

# Inseong Han

[Email](#) | [LinkedIn](#) | [GitHub](#) | Palo Alto, Stanford, CA

Data Scientist with 4+ years of experience in AI applications, combining expertise in MLOps and DevOps for robust AI systems

## Professional Experience

**Senior Data Scientist 1** | LexisNexis, San Francisco | Aug 2024 – Now  
**Data Scientist, Intern** | LexisNexis, San Francisco | Oct 2023 – June 2024

- Worked with Scala-based data pipeline handling large-scale data processing of 2.4 million entities across 11 million documents (50GB+) for search functionality development.
- Pioneered the implementation of TAG (Table Augmented Generation) by integrating SQL databases with Large Language Models, extending LLM capabilities to leverage structured database information for more accurate and contextual responses.
- Collaborated with multiple teams (30+ members) across different functions, effectively communicating technical implementations and system architecture.

**Machine Learning Engineer** | Rapport Labs, Seoul | Sep 2021 – May 2022

- Built recommender system for fashion retail (300k users, 100k items) with implicit feedback
- Automated item rank annotation, saving significant daily manual work for domain experts

**Computer Vision Engineer** | Bepro Company, Seoul | Aug 2019 – Sep 2021

- Reduced distortions in stitched images, significantly improving product satisfaction
- Implemented frame-level video synchronization module for multiple perspectives
- Optimized cloud compute costs from \$1500 to \$500 monthly
- Enhanced video stitching pipeline to 30 FPS for 4K videos (+20 FPS throughput gain)

## Personal Projects

**AI Mock Interviewer (Google Gemini Competition Entry)** | May 2024 – July 2024  
Tech Stack: Next.js, Google Gemini 1.5 Flash, Deepgram, ElevenLabs

- Architected an end-to-end AI interview system with 2-4 second response latency, enabling natural conversational flow during behavioral interviews
- Developed a comprehensive feedback system that provides timestamped audio snippets alongside specific recommendations, enabling targeted improvement
- Integrated Silero VAD (Voice Activity Detection) model to automatically determine start and end of user responses, creating more natural conversation dynamics

## Education

**Master of Science in Data Science** | University of San Francisco | July 2023 – July 2024  
**Bachelor of Science in Computer Engineering** | Seoul National University | Mar 2014 – Aug 2020

## Technical Skills

- **Machine Learning & AI:** LLM API, LangChain, Hugging Face, PyTorch
- **Programming Languages:** Python, C++
- **Data & Databases:** PostgreSQL, MongoDB, Pandas, NumPy, Matplotlib, Spark
- **Cloud & Infrastructure:** Docker, Apache Airflow, AWS, Google Cloud
- **Development Tools:** Git, VScode, Postman, Obsidian
- **Software Engineering:** System Design, Object-Oriented Programming, Design Patterns
- **Development Practices:** Agile Methodology, MLOps, DevOps