Chanhwi Kim

Info: chanwhistle.github.io

Email: chanhwi.kim@uth.tmc.edu

RESEARCH INTERESTS

Clinical/Medical NLP, Scientific AI, Multimodality, Large Language Models

EDUCATION

University of Texas Health Science Center at Houston — Houston, Texas, USA
M.S. - School of Biomedical Informatics — [Aug 2025 -]

Korea University — Seoul, Republic of Korea
M.S. - Department of Computer Science & Engineering; GPA: 4.0/4.0 — [Mar 2023 - Feb 2025]

University of Lausanne — Lausanne, Switzerland
Exchange Student; School of Biology — [Feb 2018 - Jul 2018]

Korea University — Seoul, Republic of Korea
B.S. - Division of Biotechnology; GPA: 3.6/4.0 — [Mar 2016 - Feb 2022]

ACADEMIC EXPERIENCE

Boston Children's Hospital (Harvard Medical School) — Boston, Massachusetts, USA
Visiting Scholar (Advisor: Dr. Timothy A. Miller) — [Oct 2024 – Apr 2025]

- Led a project on predicting 30-day post-discharge mortality in ICU patients CXR images and clinical text
- Contributed to a project on predicting psychiatric readmission using aspect-oriented discharge note summarization
- Funded by Korea University - KIAT Scholarship Program

Korea University — Seoul, Republic of Korea
Research Intern (Advisor: Dr. Jaewoo Kang) — [Dec 2022 – Feb 2023]

- Developed reinforcement learning techniques for CXR report summarization using VLMs

INDUSTRY EXPERIENCE

Silvia Health — Seoul, Republic of Korea
Independent Researcher — [Jun 2025 – Aug 2025]

- Focused on identifying voice-based biomarkers for Alzheimer's disease, continuing as an academic–industry collaboration
- Compared multimodal LLMs (MLLMs) with modality-specific models to analyze complementary effects on performance

Kolon Lifescience — Seoul, Republic of Korea
Clinical Trial Project Manager — [Jan 2022 – Oct 2022]

- Managed long-term follow-up studies for clinical trials and authored clinical study reports
- Coordinated with CROs, vendors, and clinical sites to ensure compliance and meet project timelines

PROJECTS

Classifying Alzheimer's Disease with Voice and Transcript Data — [Jun 2025 – Ongoing]

- Classified Alzheimer's disease and healthy controls using real-world Korean patient voice recordings and transcripts
- Compared MLLMs with modality-specific baselines such as voice and text embedding models, analyzing complementary contributions of acoustic and textual modalities
- Expanding the study to distinguish not only Alzheimer's disease but also mild cognitive disorder

Integrating Radiology Findings for Mortality Prediction — [Oct 2024 – Ongoing]

- Compared VLMs combining discharge notes with either CXR images or radiology reports for ICU mortality prediction
- Found that CXR images provided stronger and complementary signals to discharge notes compared to radiology reports, achieving higher AUROC and validation through radiologist analysis
- In preparation as *"Integration of Radiology Findings for Mortality Prediction: Radiology Reports vs. CXR Images"*

Aspect-Oriented Summarization for Psychiatric Readmission Prediction — [Oct 2024 – Aug 2025]

- Contributed to developing LLM-based aspect-oriented summarization pipelines and compared summary types for 30-day readmission prediction using real-world EHR data from four hospitals
- Published at EMNLP 2025: *"Aspect-Oriented Summarization for Psychiatric Short-Term Readmission Prediction"*

ANGEL Framework — [Oct 2023 – May 2025]

- Enhanced a generative biomedical entity linking model through preference optimization with negative samples
- Published at ACL 2025 Findings: *"Learning from Negative Samples in Biomedical Generative Entity Linking"*

Human-in-the-loop LLMs for Biomedical Hypothesis Generation — [Jul 2023 – Sep 2024]

- Designed RAG-based approaches for biomedical hypothesis generation to support human-in-the-loop LLMs

AI Platform for Precision Medicine in Diabetes Using Big Data — [May 2023 – Sep 2024]

- Finetuned LLMs to deliver domain-specific, expert-level explanations for diabetes care by integrating clinical guidelines

Publications

[P] **Kim, C.**, Yoon, W., Lee, H., Lee, J., Kang, J., *et al.* (2025). **Integration of Radiology Findings for Mortality Prediction: CXR Images vs. Radiology Reports**. *In Preparation.*

[C] Yoon, W., Ren, B., Thomas, S., **Kim, C.**, Savova, G., *et al.* (2025). **Aspect-Oriented Summarization for Psychiatric Short-Term Readmission Prediction**. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP 2025 Main).*

[C] **Kim, C.***, Kim, H.*, Park, S., Lee, J., Sung, M., *et al.* (2025). **Learning from Negative Samples in Biomedical Generative Entity Linking**. In *Findings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025 Findings).*

[W] Kim, H.*, **Kim, C.***, Lee, H., Jang, K., Lee, J., *et al.* (2024). **KU-DMIS at EHRSQL 2024: Generating SQL Query via Question Templatization in EHR**. In *Proceedings of the 6th Clinical Natural Language Processing Workshop (co-located with NAACL 2024).*

[C] Lee, S., Jang, G., **Kim, C.**, Park, S., Yoo, K., *et al.* (2023). **Enhancing Clinical Outcome Predictions through Auxiliary Loss and Sentence-Level Self-Attention**. In *Proceedings of the 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2023).*

[W] Kim, H., **Kim, C.**, Sohn, J., Beck, T., *et al.* (2023). **KU AIGEN ICL EDI@BC8 Track 3: Advancing Phenotype Named Entity Recognition and Normalization for Dysmorphology Physical Examination Reports**. In *Proceedings of the BioCreative VIII Challenge and Workshop (Presented at the AMIA 2023).*

[W] Kim, G., Kim, H., Ji, L., Bae, S., **Kim, C.**, *et al.* (2023). **KU-DMIS-MSRA at RadSum23: Pre-trained Vision-Language Model for Radiology Report Summarization**. In *Proceedings of the 22nd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks (co-located with ACL 2023).*

Honors and Awards

1st Prize at ACL 2023 BioNLP Workshop
- Developed pre-trained vision-language model for CXR report summarization

1st Prize at HackSeoul 2024 Social Responsibility Track
- AI-Powered express bus reservation system for the digitally alienated elderly

3rd Prize at BioCreative VIII Track 3
- Extracted and normalized genetic phenotypes from dysmorphology physical examination entries

4th Prize at NACCL 2024 ClinicalNLP Workshop
- Developed reliable Text-to-SQL generation models for extracting information from EHR

Top 9 at HackGlobal Grand Finals
- AI-powered platform facilitating social gatherings and community engagement for hotel guests

Teaching

KyungHee University - Language and Data Intelligence Lab　　　　　　　　　　　Suwon, Republic of Korea
Invited talk　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　[2025]
- Applying Generative AI Models to Clinical and Biomedical Domains

Korea University - Department of Computer Science & Engineering　　　　　　　Seoul, Republic of Korea
Teaching Assistant　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　[2023]
- Introduction to data science (COSE471)
- Machine learning (COSE362)