

AIWolfDial 2025

**The 3rd International Workshop of
AI Werewolf and Dialog System (AIWolfDial 2025)**
held in conjunction with
The 18th International Natural Language Generation Conference
(INLG 2025)

Proceedings of the Workshop

October 30, 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 979-8-89176-325-8

Message from the Workshop Chair

We are excited to present the Proceedings of the 3rd International AI Werewolf and Dialog System Workshop (AIWolfDial 2025) held in conjunction with the International 18th International Natural Language Generation Conference (INLG 2025). Our workshop takes place in October 30 in Hanoi, Vietnam.

Recent achievements of large language models, e.g. ChatGPT, are gathering greater attentions. However, it is not fully explored that such a huge language model could sufficiently able to handle coherent responses, longer contexts, common grounds, and logics.

The AIWolfDial 2025 international contest is held as a part of this AIWolfDial 2025 workshop. This is an international open contest for automatic players of the conversation game "Mafia", requires players not just to communicate but to infer, persuade, deceive other players via coherent logical conversations, while having the role-playing non-task-oriented chats as well. We believe that this contest reveals current issues in the recent huge language models, showing directions of next breakthrough in the NLP area.

From the viewpoint of Game AI area, players must hide information, in contrast to perfect information games such as chess or Reversi. Each player acquires secret information from other players' conversations and behavior and acts by hiding information to accomplish their objectives. Players are required persuasion for earning confidence, and speculation for detecting fabrications.

We called for papers which include the above related topics but not limited to them. After a peer review process, 3 long papers were accepted to the workshop and are included in these proceedings, with our overview paper. The accepted papers not just describe the ways they create the AIWolf player agent, but also show general ideas how to develop a LLM- and generative AI-based systems to make coherent, semantic, and characterised dialogues. Based on the subjective and win-rate evaluations, we describe overall evaluations in our overview paper and present them in the workshop.

We would like to thank the INLG conference organizing team; the workshop would not be possible without their dedication and hard work. We would also like to express our gratitude to our organizers and the program committee members for their reviewing contributions, and the local organizing team for operating the contest and subjective evaluations.

AIWolfDial 2025 workshop chair,
Yoshinobu Kano

Organizing Committee

Yoshinobu Kano, Shizuoka University

Claus Aranha, University of Tsukuba

Daisuke Katagami, Tokyo Polytechnic University

Kei Harada, The University of Electro-Communications

Michimasa Inaba, The University of Electro-Communications

Takeshi Ito, The University of Electro-Communications

Hirotsuka Osawa, Keio University

Takashi Otsuki, Yamagata University

Fujio Toriumi, The University of Tokyo

Table of Contents

<i>AIWolfDial 2025: Summary of Natural Language Division of 7th International AIWolf Contest</i>	
Yoshinobu Kano, Neo Watanabe, Yuya Harada, Yuto Sahashi, Claus Aranha, Daisuke Katagami, Kei Harada, Michimasa Inaba, Takeshi Ito, Hirotaka Osawa, Takashi Otsuki and Fujio Toriumi	1
<i>Towards a Strategic Werewolf AI Based on Expert Strategies in Five-Player Werewolf</i>	
Takuma Okada and Takeshi Ito	14
<i>Intent-driven AIWolf Agents with Hierarchical BDI Model and Personality</i>	
Yuya Harada and Yoshinobu Kano	22
<i>Influence of Utterance Impressions on Decision-Making in LLM-to-LLM Discussions</i>	
Neo Watanabe and Yoshinobu Kano	37

Conference Program

October 30th, 2025

Peer reviewed papers

AIWolfDial 2025: Summary of Natural Language Division of 7th International AI-Wolf Contest

Yoshinobu Kano, Neo Watanabe, Yuya Harada, Yuto Sahashi, Claus Aranha, Daisuke Katagami, Kei Harada, Michimasa Inaba, Takeshi Ito, Hirotaka Osawa, Takashi Otsuki and Fujio Toriumi

Towards a Strategic Werewolf AI Based on Expert Strategies in Five-Player Werewolf

Takuma Okada and Takeshi Ito

Intent-driven AIWolf Agents with Hierarchical BDI Model and Personality

Yuya Harada and Yoshinobu Kano

Influence of Utterance Impressions on Decision-Making in LLM-to-LLM Discussions

Neo Watanabe and Yoshinobu Kano

Non-archival system descriptions

CamelliaDragons

Reon Ohashi, Momoka Kato, Yugo Kato, Koki Sato, Joji Suzuki, Shinma Tsuboi, and Kazuya Tsubokura

CanisLupus

Yu Sugawara

Character-Lab

Kun Kerdthaisong, Pasin Buakhaw Peerapat, Pitikorn Khlaisamniang and Supasate Vorathammathorn

mille

Katsuki Ohto

sunamelli

Satoko Natsuori, Koya Kamada, Ryo Kamiyama, and Hiroki Nakanishi

