

Kohei Ozaki

Tokyo, Japan | eowner[at]gmail.com | ho.lc | linkedin.com/in/koheiozaki | github.com/smly

Summary

- ML Engineer and Tech Lead with 10+ years of experience in research-driven and production-grade machine learning.
- 4x Kaggle Grandmaster (20+ gold medals), known for solving complex, noisy, real-world ML problems.
- Led end-to-end AI initiatives from R&D to deployment and technical leadership across multiple industries.
- Published at CoNLL and CVPR workshops and contributed datasets and tools used by the global AI community.

Education

Nara Institute of Science and Technology , Ph.D in Computer Science	Apr 2011 – Mar 2014
• Computational Linguistics Laboratory (Supervisor: Prof. Yuji Matsumoto)	
• Completed coursework without degree.	
Nara Institute of Science and Technology , MS in Computer Science	Apr 2009 – Mar 2011
• Computational Linguistics Laboratory (Supervisor: Prof. Yuji Matsumoto)	
• Thesis: "Mutual k-Nearest Neighbor Graphs for Semi-Supervised Learning" (NAIST-IS-MT0951027)	
Tokyo University of Science , BS in Computer Science	Apr 2005 – Mar 2009
• Artificial Intelligence Laboratory (Supervisor: Prof. Ikuo Tahara)	
• Thesis: "Breaking CAPTCHAs with Directed Spanning Tree"	

Experience

AI Fellow & Senior Manager , Rist – Tokyo, Japan	Mar 2023 – Present
• Contributed to the full project lifecycle in various capacities, from initial proposal and development to serving as Project Manager and Tech Lead.	
• Led pre-sales engagements, from initial client hearings and technical surveys to proposal creation and demo development.	
Self-employed , Freelance	Dec 2017 – Oct 2023
• Advisory service and software development service.	
• Worked as a committee member of AI Edge Contest, organized by METI (Ministry of Economy, Trade and Industry) from 2017 to 2023.	
Machine Learning Engineer , Ubie – Tokyo, Japan	Mar 2022 – Feb 2023
• Designed and deployed machine learning solutions.	
• Maintained production ML systems and operational stability.	
Software Engineer , Preferred Networks – Tokyo, Japan	Jul 2019 – Feb 2022
• Contributed to the development of Optuna (a hyperparameter optimization framework) as a committer, primarily focusing on integration modules and the LightGBM Tuner.	
• Delivered data science solutions for projects in cancer screening, plant control, and retail services.	
Senior Software Engineer , Recruit Technologies – Tokyo, Japan	May 2017 - Jun 2019
• R&D on image retrieval and 3D reconstruction.	
• Developed voice recognition features for an educational application.	
Principal Software Engineer , Magne-Max Capital Management – Osaka, Japan	Oct 2016 – Apr 2017
• Development of mutual funds based on behavioral economics.	

Senior Software Engineer , Recruit Technologies – Tokyo, Japan	Nov 2015 – Sep 2016
<ul style="list-style-type: none"> • R&D on image retrieval and 3D reconstruction. • Technical advice and support for new business projects. 	
Software Engineer , FreakOut – Tokyo, Japan	Jan 2014 – Feb 2015
<ul style="list-style-type: none"> • Developed ML-based algorithms for real-time bidding in target advertising market. 	
Software Engineer , mixi – Tokyo, Japan	May 2012 – Dec 2013
<ul style="list-style-type: none"> • Research and development for recommendation on social network service. • Developed a real-time bidding system for online advertising. 	

Projects

RiichiLab – Mahjong AI Competition (github.com/smly/mjai.app)	2022 - Present
<ul style="list-style-type: none"> • Developed a Mahjong AI and a game simulator for the imperfect information game, along with a competition platform to pit AIs against each other. • Tools Used: Rust, Python, TypeScript, PyTorch, React, Docker, Google Cloud Platform 	

Kaggle Competitions (kaggle.com/confirm)	2011 - Present
<ul style="list-style-type: none"> • 4x Kaggle Grandmaster. Over 20 competition gold medals. 	

Kaggle Community Activities

<ul style="list-style-type: none"> • Authored a book about Kaggle. Organized 9 meetups. Gave 10+ talks at conferences, companies and meetups. • Google Developer Expert for Kaggle since 2025.
--

Publications

(CVPR'20 WS) Shuhei Yokoo, *Kohei Ozaki*, Edgar Simo-Serra, Satoshi Iizuka, **”Two-stage discriminative re-ranking for large scale landmark retrieval”**, Proc of the CVPR'20 workshop.

(CVPR'19 WS) *Kohei Ozaki*, Shuhei Yokoo, **”Large-scale landmark retrieval and recognition under a noisy and diverse dataset”**, Technical report for CVPR'19 workshop.

• Developed GLDv2-clean, the largest high-quality public dataset for image retrieval tasks that was later used as a seed dataset for DINOv2. github.com/cvdfoundation/google-landmark.

(CVPR'19 WS) Yusuke Niitani, Toru Ogawa, Shuji Suzuki, Takuya Akiba, Tommi Kerola, *Kohei Ozaki*, Shotaro Sano, **”Team PFDet’s methods for Open Images Challenge 2019”**, Technical report for CVPR'19 workshop.

(CoNLL'11) *Kohei Ozaki*, Masashi Shimbo, Mamoru Komachi, Yuji Matsumoto, **”Using the mutual k-nearest neighbor graphs for semi-supervised classification of natural language data”**, Proc of the CoNLL'11.

Awards (Selected)

4th place - (CVPR'23 WS) Image Matching Challenge 2023, 2023.

1st place - (CVPR'21 WS) Hotel-ID to Combat Human Trafficking, 2021.

1st place - (CVPR'19 WS) Google Landmark Recognition Challenge, 2019.

3rd place - (CVPR'19 WS) Google Landmark Retrieval Challenge, 2019.

1st place - (KDD'15 WS) ACM KDD Cup'15: Massive Open Online Courses Dropout Prediction, 2015.

IPSJ Yamashita SIG Memorial Research Award, 2012.

Lecture

”Data Science Exercises”, Kumamoto University, 2025.

- 90min lecture titled ”Introduction to LLMs” & 540min exercise on data science.

”Advanced Lectures on Information Science”, Tokyo Metropolitan University, 2019.

”Big Data Analysis”, Osaka University, 2018.