Kyojun Choo

B.S student
Soongsil University
School of Computer Science & Engineering
Email: kjun5644@soongsil.ac.kr
Home: https://chubeloop.github.io/
LinkedIn: https://www.linkedin.com/in/kyojun-choo-8047b339a/
Lab: https://nlp.ssu.ac.kr/

LIFELONG RESEARCH OBJECTIVE

My lifelong research objective is to build AI systems that people can use with confidence and trust. As AI becomes an increasingly important part of everyday life, I aim to ensure that its capabilities, limitations, and behaviors can be reliably understood and evaluated. Through research on model interpretability and trustworthy evaluation, I hope to contribute to a future where people can benefit from AI while clearly understanding when and why its outputs can be trusted.

RESEARCH INTERESTS

My research interests center on understanding and improving Large Language Models (LLMs) through interpretability and evaluation. I am particularly interested in analyzing how linguistic, task-related, and capability-related information is represented within model internals, and in using these insights to better explain model behaviors, limitations, and failure modes. At the same time, I aim to develop more reliable and meaningful evaluation methodologies that go beyond surface-level benchmark scores by examining robustness, validity, generalization, and the extent to which benchmarks truly reflect the capabilities they are intended to measure. Ultimately, I seek to connect interpretability with evaluation so that a deeper understanding of model representations can inform the design of more trustworthy, transparent, and practically useful LLM systems.

EDUCATION

- **Soongsil University** Feb. 2021 - Present
  B.S student School of Computer Science & Engineering (advisor: Chanjun Park)
  GPA: 4.14/4.5

WORK EXPERIENCES

- **Natural Language Processing Lab, Soongsil University** Nov. 2025 - Present
  Undergraduate Researcher
- **SH&M Coding Class, Spartan Future Education Institute** March. 2025 - Present
  Coding Instructor

PUBLICATIONS

Equal contribution: *, *Corresponding author*: †

International Conferences

1. **Structured but Silent: Probing Capability Requirements in LLM Hidden States**
   Kyojun Choo, Minsoo Song, Yunju Kang, Chanjun Park†
   AACL-IJCNLP 2026-Findings

Domestic Conferences

1. **행동적 타당성 감사를 통한 리더보드 강건성 분석 연구**
   추교준, 송민수, 조성현, 장지윤, 남채린, 박찬준†
   한글 및 한국어 정보처리 학술대회 (HCLT) 2026
2. **오답의 의미적 타당성 제거를 통한 LLM 벤치마크 결함 문항 식별**
   송민수, 김찬우, 추교준, 송은서, 장지윤, 남채린, 이시현, 박찬준†
   한글 및 한국어 정보처리 학술대회 (HCLT) 2026 - 우수 논문
3. **많은 정보가 더 나은 예측을 만드는가? 대규모 언어 모델의 인간 어휘 이해도 예측 분석**
   송은서, 조성현, 남채린, 추교준, 금지우, 박찬준†
   한글 및 한국어 정보처리 학술대회 (HCLT) 2026
4. **K-Press: 한국어 벤치마크 능력 프로파일링을 위한 문항 기반 능력 축 유도**
   장지윤, 조성현, 김도윤, 송지훈, 추교준, 금지우, 박찬준†
   한글 및 한국어 정보처리 학술대회 (HCLT) 2026
5. **한국어 상충 페르소나 지시에서의 LLM 해결 양상 분석: 동일 우선순위를 중심으로**
   이시현, 금지우, 송지훈, 추교준, 김찬우, 박찬준†
   한글 및 한국어 정보처리 학술대회 (HCLT) 2026

Research Projects

Soongsil University

- 2026년 인공지능 말평 운영 및 연계 행사 개최 (2026, 참여연구자 (용역과제), National Institute of the Korean Language)
- 2026년 인공지능의 한국어 능력 표준 평가 과제 구축 및 활용 (2026, 참여연구자 (용역과제), National Institute of the Korean Language)
- 우수연구 신진연구 (유형B) (2026-2029, 참여연구자 (3책5공), NRF)
- Foundation Model 운용과정에서 민감정보 추론 방지 및 리스크 평가 기술 개발 (2026-2028, 참여연구자 (3책5공), Korea Internet & Security Agency (한국인터넷진흥원))
- 토론과 소통으로 지식을 연결하는 AI 에이전트 기술 개발 (2026-2028, 참여연구자**(PM)** (3책5공), Korea Creative Content Agency(한국콘텐츠진흥원))
- AI 보안 특화 'AI 에이전트' 개발 선행 연구 (2026, 참여연구자 (용역과제), 한국정보보호학회 산하 AI보안연구회)
- 국가 · 공공기관 가이드라인 준수를 위한 AI 시스템 보안 위협 자동화 기술점검 도구 개발 (2026, 참여연구자, KIST)
- AI 연구용 컴퓨팅 지원 프로젝트 (파라미터 수준의 정보 제어 원천기술 연구 및 파라미터 스튜디오 플랫폼 개발) (2026, 참여연구자, KETI)

AWARDS/FELLOWSHIPS/RESEARCH GRANTS

Awards

- **Best Paper Award (HCLT 2026)**
  오답의 의미적 타당성 제거를 통한 LLM 벤치마크 결함 문항 식별
  송민수, 김찬우, 추교준, 송은서, 장지윤, 남채린, 이시현, 박찬준†
  한글 및 한국어 정보처리 학술대회 (HCLT) 2026

Fellowships

- Baekma Academic Excellence Scholarship (Full Tuition), Soongsil University (Mar. 2025)
- Baekma Academic Excellence Scholarship (Full Tuition), Soongsil University (Sep. 2021)

Research Grants

- VESSL AI GPU Grant Program (2026)

PATENTS

Domestic Patent Applications

- **Method and Apparatus for Determining Functional Requirements of User Queries Using Pre-Generation Hidden States of a Large Language Model**
  대규모 언어 모델의 생성 전 은닉 상태를 이용한 사용자 질의의 기능적 요구사항 판단 방법 및 장치
  Chanjun Park, Kyojun Choo
  Korean Patent Application, filed on August 19, 2026 (Application No. 10-2026-0156645).
- **APPARATUS AND METHOD FOR AUDITING PERFORMANCE RANKINGS AMONG ARTIFICIAL INTELLIGENCE MODELS USING BENCHMARK-SPECIFIC AUDIT CONDITIONS**
  벤치마크별 감사 조건을 이용한 인공지능 모델 간 성능 순위 감사 장치 및 방법
  Chanjun Park, Kyojun Choo
  Korean Patent Application, filed on Oct 6, 2026 (Application No. 10-2026-0190777).

External Activities

Teaching Assistant

- SaRangAXConnect, AI Transformation (AX) Program (September 2026)