

COLING 2025

**The 31st International Conference on Computational
Linguistics**

**Proceedings of the New Horizons in Computational
Linguistics for Religious Texts (Coling-Rel)**

Workshop chairs:

Majdi Sawalha, Sane Yagi, Faisal Alshargi,
Abdallah T. AlShdaifat, Ashraf Elnagar, Bayan Abu Shawar,
Norhan Abbas

19 January 2025

©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 979-8-89176-203-9

Message from the Organizers

This workshop is focused on the intersection of language technology and religious texts. It encourages dialogue on cutting-edge Natural Language Processing applications within this domain. Traditionally, scholars have focused on analyzing faith-defining canons, authoritative interpretations, and extracting insights from sermons, liturgies, prayers, and religious poetry.

The emergence of Large Language Models (LLMs) has revolutionized many NLP disciplines. Leveraging these powerful models for novel research avenues in religious texts promises to significantly advance this field.

This workshop explores the potential of NLP to unlock profound new understandings of religious traditions and to chart the future of this exciting research area. It brings together researchers from computational linguistics, digital humanities, and related fields to engage in insightful discussions and to explore innovative approaches.

Organizing Committee

Majdi Sawalha, The University of Jordan, Jordan and Al-Ain University, UAE

Sane Yagi, University of Sharjah, UAE

Faisal Alshargi, Amazon Robotics, USA

Abdallah T. AlShdaifat, Mohamed Bin Zayed University for Humanities, UAE

Ashraf Elnagar, University of Sharjah, UAE

Bayan Abu Shawar, Al-Ain University, UAE

Norhan Abbas, The University of Leeds, UK

Program Committee

Majdi Sawalha, The University of Jordan, Jordan and Al-Ain University, UAE (Chair)

Sane Yagi, University of Sharjah, UAE (Chair)

Abdallah T. AlShdaifat, Mohamed Bin Zayed University for Humanities, UAE

Ahmad Ruby, Uppsala University, Sweden

Akram M. Zeki, International Islamic University Malaysia, Malaysia

Ashraf Elnagar, University of Sharjah, UAE

Azzedin Mazroui, University Mohammed First, Morocco

Bassam Hammo, Princess Sumaya University for Technology, Jordan

Bayan Abu Shawar, Al-Ain University, UAE

Faisal Alshargi, Amazon Robotics, USA

Khubaib Alam, Al-Ain University, UAE

Mohammad Khair, International Computing Institute for Quran and Islamic Sciences, USA

Nizar Habash, NYU Abu Dhabi, UAE

Norhan Abbas, The University of Leeds, UK

Table of Contents

<i>Comparative Analysis of Religious Texts: NLP Approaches to the Bible, Quran, and Bhagavad Gita</i> Mahit Nandan A D, Ishan Godbole, Pranav M Kapparad and Shrutilipi Bhattacharjee	1
<i>Messages from the Quran and the Bible in Mandarin through Factor Analysis with Syntactic and Semantic Tags</i> Kuanlin Liu	11
<i>Semantic Analysis of Jurisprudential Zoroastrian Texts in Pahlavi: A Word Embedding Approach for an Extremely Under-Resourced, Extinct Language</i> Rashin Rahnamoun and Ramin Rahnamoun	23
<i>Multi-stage Training of Bilingual Islamic LLM for Neural Passage Retrieval</i> Vera Pavlova	42
<i>Automated Translation of Islamic Literature Using Large Language Models: Al-Shamela Library Application</i> Mohammad Mohammad Khair and Majdi Sawalha	53
<i>Automated Authentication of Quranic Verses Using BERT (Bidirectional Encoder Representations from Transformers) based Language Models</i> Khubaib Amjad Alam, Maryam Khalid, Syed Ahmed Ali, Haroon Mahmood, Qaisar Shafi, Muhammad Haroon and Zulqarnain Haider	59
<i>MASAQ Parser: A Fine-grained MorphoSyntactic Analyzer for the Quran</i> Majdi Sawalha, Faisal Alshargi, Sane Yagi, Abdallah T. AlShdaifat and Bassam Hammo	67
<i>Leveraging AI to Bridge Classical Arabic and Modern Standard Arabic for Text Simplification</i> Shatha Altammami	76
<i>Word boundaries and the morphology-syntax trade-off</i> Pablo Mosteiro and Damián Blasi	86

Workshop Program

Sunday, January 19, 2025

09:30–09:45 Opening remarks by the Organizing Committee

Chair: Sane Yagi

09:45–10:30 Plenary Speaker: Professor Ashraf Alnagar, College of Computing and Informatics, University of Sharjah

Chair: Sane Yagi

11:00–12:30 Session 1: Papers

Chair: Majdi Sawalha

11:00–11:30 *Comparative Analysis of Religious Texts: NLP Approaches to the Bible, Quran, and Bhagavad Gita*

Mahit Nandan A D, Ishan Godbole, Pranav M Kapparad and Shrutilipi Bhattacharjee

11:30–12:00 *Messages from the Quran and the Bible in Mandarin through Factor Analysis with Syntactic and Semantic Tags*

Kuanlin Liu

12:00–12:30 *Semantic Analysis of Jurisprudential Zoroastrian Texts in Pahlavi: A Word Embedding Approach for an Extremely Under-Resourced, Extinct Language*

Rashin Rahnamoun and Ramin Rahnamoun

12:30–14:00 Lunch Break

14:00–15:30 Session 2: Papers

Chair: Bayan Abu Shawar

14:00–14:30 *Multi-stage Training of Bilingual Islamic LLM for Neural Passage Retrieval*

Vera Pavlova

14:30–15:00 *Automated Translation of Islamic Literature Using Large Language Models: Al-Shamela Library Application*

Mohammad Mohammad Khair and Majdi Sawalha

15:00–15:30 *Automated Authentication of Quranic Verses Using BERT (Bidirectional Encoder Representations from Transformers) based Language Models*

Khubaib Amjad Alam, Maryam Khalid, Syed Ahmed Ali, Haroon Mahmood, Qaisar Shafi, Muhammad Haroon and Zulqarnain Haider

Sunday, January 19, 2025 (continued)

16:00–17:30 Session 3: Papers

Chair: Abdallah T. AlShdaifat

16:00–16:30 *MASAQ Parser: A Fine-grained MorphoSyntactic Analyzer for the Quran*
Majdi Sawalha, Faisal Alshargi, Sane Yagi, Abdallah T. AlShdaifat and Bassam Hammo

16:30–17:00 *Leveraging AI to Bridge Classical Arabic and Modern Standard Arabic for Text Simplification*
Shatha Altammami

17:00–17:30 *Word boundaries and the morphology-syntax trade-off*
Pablo Mosteiro and Damián Blasi

17:30–18:15 Session 4: Posters

Chair: Norhan Abbas

18:15–close Panel Discussion

Chair: Organizers