

Dennis Ulmer

dennisulmer.eu | [Kaleidophon](#) | bit.ly/3D5uwil | Scholar: Dennis Ulmer

Experience

Independent Identically Distributed Consulting (IID Consulting)

Amsterdam, Netherlands

FREELANCE CONSULTANT

04.2025–today

- Independent consultant helping clients research, develop and adopt specialized and tailored AI solutions, with. See iidconsulting.eu for details.
- Developed and shipped a fine-tuned language model for a specialized, challenging domain for an early-stage startup (details under NDA).

Institute for Logic, Language and Computation (ILLC)

Amsterdam, Netherlands

POSTDOCTORAL RESEARCHER

11.2024–today

- Postdoctoral Researcher at the ILLC and the University of Amsterdam under Prof. Ivan Titov. Successful supervision of three MSc student theses and publishing of multiple papers.
- Published “Anthropomorphic Uncertainty: What Verbalized Uncertainty in Language Models is Missing” in the Transactions of the Association for Computational Linguistics journal.
- Published pre-prints “Probing for Knowledge Attribution in Large Language Models” and “Calibration Is Not Enough: Evaluating Confidence Estimation Under Language Variations”.

Parameter Lab

Tübingen, Germany

RESEARCH INTERN

10.2023–02.2024

- Researched on improved calibration for large language models for question-answering under Dr. Seong Joon Oh.
- Produced open-source [open-source repository](#) and [models](#).
- Published “Calibrating Large Language Models Using Their Generations Only” and “TRAP: Targeted Random Adversarial Prompt Honeypot for Black-Box Identification” at ACL 2024 in Bangkok, Thailand.

Amazon Web Services

Seattle, United States of America

APPLIED SCIENTIST INTERN

06.2023–09.2023

- Conducted research on large language models for task-oriented dialogue under Elman Mansimov.
- Work with parameter-efficient finetuning and training acceleration techniques for large language models on AWS compute infrastructure.
- Published “Bootstrapping LLM-based Task-Oriented Dialogue Agents via Self-Talk” at ACL 2024 in Bangkok, Thailand.

Higher Technical Institute (Instituto Superior Técnico)

Lisbon, Portugal

VISITING RESEARCHER

02.2023–05.2023

- Research on Conformal Prediction for Natural Language Generation in a text generation and machine translation context, supervised by Dr. André F.T. Martins.
- Published “Non-Exchangeable Conformal Language Generation with Nearest Neighbors” at EACL 2024 and “Non-exchangeable conformal risk control” at ICLR 2024.

Pacmed B.V.

Amsterdam, Netherlands

RESEARCH INTERN

06.2020–10.2020

- Investigated uncertainty estimation techniques for neural networks on medical tabular data for safer deployment of a ML system in practice.
- Published “Trust Issues: Uncertainty Estimation Does Not Enable Reliable OOD Detection On Medical Tabular Data” at the NeurIPS 2020 Machine Learning for Health Workshop (ML4H).
- Published “Know Your Limits: Uncertainty Estimation with ReLU Classifiers Fails at Reliable OOD Detection” at Uncertainty in Artificial Intelligence (UAI 2021).
- Published all code as well as a benchmark for OOD detection in an [open-source repository](#).

Institute for Logic, Language and Computation (ILLC)

Amsterdam, Netherlands

RESEARCH ASSISTANT

12.2019–05.2020

- Conducted research in the context of the “Language in Interaction” project, bridging the gap between brain data and interpretability in Deep Natural Language Processing models in order to induce language-relevant human cognitive biases into them.
- Laid the foundation for follow-up research by preprocessing Dutch Wikipedia data as well as fMRI and EEG data.
- Performed comparative analyses between brain and neural language model activations.

MAJ // Digital

Lisbon, Portugal

SOFTWARE DEVELOPMENT INTERN

03.2017–08.2017

Developed a NLP-pipeline and backend to extract a [knowledge graph](#) from Wikipedia, using Python, Docker and Neo4j.

Advanced Linguistic Technologies GBR

Weinheim, Germany

STUDENT DEVELOPER

01.2016–01.2017

Developed a proprietary RESTful-API for better pricing in e-commerce.

Education

IT University of Copenhagen (IT-Universitetet i København)

Copenhagen, Denmark

PHD STUDIES

11.2020–05.2024

- Research on Uncertainty Quantification for Natural Language Processing and other topics as well as attending courses about Graph Theory, Algorithm Design and Bayesian Data Analysis. Supervised by Dr. Christian Hardmeier (IT University of Copenhagen) and Dr. Jes Frellsen (Technical University of Denmark).
- Published or co-authored 12 publications at a variety of venues and journals, including EMNLP, EACL, ACL, UAI, ICLR, TMLR and Nature Machine Intelligence.
- Produces multiple open-source code repositories with over 400 combined stars on Github.
- Dissertation title: “On Uncertainty in Natural Language Processing”.

University of Amsterdam (Universiteit van Amsterdam)

Amsterdam, Netherlands

MASTER STUDIES

09.2017–08.2019

- MSc Artificial Intelligence (GPA 4.0/4.0, Cum Laude). Lectures covering theoretical Machine Learning, Natural Language Processing, Computer Vision, Reinforcement Learning, Information Retrieval and other topics.
- Thesis title: “Recoding latent sentence representations - Dynamic gradient-based activation modification in RNNs”.

Heidelberg University (Ruprecht-Karls-Universität Heidelberg)

Heidelberg, Germany

BACHELOR STUDIES

10.2012–08.2016

- BSc Computational Linguistics / Geography (GPA 3.4/4.0). Attended various courses concerning linguistics, programming, logic, statistics and Machine Learning as well as human geography.
- Thesis title: “Analysis of relation prediction in German using word vector representations”.

Publications (* = equal contribution)

- 2026 **A. Kyriakou, D. Ulmer, I. Titov: Shared Doubt: Zero-shot Cross-Lingual Confidence Estimation for Language Models**, Pre-print.
- 2026 **I. Brink, A. Boer, D. Ulmer: Probing for Knowledge Attribution in Large Language Models**, Pre-print.
- 2026 **S. Weber, A. Wang, A. Gupta, A. Subramonian, D. Ulmer, E. Tanwar, G. Aich, H. Devinney, J. Hobbs, J. Mickel, J. Tint, M. Sosto, R. Groshan, S. Astarita, V. Gautam, V. Blaschke, W. Agnew, W. Y. Lee, Y. Long: Queer NLP: A Critical Survey on Literature Gaps, Biases and Trends**, Pre-print.
- 2026 **Y. Xia, D. Ulmer, T. Blevins, Y. Liu, H. Schütze, B. Roth: Calibration Is Not Enough: Evaluating Confidence Estimation Under Language Variations**, Pre-print.
- 6.2026 **D. Ulmer, A. Lorson, I. Titov, C. Hardmeier: Anthropomorphic Uncertainty: What Verbalized Uncertainty in Language Models is Missing**, Transactions of the Association for Computational Linguistics
- 09.2024 **D. Ulmer: On Uncertainty in Natural Language Processing**, PhD Dissertation, awarded as runner-up for the 2026 ACL Computational Linguistics Doctoral Dissertation Award. Copenhagen, Denmark
- 08.2024 **D. Ulmer, M. Gubri, H. Lee, S. Yoon, S. Joon Oh: Calibrating Large Language Models Using Their Generations Only**, Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL) 2024. Bangkok, Thailand
- 08.2024 **M. Gubri, D. Ulmer, H. Lee, S. Yoon, S. Joon Oh: TRAP: Targeted Random Adversarial Prompt Honey-pot for Black-Box Identification**, Findings of the Association for Computational Linguistics (ACL) 2024. Bangkok, Thailand
- 08.2024 **D. Ulmer, E. Mansimov, K. Lin, J. Sun, X. Gao, Y. Zhang: Bootstrapping LLM-based Task-Oriented Dialogue Agents via Self-Talk**, Findings of the Association for Computational Linguistics (ACL) 2024. Bangkok, Thailand
- 05.2024 **A. Farinhas, C. Zerva, D. Ulmer, A. Martíns: Non-Exchangeable Conformal Risk Control**, International Conference on Learning Representations (ICLR) 2024. Vienna, Austria
- 03.2024 **D. Ulmer, C. Zerva, A. Martíns: Non-Exchangeable Conformal Language Generation with Nearest Neighbors**, Findings of the European Chapter of the Association for Computational Linguistics (EACL) 2024. Valetta, Malta
- 10.2023 **D. Hupkes, M. Giulianelli, V. Dankers, M. Artetxe, Y. Elazar, T. Pimentel, C. Christodoulopoulos, K. Lasri, N. Saphra, A. Sinclair, D. Ulmer, F. Schottmann, K. Batsuren, K. Sun, K. Sinha, L. Khalatbari, R. Frieske, R. Cotterell, Z. Jin: State-of-the-art Generalisation Research in NLP: A Taxonomy and Review**, Nature Machine Intelligence.
- 07.2023 **J. Baan*, N. Daheim*, E. Ilia*, D. Ulmer*, H.-S. Li, R. Fernández, B. Plank, R. Sennrich, C. Zerva, W. Aziz: Uncertainty in Natural Language Generation: From Theory to Applications**, Preprint.
- 04.2023 **D. Ulmer, C. Hardmeier, J. Frellsen: Prior and Posterior Networks: A Survey on Evidential Deep Learning Methods For Uncertainty Estimation**, Transactions on Machine Learning Research.
- 12.2022 **D. Ulmer, J. Frellsen, C. Hardmeier: Exploring Predictive Uncertainty and Calibration in NLP: A Study on the Impact on Method & Data Scarcity**, Findings of 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP). Abu Dhabi, UAE

12.2022	D. Ulmer, E. Bassignana, M. Müller-Eberstein, D. Varab, M. Zhang, R. van der Goot, C. Hardmeier, B. Plank: <i>Experimental Standards for Deep Learning in Natural Language Processing Research</i> , Findings of 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP).	Abu Dhabi, UAE
04.2022	D. Ulmer, C. Hardmeier, J. Frellsen: <i>deep-significance: Easy and Meaningful Significance Testing in the Age of Neural Networks.</i> , Machine Learning Evaluation Standards Workshop at ICLR 2022	Online
04.2022	D. Ulmer, E. Bassignana, M. Müller-Eberstein, D. Varab, M. Zhang, C. Hardmeier, B. Plank: <i>Experimental Standards for Deep Learning Research: A Natural Language Processing Perspective.</i> , Machine Learning Evaluation Standards Workshop at ICLR 2022	Online
07.2021	D. Ulmer, G. Ciná. <i>Know Your Limits: Uncertainty Estimation with ReLU Classifiers Fails at Reliable OOD Detection.</i> , Uncertainty in Artificial Intelligence (UAI) 2021	Toronto, Canada
12.2020	D. Ulmer, L. Meijerink, G. Ciná. <i>Trust Issues: Uncertainty Estimation Does Not Enable Reliable OOD Detection On Medical Tabular Data.</i> , Machine Learning for Health Workshop (ML4H) at NeurIPS 2020	Vancouver, Canada
08.2019	J. Baan*, J. Leible*, M. Nikolaus*, D. Rau*, D. Ulmer*, T. Baumgärtner, D. Hupkes and E. Bruni. <i>On the Realization of Compositionality in Neural Networks.</i> , 2nd Edition of the BlackboxNLP Workshop (BlackboxNLP)	Florence, Italy
08.2019	D. Ulmer, D. Hupkes and E. Bruni. <i>Assessing incrementality in sequence-to-sequence models.</i> , 4th Workshop on Representation Learning for NLP (Repl4NLP)	Florence, Italy

Skills

Research	Uncertainty quantification for large language models, machine learning, Bayesian / frequentist statistics, deep learning
Programming	Python (profound), Java (fundamentals), R (fundamentals), MATLAB (fundamentals), Shell / bash (intermediate), \LaTeX (advanced), Javascript (fundamentals), HTML (fundamentals), CSS (fundamentals)
Machine Learning	PyTorch (profound), HuggingFace Transformers (profound), NumPy (advanced), Scikit-learn (advanced), SciPy (intermediate), Tensorflow (fundamentals)
Software Development	Git (intermediate), Docker (fundamentals), Neo4j (fundamentals), SCRUM (fundamentals), Python packaging / PyPI (intermediate)
Languages	German (native), English (fluent), Spanish (intermediate), Dutch (intermediate), Mandarin (beginner), Portuguese (beginner), Danish (beginner), Italian (beginner)
Soft Skills	Purposeful and succinct communication, openness to criticism, eagerness to receive constructive feedback, structured & independent working method, international work experience. Supervision of students and multitude of successful collaboration across distributed teams.

Talks

11.2023	IMS Stuttgart , Uncertainty Quantification for Natural Language Processing	Stuttgart, Germany
02.2023	Helsinki NLP , Uncertainty Quantification for NLP: Current State & Open Questions (Remote talk)	Helsinki, Finland
10.2022	Industry Reading Group , Uncertainty Estimation in Natural Language Processing	New York, USA
08.2019	2nd Edition of the BlackboxNLP Workshop (BlackboxNLP) , On the Realization of Compositionality in Neural Networks	Florence, Italy
01.2019	Computational Linguistics in the Netherlands (CLIN) 29 , An Incremental Encoder for Sequence-to-Sequence Modelling	Groningen, Netherlands

Awards & Achievements

- Awarded the Cohere For AI Research Grant, consisting of 5000 USD in compute credits for Cohere AI models.
- Outstanding paper award for *Experimental Standards for Deep Learning Research: A Natural Language Processing perspective* at Machine Learning Evaluation Standards Workshop at ICLR 2022.
- Outstanding reviewer at Machine Learning Evaluation Standards Workshop at ICLR 2022.
- Published open-source package [deep-significance](#) to enable easy significance testing for neural networks, starred by over 300 users on [GitHub](#), downloaded over 56000 times from [PyPI](#) (Python Package Index)¹ (May 2026), and used in at least 36 academic publications.

¹Determined via [pepy.tech](#).

Miscellaneous

UncertainNLP: First Workshop on Uncertainty-aware NLP

Valleta, Malta

CO-ORGANIZER

2023–24

Co-organizing Workshop for research related to predictive uncertainty in Natural Language Processing, co-located with the EACL 2024 conference.

Beers with NLPeers

Copenhagen, Denmark

ORGANIZER

2021–today

Initiator and organizer of a meet-up for researchers conducting Natural Language Processing Research across the different universities in Copenhagen.

Student parliament (StuRa) of Heidelberg University

Heidelberg, Germany

ELECTED STUDENT PARLIAMENT MEMBER

2014–2016

Participated in sessions and plenary discussions.

Department for Computational Linguistics, University Heidelberg

Heidelberg, Germany

ELECTED MEMBER OF THE DEPARTMENT'S STUDENT COUNCIL

2014–2016

Organized department-wide events, mediated communication between student body and faculty. Spearheaded an update to the curriculum to suit modern Machine Learning methods.