Work History

Basic Information

Name: Sou Yoshihara (　)

Service	Account
GitHub	@sousquared
Instagram	@sousquared
Twitter	@sou_squared
Zenn	@sousquared

Skills

Languages

- Japanese
  - Native
- English
  - Conversational
  - TOEIC 875 (Apr 2019)
  - TOEFL 85 (Jun 2017)

Programming Languages

- Python
- JavaScript/TypeScript
- React
- Terraform

Cloud/Infrastructure

- GCP (Google Cloud Platform)
  - Cloud Run
  - Cloud Build
  - GKE (Google Kubernetes Engine)
  - etc.

Development Experience

- PoC and production implementation using machine learning models
- API development (FastAPI)
- Clean Architecture and DDD (Domain-Driven Design) oriented coding

Others

- KaiRA (Kyoto University AI Research Association) Operating Member
  - Organizing reading sessions on artificial intelligence
- JEES SoftBank AI Talent Development Scholarship (2020-2021)

Strengths

- Proactively taking action based on my interests
- Strong presentation and explanation skills
- International network of friends around the world

Interests

- Social applications of AI technology
- Mechanisms of human cognitive abilities (neuroscience, psychology, behavioral economics)
- AI-powered art and entertainment

Work Experience

Career Summary

As an engineer, I have not only developed features but also considered which PoCs and new features would become competitive advantages for the service and contribute to revenue growth, working backwards from business and management strategies. Through meetings with business leaders and discussions with engineering managers, I have built medium- to long-term strategies and directed and managed specific execution initiatives through PDCA cycles. As a leader of an ML team of 4-5 members, I led strategic planning and research roadmaps for updating internal product search functionality and new features, achieving quantitative results such as doubling search feature usage rates. In the Auto Generation team, I participated in 4 auto-generation logic projects, achieving 90%+ quality in banner ad image auto-generation within 2-3 months. I built high-speed experiment pipelines and annotation-based improvement workflows, balancing development speed with quality improvement.

Aug 2018 - Aug 2020

Organization: Hitachi, Ltd. / Kyoto University Lab
Position: Student Researcher (Part-time)
Responsibilities

- Research and implementation of the latest machine learning papers
- Research and development of algorithms adaptable to modeling geospatial information
- Literature surveys on Graph Convolutional Networks and Relational Graph Convolutional Networks
- Implementation of parts of models devised by researchers
- Algorithm evaluation

Jan 2021

Organization: CyberAgent / Kiwami Prediction AI Division / Prediction Team
Position: ML/DS Intern
Responsibilities: Responsible for updating and accuracy evaluation of video ad score prediction models.

Apr 2022 ~ Mar 2025

Organization: CyberAgent / Kiwami AI Division / New Business Development Team
Position: ML Engineer
Responsibilities

Strategic Planning & Direction

- Led strategic planning and research roadmaps for updating internal product search functionality and new features
- Built medium- to long-term strategies through meetings with business leaders and engineering managers, directing and managing specific execution initiatives through PDCA cycles

Search Feature Improvement

- Formulated hypotheses from user usage history and improved image search quality by adding semantic search, **doubling search feature usage rates**
- Experimented and implemented logic to filter out poor-quality search results, **filtering out 40% of poor-quality search results**

ML Ops Optimization

- Conducted load testing for GPU inference batch processing using GKE and optimized resources. **Achieved 6.28x faster inference speed compared to CPU**
- Reduced processing time by **62%** through preprocessing acceleration (including cache utilization)

Cost Optimization

- Cleaned up unused GKE and VMs, **contributing to approximately $2,500 monthly cost reduction**

Organizational Development

- As a leader of an ML team of 4-5 members, promoted PoC and production development of new features utilizing AI
- As operating leader of DSOps training, planned new initiatives and built a system where new projects emerge from training (see article for details. It is written in Japanese.)

Apr 2025 ~ Present

Organization: CyberAgent / Kiwami AI Division / Auto Generation Team
Position: ML Engineer
Responsibilities

Auto-Generation Logic Development

- Participated in 4 auto-generation logic projects as a machine learning engineer
- Collaborated with 1-2 engineers and 1-2 PMs, promoting speed-focused development

Project Results

- In the 1st project, validated 2 modules and **achieved 74% success rate quality in auto-generation within 2 months**
- In the 3rd project, worked with 2 engineers on banner ad image auto-generation and **successfully developed auto-generation logic with 90% success rate quality within 3 months**
- In the 4th project, worked with 2 engineers on banner ad image auto-generation and **successfully developed auto-generation logic with 90%+ quality within 2 months**

Development Process Improvement

- Built experiment pipelines for high-speed experiments
- Designed workflows to quickly reflect annotation results into improvement logic
- Built systems for rapid verbalization of good/bad feedback from creators
- Established processes that balance development speed with quality improvement

Organizational Development

- As operating leader of DSOps training, realized a special lecture by Yuta Saito (Ph.D., Cornell University, Hanjuku Virtual Co., Ltd.)

Research

- Sou Yoshihara, Taiki Fukiage, Shin'ya Nishida, "Towards acquisition of shape bias: Training convolutional neural networks with blurred images.", VSS, Poster session, 2021.
- Sou Yoshihara, Taiki Fukiage, Shin'ya Nishida, "Does training with blurred images bring convolutional neural networks closer to humans with respect to robust object recognition and internal representations?", Front. Psychol., Vol. 14,2023
- (Japanese) Sou Yoshihara, Taiki Fukiage, Shin'ya Nishida, "Shape Bias ", VISION, Vol. 33, No.1, 1-5, 2021. Best Presentation Award, Vision Society of Japan 2020 Summer Conference code

Articles

- ” ” —— 7 DSOps (Japanese)
- Codex MCP AI Coding : Codex 3 (Japanese)
- I also write mainly technical articles on Zenn: Zenn:@sousquared (Japanese)