

Janosch Haber

SENIOR DEEP LEARNING SCIENTIST | PHD IN COMPUTER SCIENCE: COMPUTATIONAL LINGUISTICS

✉ janoschhaber@gmail.com | 🏠 janoschhaber.com | in janoschhaber

Summary

Senior Deep Learning Scientist looking for challenges combining AI research with social impact. Proven track record of academic research, hands-on expertise in developing state-of-the-art AI solutions from problem statement to commercial deployment, and on-the-job experience in leading teams and projects to reach their goals.

Experience

Senior Deep Learning Scientist

London, United Kingdom

CHATTERMILL

Jun 2023 - PRESENT

- At Chattermill, we crystallise individual, actionable insights from heaps and heaps of customer feedback data. As Senior Deep Learning Scientist I lead the development of new deep learning models to improve the quality and impact of our solutions, and research new approaches to extract valuable information from raw data.
- Skills: Natural Language Processing (NLP), Data Science, Deep Learning, Python, MLOps, CI/CD, AWS, Project Management

Lead NLP Research Scientist

London, United Kingdom

REWIRE

Jun 2022 - May 2023

- At Rewire, we developed state-of-the-art, custom-tailored AI models to detect toxic language such as hate and abuse, and other explicit content. As Lead Research Scientist I co-ordinated projects and contributed to developing context-sensitive language models all the way from initial data collection to deployment in our clients' applications. Rewire was acquired by ActiveFence in early 2023.
- Skills: Natural Language Processing (NLP), Data Science, Project Management, Deep Learning, Python, Pandas, PyTorch, MLOps, AWS

Co-Founder, Board Member and Web Designer

CORRESPONDENTS OF THE WORLD (CotW)

Jan 2020 - PRESENT

- CotW is a non-profit aimed at bringing together people from different backgrounds so they can share their perspectives and better understand one-another. I'm a co-founder and board member, balancing everyday organisational tasks with shaping our long-term mission.
- Skills: Organisation Management, Team Management, Team Leadership

Course Assistant for Natural Language Processing

London, United Kingdom

QUEEN MARY UNIVERSITY OF LONDON (QMUL)

Sep 2019 - Sep 2021

- As Course Demonstrator for the NLP module I prepared course materials, helped students with questions about their assignments (Python programming, text processing, classification) and assisted the course coordinator in scoring assessments and exams.

Student Assistant for the BSc Artificial Intelligence

Amsterdam, The Netherlands

UNIVERSITY OF AMSTERDAM (UVA)

Jan 2015 - Jun 2017

- During my Bachelors and Masters I was a Student Assistant for a number of AI Bachelor's courses. I assisted the course coordinators by helping out students with their course assignments and grading assessments and exams.

Education

PhD in Computer Science: Computational Linguistics

London, United Kingdom

QUEEN MARY UNIVERSITY OF LONDON (QMUL)

Sep 2018 - Jul 2022

- Supervised by Prof. Dr. Massimo Poesio, Dr. Julian Hough and Prof. Dr. Patrick Healey
- For my PhD I investigated nuances in the interpretation of ambiguous words - both in the human language processor and in context-sensitive AI language models. I collected a large-scale dataset of similarity ratings and with it was able to show that humans as well as current computer models do distinguish between some senses while merging others.
- My work has been published in multiple top-tier conferences (see below), my thesis can be found here.
- Research funded by the Disagreements in Language Interpretation (DALI) project, ERC Grant 695662, and in part supported by the Enrichment Scheme of The Alan Turing Institute.

Master of Science (MSc) in Artificial Intelligence

Amsterdam, The Netherlands

UNIVERSITY OF AMSTERDAM (UVA)

Sep 2016 - Jul 2018

- Cum Laude; Final Grade 8.8
- Thesis supervised by Prof. Dr. Raquel Fernández and Dr. Elia Bruni
- Thesis Title: How should we call it? - Introducing the PhotoBook Conversation Task and Dataset for Training Natural Referring Expression Generation in Artificial Dialogue Agents.
- Thesis research funded by the Facebook ParLAI Research Award.

Bachelor of Science (BSc) in Artificial Intelligence

Amsterdam, The Netherlands

UNIVERSITY OF AMSTERDAM (UVA)

Sep 2012 - Jul 2015

- Cum Laude; Final Grade 8.3
- Thesis supervised by Dr. Roberto Valenti

Scholarships & Awards

Enrichment Student

London, United Kingdom

THE ALAN TURING INSTITUTE

Oct 2021 - Jun 2022

Research Scholarship

DALI PROJECT, FUNDED BY EUROPEAN RESEARCH COUNCIL GRANT 695662

Oct 2019 - Sep 2021

Master's Thesis Research Funding

FACEBOOK PARLAI RESEARCH AWARD

Jan 2018 - Jul 2018

Full Undergraduate and Graduate Scholarships

EVANGELISCHE STUDIENSTIFTUNG VILLIGST E.V.

Feb 2013 - Jul 2018

Publications (Selection)

Polysemy - Evidence from Linguistics, Behavioural Science and Contextualised Language Models.

Janosch Haber and Massimo Poesio.
Computational Linguistics. 2024

Improving the Detection of Multilingual Online Attacks with Rich Social Media Data from Singapore

Paul Röttger Janosch Haber, Bertie Vidgen, Matthew Chapman, Vibhor Agarwal, Roy Ka-Wei Lee, Yong Keong Yap
Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL). 2023

Computational Models of Anaphora.

Massimo Poesio, Juntao Yu, Silviu Paun, Abdulrahman Aloraini, Pengcheng Lu, Janosch Haber, and Derya Cokal.
Annual Review of Linguistics. 2023

Patterns of Lexical Ambiguity in Contextualised Language Models.

Janosch Haber and Massimo Poesio.
Findings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2021

Word Sense Distance in Human Similarity Judgements and Contextualised Word Embeddings.

Janosch Haber and Massimo Poesio.
Proceedings of the Conference on Probability and Meaning (PaM2020) – SIGSEM Workshop. 2020

Assessing Polyseme Sense Similarity through Co-predication Acceptability and Contextualised Embedding Distance.

Janosch Haber and Massimo Poesio.
Proceedings of the 9th Joint Conference on Lexical and Computational Semantics (*SEM). 2020

The PhotoBook Dataset: Building Common Ground through Visually Grounded Dialogue.

Janosch Haber, Tim Baumgärtner, Ece Takmaz, Lieke Gelderloos, Elia Bruni, and Raquel Fernández.
Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL). 2019