SAMUEL STEVENS

samuel.robert.stevens@gmail.com · https://samuelstevens.me

RESEARCH VISION

My research focuses on AI for acclerating data-driven scientific discovery; specifically, I focus on developing and leveraging large pretrained models for novel discovery from learned structures. I am most proud of BioCLIP (BioCLIP 2) and sparse autoencoders for vision.

EDUCATION

Ohio State University August 2021 - Present
Ph.D. in Computer Science & Engineering *Columbus, OH*
Advisor: Professor Yu Su

Ohio State University (3.93 GPA) August 2017 - May 2021
Honors B.S. in Computer Science & Engineering, *Summa Cum Laude* *Columbus, OH*
Honors Research Distinction in Computer Science & Engineering
German Minor

SELECTED PUBLICATIONS & PREPRINTS

Interpretable and Testable Vision Features via Sparse Autoencoders
Samuel Stevens, Wei-Lun Chao, Tanya Berger-Wolf and Yu Su
arXiv Preprint, 2025

· Applies sparse autoencoders (SAEs) to vision models.

BioCLIP 2: Emergent Properties from Scaling Hierarchical Contrastive Learning
Jianyang Gu, **Samuel Stevens**, Elizabeth G Campolongo, Matthew J Thompson, Net Zhang, Jiaman Wu, Andrei Kopanev, Zheda Mai, Alexander E. White, James Balhoff, Wasila Dahdul, Daniel Rubenstein, Hilmar Lapp, Tanya Berger-Wolf, Wei-Lun Chao, Yu Su
To appear in NeurIPS, 2025

· Scales BioCLIP to 214M images and discovers emergent propoerties in learned representations.

BioCLIP: A Vision Foundation Model for the Tree of Life
Samuel Stevens, Jiaman Wu, Matthew J Thompson, Elizabeth G Campolongo, Chan Hee Song, David Edward Carlyn, Li Dong, Wasila M Dahdul, Charles Stewart, Tanya Berger-Wolf, Wei-Lun Chao and Yu Su
CVPR, 2024 **(Best Student Paper)**

· Trains and evaluates a CLIP model on 10M image-text pairs for use in ecology and biology.

Salsa Fresca: Angular Embeddings and Pre-Training for ML Attacks on Learning With Errors
Samuel Stevens, Emily Wenger, Cathy Yuanchen Li, Niklas Nolte, Eshika Saxena, Francois Charton, and Kristin Lauter
TMLR, 2025

· Improves ML attack cost on Learning With Errors by 250× through modeling improvements.

Memorization for Good: Encryption with Autoregressive Language Models
Samuel Stevens and Yu Su
arXiv Preprint, 2023

· Proposes a novel symmetric encryption algorithm based on language model memorization.

An Investigation of Language Model Interpretability via Sentence Editing
Samuel Stevens and Yu Su
EMNLP BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP, 2021

· Finds that BERT's attention weights correlate well with human rationales.

RESEARCH EXPERIENCE

Zoom May 2024 - August 2024
AI Research Scientist Intern *Remote*

· Designed and implemented abstractions for programming language model more reliably.

Meta May 2023 - August 2023
Research Scientist Intern, supervised by Kristin Lauter and Francois Charton *Seattle, WA*

· Implemented and validated simpler architecture for the ongoing SALSA project.
· Continued part-time from August 2023 to December 2023.

Self-Attention Interpretability in Sequence Classification October 2019 - April 2021
Undegraduate Honors Research Thesis *Columbus, OH*

· Fine-tuned pre-trained language models on existing sequence classification task.
· Analyzed trends in the attention maps of BERT-based models to quantify differences in interpretability.
· Identified literature gap (no large-scale quantitative measurement of BERT's interpretability on sequence level classification) and designed experiments leading to a baseline result.

OTHER PUBLICATIONS & PREPRINTS

KABR: In-situ Dataset for Kenyan Animal Behavior Recognition from Drone Videos
Maksim Kholiavchenko, Jenna Kline, Michelle Ramirez, **Samuel Stevens**, Alec Sheets, Reshma Babu, Namrata Banerji, Elizabeth Campolongo, Matthew Thompson, Nina Van Tiel, Jackson Miliko, Eduardo Bessa, Isla Duporge, Tanya Berger-Wolf, Daniel Rubenstein, Charles Stewart
WACV, 2024

The Cool and the Cruel: Separating Hard Parts of LWE Secrets
Niklas Nolte, Mohamed Malhou, Emily Wenger, **Samuel Stevens**, Cathy Li, François Charton, Kristin Lauter
AfricaCrypt, 2024

MMMU: A Massive Multi-Discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI
Xiang Yue, Yuansheng Ni, Kai Zhang, Tianyu Zheng, Ruoqi Liu, Ge Zhang, **Samuel Stevens**, Dongfu Jiang, Weiming Ren, Yuxuan Sun, Cong Wei, Botao Yu, Ruibin Yuan, Renliang Sun, Ming Yin, Boyuan Zheng, Zhenzhu Yang, Yibo Liu, Wenhao Huang, Huan Sun, Yu Su, Wenhu Chen
CVPR, 2024 **(Best Paper Finalist)**

A Simple Interpretable Transformer for Fine-Grained Image Classification and Analysis
Dipanjyoti Paul, Arpita Chowdhury, Xinqi Xiong, Feng-Ju Chang, David Carlyn, Samuel Stevens, Kaiya Provost, Anuj Karpatne, Bryan Carstens, Daniel Rubenstein, Charles Stewart, Tanya Berger-Wolf, Yu Su, Wei-Lun Chao
ICLR, 2024

Mind2Web: Towards a Generalist Agent for the Web
Xiang Deng, Yu Gu, Boyuan Zheng, Shijie Chen, **Samuel Stevens**, Boshi Wang, Huan Sun and Yu Su
NeurIPS, 2023 **(Spotlight)**

SalsaBot: Towards a Robust and Generalizable Embodied Agent
Chan Hee Song, Jiaman Wu, Ju-Seung Byeon, Zexin Xu, Vardaan Pahuja, Goonmeet Bajaj, **Samuel Stevens**, Ziru Chen and Yu Su
4th Embodied AI Workshop, CVPR 2023

Roll Up Your Sleeves: Working with a Collaborative and Engaging Task-Oriented Dialogue System
Lingbo Mo, Shijie Chen, Ziru Chen, Xiang Deng, Ashley Lewis, Sunit Singh, **Samuel Stevens**, Chang-You Tai, Zhen Wang, Xiang Yue, Tianshu Zhang, Yu Su, and Huan Sun
SIGDIAL, 2023

Bootstrapping a User-Centered Task-Oriented Dialogue System
Shijie Chen, Ziru Chen, Xiang Deng, Ashley Lewis, Lingbo Mo, **Samuel Stevens**, Zhen Wang, Xiang Yue, Tianshu Zhang, Yu Su and Huan Sun
1st Proceedings of Alexa Prize TaskBot, 2022

arXivEdits: Understanding the Human Revision Process in Scientific Writing
Chao Jiang, Wei Xu and **Samuel Stevens**
EMNLP, 2022

ENGINEERING EXPERIENCE

SpaceX (Starlink) May 2021 - August 2021
Hardware Test Associate Engineer *Seattle, WA*

- Refactored legacy code leading to $66K expected annual savings and 25% increase in first-pass yield.
- Developed data visualization for micropositioner actuator repeatability during testing.
- Developed thermal camera testing prototype, reducing test time from over 30 minutes to 30 seconds.

Microsoft May 2020 - August 2020
Software Engineering Intern *Seattle, WA (Remote)*

- Developed new Power BI feature to improve customers' ability to diagnose data-refresh problems.
- Ramped up quickly in a 100 million-line C++ codebase to begin delivering value immediately.

TicketBay January 2018 - August 2020
Lead Developer *Columbus, OH*

- Collaborated with OSU students to develop a mobile app for students to buy and sell football tickets.
- Facilitated the transfer of $165K worth of tickets between more than 7K customers.
- Led system architecture decisions in order to balance ease of development and system performance.

GE Aviation May 2018 - August 2018
Digital Technology Intern *Cincinnati, OH*

- Built an end-to-end testing solution to automate QA testing of customer-facing web application.
- Integrated test solution with existing CI/CD pipeline, leading to improved product quality with no additional developer work.

The Ohio State University May 2018 - November 2019
Lead Developer *Columbus, OH*

- Developed a webapp to distribute rich media content to support a professor's research.
- Managed 2 student developers in multiple time zones with agile methodology.

GE Aviation May 2017 - August 2017
INTERalliance Intern *Cincinnati, OH*

- In-sourced customer-facing search, targeting $300K in savings and 40K customers.
- Led MongoDB integration, giving customers access to previously unsearchable documents.

AWARDS

1st Alexa Priza Taskbot Challenge: 3rd Place 2021-2022
"The challenge is focused on developing agents that assist customers in completing tasks requiring multiple steps and decisions. It's the first conversational AI challenge to incorporate multimodal (voice and vision) customer experiences."

- Developed internal dashboard to support bug-fixes in near-real time.
- Designed and implemented automated test suite using fuzzy matching to support dynamic chatbot responses and improve deployment speed and quality.

Hack OHI/O: Awarded 3rd Place 2021
"A 48 hour hackathon that attracts over 800 participants annually for a full weekend of coding, building, learning, networking, and innovation"

- Developed an optimal trick-or-treating route planner using OpenStreetMap data and quadtrees to maximize predicted candy per mile walked.

Hack OHI/O: Awarded Best UI/UX and People's Choice 2020

- Developed a web-based, voice-powered, natural language code editor to convert natural, spoken language into Python code in real time.
- Fine-tuned pretrained transformer models and developed a custom parser to convert natural language to structured code.

Hack OHI/O: Awarded Best Hack 2019

- Developed an accessibility-focused text extraction app for visually impaired users.

· Used Tesseract OCR to extract text from images to provide content in a variety of accessible formats.

Brain Health Hack: Awarded Best Project at Large 2018

"Teams of future scientists, clinicians, engineers and coders compete to create tools that enable better care, more powerful research, or ruther empower patients to live independent and productive lives"

· Developed a Android and iOS Parkinson's Disease tracking app to improve effectiveness of medication.

· Used device accelerometers to measure a user's tremors to improve effectiveness of medication.

Hack OHI/O: Awarded Best Software Hack 2017

· Developed an iOS social media trend aggregator in 24 hours to aggregate trending topics on Twitter and Instagram.

ADDITIONAL INFO

OSU Club Powerlifting August 2021 - May 2024

OSU Club Water Polo August 2017 - May 2024

· Club Risk Manager August 2020 - May 2021

Study abroad in Dresden, Germany June 2019 - August 2019

· Awarded Huntington International Fellowship