Chanwoong Yoon

cwyoon99@gmail.com | Google Scholar | cwyoon-99.github.io | in/chanwoong-yoon

RESEARCH INTERESTS

Natural Language Processing, Retrieval-Augmented Generation, AI Safety, Model Interpretability

EDUCATION

Korea University — Seoul, Republic of Korea
M.S. - Computer Science and Engineering; GPA: 4.0/4.0 — [Mar. 2024 - Presents]

Hanyang University — Ansan, Republic of Korea
B.S. - Software, College of Computing; GPA: 3.66/4.0 — [Mar. 2017 - Feb. 2023]
(*Summa Cum Laude*)

PUBLICATIONS

1. C. Yoon, T. Lee, H. Hwang, M. Jeong, J. Kang. CompAct: Compressing Retrieved Documents Actively for Question Answering. EMNLP 2024
2. C. Yoon*, G. Kim*, B. Jeon, S. Kim, Y. Jo, J. Kang. Ask Optimal Questions: Aligning Large Language Models with Retriever's Preference in Conversational Search. Findings of NAACL 2025
3. T. Lee, C. Yoon, K. Jang, D. Lee, M. Song, H. Kim, J. Kang. ETHIC: Evaluating Large Language Models on Long-Context Tasks with High Information Coverage. NAACL 2025
4. Y. Park, C. Yoon, J. Park, M. Jeong, J. Kang. Temporal Heads: Where Language Models Recall Time-specific Information. ACL 2025
5. Y. Park, C. Yoon, J. Park, D. Lee, M. Jeong, J. Kang. ChroKnowledge: Unveiling Chronological Knowledge of Language Models in Multiple Domains. ICLR 2025
6. T. Lee, M. Song, C. Yoon, J. Park, J. Kang. The Curious Case of Analogies: Investigating Analogical Reasoning in Large Language Models. AAAI-26
7. H. Hwang, Y. Cho, C. Yoon, Y. Park, M. Song, K. Lee, G. Kim, J. Kang. Assessing LLM Reasoning Steps via Principal Knowledge Grounding. Findings of EMNLP 2025
8. J. Park, T. Lee, C. Yoon, H. Hwang, J. Kang. Outlier-Safe Pre-Training for Robust 4-Bit Quantization of Large Language Models. ACL 2025
9. J. Sohn, Y. Park, C. Yoon, S. Park, H. Hwang, M. Sung, H. Kim, J. Kang. Rationale-Guided Retrieval Augmented Generation for Medical Question Answering. NAACL 2025
10. M. Jeong, H. Hwang, C. Yoon, T. Lee, J. Kang. OLAPH: Improving Factuality in Biomedical Long-form Question Answering. preprint
11. H. Kim, D. Kim, J. Lee, C. Yoon, D. Choi, M. Gim, J. Kang. LAPIS: Language Model-Augmented Police Investigation System. CIKM 2024
12. H. Kim, H. Hwang, J. Lee, S. Park, D. Kim, T. Lee, C. Yoon, J. Sohn, D. Choi, J. Kang. Small Language Models Learn Enhanced Reasoning Skills from Medical Textbooks. NPJ Digital Medicine

RESEARCH EXPERIENCES

Georgia Institute of Technology — Atlanta, GA, USA
Visiting Researcher (Advisor: Alan Ritter) — [Oct. 2025 - Present]
- Working on fine-grained safety moderation for LLMs, focusing on token-level toxicity detection and real-time adaptive response control.
- Funded by KIAT National Fellowship (approx. $ 22,000)

DMIS Lab, Korea University — Seoul, Republic of Korea
Research Assistant (Advisor: Jaewoo Kang) — [Aug. 2023 - Sep. 2025]

Knowledge-Grounded Generation and Long-Context Reliability
- Ask Optimal Questions: Retriever-aligned Query Optimization for Conversational Search
  - Developed a novel alignment framework that enables LLMs to reformulate queries in line with retriever feedback.
  - Achieved state-of-the-art conversational query rewriting performance, significantly outperforming existing query rewriting baselines.

- CompAct: Active Context Compression for Multi-hop Question Answering
  - Developed an active strategy to compress extensive documents without losing essential information in context.
  - Demonstrated flexibility as a cost-efficient plug-in module with various off-the-shelf retrievers or readers, achieving exceptionally high performance and compression rates (47x).
- ETHIC: Evaluating Long-context Understanding via Information Coverage
  - Proposed information coverage (IC) as a metric for curating a long-context benchmark to evaluate Models' ability to fully understand high-density context.
  - revealed critical bottlenecks of LLMs in context utilization across high-density, information-rich context.

Temporal and Dynamic Knowledge Modeling
- Chroknowledge: Unveiling Chronological Knowledge of Language Models
  - Presented ChroKnowBench, a benchmark for evaluating how LLMs capture and generalize chronological knowledge across time-evolving facts.
- Temporal Heads: Specific Attention Heads for Recalling Time-specific Information
  - Identified attention heads responsible for encoding and recalling time-specific knowledge in LLMs.
  - Verified that Temporal Heads consistently respond to diverse temporal cues in text and enable controlled editing of time-sensitive knowledge.

Reasoning Enhancement and Knowledge-grounded Reasoning Evaluation
- Proposed Meerkat-7B, the first 7B-scale medical LM to pass the United States Medical Licensing Examination (USMLE), trained on reasoning chains extracted from 18 medical textbooks.
- Proposed Knowledge-grounded Evaluation framework for reasoning, a new evaluation methodology that measures the alignment of chain-of-thought reasoning with factual knowledge.

Electronics and Telecommunications Research Institute (ETRI) Daejeon, Republic of Korea
Research Intern (Advisor: Junseong Bang) [Jul. 2022 - Aug. 2022]

- Worked on a project Polbot, an intelligent conversational agent for National Police Agency.
- Improved training efficiency of Dialogue State Tracking (DST) module by applying parameter-efficient fine-tuning on the belief state generation task.

AWARDS AND HONORS

National Research Fellowship, KIAT [May 2025]
- Selected as one of only four recipients for the *National KIAT Fellowship* (USD 21,000), recognizing exceptional research potential in AI.

Outstanding Research Paper Award, Korea University [Feb. 2025]
- CompAct: Compressing Retrieved Documents Actively for Question Answering (EMNLP 2024)

Research Award, KIISE Korea Computer Congress Competition [Jul. 2022]

Merit-based Scholarships (50% Tuition), Hanyang University [Fall 2020 - Spring 2022]

TECHNICAL SKILLS

Programming: Python, C, C++
Libraries: PyTorch, NumPy, Matplotlib, Transformers, TransformerLens

RESEARCH & ANALYTICAL SKILLS

Theoretical Foundations: Linear Algebra, Probabilistic Modeling, Information Theory
Scientific Writing & Publication: publishing at top-tier conferences (ICLR, ACL, EMNLP, NAACL)
Collaboration & Mentorship: Academic collaborations as a first author and co-author (mentoring junior researchers)

SERVICE

Academic Reviewer
ICLR 2026, ARR Review (Dec. 2025), ACL 2025

Mandatory Military Service [Apr. 2018 - Mar. 2022]
Republic of Korea Air Force