David Yang

davidmyang@berkeley.edu | davidmyang.github.io | linkedin.com/in/davidmyang | github.com/davidmyang

Education

University of California, Berkeley Aug 2021 – May 2025
B.A. in Computer Science & B.A. in Linguistics; GPA: 3.9/4.0 Berkeley, CA

- **Relevant Coursework:** Language and Thought, Development of Bias, Data Science, Quantitative Linguistics, Data Structures, Efficient Algorithms, Sociolinguistics, Syntax, Historical Linguistics, Discrete Math and Probability

Papers in Submission/Preparation

Yang, D., Regier, T. (under review). *Re-examining the tradeoff between lexicon size and average morphosyntactic complexity in recursive numeral systems*.

Yang, D.*, Petrick, O.*, Petrosyan, S., & Tikhomirov G. (in prep). *nanoVR: Building and Simulating DNA nanostructures in Virtual Reality*.

Research Experience

Language and Cognition Lab Sep 2024 – Present
Advisor: Terry Regier (UC Berkeley) Berkeley, CA

- Developed language-specific grammars for a diverse set of natural and optimal artificial languages, analyzing structural variations and constraints.
- Implemented an evolutionary algorithm to examine how constraints on number grammars shape order in artificial numeral systems.
- Tested how natural languages optimize the sum of lexicon size and morphosyntactic complexity, as well as the role of the prior distribution.

Information Sciences Institute Jun 2024 - Aug 2024
Advisor: Alex Spangher (USC) Los Angeles, CA

- Designed and implemented scripts on an HPC cluster to analyze newsworthiness factors in city council meetings.
- Leveraged vLLM and topicGPT to extract and classify multi-level concepts in city policies, enabling automated policy categorization.
- Conducted statistical analysis using logistic regression models to evaluate predictive factors of newsworthiness.

Ti Lab Mar 2022 - Present
Advisor: Samson Petrosyan & Grigory Tikhmirov (UC Berkeley) Berkeley, CA

- Building a novel VR application that enables users to design and simulate custom DNA nanostructures.
- Implementing advanced features such as multi-grid movement, layered abstracted views, and multi-component structures.

Teaching Experience

CS61B: Data Structures Jan 2023 - Present
Teaching Assistant, Infrastructure Lead (5 semesters) Berkeley, CA

- Taught weekly discussion sections for 30+ students, covering implementation and theory of data structures.
- Developed course tools as part of the internal infrastructure team, improving automated grading, office hours tool, and student extensions.
- Co-taught a guest lecture on CI/CD pipelines and best practices for personal projects.

Professional Experience

Veeva Systems May 2024 - Aug 2024
Software Engineering Intern Pleasanton, CA

- Developed client-side React components for object controls to enhance usability for developers utilizing UISDK.

- Resolved UI defects and implemented unit tests for UISDK to improve stability and reliability.

Optum   Jun 2022 - Aug 2022
Software Engineering Intern   Eden Prairie, MI

- Developed a ML model using PySpark ML to predict pharmaceutical drug prices with 92% accuracy.
- Processed and analyzed 250 million rows of claim data to optimize model performance.
- Built a React-based frontend for an improved user experience and integrated Spring Boot for efficient backend API handling.

Community Service

The Music Connection   Jan 2024 - Present
Piano Tutor   Berkeley, CA

- Taught weekly piano lessons for elementary school students.

Computer Science Mentors   Aug 2022 - Dec 2023
Senior Mentor   Berkeley, CA

- Led family meetings to discuss CS pedagogy with junior mentors and taught weekly small-group sections.

Skills & Activities

Programming Languages: Python, R, Java, C#, C, HTML/CSS, SQL, RISC-V, x86, OCaml
Tools: Jupyter Notebook, Git, Pandas, NumPy, Praat, Unity, SBATCH, TensorFlow, React.js
Interests: Rockclimbing, Running, NYT word games, Piano, Saxophone, Languages
Natural Languages: English (native), Mandarin (fluent), Spanish (intermediate), Kazakh (elementary)