Nathan Wolf

978-429-5367 | nwolf3275@gmail.com | Acton, MA 01720
https://www.linkedin.com/in/nathanw0lf/ | https://timmald.github.io

PROFESSIONAL SUMMARY

Aspiring ML Researcher with strong academic background and significant experience in academic and industry research. I am specifically interested in NLP and Mechanistic Interpretability.

EDUCATION

University of Massachusetts Amherst, Amherst MA Expected May 2026
Bachelor of Science, Computer Science
Bachelor of Arts, Computational Linguistics
Commonwealth Honors College
GPA: 3.9/4.0
Relevant Coursework: Honors Machine Learning, Adv. Computational Linguistics, Natural Language Processing, Probability, Linear Algebra, Discrete Math, Data Structures, Algorithms

PUBLICATIONS

Wein, S., Serbina, A., Ji, J., Wolf, N., DeGraaff, J.,Kini, P., Leonor Pacheco, M. "Lost in Translation, and Found: Detecting and Interpreting Translation Effects" Submitted to ACL, 2025.;

Staub, A., Wolf, N., Dillon, B. “Linking syntactic computation to eye movements in reading: Evidence from agreement” Submitted to *Open Mind: Discoveries in Cognitive Science*, 2025.

RELEVANT EXPERIENCE

NLP Researcher Sep 2025 - Present
University of Massachusetts Amherst, Amherst, MA

- Honors thesis advised by Dr. Katrin Erk
- Investigated the impact of morphologically-aligned tokenization on language model performance in fusional vs. agglutinative languages
- Built efficient tokenization and pretraining workflows to work with limited compute resources
- Received competitive $850 Honors Research Grant

Research Assistant, NLP Group Sep 2025 - Present
Amherst College, Amherst, MA

- Advised by Dr. Shira Wein
- Conducted interpretability research on a model distinguishing text translated into English from text originally written in English
- Used attention visualization and Integrated Gradients saliency data to find key linguistic features the model was relying on to identify translated English.
- Co-authored paper submitted to ACL

Research Assistant, Psycholinguistics Lab Feb - May 2025
University of Massachusetts Amherst, Amherst, MA

- Honors Independent Study advised by Dr. Brian Dillon
- Studied the influence of LLM surprisal on reading times in eye-tracking data
- Performed linear mixed-effects analysis on eye-tracking data to estimate effect sizes
- Co-authored paper submitted to Open Mind: Discoveries in Cognitive Science

MassAI Co-President / Project Leader / Instructor Jul 2023 - Present
University of Massachusetts Amherst, Amherst, MA

- Taught a team of first-year students to build an unsupervised email classifier as their first ML project
- Adapted research literature on document clustering and tested potential methodologies to help guide the team
- Built reinforcement learning (RL) gym and agents to provide MassAI with an RL competition event
- Organized and led interactive workshops to teach students about Machine Learning with engaging hands-on exercises

Software Engineer Intern, R&D Summer 2024 and Summer 2025
Shell TechWorks, Boston, MA

- Researched methodologies for efficient fine-tuning of LLMs
- Built multi-agent RAG workflow to analyze the investability of startup pitches
- Built hybrid search RAG workflow to query Elasticsearch database with natural language questions, achieved 90% accuracy

SKILLS

Python | R | C/C++ | JS | Unix Shell (Bash) | Pytorch | Tensorflow | Huggingface | NumPy | Pandas | Langchain | Elasticsearch

AWARDS

- Honors Research Grant, Commonwealth Honors College 2025
- Best Use of AI, HackHer413 2025
- UMass Amherst Dean's List, all semesters 2022 - Present
- Best Healthcare App, HackUMass X 2022
- AP Scholar with Distinction 2022
- Eagle Scout 2021

PERSONAL PROJECTS

MorphSeg Nov 2025

- Trained morpheme segmentation models based on Tü Seg architecture
- Published model as a Python package and a SpaCy module for use in NLP research
- Created and hosted morpheme segmentation API

GRPO Experiments Jun 2025

- Experimented with GRPO for Reinforcement Learning LLM fine-tuning
- Fine-tuned LLM to prioritize the 1000 most frequent english words
- Fine-tuned LLM to rhyme and use consistent number of syllables per line

Video Script Generator Feb - May 2025

- Fine-tuned LLM to generate original video scripts
- Created custom training corpus
- Experimented with methodologies and hyperparameters for best results

Unnatural Language Processing Feb - Aug 2023

- Trained RNN-based model to predict phonemes in English text
- Deployed model to create a text-to-speech site