

Khai-Nguyen Nguyen

knguyen07@wm.edu | (+1) 272-788-0512 | Google Scholar | Personal website | LinkedIn

EDUCATION	College of William and Mary <i>Graduate Student in Computer Science</i>	Virginia, USA 2023 - current
	• GPA: 3.96/4.00	
Bucknell University <i>B.Sc. in Computer Science and Engineering, minor in Statistics</i>	Pennsylvania, USA 2019 - 2023	
• GPA: 3.94/4.00 (<i>Summa Cum Laude</i>)		
RESEARCH INTERESTS	Natural Language Processing, Trustworthy AI, Multimodal AI, AI for Healthcare, AI for Software Engineering	
SELECTED PUBLICATIONS <small>*: equal contribution</small>	<ol style="list-style-type: none">1. Vision Language Models are Biased [pdf] An Vo*, Khai-Nguyen Nguyen*, Mohammad Reza Taesiri, Vy Tuong Dang, Anh Totti Nguyen, Daeyoung Kim <i>AI4Math Workshop @ ICML 2025 Submitted to ICLR 2026</i>2. Sentiment Reasoning for Healthcare [pdf] Khai-Nguyen Nguyen*, Khai Le-Duc*, Bach Phan Tat, Duy Le, Truong-Son Hy <i>ACL 2025, Industry Track (Oral)</i>3. Medical Spoken Named Entity Recognition [pdf] Khai Le-Duc, David Thulke, Hung-Phong Tran, Long Vo-Dang, Khai-Nguyen Nguyen, Truong-Son Hy, Ralf Schluter <i>NAACL 2025, Industry Track</i>4. Resource-Efficient & Effective Code Summarization [pdf] Saima Afrin, Joseph Call, Khai-Nguyen Nguyen, Oscar Chaparro, Antonio Mastropaoolo <i>ACM International Conference on AI Foundation Models and Software Engineering (FORGE 2025)</i>5. Real-time Speech Summarization for Medical Conversations [pdf] Khai Le-Duc*, Khai-Nguyen Nguyen*, Long Vo-Dang, Truong-Son Hy <i>Interspeech 2024 (Oral)</i>6. Getting away with network pruning: From sparsity to geometry and linear regions [pdf] Jeffrey Cai*, Khai-Nguyen Nguyen*, Nishant Shrestha, Aidan Good, Ruisen Tu, Xin Yu, Shandian Zhe, Thiago Serra <i>CPAIOR 2023 Workshop on Sparsity in Neural Networks, ICLR 2023</i>	
PREPRINTS <small>*: equal contribution</small>	<ol style="list-style-type: none">1. Beyond Tokens: Feature-Level Attribution for Explainable Neural Code Models Khai-Nguyen Nguyen*, Antonio Vitale*, Simone Scalabrino, Antonio Mastropaoolo <i>Submitting to FSE 2026</i>2. S-Chain: Structured Visual Chain-of-Thought for Medicine [pdf] Khai Le-Duc, Phuong T.H. Trinh, Duy Minh Ho Nguyen, Tien-Phat Nguyen, Nghiem Tuong Diep, An Ngo, Tung Vu, Trinh Vuong, Anh-Tien Nguyen, Nguyen Dinh Mau, Van Trung Hoang, Khai-Nguyen Nguyen, Hy Nguyen, Chris Ngo, Anji Liu, Nhat Ho, Anne-Christin Hauschild, Khanh Xuan Nguyen, Thanh Nguyen-Tang, Pengtao Xie, Daniel Sonntag, James Zou, Mathias Niepert, Anh Totti Nguyen <i>Submitting to ICLR 2026</i>	

RESEARCH EXPERIENCES

AURA Lab, College of William and Mary, Virginia, USA <i>Research Assistant. Advisor: Prof. Antonio Mastropaoletti</i>	2025.05 - current
 <i>Student Collaborator.</i>	2024.10 - 2025.01
Nguyen Lab, Auburn University, Alabama, USA <i>Remote Collaborator. Advisor: Prof. Anh Totti Nguyen</i>	2025.01 - 2025.08
FPT Software AI Center, Vietnam <i>Remote Research Intern. Supervisor: Prof. Truong-Son Hy</i>	2024.05 - 2024.09
Analytics Lab, Bucknell University, Pennsylvania, USA <i>Undergraduate Research Assistant. Advisor: Prof. Thiago Serra</i>	2022.05 - 2023.01

INDUSTRY EXPERIENCES

Machine Learning Intern, CodaMetrix Boston, USA <ul style="list-style-type: none">Developed a LLM judge to evaluate and refine medical entity annotations, improving performance by 2–4% F1 over human baselineDesigned and deployed entity extraction LLMs with DSPy prompt optimization on Databricks, achieving an overall F1 of 70% on Llama-3.1-70B	2025.05 - 2025.08
Machine Learning Intern, CodaMetrix Boston, USA <ul style="list-style-type: none">Trained a BERT-based LLM for ICD-10 multilabel classification, outperforming the deployed system by 4% accuracyAccelerate training by 4 times on Databricks using distributed and parallel training	2024.05 - 2024.08
Developer, Cazoodle Illinois, USA <ul style="list-style-type: none">Developed a query expansion system that utilize LLMs to comprehend user intent for targeted keyword recommendationBuilt and deployed a feature-rich Django interface using ElasticSearch BM25 search engine for comprehensive end-to-end testing	2023.05 - 2023.08

PRESENTATIONS

Sentiment Reasoning for Healthcare <i>Oral presentation at ACL 2025 Industry Track</i>
Medical Spoken Named Entity Recognition <i>Poster presentation at NAACL 2025 Industry Track</i>
Real-time Speech Summarization for Medical Conversations <i>Oral presentation at Interspeech 2024 & Poster presentation at MASC-SLL 2024</i>
Like a bilingual baby: The advantage of visually grounding a bilingual language model <i>Poster presentation at Susquehanna Valley Undergraduate Research Symposium 2022</i>

AWARDS AND HONORS

• Recipient Computer Science Fellowship, College of William and Mary <i>Awarded \$500 by the CSCI admissions committee</i>	2023
• Dean's List Bucknell University <i>Achieved a GPA of 3.6 or higher every semester</i>	2019 – 2023
• Recipient, Program for Undergraduate Research Grant, Bucknell University <i>Awarded \$4000 for summer research</i>	2021
• Honorable Mention, Mathematical Contest in Modeling <i>Ranked top 25% (3525/13753) among the participants</i>	2020

TEACHING EXPERIENCES

William and Mary: CSCI 140: Intro to Data Science CSCI 304: Computer Organization
Bucknell University: CSCI 204: Data Structures and Algorithms MATH 201, 202: Calculus I, II