Jae-Won **Chung**

+1 (734) 496-1803 | jwnchung@umich.edu | jaewonchung.me | jaywonchung | jae-won-chung-cs

Summary

I am a fifth year PhD candidate in CSE at the University of Michigan, advised by Professor Mosharaf Chowdhury. I build **efficient software systems for deep learning**, with a recent focus on the efficient management of not only time, but also **energy**. My research views energy as a first-class systems resource that is worth carefully optimizing and allocating based on precise measurement and understanding.

I am passionate about **open-source** software and making real-world impact with my research. My open-source works, including the Zeus library, have received wide recognition from academia and industry from for instance Google, PyTorch Foundation, and GitHub. I created and lead the ML.ENERGY initiative as part of my research and open-source efforts, which is now a cross-institutional effort.

Education

University of Michigan   *Ann Arbor, MI, USA*

PH.D. CANDIDATE IN COMPUTER SCIENCE AND ENGINEERING   *Sep 2021 - present*

University of Michigan   *Ann Arbor, MI, USA*

M.S. IN COMPUTER SCIENCE AND ENGINEERING   *Sep 2021 - Apr 2023*

Seoul National University   *Seoul, South Korea*

B.S. IN ELECTRICAL AND COMPUTER ENGINEERING   *Mar 2015 - Aug 2021*

- GPA: 4.04/4.3 (overall) 4.15/4.3 (major), Summa Cum Laude. Period includes two years of military service.

Publications

Peer-reviewed conference publications   * Equal contribution

- **The ML.ENERGY Benchmark: Toward Automated Inference Energy Measurement and Optimization,** Jae-Won Chung, Jeff J. Ma, Ruofan Wu, Jiachen Liu, Oh Jun Kweon, Yuxuan Xia, Zhiyu Wu, Mosharaf Chowdhury, **NeurIPS Datasets & Benchmarks track (spotlight)**, 2025 (Spotlight acceptance rate = 2.81%)

- **Reducing Energy Bloat in Large Model Training,** Jae-Won Chung, Yile Gu, Insu Jang, Luoxi Meng, Nikhil Bansal, Mosharaf Chowdhury, **SOSP**, 2024 (Acceptance rate = 17.34%)

- **Zeus: Understanding and Optimizing GPU Energy Consumption of DNN Training,** Jie You*, Jae-Won Chung*, Mosharaf Chowdhury, **NSDI**, 2023 (Acceptance rate = 18.38%)

- **ShadowTutor: Distributed Partial Distillation for Mobile Video DNN Inference,** Jae-Won Chung, Jae-Yun Kim, Soo-Mook Moon, International Conference on Parallel Processing (**ICPP**), 2020 (Acceptance rate = 28.99%)

Preprints and workshop publications

- **Cornserve: Efficiently Serving Any-to-Any Multimodal Models,** Jeff J. Ma*, Jae-Won Chung*, Akshay Jajoo, Myungjin Lee, Mosharaf Chowdhury, Preprint, 2025

- **Toward Cross-Layer Energy Optimizations in AI Systems,** Jae-Won Chung, Nishil Talati, Mosharaf Chowdhury, **DOE ASCR Energy-Efficient Computing for Science Workshop**, 2024

- **Andes: Defining and Enhancing Quality-of-Experience in LLM-Based Text Streaming Services,** Jiachen Liu, Jae-Won Chung, Zhiyu Wu, Fan Lai, Myungjin Lee, Mosharaf Chowdhury, Preprint, 2024

- **Chasing Low-Carbon Electricity for Practical and Sustainable DNN Training,** Zhenning Yang, Luoxi Meng, Jae-Won Chung, Mosharaf Chowdhury, **ICLR Workshop**: Tackling Climate Change with Machine Learning, 2023

Open Source Projects

- **Zeus** (☆321 ⑂ 40), <u>A framework for deep learning energy measurement and optimization.</u> **PyTorch ecosystem project.** Provides a unified abstraction over various CPU, GPU, and SoC vendors, and an array of time & energy optimizers.
- **Cornserve** (☆80 ⑂ 5), <u>An efficient serving platform for any-to-any multimodal models.</u> Takes a microservice approach to serving complex-structured multimodal input & output models.
- **BERT4Rec-VAE-Pytorch** (☆407 ⑂ 94), <u>A PyTorch framework for recommendation model training</u>, with abstract classes for pluggable model, dataset, and samplers. BERT4Rec and Netflix VAE models implemented.
- **Reason** (☆195 ⑂ 4), <u>A shell for managing research papers, written in Rust.</u> Supports importing papers from file and URL, attaching markdown notes, and creating an HTML book with notes. Uses `serde` to persist data in human-readable and cloud sync-friendly format.
- **Pegasus** (☆32 ⑂ 3), <u>An SSH command runner with a focus on simplicity, written in Rust.</u> Runs multiple commands asynchronously using the `tokio` runtime and streams stdout and stderr back to the user. Battle-tested through multiple research projects and benchmarking.

Number of stars and forks are up-to-date as of December 16th, 2025.

Honors & Awards

Aug 2025	**GitHub Secure Open Source Fund,** $10,000 for the development of the <u>Zeus</u> project	*GitHub*
May 2024	**PyTorch Ecosystem Project,** <u>Zeus</u> was included in the PyTorch Ecosystem	*PyTorch Foundation*
Jan 2024	**Research award,** $20,000 for the development of the <u>ML.ENERGY Initiative</u>	*Salesforce*
Jan 2024	**Mozilla Technology Fund 2024,** $50,000 for the development of the <u>Zeus</u> project	*Mozilla*
Nov 2022	**Carbon Hack '22 Second Best Solution,** <u>Carbon-Aware DNN Training with Zeus</u>, $25,000	*Green Software Foundation*
Jul 2021	**Kwanjeong Overseas Scholarship,** $25,000	*Kwanjeong Educational Foundation*
Mar 2019	**Kwanjeong Undergraduate Scholarship,** $20,000 over two years	*Kwanjeong Educational Foundation*

Talks

Dec 2025	<u>**Energy and Power as First-Class ML Design Metrics (w/ NVIDIA, Google, Meta)**</u>	*NeurIPS 25 Tutorial*
Oct 2025	**Energy as a First-Class ML Design Metric**	*UW–Madison MadSystems Seminar*
Sep 2025	**Power and Energy as First-Class AI Design Metrics**	*KPAI (Bay Area Korean AI Meetup)*
Jun 2025	**Energy as a First-Class Resource in Machine Learning Systems**	*Pruna AI*
May 2025	**Energy-Efficient Systems for Machine Learning**	*<u>Harvard Power and AI Initiative</u>*
Nov 2024	**Energy-Efficient Systems for Machine Learning**	*SOSP 24 Doctoral Workshop*
Apr 2024	**Power and Energy Considerations in Machine Learning Systems**	*University of Michigan (EECS 598)*
Oct 2023	**Energy-Efficient Software Systems for Machine Learning**	*Seoul National University*
Oct 2023	<u>**Energy-Efficient Deep Learning with PyTorch and Zeus**</u>	*PyTorch Conference*
Sep 2023	**Energy-Efficient Deep Learning with Zeus**	*Massachusetts Institute of Technology*

Selected Media Coverage

My research and open-source works have been covered by various media outlets, including MIT Technology Review, ArsTechnica, and Science News.

Jul 2025	<u>**How much energy does your AI prompt use? It depends.**</u>	*Science News*
May 2025	<u>**We did the math on AI's energy footprint. Here's the story you haven't heard.**</u>	*MIT Technology Review*
May 2025	<u>**AI Consumes Lots of Energy. Can It Ever Be Sustainable?**</u>	*The New Stack*
Mar 2025	<u>**Can we make AI less power-hungry? These researchers are working on it.**</u>	*ArsTechnica (front page)*
Nov 2024	<u>**Up to 30% of the power used to train AI is wasted: A software tool could help fix that.**</u>	*Tech Xplore*
Apr 2023	<u>**University of Michigan's 'Zeus' Framework Downsizes AI's Massive Carbon Footprint.**</u>	*HPCWire*
Apr 2023	<u>**Researchers claim they can cut AI training energy demands by 75%.**</u>	*DatacenterDynamics*

Service

- **Systems/Software Reading Group,** Paper reading group inside Michigan CSE, Organizer from September 2022 to December 2025

Teaching

- **CSE585: Systems for Generative AI (UMich, Fall 25),** GSI, three lectures on GenAI and GenAI systems fundamentals.
- **Operating Systems (SNU, Spring 21),** Lead TA, managed Linux kernel hacking projects and led student team design reviews.
- **Computer Architecture (SNU, Fall 20),** Peer tutor, provided 30 hours of online lecture. **Best Tutor Award!**