

Richard Diehl Martinez

AI Research Fellow at the University of Cambridge

My research focuses on designing novel training techniques and developing analytical frameworks to improve the performance and computational efficiency of language models.

- ✉ richarddiehl@gmail.com
- ☎ +1 (857) 253-9116
- 🌐 [LinkedIn](#)
- 🐙 [GitHub](#)
- 🌐 [Website](#)

Education

-  **PhD, Computer Science**
University of Cambridge 2021 - 2025
Focus: Natural Language Processing
-  **MSc, Computer Science**
Stanford University 2018 - 2020
Focus: Natural Language Processing
-  **BSc, Management Science & Engineering**
Stanford University 2014 - 2018
Focus: Operations Research

Awards

- Gates Cambridge Scholar** (Cambridge)
- Best Undergraduate Thesis** (Stanford)
- Graduate Teaching Award** (Stanford)

Skills

Programming: Python, R, C, JavaScript, React/Next.js
Frameworks: PyTorch/Tensorflow, HuggingFace, PyTorch Lightning, CUDA (Triton), Docker, Ray
Languages: English (native), German (native), Spanish (native), French (B2), Chinese (HSK4)

Experience

Industry

-  **Founder**
Pico Labs (picolm.io) 2024 - Present
Building open-source tools and frameworks for small language model research and development. The open-source repositories have received over 400 stars on GitHub.
-  **Applied Research Scientist**
Amazon Alexa 2020 - 2021
Developed neural language models to personalize and improve Alexa's Automatic Speech Recognition (ASR) capabilities, increasing production accuracy by up to 10% on certain test sets. Spearheaded a project to continuously augment training data with real-time trending topics.

Teaching

-  **Stanford CS 221 - AI: Principles and Techniques**
Graduate Teaching Assistant 2018 - 2019
Led discussion and review sections for over 300 students and provided mentorship on final class research projects. Awarded a best graduate teaching assistant award.

Research

-  **Researcher, Cambridge NLP Group**
University of Cambridge 2021 - 2025
Supervised by Paula Buttery. Researched parameter-efficient pre-training techniques to bridge the performance gap between small and large language models. Published 11 papers at major conferences including ACL and EMNLP, with a best-paper award at CoNLL.
-  **Researcher, Stanford NLP Group**
Stanford University 2018 - 2020
Supervised by Dan Jurafsky. Implemented generative language modeling techniques to detect and correct bias in news and online media. Results published at AAAI 2020, as a keynote presentation paper.
-  **Researcher, Stanford AI Lab (SAIL)**
Stanford University 2017 - 2018
Supervised by Silvio Savarese. Designed the pedestrian emotion detection pipeline used in JackRabbit, a self-navigating autonomous robot designed for indoor environments.

Selected Publications

[Full list: Google Scholar](#)

Investigating ReLoRA: Effects on the Learning Dynamics of Small Language Models Yuval Weiss, David Demetri Africa, Paula Buttery, Richard Diehl Martinez	<i>BlackboxNLP 2025</i>
Meta-Pretraining for Zero-Shot Cross-Lingual NER in Low-Resource Philippine Languages David Demetri Africa, Yuval Weiss, Paula Buttery, Richard Diehl Martinez	<i>MRL 2025</i>
Pico: A Modular Framework for Hypothesis-Driven Small Language Model Research Richard Diehl Martinez, David Demetri Africa, Yuval Weiss, Suchir Salhan, Ryan Daniels, Paula Buttery	<i>EMNLP 2025</i>
Mitigating Frequency Bias and Anisotropy in Language Model Pre-Training with Syntactic Smoothing Richard Diehl Martinez, Zebulun Goriely, Andrew Caines, Paula Buttery, Lisa Beinborn	<i>EMNLP 2024</i>
Tending Towards Stability: Convergence Challenges in Small Language Models Richard Diehl Martinez, Pietro Lesci, Paula Buttery	<i>EMNLP Findings 2024</i>
CLIMB: Curriculum Learning for Infant-inspired Model Building Richard Diehl Martinez, Zeb Goriely, Hope McGovern, Andrew Caines, Paula Buttery, Lisa Beinborn	<i>CoNLL 2023 (Best Paper)</i>
Automatically Neutralizing Subjective Bias in Text Reid Pryzant, Richard Diehl Martinez, Nathan Dass, Sadao Kurohashi, Dan Jurafsky, Diyi Yang	<i>AAAI 2020 (Keynote)</i>