

EMNLP 2023

**The Second Workshop on Pattern-based Approaches to NLP
in the Age of Deep Learning
(Pan-DL)**

Proceedings of the Workshop

December 6, 2023

©2023 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 979-8-89176-046-2

Message from the Organizers

Welcome to the second edition of the Workshop on Pattern-based Approaches to NLP in the Age of Deep Learning (Pan-DL)! Our workshop is being organized in a hybrid format on December 6, 2023, in conjunction with the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP).

In the past year, the natural language processing (NLP) field (and the world at large!) has been hit by the large language model (LLM) "tsunami." This happened for the right reasons: LLMs perform extremely well in a multitude of NLP tasks, often with minimal training and, perhaps for the first time, have made NLP technology extremely approachable to non-expert users. However, LLMs are not perfect: they are not really explainable, they are not pliable, i.e., they cannot be easily modified to correct any errors observed, and they are not efficient due to the overhead of decoding. In contrast, rule-based methods are more transparent to subject matter experts; they are amenable to having a human in the loop through intervention, manipulation and incorporation of domain knowledge; and further the resulting systems tend to be lightweight and fast. This workshop focuses on all aspects of rule-based approaches, including their application, representation, and interpretability, as well as their strengths and weaknesses relative to state-of-the-art machine learning approaches.

Considering the large number of potential directions in this neuro-symbolic space, we emphasized inclusivity in our workshop. We received 19 submissions and accepted 10 for oral presentation. This resulted in an overall acceptance rate of 52%. Our workshop also includes 6 presentations of papers that were accepted in Findings of EMNLP.

In addition to the oral presentations of the accepted papers, our workshop includes a keynote talk by Yunyao Li, who has made many important contributions to the field of symbolic approaches for natural language processing. Further, the workshop contains a panel that will discuss the merits and limitations of rules in the new LLM era. The panelists will be academics with expertise in both neural- and rule-based methods, industry experts that employ these methods for commercial products, and subject matter experts that have used rule-based methods for domain-specific applications.

We thank Yunyao Li and the panelists for their important contribution to our workshop!

Finally, we are thankful to the members of the program committee for their insightful reviews! We are confident that all submissions have benefited from their expert feedback. Their contribution was a key factor for accepting a diverse and high-quality list of papers, which we hope will make the first edition of the Pan-DL workshop a success, and will motivate many future editions.

Pan-DL 2023 Organizers
December 6, 2023

Organizing Committee

Laura Chiticariu, IBM, USA
Gus Hahn-Powell, University of Arizona, USA
Dayne Freitag, SRI International, USA
Ellen Riloff, University of Arizona, USA
Clayton T. Morrison, University of Arizona, USA
Rebecca Sharp, Lex Machina, USA
Marco Valenzuela-Escárcega, Lex Machina, USA
Mihai Surdeanu, University of Arizona, USA
Enrique Noriega-Atala, University of Arizona, USA

Program Committee

Keith Alcock, University of Arizona, USA
Dillon Mee, University of Arizona, USA
Simon Mille, Dublin City University, Ireland
Mithun Paul, University of Arizona, USA
Haris Riaz, University of Arizona, USA
Arie Sutiono, University of Arizona, USA
Zheng Tang, Samsung Research America, USA
Robert Vacareanu, University of Arizona, USA
Hoang Van, University of Colorado Boulder, USA
Vikas Yadav, Samsung Research America, USA

Table of Contents

<i>Nearest Neighbor Search over Vectorized Lexico-Syntactic Patterns for Relation Extraction from Financial Documents</i> Pawan Rajpoot and Ankur Parikh	1
<i>LEAF: Linguistically Enhanced Event Temporal Relation Framework</i> Stanley Lim, Da Yin and Nanyun Peng	6
<i>A Graph-Guided Reasoning Approach for Open-ended Commonsense Question Answering</i> Zhen Han, Yue Feng and Mingming Sun	20
<i>Generating Irish Text with a Flexible Plug-and-Play Architecture</i> Simon Mille, Elaine Uí Dhonnchadha, Lauren Cassidy, Brian Davis, Stamatia Dasiopoulou and Anya Belz	25
<i>Symbolic Planning and Code Generation for Grounded Dialogue</i> Justin Chiu, Wenting Zhao, Derek Chen, Saujas Vaduguru, Alexander Rush and Daniel Fried. . .	43
<i>Towards Zero-Shot Frame Semantic Parsing with Task Agnostic Ontologies and Simple Labels</i> Danilo Neves Ribeiro, Jack Goetz, Omid Abdar, Mike Ross, Annie Dong, Kenneth Forbus and Ahmed Mohamed	54
<i>Co-evolving data-driven and NLU-driven Synthesizers for Generating Code in Domain Growth and Data Scarcity</i> Jiasheng Gu, Zifan Nan, Zhiyuan Peng, Xipeng Shen and Dongkuan Xu	64
<i>Complementary Roles of Inference and Language Models in QA</i> Liang Cheng, Mohammad Javad Hosseini and Mark Steedman	75
<i>Controlled Data Augmentation for Training Task-Oriented Dialog Systems with Low Resource Data</i> Sebastian Steindl, Ulrich Schäfer and Bernd Ludwig	92
<i>A Hybrid of Rule-based and Transformer-based Approaches for Relation Extraction in Biodiversity Literature</i> Roselyn Gabud, Portia Lapitan, Vladimir Mariano, Eduardo Mendoza, Nelson Pampolina, Maria Art Antonette Clariño and Riza Batista-Navarro	103

Conference Program

Wednesday, December 6, 2023

9:00–9:15 *Introduction*

9:15–10:30 *Panel*

10:30–10:45 *Coffee Break*

10:45–12:00 *Keynote*
Yunyao Li, Director of Machine Learning, Adobe Experience Platform

12:00–13:00 *Lunch*

13:00–14:15 **Session One**

13:00–13:15 *Nearest Neighbor Search over Vectorized Lexico-Syntactic Patterns for Relation Extraction from Financial Documents*
Pawan Rajpoot and Ankur Parikh

13:15–13:30 *LEAF: Linguistically Enhanced Event Temporal Relation Framework*
Stanley Lim, Da Yin and Nanyun Peng

13:30–13:45 *A Graph-Guided Reasoning Approach for Open-ended Commonsense Question Answering*
Zhen Han, Yue Feng and Mingming Sun

13:45–14:00 *Generating Irish Text with a Flexible Plug-and-Play Architecture*
Simon Mille, Elaine Uí Dhonnchadha, Lauren Cassidy, Brian Davis, Stamatia Dasiopoulou and Anya Belz

14:00–14:15 *Findings Talk - A Lightweight Method to Generate Unanswerable Questions in English*
Vagrant Gautam

14:15–14:25 *Short Break*

Wednesday, December 6, 2023 (continued)

14:25–15:30 Session 2

14:25–14:40 *Symbolic Planning and Code Generation for Grounded Dialogue*
Justin Chiu, Wenting Zhao, Derek Chen, Saujas Vaduguru, Alexander Rush and Daniel Fried

14:40–14:55 *Towards Zero-Shot Frame Semantic Parsing with Task Agnostic Ontologies and Simple Labels*
Danilo Neves Ribeiro, Jack Goetz, Omid Abdar, Mike Ross, Annie Dong, Kenneth Forbus and Ahmed Mohamed

14:55–15:10 *Co-evolving data-driven and NLU-driven Synthesizers for Generating Code in Domain Growth and Data Scarcity*
Jiasheng Gu, Zifan Nan, Zhiyuan Peng, Xipeng Shen and Dongkuan Xu

15:10–15:25 *Findings Talk - What Makes Chain-of-Thought Prompting Effective? A Counterfactual Study*
Aman Maadan

15:30–15:45 Coffee Break

15:45–17:00 Session Three

15:45–16:00 *Complementary Roles of Inference and Language Models in QA*
Liang Cheng, Mohammad Javad Hosseini and Mark Steedman

16:00–16:15 *Controlled Data Augmentation for Training Task-Oriented Dialog Systems with Low Resource Data*
Sebastian Steindl, Ulrich Schäfer and Bernd Ludwig

16:15–16:30 *A Hybrid of Rule-based and Transformer-based Approaches for Relation Extraction in Biodiversity Literature*
Roselyn Gabud, Portia Lapitan, Vladimir Mariano, Eduardo Mendoza, Nelson Pampolina, Maria Art Antonette Clariño and Riza Batista-Navarro

16:30–16:45 *Findings Talk - Solving the Right Problem is Key for Translational NLP*
Bernal Jimenez Gutierrez

16:45–17:00 *Findings Talk - TR-Rules: Rule-based Model for Link Forecasting on Temporal Knowledge Graph Considering Temporal Redundancy*
Ningyuan Li

17:00–17:10 Short Break

Wednesday, December 6, 2023 (continued)

17:10–17:40 Session Four

17:10–17:25 *Findings Talk - BanLemma: A Word Formation Dependent Rule and Dictionary Based Bangla Lemmatizer*
Sadia Afrin

17:25–17:40 *Findings Talk - Evaluating Emotion Arcs Across Languages: Bridging the Global Divide in Emotion Research*
Daniela Teodorescu

17:40–17:45 *Best Paper Award*

17:45–18:00 *Closing*

