

NAACL-HLT 2021

**Deep Learning Inside Out (DeeLIO):
The 2nd Workshop on Knowledge Extraction and Integration
for Deep Learning Architectures**

Proceedings of the Workshop

June 10 2021

©2021 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 978-1-954085-30-5

Introduction

Welcome to the Second Workshop on Knowledge Extraction and Integration for Deep Learning Architectures (DeeLIO)! Following the first successful edition of the workshop at EMNLP 2020, DeeLIO 2021 continues to bring together the knowledge interpretation, extraction and integration lines of research in deep learning, and to cover the area in between. Now in its second year, DeeLIO has already been established as a forum for the exchange of ideas on these topics, fostering collaboration within these research fields.

This volume includes the 14 papers presented at the workshop. DeeLIO 2021 was co-located with the Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2021) and was held on June 10, 2021 as an online workshop.

We received an exceptional batch of high-quality research papers, and decided to finally accept 14 out of 22 submissions (acceptance rate 63.6%), which were presented at the workshop. It is again great to see that the accepted papers cover both thematic aspects of DeeLIO: the extraction of linguistic knowledge from deep neural models as well as the integration of knowledge from external resources, and all this for different languages and applications. All papers were presented as posters during several online poster sessions with live interactions and Q&A sessions.

We take this opportunity to thank the DeeLIO program committee for their help and thorough reviews. We also thank the authors who presented their work at DeeLIO, and the workshop participants for the valuable feedback and discussions. Encouraged by the great research presented at the workshop and all the positive feedback received, we hope to continue with the DeeLIO organization in the years to come. Finally, we are deeply honored to have three excellent talks from our invited speakers Sebastian Riedel, Vered Shwartz, and Lena Voita.

The DeeLIO workshop organizers,
Eneko Agirre, Marianna Apidianaki, and Ivan Vulić

Organizers:

Eneko Agirre (University of the Basque Country)
Marianna Apidianaki (University of Helsinki)
Ivan Vulić (University of Cambridge)

Invited Speakers:

Sebastian Riedel (Facebook AI Research and University College London)
Vered Shwartz (Allen Institute for AI and University of Washington)
Lena Voita (University of Edinburgh and University of Amsterdam)

Program Committee:

Mohit Bansal (UNC Chapel Hill)
Davide Buscaldi (Université Paris 13)
Ilias Chalkidis (University of Copenhagen)
Anne Cocos (Ask Iggy)
Valeria de Paiva (Topos Institute)
Luis Espinosa-Anke (University of Cardiff)
Francis Ferraro (University of Maryland Baltimore County)
Goran Glavaš (University of Mannheim)
Eduard Hovy (Carnegie Mellon Institute)
Aishwarya Kamath (New York University)
Giannis Karamanolakis (Columbia University)
Anne Lauscher (University of Mannheim)
Carolin Lawrence (NEC Labs)
Qianchu Liu (University of Cambridge)
Olga Majewska (University of Cambridge)
Prodromos Malakasiotis (Athens University of Economics and Business)
Stephen Mayhew (Duolingo)
Sebastian Padó (University of Stuttgart)
Simone Paolo Ponzetto (University of Mannheim)
Roi Reichart (Technion, IIT)
Steven Schockaert (University of Cardiff)
Sabine Schulte im Walde (University of Stuttgart)
Aitor Soroa (University of the Basque Country)
Tim Van de Cruys (KU Leuven)
John Wieting (Carnegie Mellon University)
Deyi Xiong (Tianjin University)
Dian Yu (Tencent AI Lab)

Table of Contents

<i>Transformer visualization via dictionary learning: contextualized embedding as a linear superposition of transformer factors</i>	
Zeyu Yun, Yubei Chen, Bruno Olshausen and Yann LeCun	1
<i>Reconstructing Implicit Knowledge with Language Models</i>	
Maria Becker, Siting Liang and Anette Frank	11
<i>Investigating the Effect of Background Knowledge on Natural Questions</i>	
Vidhisha Balachandran, Bhuwan Dhingra, Haitian Sun, Michael Collins and William Cohen ...	25
<i>Augmenting Topic Aware Knowledge-Grounded Conversations with Dynamic Built Knowledge Graphs</i>	
Junjie Wu and Hao Zhou	31
<i>What Makes My Model Perplexed? A Linguistic Investigation on Neural Language Models Perplexity</i>	
Alessio Miaschi, Dominique Brunato, Felice Dell’Orletta and Giulia Venturi	40
<i>How Do BERT Embeddings Organize Linguistic Knowledge?</i>	
Giovanni Puccetti, Alessio Miaschi and Felice Dell’Orletta	48
<i>ERNIE-NLI: Analyzing the Impact of Domain-Specific External Knowledge on Enhanced Representations for NLI</i>	
Lisa Bauer, Lingjia Deng and Mohit Bansal	58
<i>Enhancing Multiple-Choice Question Answering with Causal Knowledge</i>	
Dhairya Dalal, Mihael Arcan and Paul Buitelaar	70
<i>Low Anisotropy Sense Retrofitting (LAsER) : Towards Isotropic and Sense Enriched Representations</i>	
Geetanjali Bihani and Julia Rayz	81
<i>KW-ATTN: Knowledge Infused Attention for Accurate and Interpretable Text Classification</i>	
Hyeju Jang, Seojin Bang, Wen Xiao, Giuseppe Carenini, Raymond Ng and Young ji Lee	96
<i>Multi-input Recurrent Independent Mechanisms for leveraging knowledge sources: Case studies on sentiment analysis and health text mining</i>	
Parsa Bagherzadeh and Sabine Bergler	108
<i>What BERTs and GPTs know about your brand? Probing contextual language models for affect associations</i>	
Vivek Srivastava, Stephen Pilli, Savita Bhat, Niranjana Pedanekar and Shirish Karande	119
<i>Attention vs non-attention for a Shapley-based explanation method</i>	
Tom Kersten, Hugh Mee Wong, Jaap Jumelet and Dieuwke Hupkes	129
<i>Predicting Numerals in Natural Language Text Using a Language Model Considering the Quantitative Aspects of Numerals</i>	
Taku Sakamoto and Akiko Aizawa	140

