

# Yuki Okuda

## Machine Learning Engineer

Sapporo, Japan | yanagi.ayase@gmail.com | GitHub: yagays | LinkedIn: yukiokuda | Web: <https://yag.xyz>

### Summary

Machine Learning Engineer with 10+ years of experience in ML, NLP, and data analytics. Proven track record in healthcare AI, natural language processing, and building data analytics platforms. Currently focused on LLM application development including RAG systems and AI agents.

### Experience

#### NLPeanuts — ML Engineer Nov 2025 – Present

- Developing RAG (Retrieval-Augmented Generation) systems for enterprise applications
- Building LLM-based conversational AI agents

#### Ubie, Inc. — ML Engineer / BI Engineer Jul 2020 – Sep 2025

##### BI Engineer (2022–2025)

- Built and maintained company-wide data analytics platform using dbt and Lightdash
- Designed and implemented A/B testing infrastructure to support data-driven product decisions
- Conducted funnel analysis and user behavior analytics to drive toC platform improvements

##### ML Engineer (2020–2024)

- Developed OCR and information extraction system for medication record books (Okusuri Techo)
- Built automated medical record generation system from speech recognition during medical interviews
- Applied ML techniques to develop a symptom checker for healthcare applications

#### Sansan, Inc. — Researcher, DSOC Jan 2018 – Jun 2020

- Developed NLP-based business card data processing systems
- Built NLP pipeline for news delivery system: language detection, company name extraction, and genre classification
- Developed similar person search by applying Graph Embedding to business card exchange networks

#### Recruit Technologies Co., Ltd. — Data Scientist Apr 2015 – Dec 2017

- Developed recommendation and behavior prediction models for Rikunabi (major job recruiting platform)
- Served as product owner and public spokesperson for “A3RT” — a suite of public ML/AI APIs
- Planned and developed an automated content generation system using seq2seq models

### Education

#### Nara Institute of Science and Technology (NAIST) — M.Eng., Information Science 2013 – 2015

### Skills

**Programming:** Python, React, SQL  
**ML/AI:** NLP, LLM, RAG, Deep Learning (PyTorch), scikit-learn  
**Data Engineering:** dbt, BigQuery, Looker Studio, Lightdash, Redash  
**Tools:** Git, Docker, AWS, GCP  
**Languages:** Japanese (native), English (TOEIC 875)

### Talks & Publications

- **PyCon JP 2021:** “Introduction to Document Intelligence with Python”  
[https://speakerdeck.com/yag\\_ays/introduction-to-document-intelligence-with-python](https://speakerdeck.com/yag_ays/introduction-to-document-intelligence-with-python)
- **NLP 2020:** “Corporate Keyword Extraction from News Articles”  
[https://www.anlp.jp/proceedings/annual\\_meeting/2020/pdf\\_dir/P2-13.pdf](https://www.anlp.jp/proceedings/annual_meeting/2020/pdf_dir/P2-13.pdf)
- **IBIS 2018:** “Japanese Name Segmentation Using Bidirectional LSTM”  
<https://buildersbox.corp-sansan.com/entry/2018/11/16/113000>