The 31st International Conference on Computational Linguistics

1st International Workshop on Nakba Narratives as Language Resources (Nakba-NLP 2025)

Proceedings of the workshop
Edited by:
Mustafa Jarrar, Nizar Habash, Mo El-Haj, Amal Haddad Haddad, Zeina Jallad, Camille Mansour, Diana Allan, Paul Rayson, Tymaa Hammouda, Sanad Malaysha

January 20, 2025
https://sina.birzeit.edu/nakba-nlp/

©2025 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 979-8-89176-216-9

Preface

Welcome to the 1st International Workshop on Nakba Narratives as Language Resources (Nakba-NLP 2025), co-located with the 31st International Conference on Computational Linguistics (COLING 2025). Nakba-NLP 2025 was held virtually on January 20, 2025.

The narratives of the (ongoing) Palestinian Nakba hold immense historical, cultural, literary, and academic value. Preserving this content and empowering it with AI tools is vital to ensure its accessibility and usability for present and future generations. Nakba narratives and testimonies exist in diverse formats such as manuscripts, books, audio recordings, novels, and films, making their conversion into machine-understandable formats a significant challenge. Establishing accessible archives and well-annotated collections is crucial for researchers and historians to validate and share meaningful information.

This workshop aims to explore how artificial intelligence, natural language processing, and corpus linguistics can assist in understanding, disseminating, and preserving Nakba narratives and testimonies. The goal is to create accessible, comprehensive, and well-annotated collections that empower researchers and historians to validate and share critical insights derived from these data. The workshop targets datasets and narratives in Arabic, English, and other languages.

This year's workshop featured a total of 18 submissions, 14 submissions were accepted. The accepted papers are authored by 43 scholars representing diverse perspectives from 11 different countries, including the Egypt, France, Germany, Lebanon, Malta, Netherlands, Palestine, Spain, Turkey, USA, and UK. This highlights the international reach and collaborative spirit of the workshop. Each submission to the workshop was evaluated by at least two reviewers who were members of the Program Committee.

The Nakba-NLP 2025 workshop featured a keynote address and a panel discussion. Ilan Pappé, a renowned historian and political scientist from the University of Exeter, specializing in the Nakba, delivered the keynote address titled "*The Words Laundrette: Unmasking Bias and Propaganda in the Discourses on the Ongoing Nakba*". In his address, Pappé delved into how language is used, shaped, and contested within political and cultural narratives surrounding the Nakba—providing critical insights directly applicable to the role of NLP in analyzing such complex discourses.

The panel discussion, "*Digital Archives and Cultural Heritage in the LLM Era*," brought together leading experts, including Dawn Knight, Antonio Moreno Sandoval, Muhammad Abdul-Mageed, and Mustafa Jarrar. Their thought-provoking dialogue explored the rapidly evolving interplay between digital archives, cultural heritage, Large Language Models and Multimodal Large Models, highlighting the field's pressing challenges and transformative opportunities.

We would like to thank everyone who submitted a paper to the workshop and to all the members of the Program Committee.

Workshop organizers

Workshop Page: `https://sina.birzeit.edu/nakba-nlp`

Organizing Committee

Nakba-NLP 2025:

Organising Committee & Workshop Chairs:

Mustafa Jarrar (Birzeit University, Palestine)
Nizar Habash (New York University Abu Dhabi, UAE)
Mo El-Haj (Lancaster University, UK)
Amal Haddad Haddad (Universidad de Granada, Spain)
Zeina Jallad (Harvard Law School, USA)
Camille Mansour (Institute for Palestine Studies, Lebanon)
Diana Allan (McGill University, Canada)
Paul Rayson (Lancaster University, UK)
Tymaa Hammouda (Birzeit University, Palestine)
Sanad Malaysha (Birzeit University, Palestine)

Programme Committee:

Abdelkader El Mahdaouy (Mohamed VI polytechnic University)
Abdellah El Mekki (Mohammed VI Polytechnic University)
Abdelrahim Qaddoumi (NYU)
Abdulrahman Abdulsalam (University of Utah)
Abed Alhakim Freihat (University of Trento)
Adnan Yahya (Birzeit University)
Ala Alazzeh (Birzeit University)
Ali AlKhathlan, (King Abdulaziz University)
Almoataz B. Al-Said, (Cairo University)
Amr Keleg (The University of Edinburgh)
Areej Jaber (Technical University Khadouri)
Ashraf Elnagar (University of Sharjah)
Ayah Soufan (Strathclyde University)
Azzeddine Mazroui (University Mohammed First)
Badr AlKhamissi (EPFL)
Baker Abdalhaq (Annajah National University)
Basem Ezbidi (Birzeit university)
Bassam Haddad (University of Petra)
Bayan AbuShawar (Al Ain University)
Dana Abdulrahim (University of Bahrain)
Dima Taji (Charles University)
ElMoatez Billah Nagoudi (The University of British Columbia)
Eyad Elyan (Robert Gordon University)
Fadhl Eryani (University of Tübingen)
Fadi Zaraket (American University of Beirut)
Faisal Awartani (Insights for Research Polling and Training)
Fatemah Husain (Kuwait University)
Fatima Haouari (Qatar University)
Fethi Bougares (LIUM- Le Mans Université)
Fouzi Harrag (Ferhat Abbas University)
Ghassan Mourad (libanese university)
Go Inoue (Mohamed bin Zayed University of Artificial Intelligence)
Habiba Dahmani (Mohamed Boudiaf University)
Haithem Afli (ADAPT Centre, Munster Technological University)

Hamada Nayel (Benha University)
Ibrahim Abu Farha (University of Sheffield)
Imed Zitouni (Google)
Imene Bensalem (Constantine 2 University)
Injy Hamed (Institute for Natural Language Processing, University of Stuttgart)
Kamel Gaanoun (National Institute of Statistics and Applied Economics)
Kamel Smaili (LORIA)
Kaoukab Chebaro (Columbia University)
Karim Bouzoubaa (Mohammed V University in Rabat)
Khaled Shaalan (The British University in Dubai)
Khalid Choukri (ELRA/ELDA) Khalil Mrini (Bytedance)
Khloud Al Jallad (SySSR)
Labib Arafeh (AlQuds University)
Lama Nachman (Intel Labs)
Lamia Hadrich-Belguith (ANLP Research Group, MIRACL Lab, FSEGS, Sfax University)
Majdi Sawalha (The University of Jordan)
Manar Alkhatib (British University in Dubai)
Maram Hasanain (Qatar Computing Research Institute)
Mo El-Haj (Lancaster University)
Mohamed Lichouri (Centre de Recherche Scientifique et Technique pour le Développement de la Langue Arabe (CRSTDLA)
Mohamed Yahya, (NLP Researcher)
Mohammad Abuoudeh (Al Hussein Bin Talal University)
Mohammed Attia (Google Inc.)
Mohammed Khalilia (Birzeit University)
Mohammed Salah Al-Radhi (Budapest University of Technology and Economics)
Mona Baker (University of Oslo)
Moustafa Al-Hajj (Lebanese University)
Muhammad Abdul-Mageed (The University of British Columbia)
Munir Fakher Eldin (Birzeit University)
Munther Dahleh (MIT)
Nada Ghneim (Damascus University)
Omar Shehabi (Yale Law School)
Omar Tesdell (Birzeit University)
Omar Trigui (University of Sousse in Tunisia)
Owen Rambow (Stony Brook University)
Radi Jarrar (Birzeit University)
Radwan Tahboub (Palestine Polytechnic University)
Raia Abu Ahmad (DFKI)
Reem Suwaileh (Qatar University)
Saad Ezzini (Lancaster University)
Sabri Boughorbel (Qatar Computing Research Institute)
Sahar Ghannay (CNRS, LISN)
Salima Harrat (ENS Bouzaréah, Algiers)
salima mdhaffar (LIA - University of Avignon)
Samhaa R. El-Beltagy (Newgiza University/Optomatica)
Sari Hanafi (American University of Beirut)
Seif Mechti (ISSEPS)
Serin Atiani (Princess Sumaya University for Technology)
Shady Elbassuoni (American University of Beirut)
Sultan Alrowili (University of Delaware)

Susan Akram (Boston University)
Tamer Elsayed (Qatar University)
Thaher Gharabeh (University of Granada)
Violetta Cavalli-Sforza (Al Akhawayn University)
Wajdi Zaghouani (Northwestern University Qatar)
Walid Magdy (The University of Edinburgh)
Wassim El-Hajj (American University of Beirut)
Watheq Mansour (The University of Queensland)
Wissam Antoun (Inria)
Yahya Mohamed Elhadj (Arab Center for Research and Policy) Studies

Table of Contents

Deciphering Implicatures: On NLP and Oral Testimonies
Zainab Sabra 1

A cultural shift in Western perceptions of Palestine
Terry Regier and Muhammad Ali Khalidi 9

Cognitive Geographies of Catastrophe Narratives: Georeferenced Interview Transcriptions as Language Resource for Models of Forced Displacement
Annie K. Lamar, Rick Castle, Carissa Chappell, Emmanouela Schoinoplokaki, Allene M. Seet, Amit Shilo and Chloe Nahas 18

Sentiment Analysis of Nakba Oral Histories: A Critical Study of Large Language Models
Huthaifa I. Ashqar 30

The Nakba Lexicon: Building a Comprehensive Dataset from Palestinian Literature
Izza AbuHaija, Salim Al Mandhari, Mo El-Haj, Jonas Sibony and Paul Rayson 37

Arabic Topic Classification Corpus of the Nakba Short Stories
Osama Hamed and Nadeem Zaidkilani 48

Exploring Author Style in Nakba Short Stories: A Comparative Study of Transformer-Based Models
Osama Hamed and Nadeem Zaidkilani 56

Detecting Inconsistencies in Narrative Elements of Cross Lingual Nakba Texts
Nada Hamarsheh, Zahia Elabour, Aya Murra and Adnan Yahya 63

Multilingual Propaganda Detection: Exploring Transformer-Based Models mBERT, XLM-RoBERTa, and mT5
Mohamed Ibrahim Ragab, Ensaf Hussein Mohamed and Walaa Medhat 75

Collective Memory and Narrative Cohesion: A Computational Study of Palestinian Refugee Oral Histories in Lebanon
Ghadir A. Awad, Tamara N. Rayan, Lavinia Dunagan and David Gamba 83

The Missing Cause: An Analysis of Causal Attributions in Reporting on Palestine
Paulina Garcia Corral, Hannah Bechara, Krishnamoorthy Manohara and Slava Jankin 103

Bias Detection in Media: Traditional Models vs. Transformers in Analyzing Social Media Coverage of the Israeli-Gaza Conflict
Marryam Yahya Mohammed, Esraa Ismail Mohamed, Mariam Nabil Esmat, Yomna Ashraf Nagib, Nada Ahmed Radwan, Ziad Mohamed Elshaer and Ensaf Hussein Mohamed 114

NakbaTR: A Turkish NER Dataset for Nakba Narratives
Esma Fatıma Bilgin Tasdemir and Şaziye Betül Özateş 122

Integrating Argumentation Features for Enhanced Propaganda Detection in Arabic Narratives on the Israeli War on Gaza
Sara Nabhani, Claudia Borg, Khalid Al Khatib and Kurt Micallef 127

Time Activity & Authors

09:00 - 09:20 **Opening Session**: Welcome by Workshop Chairs Mustafa Jarrar and Camille Mansour

09:20 - 09:40 **Talk 1**: Integrating Argumentation Features for Enhanced Propaganda Detection in Arabic Narratives on the Israeli War on Gaza.
Authors: Sara Nabhani, Claudia Borg, Kurt Micallef, Khalid Al-Khatib.
Paper ID: 18

09:40 - 10:00 **Talk 2**: Multilingual Propaganda Detection: Exploring Transformer-Based Models mBERT, XLM-RoBERTa, and mT5.
Authors: Mohamed Ibrahim Ragab, Ensaf Hussein Mohamed and Walaa Medhat.
Paper ID: 12

10:00 - 10:20 **Talk 3:** Bias Detection in Media: Traditional Models vs. Transformers in Analyzing Social Media Coverage of the Israeli-Gaza Conflict.
Authors: Marryam Yahya Mohammed, Esraa Ismail Mohamed, Mariam Nabil Esmat, Yomna Ashraf Nagib, Nada Ahmed Radwan, Ziad Mohamed Elshaer and Ensaf Hussein Mohamed.
Paper ID: 16

10:20 - 10:40 **Talk 4**: The Missing Cause: An Analysis of Causal Attributions in Reporting on Palestine.
Authors: Paulina Garcia Corral, Hannah Bechara, Krishnamoorthy Manohara and Slava Jankin
Paper ID: 15

10:40 - 11:00 **Talk 5**: Deciphering Implicatures: On NLP and Oral Testimonies.
Authors: Zainab Sabra
Paper ID: 4

11:00 - 11:15 Coffee Break

11:15 - 12:30 **Panel Discussion**: Digital Archives and Cultural Heritage in the LLMs Era. **Panel Chair:** Mo El-Haj
Muhammad Abdul-Mageed, The University of British Columbia, Canada.
Antonio Moreno Sandoval, Autonomous University Madrid, Spain.
Dawn Knight, Cardiff University, UK.
Mustafa Jarrar, Birzeit University, Palestine.

12:30 - 13:00 Lunch Break

13:00 - 13:15 **Talk 6:** Sentiment Analysis of Nakba Oral Histories: A Critical Study of Large Language Models.
Authors: Huthaifa I. Ashqar
Paper ID: 7

13:15 - 13:30 **Talk 7:** Arabic Topic Classification Corpus of the Nakba Short Stories.
Authors: Osama Hamed and Nadeem Zaidkilani.
Paper ID: 9

13:30 - 13:45 **Talk 8**: Exploring Author Style in Nakba Short Stories: A Comparative Study of Transformer-Based Models.
Authors: Osama Hamed and Nadeem Zaidkilani.
Paper ID: 10

13:45 - 14:00 Coffee Break

14:00 - 14:20 **Talk 9:** NakbaTR: A Turkish NER Dataset for Nakba Narratives.
Authors: Esma Fatıma Bilgin Tasdemir and Şaziye Betül Özateş.
Paper ID: 17

14:20 - 14:40 **Talk 10:** The Nakba Lexicon: Building a Comprehensive Dataset from Palestinian Literature.
Authors: Izza AbuHaija, Salim Al Mandhari, Mo El-Haj, Jonas Sibony and Paul Rayson.
Paper ID: 8

14:40 - 15:00 **Talk 11:** Cognitive Geographies of Catastrophe Narratives: Georeferenced Interview Transcriptions as Language Resource for Models of Forced Displacement.
Authors: Annie K. Lamar, Rick Castle, Carissa Chappell, Emmanouela Schoinoplokaki, Allene M. Seet, Amit Shilo and Chloe Nahas.
Paper ID: 6

15:00 - 15:20 **Talk 12:** Collective Memory and Narrative Cohesion: A Computational Study of Palestinian Refugee Oral Histories in Lebanon.
Authors: Ghadir A. Awad, Tamara N. Rayan, Lavinia Dunagan and David Gamba.
Paper ID: 13

15:20 - 15:40 **Talk 13:** A cultural shift in Western perceptions of Palestine.
Authors: Terry Regier and Muhammad Ali Khalidi.
Paper ID: 5

15:40 - 16:00 **Talk 14:** Detecting Inconsistencies in Narrative Elements of Cross Lingual Nakba Texts.
Authors: Nada Hamarsheh, Zahia Elabour, Aya Murra and Adnan Yahya.
Paper ID: 11

16:00 - 17:00 **Keynote**: The Words Laundrette: Unmasking Bias and Propaganda in the Discourses on the Ongoing Nakba By Ilan Pappe.

17:00 - 17:15 Closing Remarks and Wrap-Up