Kai Yao

AI Researcher & ML Systems Engineer

PhD Viva Passed (Oct 2026), University of Edinburgh • Conferral Jan 2027

Applied ML • Model Optimization & Deployment • Trustworthy AI

kai.yao@ed.ac.uk | kaiyao17@gmail.com

+86 17302151562 | +44 7867 763608

kaikaiyao.github.io | github.com/kaikaiyao

LinkedIn | Google Scholar

PROFILE

AI researcher and former Intel/Huawei engineer combining three years of production ML systems and partner-facing deployment with a University of Edinburgh PhD (viva passed Oct 2026; seven first-author AI papers, including two at NeurIPS 2026). Track record delivering open-source developer tools, automated inference optimization across PyTorch, Hugging Face, Alibaba Cloud, and AWS, and rigorous evaluation systems for deployed AI models.

INDUSTRY EXPERIENCE

Intel Corp. | *AI Frameworks Engineer* — Shanghai, China • Jun 2021 – Apr 2023

- **Neural Coder Technical Lead:** Architected and led Neural Coder inside open-source Intel Neural Compressor (40 upstream commits), using Python static analysis (syntax parsing, type inference, call-graph analysis) to insert and benchmark PyTorch optimizations automatically; featured live in CEO Pat Gelsinger's *Intel Innovation 2022* keynote (10×+ ResNet50 inference speedup on 4th Gen Xeon, accuracy maintained).
- **Hugging Face & Developer Tooling:** Delivered Optimum Intel INT8 quantization across 18 Hugging Face Transformers PyTorch tasks (∼1.7× inference speedup on 3rd Gen Xeon, typically <1% accuracy drop), plus JupyterLab and VS Code extensions emitting inspectable code patches.
- **Cloud Partner & Customer Deployment:** Partnered directly with **Alibaba Cloud** (integrating Neural Coder and BladeDISC into PAI-DSW notebooks) and **AWS** to run cross-platform benchmarks, isolate performance bottlenecks, and deliver actionable deployment guidance.
- **Generative AI Optimization & Recognition:** Built PyTorch adapters and INT8 inference workflows across Intel Neural Compressor and Intel Extension for PyTorch, including CPU acceleration for Stable Diffusion; authored four Intel Analytics Software technical articles and earned three Intel division awards.

Huawei Technologies Co., Ltd. | *AI Algorithm Engineer* — Shanghai, China • Oct 2020 – Jun 2021

- Designed and tuned CNN and RNN models for 5G MU-MIMO wireless systems under production latency and hardware constraints.
- Applied architecture optimization and weight compression to improve training and inference efficiency; collaborated with deployment and validation teams to integrate algorithms into 5G features and verify end-to-end performance.

DOCTORAL RESEARCH & APPLIED ML ENGINEERING

University of Edinburgh | *PhD Researcher, School of Informatics (Advisor: Dr. Marc Juarez)* — Edinburgh, UK • May 2023 – Present

- **Post-Deployment Model Auditing (FARE, NeurIPS 2026):** Designed a black-box, output-only verification system that detects silent model, checkpoint, or inference-pipeline substitutions after certification from generated outputs alone. Achieved **92.6%–99.5% mean TPR at 1% FPR** across four swap scenarios (92.6% vs. 18.3% best baseline on Stable Diffusion version swaps) and retained **92.4%–99.4% TPR** under white-box PGD attacks.
- **Provenance Bounds & API Security (NeurIPS 2026):** Derived minimax total-variation limits on passive image provenance under editing and demonstrated how public verification interfaces (scores, labels, binary outputs, abstentions) leak signal enabling surrogate evasion—yielding concrete guidelines for deployed verification APIs.
- **Reproducible Robustness Benchmarking (IEEE SaTML 2026 Oral; ACM AISec 2026):** Engineered end-to-end PyTorch evaluation pipelines on Edinburgh's Kubernetes/NVIDIA GPU cluster covering 14 fingerprinting methods across 12 GAN/VAE/diffusion generators and 5 attack strategies (*Smudged Fingerprints*, featured in University of Edinburgh News, DIGIT, and Herald Scotland), plus a 55-transformation benchmark for local invisible image watermarks.
- **Keyed Attribution, Fairness & LLM Auditing (Preprint; SaTML 2025 Oral; AIES 2026):** Built *SPRINT* for keyed passive attribution (98.83% clean accuracy; <1% adaptive attack success with private keys); synthesized 20 differential-privacy fairness studies into a diagnostic taxonomy (*SaTML '25*); audited 30 university AI policies with an automated rubric across four open-weight LLMs (*AIES '26*).

EDUCATION

University of Edinburgh | PhD in Cyber Security, Privacy and Trust (Full PhD Scholarship) May 2023 – Jan 2027 (expected)

PhD viva passed Oct 2026; formal degree conferral expected Jan 2027. School of Informatics, Edinburgh, UK.

Johns Hopkins University | MS in Mechanical Engineering (Departmental Research Fellowship) Sep 2017 – May 2020

Focus: applied deep learning and computer vision for quantitative microscopy and biomedical imaging. Baltimore, MD, USA.

Fudan University | BS in Theoretical and Applied Mechanics (Outstanding Graduate of the Year) Sep 2012 – Jun 2017

Includes one-year exchange at **Osaka University** (2014–2015, Japanese Government JASSO Full Scholarship).

EARLIER APPLIED COMPUTER VISION & SCIENTIFIC ML

Johns Hopkins University | *Graduate Researcher* Baltimore, MD, USA • 2018 – 2020

- Built end-to-end microscopy preprocessing, U-Net/CNN training, and experimental validation pipelines for single-cell classification, unsupervised phenotyping, and dynamic volume prediction (*CTRL*, 5–6% mean error vs. physical measurements).
- Co-first author on four peer-reviewed papers in *Journal of Cell Science* (2020), *Journal of Cell Biology* (2019), *Scientific Reports* (2019), and *Bio-protocol* (2020); co-authored computational stent-deployment validation in *BioMedical Engineering OnLine* (2016).

SELECTED PUBLICATIONS & OPEN-SOURCE WORK

NeurIPS 2026	**Yao K**, Juarez M. *FARE: Forensic Acceptance Region Estimation for Catching Bait-and-Switch Image Generators.* [arXiv] [Code]
NeurIPS 2026	**Yao K**. *Can Pixels Alone Reveal Image Origin? Minimax Limits and Learnable Interfaces for Passive Provenance.* [arXiv] [Code]
IEEE SaTML 2026	**Yao K**, Juarez M. *Smudged Fingerprints: A Systematic Evaluation of the Robustness of AI Image Fingerprints* (Oral). [arXiv] [Code]
ACM AISec 2026	**Yao K**[*], Szilágyi B[*], et al. *What Breaks Local Watermarks? A Robustness Benchmark for Local Invisible Image Watermarking.* [arXiv]
IEEE SaTML 2025	**Yao K**, Juarez M. *SoK: What Makes Private Learning Unfair?* (Oral). [IEEE] [arXiv]
AAAI/ACM AIES 2026	**Yao K**. Two papers on AI governance and capability assessment in AI-mediated education. [Paper 1] [Paper 2]
Preprint (2026)	**Yao K**, Juarez M. *SPRINT: Robust Model Attribution of Generated Images via Secret Pixel Reconstruction.* [arXiv]
Open Source (Intel)	Neural Coder & PyTorch quantization components in *Intel Neural Compressor*. [Commits] [Technical Blogs]

[*] *Equal contribution / co-first authorship.*

TECHNICAL STACK & LANGUAGES

Languages	Python (primary), C++, MATLAB, Java, SQL
ML & Vision	PyTorch, TensorFlow, Hugging Face Transformers, NumPy, Pandas, CNNs, U-Net, Diffusion, GANs, VAEs
Systems & Deployment	INT8 quantization, model compression, inference optimization, static program analysis, performance benchmarking, Intel Neural Compressor, IPEX, Linux, Git, Docker, Kubernetes
Evaluation & Security	Threat modelling, adversarial robustness testing, black-box model auditing, watermarking & fingerprinting, fixed-FPR statistical calibration, AUROC/TPR/FPR analysis, differential privacy & fairness
Spoken Languages	Mandarin Chinese (native), English (full professional proficiency; US MS & UK PhD), Japanese (JLPT N1)

TALKS, STAKEHOLDER COMMUNICATION & SERVICE

- **External Keynote & Partner Communication:** Delivered invited keynote on black-box model fingerprinting and IP protection at *Kraus & Lederer Law Firm* (Munich, 2026); presented oral talks at IEEE SaTML 2025 & 2026; led customer-facing technical integration with engineering teams at Alibaba Cloud, AWS, and Hugging Face.
- **Teaching & Mentoring:** Graduate Guest Lecturer (*Image Provenance in the AI Era*) and TA at the University of Edinburgh (2023–2026); TA at Johns Hopkins University (2019–2020); mentored five undergraduate and postgraduate researchers.
- **Program Committee & Reviewing:** Program Committee for IEEE SaTML 2027 and ACM AISec 2026; Reviewer for AAAI (2026, 2027), PETS/PoPETs (2027), and ACM DTRAP.