciplinary collaboration among language rights, language ideology and conversational AI communities"*
- 16:40 - 16:50 *Closing Words*