Kento Nishi

kentonishi@college.harvard.edu | GitHub | LinkedIn | Google Scholar | Garmin

Education

Harvard University — Bachelor's & Master's in Computer Science Aug. 2022 – Present

- Concurrent four-year AB/SM Program, w/ Honors. GPA: 3.971. Graduating May 2026.

Skills

Programming: Python, TypeScript/JavaScript, C++, Java, Svelte, OpenGL, Bash, Git, Docker, LaTeX.
Research: PyTorch, CNNs, transformers, diffusion models, mechanistic interpretability, representation learning, explainability, multi-modal learning, data augmentation, LLMs, audio processing, GPU/CUDA.
Misc.: bilingual English/Japanese (professional-level), teaching, distance running, music composition.

Affiliations

Physics of AI (PAI) Group @ Harvard/NTT, w/ H. Tanaka & E. S. Lubana. Mar. 2023 – Present

Visual Computing Group (VCG) @ Harvard, w/ H. Pfister. Aug. 2022 – Present

Four Eyes Lab @ UC Santa Barbara, w/ T. Höllerer. Jun. 2020 – Aug. 2022

Grants

Ezoe Memorial Recruit Foundation Scholarship Apr. 2023 – Present

- $95,000/year, covering tuition and living expenses. Supports BA, MS, and PhD.
- Oldest & most selective scholarship program in Japan (approx. 6 recipients/year).

Employment

Comcast — Applied AI Labs, Speech AI Team

Part-Time Contractor (post-internship extension, remote) Sep. 2025 – Present

- Continuing research to mechanistically probe biases in Text-to-Speech diffusion models.
- Aiming for a tier-1 conference paper submission and a proprietary patent filing.

Grad-Level Summer Research Intern (Washington DC, in-person) Jun. – Aug. 2025

- Worked on production TTS models that serve over 50 million daily users.
- Uncovered and diagnosed train-test distribution gaps in time-domain diffusion models.

Harvard SEAS

CS 79 Course TA Aug. 2025 – Present

- Teaching Assistant for CS 79: Design of Useful and Usable Interactive Systems by K. Gajos.
- Serving as the Studio Planning Lead; sole undergraduate member of course staff.

Distinctions

Tier-1 Conference Reviewer: ICLR 2026 (invitation-only). 2025

Harvard AI Safety Team Program Director: Research Compute Lead at AISST. 2023 – Present

John Harvard Scholar: top 5% GPA in the Harvard College class of 2026. 2023, 2025

Advanced Half-Marathon Runner: 1h25m10s PR (Nov. 2024). 4x race finisher. 2021 – Present

Regeneron STS Top 300 Scholar: the oldest, most prestigious high school STEM award. 2022

Research Themes

Diffusion models, sequential tasks, control & alignment

- Developing a theory on train/test-time gaps in diffusion models w/ time-domain data (paper soon).
- Began in **MATS 2025** (ML Alignment & Theory Scholars Program); $12K grant & $10K for GPUs.
- Extended into mech-interp work at Comcast, debugging production Text-to-Speech diffusion models.
- Advised by H. Tanaka (PAI Group), H. Pfister (Harvard VCG), K. Kumar & R. Tang (Comcast).

Transformers, representation learning, model mechanisms

- Explaining deep learning mechanisms w/ interpretable & controllable tasks on synthetic graphs.
- Published **first-author ICML 2025** paper; co-authored **2x ICLR** and **2x NeurIPS** papers.
- Funded by PRISE 2024 (Harvard Prog. for Research in Science and Eng.); $3K + housing & meals.
- Funded by CBS-NTT Fellowship (Physics of Intelligence Program); $5K summer 2023 stipend.
- Advised by H. Tanaka & E. S. Lubana (PAI Group).

Computer vision, semi-supervised learning, model robustness

- Designing compute-efficient ways to train accurate vision models under partial/noisy supervision.
- Published **first-author CVPR 2024** paper as a sophomore undergrad.
- At age 16, published **first-author CVPR 2021** paper & **first-author AAAI 2021** short paper.
- Won Regeneron STS Top 300 Award & gave invited talk at Forum on Information Technology in 2022.
- Advised by H. Pfister (Harvard VCG) since 2022; prev. T. Höllerer (UCSB Four Eyes Lab).

Technical Projects

LiveTL Apps Nov. 2020 – Present

- Lead dev of three modular addons that improve YouTube and Twitch (LiveTL, HyperChat, YtcFilter).
- 100K+ total users; 900+ repo stars; 20+ code contributors. Free, open-source, and cross-platform.

holoEN Christmas Advent Calendar Nov./Dec., 2022 – Present

- Full-stack dev of the event platform (holoen-advent.com); officially commissioned by Cover Corp.
- 250K+ total users; a beloved yearly tradition in the hololive English community.

Exio UI Elements — npm May 2024

- A customizable, framework-agnostic web UI library used across my open-source websites and apps.

Torch Pitch Shift — PyPI Jun. 2021

- The first Python library for pitch-shifting on GPU at the time; later added to PyTorch upstream.
- 750K+ downloads/month; 135+ GitHub stars; used by torch-audiomentations (1.1K+ stars).

All Publications

Representation Shattering in Transformers: A Synthetic Study with Knowledge Editing

ICML 2025, as **first author**.
Our interpretable task reveals why edits to model weights can destroy representation geometries.

In-Context Learning of Representations

ICLR 2025, as co-author.
We explain how LLMs reorganize representations in-context to align with task-specific structures.

Structured In-Context Task Representations

NeurIPS 2024 NeurReps Workshop, as co-author.
A precursor to "In-Context Learning of Representations."

Stepwise Inference in Transformers: Exploring a Synthetic Graph Navigation Task

NeurIPS 2023 R0-FoMo Workshop, as co-author.
Our synthetic DAG navigation task clarifies when chain-of-thought might help autoregressive models.

Joint-Task Regularization for Partially Labeled Multi-Task Learning

CVPR 2024, as **first author**.
We propose a regularizer to train models using partially labeled data with linear complexity.

Augmentation Strategies for Learning with Noisy Labels

CVPR 2021, as **first author**.
Our decoupled augmentation strategy improves model robustness under noisy-label scenarios.

Improving Label Noise Robustness with Data Augmentation and Semi-Supervised Learning

AAAI 2021 Student Abstract Track, as **first author**.
A precursor to "Augmentation Strategies for Learning with Noisy Labels."

References

Available upon request.