Ercong Nie

nie@cis.lmu.de @NielKlug LinkedIn Scholar
https://www.cis.lmu.de/~nie Oettingenstr. 67, 80538 Munich, Germany

Status: July 2024

Research Interest

- **Multilingual NLP:** multilinguality of LLMs [5], cross-lingual transfer [4], [9], [13], historical language processing [11].
- **NLP Efficient Methods:** prompt-based learning [8], [10], [12], low-resource learning [3], parameter-efficient fine-tuning [6], [7].
- **Human-Centric NLP:** NLP inspired by human language processing [1], [14], computational neurolinguistics [2].

Education

2022 – now **Ph.D. University of Munich (LMU)** Natural Language Processing
Supervised by Dr. Helmut Schmid and Prof. Hinrich Schütze at Center for Information and Language Processing (CIS).

2020 – 2022 **M.Sc. University of Munich (LMU)** Computational Linguistics and Informatics
Thesis title: *Zero-Shot Learning on Low-Resource Languages by Cross-Lingual Retrieval*.

2015 – 2019 **B.A. Shanghai Jiao Tong University (SJTU)** German Linguistics and Finance
Thesis title: *A corpus-based study of metaphors in the football language*.

2018 **Exchange, University of Heidelberg (RKU)** German Linguistics
Study courses: *Language typology, Psycolinguistics, Second Language Acquisition, etc.*

Experience

2024 – now **Principle Investigator, *Software Campus Doctoral Research Project Funded by BMBF***
Leading a doctoral research project funded by German Federal Ministry of Education and Research (***BMBF***) on the topic of retrieval-augmented generation (RAG) with LLMs in cooperation with the industrial partner ***Bosch***.

2022 – now **Doctoral Researcher, *Munich Center for Machine Learning (MCML)***
Conducted advanced NLP research within the framework of MCML.

2024 **Visiting Researcher, *University of Copenhagen***
Research stay at Pioneer Centre for Artificial Intelligence & CopeNLU lab led by Prof. Isabelle Augenstein.

2024 **Guest PhD, *Technical University of Dresden***
Guest PhD student at the Chair of Scalable Software Architectures for Data Analytics led by Prof. Michael Färber at ScaDS.AI / TU Dresden.

2021 – 2024 **Teaching Assistant, *University of Munich (LMU)***
Leading tutorials and exercises of Bachelor and Master courses covering a wide range of NLP-related topics, such as statistical methods, finite state technologies, deep learning foundations, human-centric NLP, etc.

2021 – 2022 **Working Student, *Infineon Technologies AG***
Worked at the *Supply Chain Innovation Team*, developed a document retrieval system based on sentence transformers and knowledge graph.

2021 **Research Assistant, *University of Munich (LMU)***
Assisted in the DFG project *ReMLAV: Relational Machine Learning for Argument Validation*, worked on the construction of a claim detection dataset.

Selected Publications

* Equal contribution, † Corresponding author.

ECML-PKDD 2024 — **A Unified Data Augmentation Framework for Low-Resource Multi-Domain Dialogue Generation** [3]
Yongkang Liu*, Ercong Nie*, Shi Feng, Zheng Hua, Zifeng Ding, and Daling Wang, Yifei Zhang, Hinrich Schütze

ACL 2024 — **GNNavi: Navigating the Information Flow in Large Language Models by Graph Neural Network** [6]
Shuzhou Yuan, Ercong Nie, Michael Färber, Helmut Schmid, and Hinrich Schütze

LREC-COLING 2024 — **Decoding Probing: Revealing Internal Linguistic Structures in Neural Language Models using Minimal Pairs** [2]
Linyang He, Peilin Chen, Ercong Nie, Yuanning Li, and Jonathan R. Brennan

EACL 2024 — **ToPro: Token-Level Prompt Decomposition for Cross-Lingual Sequence Labeling Tasks** [4]
Bolei Ma*, Ercong Nie*, Shuzhou Yuan, Helmut Schmid, Färber Michael, and Frauke Kreuter, Hinrich Schütze

EMNLP 2023 — **Unleashing the Multilingual Encoder Potential: Boosting Zero-Shot Performance via Probability Calibration** [12]
Ercong Nie, Helmut Schmid, and Hinrich Schütze

ACL 2023 — **Cross-Lingual Retrieval Augmented Prompt for Low-Resource Languages** [13]
Ercong Nie*, Sheng Liang*, Helmut Schmid, and Hinrich Schütze

Skills

Natural Languages — Chinese Mandarin (Native), English (Fluent), German (Fluent)
Programming Languages — Python (Proficient), Bash (Good), C++ (Basic), R (Basic)
Deep Learning — PyTorch, HuggingFace, Numpy, Scikit-learn, Keras,…
Tools — LaTeX, HTML, Linux, Jupyter Notebook, Git, VSCode, PyCharm, Slurm,…
Misc. — academic research and teaching, student supervision, project management.

Honors and Grants

2024 — **BMBF Software Campus Doctoral Research Grant (115,000 EUR)**. German Federal Ministry of Education and Research (BMBF), Germany.

DDSA Research Visit Grant (15,000 DKK). Danish Data Science Academy (DDSA), Denmark.

2022 — **LMU-CSC Scholarship for Doctoral Study**. University of Munich (LMU) and China Scholarship Council (CSC), Germany & China.

2019 — **Outstanding Graduate**. Shanghai Jiao Tong University (SJTU), China.

2016 – 2018 — **Annual Academic Excellence Scholarship**. Shanghai Jiao Tong University (SJTU), China.

2016, 2018 — **Merit Student Award**, Shanghai Jiao Tong University (SJTU), China.

Teaching

Tutoring

SS 2024, SS 2023 — **Human-centric Natural Language Processing Seminar** (Master)

WS 2023 – 24 — **Statistical Methods in the Language Processing** (Bachelor)

WS 2023, WS 2022 — **Foundations of Advanced Natural Language Processing** (Master)

WS 2021 – 22 — **Finite State Technologies** (Master)

Student Supervision

SS 2024

- **Master Thesis (with Bosch), *University of Munich (LMU)*** NLP
  Enhancing Retrieval Augmented Generation (RAG) for Domain-Specific Content.
- **Master Thesis, *Technical University of Munich (TUM)*** Computer Science
  Multilingual Knowledge Incorporation for Large Language Models.
- **Course Project, *University of Munich (LMU)*** Applied Deep Learning
  KDD CUP 24: Meta Comprehensive Retrieval-Augmented Generation Benchmark.

WS 2023 – 24

- **Master Thesis, *University of Munich (LMU)*** Computational Linguistics
  Enhancing Reasoning and Safety: Integrating Classical Rule-Based AI with Large Language Models.
- **Course Project, *University of Munich (LMU)*** Applied Deep Learning
  SemEval'24 Task 1: Measuring the Semantic Textual Relatedness.
- **Course Project, *University of Munich (LMU)*** Applied Deep Learning
  SemEval'24 Task 8: Machine-Generated Text Detection.

SS 2023

- **Master Thesis, *University of Munich (LMU)*** Computational Linguistics
  Prompt-based finetuning of multilingual models for zero-shot cross-lingual transfer.
- **Master Thesis, *University of Munich (LMU)*** Computational Linguistics
  Extend the cross-lingual retrieval-augmented prompting method to new tasks and settings.
- **Course Project, *University of Munich (LMU)*** Applied Deep Learning
  BabyLM Challenge: Sample-efficient pretraining on a developmentally plausible corpus.

Publication List

1. H. Chen, J. Büssing, D. Rügamer, and **E. Nie**†, "Team mgtd4adl at semeval-2024 task 8: Leveraging (sentence) transformer models with contrastive learning for identifying machine-generated text," in *Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024)*, Mexico City, Mexico: Association for Computational Linguistics, Jun. 2024, pp. 1722–1729.

2. L. He, P. Chen, **E. Nie**, Y. Li, and J. R. Brennan, "Decoding probing: Revealing internal linguistic structures in neural language models using minimal pairs," in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation*, Torino, Italy: ELRA Language Resources Association and International Committee on Computational Linguistics, May 2024. URL: https://aclanthology.org/2024.lrec-main.402.pdf.

3. Y. Liu*, **E. Nie***, S. Feng, *et al.*, "A unified data augmentation framework for low-resource multi-domain dialogue generation," in *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, Vilnus, Lithunia: Springer, Sep. 2024. URL: https://arxiv.org/abs/2406.09881.

4. B. Ma*, **E. Nie***, S. Yuan, *et al.*, "Topro: Token-level prompt decomposition for cross-lingual sequence labeling tasks," in *Proceedings of the 18th Conference of the European Chapter of the Association for*

Computational Linguistics, Malta, Malta: Association for Computational Linguistics, Mar. 2024. URL: https://aclanthology.org/2024.eacl-long.164v2.pdf.

5 **E. Nie**, S. Yuan, B. Ma, *et al.*, "Decomposed prompting: Unveiling multilingual linguistic structure knowledge in english-centric large language models," *preprint*, Mar. 2024. URL: https://arxiv.org/pdf/2402.18397.pdf.

6 S. Yuan, **E. Nie**, M. Färber, H. Schmid, and H. Schütze, "Gnnavi: Navigating the information flow in large language models by graph neural network," in *Findings of the Association for Computational Linguistics: ACL 2024*, Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024. URL: https://arxiv.org/pdf/2402.11709.pdf.

7 S. Yuan*, **E. Nie***, B. Ma, and M. Färber, "Why lift so heavy? slimming large language models by cutting off the layers," *preprint*, Feb. 2024. URL: https://arxiv.org/pdf/2402.11700.pdf.

8 X. Li, **E. Nie**, and S. Liang, "Crosslingual retrieval augmented in-context learning for bangla," in *Proceedings of the First Workshop on Bangla Language Processing*, Singapore, Singapore: Association for Computational Linguistics, Dec. 2023. URL: https://aclanthology.org/2023.banglalp-1.15.pdf.

9 X. Li, **E. Nie**, and S. Liang, "From classification to generation: Insights into crosslingual retrieval augmented icl," in *NeurIPS 2023 Workshop on Instruction Tuning and Instruction Following*, New Orleans, Louisiana, United States, Dec. 2023. URL: https://openreview.net/pdf?id=KLPLCXo4aD.

10 B. Ma, **E. Nie***, H. Schmid, and H. Schuetze, "Is prompt-based finetuning always better than vanilla finetuning? insights from cross-lingual language understanding," in *Proceedings of the 19th Conference on Natural Language Processing (KONVENS 2023)*, M. Georges, A. Herygers, A. Friedrich, and B. Roth, Eds., Ingolstadt, Germany: Association for Computational Lingustics, Sep. 2023, pp. 1–16. URL: https://aclanthology.org/2023.konvens-main.1.

11 **E. Nie**, H. Schmid, and H. Schütze, "Cross-lingual constituency parsing for middle high german: A delexicalized approach," in *Proceedings of the Ancient Language Processing Workshop*, Varna, Bulgaria: INCOMA Ltd., Shoumen, Bulgaria, Sep. 2023. URL: https://aclanthology.org/2023.alp-1.8/.

12 **E. Nie**, H. Schmid, and H. Schütze, "Unleashing the multilingual encoder potential: Boosting zero-shot performance via probability calibration," in *Findings of the Association for Computational Linguistics: EMNLP 2023*, Singapore, Singapore: Association for Computational Linguistics, Dec. 2023. URL: https://aclanthology.org/2023.findings-emnlp.1056/.

13 **E. Nie***, S. Liang*, H. Schmid, and H. Schütze, "Cross-lingual retrieval augmented prompt for low-resource languages," in *Findings of the Association for Computational Linguistics: ACL 2023*, Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 8320–8340. DOI: 10.18653/v1/2023.findings-acl.528.

14 Z. Zhang*, H. Yang*, B. Ma*, D. Rügamer, and **E. Nie**†, "Baby's cothought: Leveraging large language models for enhanced reasoning in compact models," in *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, Singapore, Singapore: Association for Computational Linguistics, Dec. 2023. URL: https://aclanthology.org/2023.conll-babylm.13/.

15 I. Ziegler, B. Ma, **E. Nie**, *et al.*, "What cleaves? is proteasomal cleavage prediction reaching a ceiling?" In *NeurIPS 2022 Workshop on Learning Meaningful Representations of Life*, online, Dec. 2022. URL: https://openreview.net/forum?id=bUyk2atqXqt.