



LREC 2026

**NLP4Ecology @ LREC 2026**

**Workshop Proceedings**

**Editors**

**Francesca Grasso, Valerio Basile, Cristina Bosco,  
Muhammad Okky Ibrohim, Maria Skeppstedt, and  
Manfred Stede**

12 May 2026

NLP4Ecology 2026

The 2nd Workshop on Ecology, Environment, and Natural Language Processing

©ELRA Language Resources Association (ELRA), 2026

These proceedings are licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ISBN 978-2-493814-53-1

## Preface

We are pleased to welcome you to NLP4Ecology 2026, the 2nd edition of the International Workshop on Ecology, Environment, and Natural Language Processing. Following a successful first edition, which confirmed a rapidly growing community interested in the intersection of NLP and ecology, this year's workshop takes place on May 12th, 2026, co-located with LREC 2026 in Palma de Mallorca, Spain. The larger venue and its broader audience provide an important opportunity to further expand the visibility and impact of this initiative.

NLP4Ecology is designed to be a venue for publication and exchange between the Natural Language Processing (NLP) community and researchers and practitioners from a wide range of disciplines. It aims to explore how computational linguistics and NLP tools, methods, and applications can contribute to addressing urgent environmental issues - not only climate change, which has received the most research attention so far due to its visibility and quantifiability, but also broader ecological crises affecting humans, non-human animals, and ecosystems worldwide. Tackling these issues requires interdisciplinary collaboration, and the NLP community has an important role to play in this effort.

The workshop continues to stimulate an interdisciplinary space at the intersection of computational linguistics and environmental and ecological research. The contributions in this edition are primarily rooted in NLP and applied computational methods, while also attracting (eco-)linguistic perspectives on environmental discourse, including the analysis of narratives, constructions, and framing. In this context, the workshop aims to support the development of new research directions and collaborations, and to encourage critical reflection on the role of language technologies in actively shaping environmental knowledge and discourse.

This year's program reflects both continuity and growth. We received a total of 22 submissions, all of which underwent a double-blind review process. In the end, 14 papers were accepted, corresponding to an overall acceptance rate of 63%. We are deeply grateful to the members of our Program Committee for their careful and constructive evaluations, which were essential in shaping a diverse and high-quality program.

The workshop features a keynote lecture by **Leonie Bossert**, whose work at the intersection of environmental ethics and philosophy of technology offers an enlightening perspective on questions of responsibility, technology, and human-nature relations.

The accepted papers reflect a range of approaches to environmental language and data. Several contributions focus on sustainability communication and ESG reporting, addressing challenges such as greenwashing detection, readability, and alignment with reporting standards. Others examine environmental discourse and narratives, including climate disinformation, media coverage of extreme events, and the analysis of historical and multilingual corpora. A number of papers investigate the role of large language models in ecological contexts, with particular attention to evaluation, interpretability, and explanatory capabilities, as well as their alignment with ecological values. Additional contributions address applied NLP tasks, including corpus construction in lower-resource settings, biodiversity data processing, and retrieval-based approaches to environmental information. The contributions draw on a variety of data sources, including social media content, news corpora, sustainability and ESG reports, biodiversity databases, and historical travel narratives. This diversity of data highlights the wide range of contexts in which environmental language is produced and interpreted, and supports a more comprehensive understanding of how ecological issues are communicated across domains.

Regarding linguistic and geographical diversity, the accepted papers reflect a broad range of languages and research contexts, including English, German, Russian, and Vietnamese, as

well as multilingual evaluation settings. The contributing authors are affiliated with institutions across multiple countries, including Italy, Germany, the Netherlands, the United Kingdom, Spain, Greece, Cyprus, and the Philippines, among others.

Organizing a workshop of this kind requires the support and collaboration of many people. We would like to sincerely thank our Program Committee members again for their time and expertise, as well as all authors for their submissions and contributions. Finally, we thank all participants for being part of this second edition and for helping to build a growing research community at the intersection of NLP and ecology.

#### **NLP4Ecology 2026 Organizing Committee**

Francesca Grasso  
Valerio Basile  
Cristina Bosco  
Muhammad Okky Ibrohim  
Maria Skeppstedt  
Manfred Stede

May 2026

# Organizing Committee

## Workshop Chairs

- Francesca Grasso, University of Turin, Italy
- Valerio Basile, University of Turin, Italy
- Cristina Bosco, University of Turin, Italy
- Muhammad Okky Ibrohim, Universitas Indonesia, Indonesia, and University of Tartu, Estonia
- Maria Skeppstedt, Uppsala University, Sweden
- Manfred Stede, University of Potsdam, Germany

## Program Committee

- Luca Brigada Villa, University of Pavia, Italy
- Tyler A. Chang, University of California San Diego; Google DeepMind, USA
- Giacomo De Luca, University of Rome "Tor Vergata", Italy
- Steffen Frenzel, University of Potsdam, Germany
- Sara Gemelli, University of Bergamo, Italy
- Anna Glazkova, University of Tyumen, Russia
- Amanda Starling Gould, Duke University, USA
- Frederik Hudi, Nara Institute of Science and Technology, Japan
- Veronika Laippala, University of Turku, Finland
- Stefano Locci, University of Turin, Italy
- Stella Markantonatou, Athena Research Center, Greece
- Luca Moroni, Sapienza University of Rome, Italy
- Robert Östling, Stockholm University, Sweden
- Lucia Passaro, University of Pisa, Italy
- Melanie Siegel, Darmstadt University of Applied Sciences, Germany
- Maximos Skandalis, CNRS & University of Montpellier, France
- Ivan Spada, Fondazione Bruno Kessler (FBK), Italy
- Masashi Takeshita, Nagoya University, Japan
- Fabio Massimo Zanzotto, University of Rome "Tor Vergata", Italy

## Invited Speaker

- Leonie N. Bossert, University of Vienna, Austria



# Table of Contents

|                                                                                                                                                                                                                                                                                     |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Retrieving Climate Change Disinformation by Narrative</i><br>Max Upravitelev, Veronika Solopova, Charlott Jakob, Premtim Sahitaj, Sebastian Möller<br>and Vera Schmitt .....                                                                                                     | 1   |
| <i>Unsupervised GRI-TCFD Alignment with LLM-Assisted Validation for Climate Disclosure and<br/>Greenwashing Risk Analysis</i><br>Seyed Alireza Mousavian Anaraki, Danilo Croce, Roberta Costa, Luigi Tiburzi, Armando<br>Calabrese and Roberto Basili .....                         | 15  |
| <i>Towards Empowering Consumers through Sentence-level Readability Scoring in German ESG<br/>Reports</i><br>Benjamin Josef Schüßler and Jakob Prange .....                                                                                                                          | 26  |
| <i>Disambiguating Geographic Names in Biodiversity Occurrence Data: A Retrieval-Augmented<br/>Generation Approach</i><br>Yanni Jose C. Ella, Monica Ashley R. Laviste, John Michael L. Lastimoso, Wilfred John E.<br>Santiañez, Riza Batista-Navarro and Roselyn Santos Gabud ..... | 42  |
| <i>Sentiment and Stance in EFL Responses to AI-Generated Environmental Content</i><br>Andry Sophocleous .....                                                                                                                                                                       | 53  |
| <i>What Stories Do Language Models Tell About Nature? A Multi Layer Evaluation Framework for<br/>Ecological Alignment</i><br>Jorge Vallego, Eleanor Tiernan, Mah Rukh, Mariana Roccia and Sabina Fiebig Lord ...                                                                    | 60  |
| <i>Ecological Discourse Modeling in a Low-Resource Setting: A Longitudinal Vietnamese Climate<br/>Corpus with Comparative Topic Modeling</i><br>Huyen Phuong Nguyen .....                                                                                                           | 69  |
| <i>Greench-v1: distilling SLMs on Greenwashing Detection</i><br>Federico Raspanti, Alessandro Pietro Bardelli Bardelli, Simona Scala, İrem Demirtaş, Mar-<br>ilena Di Bari and Michele Filannino .....                                                                              | 79  |
| <i>Analyzing Environmental Discourse through Construction-Based Pattern Extraction</i><br>Elisa Chierchiello, Eliana Di Palma, Ludovica Pannitto and Cristina Bosco .....                                                                                                           | 87  |
| <i>Mapping the Historical Ecology of the Cyclades: A Diachronic Natural Language Processing<br/>Analysis of Travel Narratives (1700–1920)</i><br>Aikaterini Christopoulou, Vassilis Detsis and Basilis Gatos .....                                                                  | 105 |
| <i>Retrieving Floods without Floodlights: Topic Models as Binary Classifiers for Extreme Climate<br/>Events in German News</i><br>Brielen Madureira, Mariana Madruga de Brito and Andreas Niekler .....                                                                             | 115 |
| <i>Why Is This Green? LLM-Based Explanations of Implicit Green Practices in Social Media</i><br>Anna Glazkova, Olga Zakharova and Daria Lebedeva .....                                                                                                                              | 135 |
| <i>Introducing a Green Leaderboard for Sustainable Risk Prediction in Streaming NLP Shared<br/>Tasks.</i><br>Alba María Mármol-Romero, Adrián Moreno Muñoz and Arturo Montejo-Raez .....                                                                                            | 144 |

*Not Everything Is Greenwashing: Limitations of Automatic Analysis of Sustainability Reports, and a Proposal*

Maria Pilar Uribe Silva, Rik van Noord and Malvina Nissim ..... 154

# Workshop Program

Tuesday, May 12, 2026

|             |                                                                                                                                                                                                                                                                        |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 14:10–14:55 | <b>Invited Talk: Leonie N. Bossert</b>                                                                                                                                                                                                                                 |
| 14:55–15:55 | <b>Session 1</b>                                                                                                                                                                                                                                                       |
| 14:55–15:10 | <i>Retrieving Climate Change Disinformation by Narrative</i><br>Max Upravitelev, Veronika Solopova, Charlott Jakob, Premtim Sahitaj, Sebastian Möller and Vera Schmitt                                                                                                 |
| 15:10–15:25 | <i>Unsupervised GRI-TCFD Alignment with LLM-Assisted Validation for Climate Disclosure and Greenwashing Risk Analysis</i><br>Seyed Alireza Mousavian Anaraki, Danilo Croce, Roberta Costa, Luigi Tiburzi, Armando Calabrese and Roberto Basili                         |
| 15:25–15:40 | <i>Towards Empowering Consumers through Sentence-level Readability Scoring in German ESG Reports</i><br>Benjamin Josef Schübler and Jakob Prange                                                                                                                       |
| 15:40–15:55 | <i>Disambiguating Geographic Names in Biodiversity Occurrence Data: A Retrieval-Augmented Generation Approach</i><br>Yanni Jose C. Ella, Monica Ashley R. Laviste, John Michael L. Lastimoso, Wilfred John E. Santiañez, Riza Batista-Navarro and Roselyn Santos Gabud |
| 16:00–16:50 | <b>Poster Session</b>                                                                                                                                                                                                                                                  |
| 16:00–16:50 | <i>Sentiment and Stance in EFL Responses to AI-Generated Environmental Content</i><br>Andry Sophocleous                                                                                                                                                                |
| 16:00–16:50 | <i>What Stories Do Language Models Tell About Nature? A Multi Layer Evaluation Framework for Ecological Alignment</i><br>Jorge Vallego, Eleanor Tiernan, Mah Rukh, Mariana Roccia and Sabina Fiebig Lord                                                               |
| 16:00–16:50 | <i>Ecological Discourse Modeling in a Low-Resource Setting: A Longitudinal Vietnamese Climate Corpus with Comparative Topic Modeling</i><br>Huyen Phuong Nguyen                                                                                                        |
| 16:00–16:50 | <i>Greench-v1: distilling SLMs on Greenwashing Detection</i><br>Federico Raspanti, Alessandro Pietro Bardelli Bardelli, Simona Scala, İrem Demirtaş, Marilena Di Bari and Michele Filannino                                                                            |

- 16:00–16:50 *Analyzing Environmental Discourse through Construction-Based Pattern Extraction*  
Elisa Chierchiello, Eliana Di Palma, Ludovica Pannitto and Cristina Bosco
- 16:00–16:50 *Mapping the Historical Ecology of the Cyclades: A Diachronic Natural Language Processing Analysis of Travel Narratives (1700–1920)*  
Aikaterini Christopoulou, Vassilis Detsis and Basilis Gatos
- 17:00–17:55      Session 2**
- 17:00–17:15 *Retrieving Floods without Floodlights: Topic Models as Binary Classifiers for Extreme Climate Events in German News*  
Brielen Madureira, Mariana Madruga de Brito and Andreas Niekler
- 17:15–17:25 *Why Is This Green? LLM-Based Explanations of Implicit Green Practices in Social Media*  
Anna Glazkova, Olga Zakharova and Daria Lebedeva
- 17:25–17:40 *Introducing a Green Leaderboard for Sustainable Risk Prediction in Streaming NLP Shared Tasks.*  
Alba María Mármol-Romero, Adrián Moreno Muñoz and Arturo Montejo-Raez
- 17:40–17:55 *Not Everything Is Greenwashing: Limitations of Automatic Analysis of Sustainability Reports, and a Proposal*  
Maria Pilar Uribe Silva, Rik van Noord and Malvina Nissim