Geewook Kim (김기욱)

Applied Research Scientist at NAVER Cloud | Ph.D. Candidate at KAIST AI

gwkim.rsrch@gmail.com https://geewook.kim https://linkedin.com/in/gwkim-rsrch

Research Interests

My research focuses on resource-efficient multimodal generative AI spanning **text**, **image**, **audio**, and **video**. I investigate the principles driving capabilities, limitations, and generalization of generative models in resource-constrained settings, and explore principled approaches to building AI systems that are trustworthy, transparent, and broadly beneficial in real-world applications.

Key Areas:

- Resource-efficient multimodal generative models (**text, image, audio, video**)
- Understanding generalization and constraints of modern generative AI in practical settings
- Trustworthy and responsible deployment of AI systems for real-world adoption

I bridge fundamental research and practical system building, as exemplified by open-source projects such as Donut (6.7K+ GitHub stars), contributing both novel methodologies and production-ready tools.

Employment

Tech Lead / Applied Research Scientist, **NAVER Cloud Corp.**, Korea — May 2023–Present

- As a technical leader of both Semantic Perception and Vision Understanding teams, I lead applied research and engineering for NAVER's LLM-based multimodal products (HyperCLOVA X Vision & Video, https://clova.ai/hyperclova), directly managing up to 8 researchers and engineers. During this period, I have published multiple papers at top-tier venues including **AAAI**, **ICLR**, **ICCV**, **ECCV**, **NAACL**, and **EMNLP** (see Publications). Selected industry achievements:

Tech Report	"HyperCLOVA X 32B Think" and "HyperCLOVA X 8B Omni" Technical Reports [Think][Omni]
Tech Blog	"HyperCLOVA X Vision: 눈을 뜨다 (Introducing HyperCLOVA X Vision)" [Link]
Tech Blog	"HyperCLOVA X Video: Seeing through motion" [Link]
Conference Talk	"[DAN 24] HyperCLOVA X Vision: Open Your Eyes, CLOVA X!" [Session Link][Video Link]
Open Model	"HyperCLOVAX-SEED-Think-32B" and "HyperCLOVAX-SEED-Omni-8B" [Think-32B][Omni-8B]

Lecturer (Part-Time), **University of Seoul**, Korea — Spring 2025

- I designed and delivered "AI Engineering in Production (산업AI공학)", a graduate course covering recent advances in generative AI models—particularly Large Language Models (LLMs)—and their industrial applications.
  Course Website: https://geewook.kim/lecture/uos25spring-91035

Applied Research Scientist, **NAVER Corp.**, Korea — Apr. 2020–Apr. 2023

- I conducted research and developed systems for NAVER's Document AI product family, including OCR [Link] and Information Extraction [Link]. This work was featured in the conference talk "이곳에 방문하셨군요!:

사진 속 영수증의 가게 찾기 (Identifying a store from a receipt image)" [Link]. I also initiated and led several research and open-source projects, including Donut and Webvicob.

Research part-timer/trainee, Mathematical Statistics Team, Sep. 2017–Feb. 2020
RIKEN Center for Advanced Intelligence Project, Japan
Supervisor: Prof. Hidetoshi Shimodaira

Research Intern, **CLOVA OCR, NAVER Corp.** [Link], Korea Aug. 2018–Oct. 2018, Aug. 2019–Sep. 2019
Supervisor: Dr. Hwalsuk Lee

ML Engineering Intern, **Recruit Holdings Co., Ltd.**, Tokyo, Japan 2017 (2 Months)

Software Engineering Intern, **Abeja, Inc.**, Tokyo, Japan 2016 (2 Months)

Planning Intern, **SoftBank Corp.**, Tokyo, Japan 2015 (1 Month)

Education

Korea Advanced Institute of Science & Technology (KAIST) Sep. 2023–Aug. 2026 (Expected)
Doctor of Philosophy in Artificial Intelligence, Language & Knowledge Lab. [Link]
Kim Jaechul Graduate School of Artificial Intelligence
(Supervisor : Prof. Minjoon Seo)

Kyoto University Apr. 2018–Mar. 2020
Master of Informatics, Statistical Intelligence Lab. [Link]
Graduate School of Informatics
(Supervisor : Prof. Hidetoshi Shimodaira)

Kyoto University Apr. 2014–Mar. 2018
Bachelor of Engineering, Applied Mathematics and Physics Course [Link]
School of Informatics and Mathematical Science
(Supervisor : Prof. Hidetoshi Shimodaira)

Publications

Google Scholar: https://scholar.google.com/citations?user=1a2QbgEAAAAJ

18 peer-reviewed international conference papers, 4 workshop papers, 4 technical reports, and 8 patents. Two papers have each received 800+ Google Scholar citations, with open-source projects accumulating 10K+ GitHub stars combined.

Peer-Reviewed International Conferences

[C18] Hyunji Lee, Seunghyun Yoon, Yunjae Won, Hanseok Oh, **Geewook Kim**, Trung Bui, Franck Dernoncourt, Elias Stengel-Eskin, Mohit Bansal, and Minjoon Seo, "*Instruction Tuning with and without Context: Behavioral Shifts and Downstream Impact*", Proceedings of the 2026 Conference of the European Chapter of the Association for Computational Linguistics, 2026. — EACL 2026

[C17] **Geewook Kim** and Minjoon Seo, "*State-Space Hierarchical Compression with Gated Attention and Learnable Sampling for Hour-Long Video Understanding in Large Multimodal Models*", Proceedings of the Fortieth AAAI Conference on Artificial Intelligence, 2026. — AAAI 2026
Oral Presentation (1089/23678=4.60%)

[C16] Gio Paik, **Geewook Kim**, and Jinbae Lim, "*MMRefine: Unveiling the Obstacles to Robust Refinement in Multimodal Large Language Models*", Findings of the Association for Computational Linguistics, 2025. — ACL Findings 2025

[C15] Sanghee Park*, **Geewook Kim***† (co-first and corresponding author), "*Evaluating Multimodal Generative AI with Korean Educational Standards*", Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies, 2025. — NAACL-HLT 2025

[C14] Seongyun Lee* , **Geewook Kim*** (co-first author), Jiyeon Kim* (co-first author), Hyunji Lee, Hoyeon Chang, Sue Hyun Park, and Minjoon Seo, "*How Does Vision-Language Adaptation Impact the Safety of Vision Language Models?*", Proceedings of the Thirteenth International Conference on Learning Representations, 2025. — ICLR 2025

[C13] **Geewook Kim** and Minjoon Seo, "*On Efficient Language and Vision Assistants for Visually-Situated Natural Language Understanding: What Matters in Reading and Reasoning*", Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, 2024. — EMNLP 2024

[C12] Seongyun Lee, Seungone Kim, Sue Hyun Park, **Geewook Kim**, and Minjoon Seo, "*Prometheus-Vision: Vision-Language Model as a Judge for Fine-Grained Evaluation*", Findings of the Association for Computational Linguistics, 2024. — ACL Findings 2024

[C11] Yamato Okamoto, Youngmin Baek, **Geewook Kim**, Ryota Nakao, DongHyun Kim, Moon Bin Yim, Seunghyun Park, and Bado Lee, "*CREPE: Coordinate-Aware Cost-Efficient Document Parsing End-to-End Model*", Proceedings of the International Conference on Document Analysis and Recognition, 2024. — ICDAR 2024

[C10] **Geewook Kim**†, Hodong Lee, Daehee Kim, Haeji Jung, Sanghee Park, Yoonsik Kim, Sangdoo Yun, Taeho Kil, Bado Lee, and Seunghyun Park, "*Cream: Visually-Situated Natural Language Understanding with Contrastive Reading Model and Frozen Large Language Models*", Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, 2023. — EMNLP 2023

[C9] Daehee Kim, Yoonsik Kim, DongHyun Kim, Yumin Lim, **Geewook Kim**, and Taeho Kil, "*SCOB: Universal Text Understanding via Character-wise Supervised Contrastive Learning with Online Text Rendering for Bridging Domain Gap*", IEEE/CVF International Conference on Computer Vision, 2023. — ICCV 2023

[C8] Donghyun Kim, Teakgyu Hong, Moonbin Yim, Yoonsik Kim, and **Geewook Kim**† (corresponding author), "*On Web-based Visual Corpus Construction for Visual Document Understanding*", International Conference on Document Analysis and Recognition, 2023. — ICDAR 2023

[C7] **Geewook Kim**†, Teakgyu Hong, Moonbin Yim, Jeongyeon Nam, Jinyoung Park, Jinyeong Yim, Wonseok Hwang, Sangdoo Yun, Dongyoon Han, and Seunghyun Park, "*OCR-free Document Understanding Transformer*", European Conference of Computer Vision, 2022.
810+ citations, **6.7K+ stars at GitHub** — ECCV 2022

[C6] Wonseok Hwang, Hyunji Lee, Jinyeong Yim, **Geewook Kim**, and Minjoon Seo, "*Cost-effective End-to-end Information Extraction for Semi-structured Document Images*", Empirical Methods in Natural Language Processing, 2021. — EMNLP 2021

[C5] Sungrae Park, **Geewook Kim**, Junyeop Lee, Junbum Cha, Ji-Hoon Kim, and Hwalsuk Lee, "*Scale down Transformer by Grouping Features for a Lightweight Character-level Language Model*", International Conference on Computational Linguistics, 2020. — COLING 2020

[C4] Jeonghun Baek, **Geewook Kim**, Junyeop Lee, Sungrae Park, Dongyoon Han, Sangdoo Yun, Seong Joon Oh, Hwalsuk Lee, "*What is wrong with scene text recognition model comparisons? dataset and model analysis*", International Conference on Computer Vision, 2019.
825+ citations, **3.9K+ stars at GitHub**
Oral Presentation (187/4304=4.3%) — ICCV 2019

[C3] **Geewook Kim**, Akifumi Okuno, Kazuki Fukui, and Hidetoshi Shimodaira, "*Representation Learning with Weighted Inner Product for Universal Approximation of General Similarities*", International Joint Conference on Artificial Intelligence, 2019. (Presented as both **oral** and **poster**). — IJCAI 2019

[C2] **Geewook Kim**, Kazuki Fukui, Hidetoshi Shimodaira, "*Segmentation-free Compositional n-gram Embedding*", North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019. — NAACL-HLT 2019

[C1] Akifumi Okuno, **Geewook Kim**, and Hidetoshi Shimodaira, "*Graph Embedding with Shifted Inner Product Similarity and Its Improved Approximation Capability*", International Conference on Artificial Intelligence and Statistics, 2019. — AISTATS 2019

Peer-Reviewed Workshop Papers

[W4] **Geewook Kim***, Shuhei Yokoo* (co-first author), Sukmin Seo, Atsuki Osanai, Yamato Okamoto, and Youngmin Baek, "*On Text Localization in End-to-End OCR-Free Document Understanding Transformer Without Text Localization Supervision*", International Conference on Document Analysis and Recognition Workshops, 2023. — ICDAR-W 2023

[W3] **Geewook Kim**, Wonseok Hwang, Minjoon Seo, and Seunghyun Park, "*Semi-Structured Query Grounding for Document-Oriented Databases with Deep Retrieval and Its Application to Receipt and POI Matching*", AAAI Workshop on Knowledge Discovery from Unstructured Data in Financial Services, 2022. — AAAI-W 2022

[W2] Masahiro Naito, Sho Yokoi, **Geewook Kim**, Hidetoshi Shimodaira, "*Revisiting Additive Compositionality: AND, OR and NOT Operations with Word Embeddings*", ACL-IJCNLP 2021 Student Research Workshop, 2021. — ACL-SRW 2021

[W1] **Geewook Kim**, Kazuki Fukui, and Hidetoshi Shimodaira, "*Word-like Character n-gram Embedding*", EMNLP Workshop on Noisy User-generated Text, 2018. — EMNLP-W 2018

Tech Reports

[T4] NAVER Cloud HyperCLOVA X Team, "*HyperCLOVA X 32B Think Technical Report*", 2026. — Tech Report 2026

[T3] NAVER Cloud HyperCLOVA X Team, "*HyperCLOVA X 8B Omni Technical Report*", 2026. — Tech Report 2026

[T2] HyperCLOVA X Team, "*HyperCLOVA X THINK Technical Report*", 2025. — Tech Report 2025

[T1] HyperCLOVA X Team, "*HyperCLOVA X Technical Report*", 2024. — Tech Report 2024

Preprints, Tech Blogs, and Domestic Publications

[O6] **Geewook Kim** and Minjoon Seo, "*Does Audio Matter for Modern Video-LLMs and Their Benchmarks?*", arXiv preprint, 2025. — Preprint 2025

[O5] NAVER Cloud Vision Understanding Team, "*HyperCLOVA X Video: Seeing through motion*", CLOVA AI Tech Blog, 2025. — Tech Blog 2025

[O4] NAVER Cloud Vision Understanding Team, "*Introducing HyperCLOVA X Vision*", CLOVA AI Tech Blog, 2024. — Tech Blog 2024

[O3] Morihiro Mizutani, Akifumi Okuno, **Geewook Kim**, Hidetoshi Shimodaira, "*Stochastic Neighbor Embedding of Multimodal Relational Data for Image-Text Simultaneous Visualization*", arXiv preprint, 2020. — Preprint 2020

[O2] **Geewook Kim**, Sho Yokoi, and Hidetoshi Shimodaira, "単語埋め込みの二種類の加法構成性 (**Two Types of Additive Compositionality in Word Embedding**)", Annual Meeting of the Association for Natural Language Processing (in Japanese), 2020. — ANLP 2020

[O1] **Geewook Kim**, Akifumi Okuno, and Hidetoshi Shimodaira, "擬ユークリッド空間への単語埋め込み (*Embedding Words into Pseudo-Euclidean Space*)", Annual Meeting of the Association for Natural Language Processing (in Japanese), 2019. (Selected to receive both awards: **Young Researcher Award** [Link] and **Best Poster Award**). — ANLP 2019

Patents

[P8] **Geewook Kim**, Hodong Lee, Daehee Kim, Sanghee Park, Yoonsik Kim, Sangdoo Yun, Bado Lee, and Seunghyun Park, "*METHOD AND SYSTEM FOR ENHANCING IMAGE UNDERSTANDING*", KR Patent KR1020230117309A, filed in September 2023. — Sep. 2023

[P7] **Geewook Kim** and Youngmin Baek, "*METHOD, COMPUTER DEVICE, AND COMPUTER PROGRAM FOR TEXT LOCALIZATION IN END-TO-END DOCUMENT UNDERSTANDING MODEL*", KR Patent KR1020230075133A, filed in June 2023. — Jun. 2023

[P6] Daehee Kim, Yoonsik Kim, DongHyun Kim, Yumin Lim, **Geewook Kim**, and Taeho Kil, "*METHOD AND SYSTEM FOR PERFORMING OCR USING CHARACTER-WISE SUPERVISED CONTRASTIVE LEARNING MODEL*", KR Patent KR1020230037945A, filed in March 2023. — Mar. 2023

[P5] **Geewook Kim**, Teakgyu Hong, Moonbin Yim, and Seunghyun Park, "*METHOD AND SYSTEM FOR EXTRACTING STRUCTURED INFORMATION FROM SEMI-STRUCTURED DOCUMENTS THROUGH DEEP-LEARNING*", KR Patent KR102784722B1, filed in January 2022 and granted in March 2025. — Jan. 2022

[P4] **Geewook Kim**, Wonseok Hwang, and Minjoon Seo, "*METHOD AND SYSTEM FOR DATA SEARCHING*", KR Patent KR102684423B1, filed in August 2021 and granted in July 2024 (also granted as JP Patent JP7367139B2, filed in July 2022 and granted in October 2023). — Aug. 2021

[P3] **Geewook Kim**, Seung Shin, Youngmin Baek, Hyosun Wang, Jungun Kim, and Seungbeom Choi, "*CHARACTER RECOGNITION METHOD AND SYSTEM ROBUST TO ERRORS OF CHARACTER RECOGNITION THAT RECOGNIZE INFORMATION INCLUDED IN TABLES*", KR Patent KR102697516B1, filed in July 2021 and granted in August 2024 (also granted as JP Patent JP7398526B2, filed in July 2022 and granted in December 2023). — Jul. 2021

[P2] Wonseok Hwang, Jinyeong Yim, **Geewook Kim**, Minjoon Seo, and Hyunji Lee, "*METHOD AND SYSTEM FOR EXTRACTING INFORMATION FROM SEMI-STRUCTURED DOCUMENTS*", KR Patent KR102649429B1, filed in May 2021 and granted in March 2024. — May 2021

[P1] Seonghyeon Kim, **Geewook Kim**, Jaeheung Surh, Daehyun Nam, Seungbeom Choi, Seung Shin, Youngmin Baek, and Hyosun Wang, "*METHOD AND SYSTEM FOR RECOGNIZING TABLES*", KR Patent KR102699224B1, filed in March 2021 and granted in August 2024. — Mar. 2021

Teaching

Lecturer, **AI Engineering in Production** — Spring 2025
University of Seoul
Graduate course, 3 credits, 28 students enrolled
Course Website: https://geewook.kim/lecture/uos25spring-91035

Invited Talks & Seminars

HyperCLOVA X Vision: Open Your Eyes, CLOVA X! — Nov. 2024
TEAM NAVER CONFERENCE DAN 24 [Session Link][Slide]

HyperCLOVA X Vision: Open Your Eyes, CLOVA X! — Oct. 2024
NAVER ENGINEERING DAY

Vision-Language Models for Context-Rich Image Understanding Tasks — Apr. 2024
University of Seoul [Slide][Session Link]

Fine-Grained Evaluation of Vision-Language Models through VLM as a Judge — Feb. 2024
NAVER Tech Meetup [Slide]

Recent Advances in Document AI — Mar. 2023
Korea University

Recent Advances in Document AI — Dec. 2022
Kookmin University

OCR-Free Document Understanding Transformer — Nov. 2022
Microsoft [Slide]

Identifying a store from a receipt image — 2021
DEVIEW Conference [Video][Slide][Session Link]

Representation Learning with Weighted Inner Product for Universal Approximation of General Similarities 2019
Michinoku Communication Science Seminar, Tohoku University [Session Link]

Academic Services

Journal Reviewer:
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) (since 2025),
ACM Computing Surveys (CSUR) (since 2025), *Expert Systems with Applications (ESWA)* (since 2025),
Journal of Medical Internet Research (JMIR) (since 2026), *IEEE Access* (since 2023)

Conference Reviewer:
ICML (since 2026), *BMVC* (since 2026), *AAAI* (since 2025), *ICCV* (since 2025), *CVPR* (since 2024),
ACL ARR (*ACL*, *EMNLP*, *NAACL*, *COLING*) (since 2024),
Industry Track of *ACL*, *EMNLP*, *NAACL*, *EACL*, *COLING* (since 2022)

Selected Honors and Awards

Seiwa International Students Scholarship [Link] 2019

Korea-Japan Joint Government Scholarship [Link] 2013–2018
Admission and tuition fees, and living costs covered for a year of preliminary education and four years of Bachelor's studies

Young Researcher Award at Annual Meeting of the Association for Natural Language Processing [Link]
2019

❖ Last updated on February 15, 2026 ❖