Jenny Jiwon Kim

jiwonk7@illinois.edu jennyjiwonk.github.io jenny-jiwon-kim

Interests

Human-Centered AI for Mental Health and Social Wellbeing

LLM Alignment, LLM-as-a-Judge, Agentic AI, Human–AI Interaction, AI Safety, AI-Mediated Support and Writing, Model Post-Training, Machine Teaching, Preference Optimization, Social Computing, Natural Language Processing

Education

University of Illinois Urbana-Champaign *Aug 2025 – Present*
Ph.D. in Computer Science, Advisor: Prof. Koustuv Saha *Urbana, IL, USA*

Seoul National University *Mar 2019 – Aug 2025*
B.S. in Computer Science and Engineering & Bachelor of Business Administration *Seoul, Korea*

Selected Publications

LLUMI: Improving LLM Writing Assistance for Mental Health Support with Online Community Feedback
Jiwon Kim, M. Ajit, S. Gong, et al. *EMNLP 2026 (Main Oral)*

AI Psychosis: Does Conversational AI Amplify Delusion-Related Language?
S. R. Shimgekar, V. Gunda, **Jiwon Kim**, et al. *EMNLP 2026 (Main)*

AI Content Moderation in Therapy Conversations
Jiwon Kim, Claire Wang, Taeung Yoon, et al. *ICWSM Digital Minds 2026*

PAIR-SAFE: A Paired-Agent Approach for Runtime Auditing and Refining AI-Mediated Mental Health Support
Jiwon Kim, V. J. Rodriguez, D. W. Yoo, et al. *Under review (ACL ARR)*

PeerPen: AI-Assisted Writing for Online Mental Health Peer Support
Jiwon Kim, S. Gong, M. Ajit, et al. *Under review (CHI)*

Work Experience

OnCARE Lab *Aug 2025 – Present*
Graduate Research Assistant, UIUC *Urbana, IL*

- **LLM Alignment and Safety Evaluation**: Fine-tuned open-source LLMs on expert-annotated counseling data and developed an LLM-as-a-Judge framework for automated evaluation, safety guardrails, and runtime refinement.
- **AI-Mediated Writing for Online Communities**: Developed an AI writing assistant for peer-support communities; applied Direct Preference Optimization to align model with community preferences using large-scale Reddit data.

Dalpha AI *Dec 2024 – May 2025*
AI Researcher, Core AI Team *Seoul, Korea*

- **Multimodal AI Search Engine**: Engineered an end-to-end AI search pipeline integrating text and image embeddings, multimodal retrieval, semantic search, and ranking optimization, achieving a **23% increase in CTR**.
- Developed logo detection tools using methods in computer vision, object detection, and automated visual analytics.

Human-Oriented Language Intelligence Lab *Mar 2024 – Dec 2024*
Research Intern, Seoul National University *Seoul, Korea*

- **Personalized LLMs**: Developed a value-aligned LLM chatbot for personalized persuasion using preference modeling, retrieval-augmented generation, and contrastive learning, achieving a **48% reduction in perplexity**.

Real-Time Ubiquitous Systems Lab *Jul 2023 – Jan 2024*
Research Intern, Seoul National University *Seoul, Korea*

- **Digital Phenotyping**: Engineered a mobile sensing application for adolescent mental health risk prediction using multimodal sensor data, **increasing 58% in recall**; deployed in a 3-month field study with **300+ participants**.

Cclim, Social Venture *Jan 2021 – Dec 2021*
Data Analytics Manager *Seoul, Korea*

- **AdTech Engineering and Ad Optimization**: Built AI video analytics tool using vision APIs for audience measurement and foot-traffic estimation, securing **12 advertising contracts** and coverage in **25 news articles**.
- Ran Google Ads campaigns through data-driven audience targeting and ad optimization, **increasing CTR by 14%**.

Programming Experience

Large-Scale Data Analysis with NAVER *Oct 2024 – Dec 2024*

- Analyzed large-scale search and e-commerce log data using SQL across a platform processing **200M+ queries daily**, identifying user and seller behavior patterns, multi-homing dynamics, and data-driven growth opportunities.

Upstage AI Full Stack LLM Project *Apr 2024 – Jun 2024*

- Led a software development team to build and deploy a personalized AI consultant powered by Solar LLMs, integrating Retrieval-Augmented Generation (RAG) and user profiling for adaptive, context-aware interactions; launched an experimental web application with **200+ participants** over 3 days.

Drug-Likeness Evaluation Project with Aigen Drug *Jul 2023 – Dec 2023*

- Developed a novel drug-likeness metric using chemical embedding spaces and Graph Transformers, achieving **86% accuracy** in molecular property prediction; optimized MolCLR inference through model optimization and Docker-based deployment, reducing latency from **3s to 200ms**.

Cardiovascular Disease Detection Competition *Apr 2021 – Jun 2021*

- Developed cardiovascular risk prediction models using empirical mode decomposition (EMD) and adaptive thresholding for peak detection; applied Fourier transforms and entropy-based feature extraction for analysis.

Samsung DS2 Competition *Apr 2021 – Jun 2021*

- Analyzed dynamic sensor data using time-series pattern mining, graph-based signal processing, and graph integrals to extract spatiotemporal patterns and improve data-driven insights.

Awards

Upstage AI Innovation *2024*
Presented a creative AI chat service using Solar LLM, selected as a top 10 finalist among 200+ teams

1st place, Korea National Competition of Enactus *2021*
Won first place competing against participants from over 25 universities across Korea

International Social Entrepreneurship Project Presentation *2021*
Presented on the impact of a social venture, representing Korea among 72,000+ participants from 36 countries

Teaching

CS 107 Data Science Discovery *Aug 2026 - Present*
Teaching Assistant

CS 416 Data Visualization *Jun 2026 - Aug 2026*
Teaching Assistant

Technical Skills

Programming Languages: Python, C/C++, React, Java, Javascript, Typescript, SQL, HTML/CSS

Frameworks: PyTorch, TensorFlow, Hugging Face, Pandas, Matplotlib, Scikit-learn, BokehJS, ReactJS, Kotlin

Language: Bilingual Proficiency in English (**IBT TOEFL: 117/120**)