Maartje ter Hoeve

Online version
Last updated: January 2026

Research Interests

Advancing NLP equally for everyone. Human-Centered NLP, Multilinguality, Low-resource, Personalization, Explainability, Linguistics

Professional Experience

Oct 2025 – **Present** **Apple ML Research**, *Senior Machine Learning Researcher*.
Leading research efforts on Human-Centered NLP
- Focused on three pillars: Multilinguality, Low-resource NLP, and Personalization.
- Managing and mentoring Ph.D. Research Interns and junior Machine Learning Researchers, while maintaining significant IC-level technical contributions.
- Outcomes published in top-tier ML and NLP conferences.

Jun 2023 – Oct 2025 **Apple ML Research**, *Machine Learning Researcher*.

Apr 2022 – Aug 2022 **Apple ML Research**, *Research Intern*.
Advisor: Dr. Natalie Schluter

Jun 2021 – Oct 2021 **Facebook AI Research**, *Research Intern*.
Advisor: Prof. Dr. Emmanuel Dupoux

May 2019 – Aug 2019 **Microsoft Research AI**, *Research Intern*.
Advisors: Dr. Robert Sim & Dr. Ryen White

Jan 2017 – Sep 2017 **Blendle**, *Research Intern*.
Advisors: Prof. Dr. Maarten de Rijke & Dr. Anne Schuth

May 2016 – Jun 2016 **ASML**, *Research Intern*.
Advisor: Dr. Ivan Titov

Education

2018 – 2023 **University of Amsterdam**, *PhD, Human-Centered NLP*.
Advisors: Prof. Dr. Maarten de Rijke (University of Amsterdam) & Dr. Julia Kiseleva (Microsoft Research)
Ph.D. Committee: Prof. Dr. Maarten de Rijke, Dr. Julia Kiseleva, Prof. Dr. Raquel Fernández Rovira, Prof. Dr. Evangelos Kanoulas, Prof. Dr. Christof Monz, Prof. Dr. Barbara Plank, Dr. Robert Sim

2015 – 2017 **University of Amsterdam**, *MSc, Artificial Intelligence*, Cum Laude / With Distinction, GPA: 8.7/10.

2014 – 2015 **University of Amsterdam**, *Pre-master, Artificial Intelligence*, GPA: 9.0/10.

Sep 2013 – Dec 2013 **University of Edinburgh**, *Erasmus Exchange, focused on Informatics & Linguistics*, GPA: 9.0/10.

2011 – 2014 **University of Amsterdam**, *BA Linguistics (Honors)*, Cum Laude / With Distinction, GPA: 8.9/10.

Selected Activities and Awards

Community Service **Area Chair**, NAACL 2024, EMNLP 2024.

Reviewer, ICLR 2024, NeurIPS 2024, ACL ARR 2024, WiML 2024, ACL ARR 2023, ACL ARR 2022, EMNLP 2022, CoNLL 2021, TOIS 2021, ACL 2020, CIKM 2020, EACL 2020, TOIS 2020, NAACL 2019, ICML 2019, TOIS 2019, CIKM 2018, NeurIPS 2018, WWW 2017, WSDM 2017.

Mentor, WiML NeurIPS 2025, WiNLP EMNLP 2024, Black in AI NeurIPS 2025.

Co-founder, organizer, and mentor at Inclusive AI, Mentorship program at the University of Amsterdam.

Workshop organizer **PALS: Exploring Active and Passive LLM Personalization**, EMNLP 2025.

Challenge organizer **IGLU: Interactive Grounded Language Understanding in a Collaborative Environment**, NeurIPS 2021, 2022.

Board Positions **Project Officer**, Extensus Alumni Network, 2015-2019.

President, Extensus Honours Association, 2014-2015.

Participant **Microsoft Research AI Breakthroughs workshop**, 2020.

Data Science Summer School, Paris 2018.

Google NLP Summit, Zurich 2017.

Awards **SAC Highlight (Top 1%)**, EMNLP 2025.

ASML Technology Scholarship, 2015-2017.

Misc **Organizer of Search Engines Amsterdam Meetup (SEA)**, 2020-2021.

Selected Invited Talks

2025 NeurIPS: WiML

2024 EMNLP: Workshop on Customizable NLP

2024 Guest Lecture at Georgia Tech

2022 Search Engines Amsterdam

2022 DeepMind

2022 Spotify Research

2021 Guest in Data Skeptic Podcast

2020 ICAI Amsterdam Meetup

Languages

Dutch. *Native*

English. *Full professional proficiency*

German. *Professional working proficiency.* I took a two-week intensive language course at the Goethe Institut in Berlin in 2023, and evening classes at the German School of Silicon Valley in 2024. Both at C1 level.

French. *Limited working proficiency.* I took a two-week intensive language course at Lyon Bleu in Lyon in 2011, and evening classes at Langue Appart' in Paris in 2022. Both at A2 level.

Teaching

Ph.D. Research Interns

2025 **Arnav Arora**, Apple MLR.

2025 **Anastasiia Sedova**, Apple MLR.

2024 **Verena Blaschke**, Apple MLR.
Topic: Cross-lingual Transfer

2024 **Lucie Charlotte Magister**, Apple MLR.
Topic: Towards LLM Personalization: Learning to Remember User Conversations

Master Thesis Supervision

2022 **Rahel Habacker**, *MSc Artificial Intelligence*, University of Amsterdam.
Topic: Interactive Language Modeling

2021 **Michael Neely**, *MSc Artificial Intelligence*, University of Amsterdam.
Topic: Question Driven Automatic Summarization

2021 **Justine Winkler**, *MSc Artificial Intelligence*, Radboud University Nijmegen.
Topic: Modeling Background Knowledge for Automatic Summarization

2020 **Henning Bartsch**, *MSc Artificial Intelligence*, University of Amsterdam.
Topic: Improving User Representations for News Recommendation

2019 **Joris Baan**, *MSc Artificial Intelligence*, University of Amsterdam.
Topic: Transparency and Interpretability in Abstractive Summarization

2019 **David Stap**, *MSc Artificial Intelligence*, University of Amsterdam.
Topic: Conditional Image Generation and Manipulation

2018 **David Ruhe**, *MSc Data Science*, University of Amsterdam.
Topic: Adversarial Examples for Capsule Networks

2018 **Dennis Craandijk**, *MSc Data Science*, University of Amsterdam.
Topic: Automatic Unseen Event Extraction from Text

2018 **Mink Rohmer**, *MSc Data Science*, University of Amsterdam.
Topic: Classifying the Severity of Written Complaints

Courses (TA and Lectures)

2021 **Fairness, Accountability and Transparency**, *MSc Artificial Intelligence*, University of Amsterdam.

2019 **Information Retrieval**, *MSc Data Science*, University of Amsterdam.

2018 **Information Retrieval 2**, *MSc Artificial Intelligence*, University of Amsterdam.

2016 **Natural Language Processing**, *BSc Artificial Intelligece*, University of Amsterdam.

2015 **Logic Programming and Search**, *BSc Artificial Intelligence*, University of Amsterdam.

2015 **Introduction to AI**, *BSc Artificial Intelligence*, University of Amsterdam.

Selected Publications

🎓 `https://scholar.google.com/citations?user=U4ATzGYAAAAJ`

Skyler Seto, **Maartje ter Hoeve**, Maureen de Seyssel, David Grangier. *Assessing the Role of Data Quality in Training Bilingual Language Models.* EMNLP 2025.

Maureen de Seyssel, Jie Chi, Skyler Seto, **Maartje ter Hoeve**, Masha Fedzechkina, Natalie Schluter. *Discriminating Form and Meaning in Multilingual Models with Minimal-Pair ABX Tasks*. EMNLP 2025. **SAC Highlight, Top 1%**.

Eileen Pan, Anna Seo Gyeong Choi, **Maartje ter Hoeve**, Skyler Seto, Allison Koenecke. *Analyzing Dialectical Biases in LLMs for Knowledge and Reasoning Benchmarks.* EMNLP 2025.

Andrew Silva, Katherine Metcalf, Lucie Charlotte Magister, Clemencia Siro, **Maartje ter Hoeve**. *Proceedings of the First Workshop on Tailoring AI: Exploring Active and Passive LLM Personalization (PALS 2025).* EMNLP 2025.

Skyler Seto*, **Maartje ter Hoeve***, He Bai, Natalie Schluter, David Grangier. *Training Bilingual LMs with Data Constraints in the Targeted Language.* ACL, 2025.

Verena Blaschke, Masha Fedzechkina, **Maartje ter Hoeve**. *Analyzing the Effect of Linguistic Similarity on Cross-Lingual Transfer: Tasks and Experimental Setups Matter.* ACL, 2025.

Lucie Charlotte Magister, Katherine Metcalf, Yizhe Zhang, **Maartje ter Hoeve**. *On the Way to LLM Personalization: Learning to Remember User Conversations.* L2M2 workshop at ACL, 2025

Rita Ramos, Everlyn Asiko Chimoto, **Maartje ter Hoeve**, Natalie Schluter. *GrammaMT: Improving Machine Translation with Grammar-Informed In-Context Learning.* ACL, 2025.

Hellina Hailu Nigatu, Min Li, **Maartje ter Hoeve**, Saloni Potdar, Sarah Chasins. *mRAKL: Multilingual Retrieval-Augmented Knowledge Graph Construction for Low-Resourced Languages.* ACL, 2025.

Yong Lin, Skyler Seto, **Maartje ter Hoeve**, Katherine Metcalf, Barry-John Theobald, Xuan Wang, Yizhe Zhang, Chen Huang, Tong Zhang. *On the Limited Generalization Capability of the Implicit Reward Model Induced by Direct Preference Optimization.* EMNLP Findings, 2024.

Maartje ter Hoeve. *New Directions in Human-Centered Language Technology. Understanding and Improving NLP Models.* PhD Thesis, 2023.

Aarohi Srivastava, Abhinav Rastogi, ..., **Maartje ter Hoeve**, ..., Zirui Wang, Ziyi Wu. *Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models.* TMLR, 2023.

Maartje ter Hoeve, Evgeny Kharitonov, Dieuwke Hupkes, Emmanuel Dupoux. *Towards Interactive Language Modeling.* SPA-NLP workshop at ACL 2022 & InterNLP workshop at NeurIPS 2022.

Julia Kiseleva, Alexey Skrynnik, Artem Zholus, Shrestha Mohanty, Negar Arabzadeh, Marc-Alexandre Côté, Mohammad Aliannejadi, Milagro Teruel, Ziming Li, Mikhail Burtsev, **Maartje ter Hoeve**, Zoya Volovikova, Aleksandr Panov, Yuxuan Sun, Kavya Srinet, Arthur Szlam, Ahmed Awadallah *IGLU 2022: Interactive Grounded Language Understanding in a Collaborative Environment at NeurIPS 2022.* Competition at NeurIPS 2022.

Maartje ter Hoeve, Julia Kiseleva, Maarten de Rijke. *What Makes a Good and Useful Summary? Incorporating Users in Automatic Summarization Research.* NAACL 2022.

Ana Lucic, **Maartje ter Hoeve**, Gabriele Tolomei, Maarten de Rijke, Fabrizio Silvestri. *CF-GNNExplainer: Counterfactual Explanations for Graph Neural Networks.* AISTATS 2022.

Julia Kiseleva, Ziming Li, Mohammad Aliannejadi, Shrestha Mohanty, **Maartje ter Hoeve**, Mikhail Burtsev, Alexey Skrynnik, Artem Zholus, Aleksandr Panov, Kavya Srinet, Arthur Szlam, Yuxuan Sun, Katja Hofmann, Michel Galley, Ahmed Awadallah. *Interactive Grounded Language Understanding in a Collaborative Environment: IGLU 2021.* NeurIPS 2021 Competitions and Demonstrations Track. PMLR.

Maartje ter Hoeve, Robert Sim, Elnaz Nouri, Adam Fourney, Maarten de Rijke, Ryen White. *Conversations with Documents: An Exploration of Document-Centered Assistance.* CHIIR 2020.

David Stap, Maurits Bleeker, Sarah Ibrahimi, **Maartje ter Hoeve**. *Conditional Image Generation and Manipulation for User-Specified Content.* AICCW workshop at CVPR 2020.

Joris Baan, **Maartje ter Hoeve**, Marlies van der Wees, Anne Schuth, Maarten de Rijke. *Do Transformer Attention Heads Provide Transparency in Abstractive Summarization?* FACTS-IR workshop at SIGIR 2019.

Maartje ter Hoeve, Mathieu Heruer, Daan Odijk, Anne Schuth, Martijn Spitters, Ron Mulder, Nick van der Wildt and Maarten de Rijke. *Do News Consumers Want Explanations for Personalized News Rankings?*. FATREC Workshop on Responsible Recommendation at RecSys 2017.