

EthicalLLMs 2025

**Proceedings of the
The First Workshop on Ethical Concerns in Training,
Evaluating and Deploying Large Language Models**

associated with
**The 15th International Conference on
Recent Advances in Natural Language Processing
RANLP'2025**

Edited by Damith Premasiri, Tharindu Ranasinghe and Hansi Hettiarachchi

13 September, 2025
Varna, Bulgaria

The First Workshop on Ethical Concerns in Training, Evaluating
and Deploying Large Language Models
Associated with the International Conference
Recent Advances in Natural Language Processing
RANLP'2025

PROCEEDINGS

Varna, Bulgaria
13 September 2025

Online ISBN 978-954-452-103-5

Designed by INCOMA Ltd.
Shoumen, BULGARIA

Preface

We are delighted to present the proceedings of the First Workshop on Ethical Concerns in Training, Evaluating and Deploying Large Language Models (EthicalLLMs 2025), held in conjunction with RANLP 2025 in Varna, Bulgaria.

Large language models (LLMs) have made tremendous strides in recent years, enabling new applications in summarization, question answering, generation, reasoning, and more. Yet alongside their technical capabilities, there are growing ethical challenges that demand careful scrutiny. Issues such as bias, transparency, accountability, cultural sensitivity, and fairness are all deeply enmeshed in the development and deployment of LLMs. The EthicalLLMs workshop was conceived to create a space for researchers, practitioners, and ethicists to explore these issues, exchange methodologies, and forge directions toward more responsible LLM design.

We are pleased to include in these proceedings a variety of contributions that reflect the rich and urgent research in this space. The papers address theoretical, empirical, and applied work, and together they reflect diverse perspectives on ensuring that LLMs are built and deployed with due regard for ethical, social, and human-centred concerns.

This workshop would not have been possible without the dedication and support of many people. We are deeply grateful to all the authors who submitted their work, helping to shape the conversation around responsible LLMs. We extend warm thanks to the programme committee members for their thoughtful reviews, constructive feedback, and commitment to maintaining high standards of scholarship. We especially thank Prof Paul Rayson for accepting our invitation to deliver the keynote address, enriching our discussions with his insights into ethics and language technologies. Finally, we thank the RANLP organizing team and supporting institutions for enabling this workshop to take place.

We are hopeful that the ideas presented here will stimulate further research, dialogue, and action toward more ethical, equitable, and trustworthy language models.

Organizing Committee

Damith Premasiri, Lancaster University, UK

Tharindu Ranasinghe, Lancaster University, UK

Hansi Hettiarachchi, Lancaster University, UK

Program Committee

Deshan Sumanathilaka, Swansea University, UK

Daisy Lal, Lancaster University, UK

Salmane Chafik, Mohammed VI Polytechnic University, Morocco.

Isaac Omolayo

Table of Contents

<i>TextBandit: Evaluating Probabilistic Reasoning in LLMs Through Language-Only Decision Tasks</i>	
Arjun Damerla, Jimin Lim, Yanxi Jiang, Nam Nguyen Hoai Le and Nikil Selladurai	1
<i>CoVeGAT: A Hybrid LLM & Graph-Attention Pipeline for Accurate Citation-Aligned Claim Verification</i>	
Max Bader, Akshatha Arunkumar, Ohan Ahmad, Maruf Hassen, Charles Duong and Kevin Zhu .	9
<i>TVS Sidekick: Challenges and Practical Insights from Deploying Large Language Models in the Enterprise</i>	
Paula Reyero Lobo, Kevin Johnson, Bill Buchanan, Matthew Shardlow, Ashley Williams and Sam Attwood	17

Conference Program

9.00–9.10 **Welcome to the workshop**

9.10–10.00 **Keynote speech by Prof Paul Rayson**

10.00–10.30 *TextBandit: Evaluating Probabilistic Reasoning in LLMs Through Language-Only Decision Tasks*

Arjun Damerla, Jimin Lim, Yanxi Jiang, Nam Nguyen Hoai Le and Nikil Selladurai

10.30–11.00 **Coffee Break**

11.00–11.30 *CoVeGAT: A Hybrid LLM & Graph-Attention Pipeline for Accurate Citation-Aligned Claim Verification*

Max Bader, Akshatha Arunkumar, Ohan Ahmad, Maruf Hassen, Charles Duong and Kevin Zhu

11.30–12.00 *TVS Sidekick: Challenges and Practical Insights from Deploying Large Language Models in the Enterprise*

Paula Reyero Lobo, Kevin Johnson, Bill Buchanan, Matthew Shardlow, Ashley Williams and Sam Attwood

End of the workshop

