

David (Dowon) Baek

☎ (+1) 617-583-2851 | ✉ dbaek@mit.edu | 🌐 david-baek | in dbaek-ai

EDUCATION

Massachusetts Institute of Technology (MIT)

Cambridge, MA, USA

M.S. in Electrical Engineering & Computer Science (EECS), GPA: 5.0/5.0

Sep 2023 - Present

- Advisor: Max Tegmark
- Research Area: LLM Interpretability, Representation Learning, AI Safety

Seoul National University (SNU)

Seoul, Korea

B.S. in Physics and Computer Science, Summa Cum Laude, GPA: 4.23/4.3

Mar 2017 – Aug 2023

- Presidential Award (Ranked **1st** among graduating cohort in College of Natural Sciences)
- Includes two years on leave for compulsory military service (2020–21, Job: Cyber Security Specialist)

PUBLICATIONS

1. D. Baek*, Y. Li*, E. Michaud*, J. Engels, X. Sun, M. Tegmark, “The Geometry of Concepts: Sparse Autoencoder Feature Structure,” [\[arXiv\]](#).
2. D. Baek, Y. Li, M. Tegmark, “Generalization from Starvation: Hints of Universality in LLM Knowledge Graph Learning,” [\[arXiv\]](#).
3. D. Baek, Z. Liu, M. Tegmark, “GenEFT: Understanding Statics and Dynamics of Model Generalization via Effective Theory,” *ICLR 2024 Workshop on Bridging the Gap Between Practice and Theory in Deep Learning*, [\[arXiv\]](#).
4. S. H. Park, D. Baek, I. Park, S. Hahn, “Design of Scalable Superconducting Quantum Circuits using Flip-chip Assembly,” *IEEE Transactions on Applied Superconductivity*, 33(5), pp.1-6, 2023 [\[Link\]](#).

EXPERIENCE

Tegmark AI Safety Group

Dec 2023 - Present

Graduate Research Assistant (Advisor: Prof. Max Tegmark)

Cambridge, MA, USA

- Conducted research on the geometric structure of knowledge representations in Large Language Models (LLMs), and investigated hints of representation universality by stitching LLMs of size 1B–8B using PyTorch and Transformers library.
- Analyzed geometric structures reflecting semantic relations between concepts in Sparse Autoencoder (SAE) point cloud.
- Proposed and empirically verified physics-inspired effective theory of neural network generalization, using Python/PyTorch.

Applied Superconductivity Laboratory

Feb 2022 – Feb 2023

Undergraduate Research Assistant (Advisor: Prof. Seungyong Hahn)

Seoul, Korea

- Proposed highly scalable superconducting quantum circuits using flip-chip design and multi-path couplers using Python.
- Studied neural network-based control pulse and geometry optimization strategies for superconducting qubits using Python.

PROJECTS (SELECTED)

Automated System for Effectively Managing Leaves of Soldiers

Sep 2020 - Nov 2020

- Built web application using Node.js, Express, MongoDB, Passport.js for backend and Vue/Vuetify, Chart.js for frontend.
- Interactive user dashboard provides real-time insights into remaining leave balances, leave application status, detailed leave history, a personalized calendar, and visual leave trends through intuitive charts, offering users a seamless experience.

Wanderlust: Community for Hikers

Sep 2019 - Dec 2019

- Built a mobile application that allows users to create, rate, and share hiking routes using Flutter, Google Maps, and Firebase.

HONORS & AWARDS (SELECTED)

- Silver Medal, University Physics Competition, 2018
- Finalist, Samsung Collegiate Programming Cup (SCPC), 2018
- Silver Medal, Korean Mathematical Olympiad (High School Division), 2016
- Silver Medal, International Junior Science Olympiad (IJSO), 2014

TECHNICAL SKILLS

Programming: Python, C/C++, Java, Matlab, Mathematica, L^AT_EX, HTML, Javascript

Libraries: PyTorch, Tensorflow[†], Numpy, Pandas, Matplotlib, Scikit-learn, Scipy, QuTiP, Vue.js/Vuetify, etc.

COMMUNITY SERVICE & TEACHING EXPERIENCE (SELECTED)

- Chair of Publicity & Communications Committee @ Ashdown House (MIT Graduate Housing) Nov 2023 - Present
- Vice President of Publicity @ MIT EECS Graduate Student Association Jan 2024 - Present
- Teaching Assistant, Basic Mathematics and Programming Practice for Machine Learning @ SNU Mar 2023 - Jun 2023