

Minbyul Jeong

AI Research Engineer
Upstage AI
338, Gwanggyojungang-ro, Suji-gu,
Yongin-si, Gyeonggi-do, Republic of
Korea, 16942

Home: <http://minstar.github.io/>
Email: wjdalsquf@gmail.com
Github: /minstar
Scopus: 57209640341

Industrial Experience

Upstage <i>Principal Researcher Scientist of Large Language Model Team</i>	Aug. 2024 – Current Seoul
--	------------------------------

Education

Korea University <i>Ph.D (MS Integrated) Student, Computer Science & Engineering</i>	Mar. 2019 – Aug. 2024 Seoul
Korea University <i>B.Sc., Computer Science & Engineering</i>	Mar. 2015 – Feb. 2019 Seoul

Awards & Honors

BioASQ 9B <i>A challenge on large-scale biomedical semantic indexing and question answering</i>	1st place 2021
BioASQ 8B <i>A challenge on large-scale biomedical semantic indexing and question answering</i>	1st place 2020
BioASQ 7B <i>A challenge on large-scale biomedical semantic indexing and question answering</i>	1st place 2019

Preprints

[P11] Minbyul Jeong, "Ko-WideSearch: A Korean Breadth-Search Benchmark for Exhaustive Set Enumeration by Web Agents". arXiv preprint. 2026 [[Project Page](#)] [[Paper](#)]

[P10] Minbyul Jeong, "OpenBioRQ: Unsolved Biomedical Research Questions for Agents". arXiv preprint. 2026 [[Project Page](#)] [[Paper](#)]

[P9] Minbyul Jeong, "Healthcare AI GYM for Medical Agents". arXiv preprint. 2026 [[Project Page](#)] [[Paper](#)]

[P8], Solar AI Tech Team, "Solar Open Technical Report". arXiv preprint. 2026 [[Paper](#)]

[P7] Jungho Cho, Minbyul Jeong, and Sungrae Park, "User-Oriented Multi-Turn Dialogue Generation with Tool Use at scale". arXiv preprint. 2026 [[Paper](#)]

[P6] Yein Park, Minbyul Jeong[†], and Jaewoo Kang[†], "THINKING SPARKS!: EMERGENT ATTENTION HEADS IN REASONING MODELS DURING POST TRAINING". arXiv preprint. 2025. [[Paper](#)]

[P5] Yongwoo Song, Minbyul Jeong[†], and Mujeen Sung[†], "Trustworthy Agents for Electronic Health Records through Confidence Estimation". arXiv preprint. 2025. [[Paper](#)]

[P4] Minbyul Jeong[†], Jungho Cho, Minsoo Khang, Dawoon Jung, and Teakgyu Hong, "System Message Generation for User Preferences using Open-Source Models". arXiv preprint. 2025. [[Paper](#)]

[P3] Minbyul Jeong, Hyeon Hwang, Chanwoong Yoon, Taewhoo Lee, and Jaewoo Kang, "OLAPH: Improving Factuality in Biomedical Long-form Question Answering". arXiv preprint. 2024. [\[Github\]](#) [\[Paper\]](#) [\[Youtube\]](#)

[P2] Minbyul Jeong, Jaewoo Kang, "Regularization for Long Named Entity Recognition". arXiv preprint. 2021. [\[Github\]](#) [\[Paper\]](#)

[P1] Raehyun Kim, Chan Ho So, **Minbyul Jeong**, Sanghoon Lee, Jinkyu Kim, Jaewoo Kang, "HATS: A Heirarchical Graph Attention Network for Stock Movement Prediction". arXiv preprint. 2019. [\[Github\]](#) [\[Paper\]](#)

Publications

[C15] Yein Park, Chanwoong Yoon, Jungwoo Park, **Minbyul Jeong**[†] and Jaewoo Kang[†], "Does Time Have Its Place? Temporal Heads: Where Language Models Recall Time-specific Information". In: Proceedings of ACL 2025. (AI Top-tier Conference). [\[Github\]](#) [\[Paper\]](#)

[C14] Yein Park, Chanwoong Yoon, Jungwoo Park, Donghyeon Lee, **Minbyul Jeong**[†] and Jaewoo Kang[†], "ChroKnowledge: Unveiling Chronological Knowledge of Language Models in Multiple Domains". In: Proceedings of ICLR 2025. (AI Top-tier Conference). [\[Github\]](#) [\[Paper\]](#) [\[Homepage\]](#)

[C13] Chanwoong Yoon, Taewhoo Lee, Hyeon Hwang, **Minbyul Jeong**[†], and Jaewoo Kang[†], "CompAct: Compressing Retrieved Documents Actively for Question Answering". In: Proceedings of EMNLP 2024. (NLP Top-tier Conference). [\[Github\]](#) [\[Paper\]](#)

[C12] Minbyul Jeong, Jiwoong Sohn, Mujeen Sung, Jaewoo Kang, "Improving Medical Reasoning through Retrieval and Self-Reflection with Retrieval-Augmented Large Language Models". In: Proceedings of ISMB 2024. (Bioinformatics 2024, 40, Supplementary_1, i119-i129) [\[Github\]](#) [\[Paper\]](#)

[C11] Minbyul Jeong, Jaewoo Kang, "Consistency enhancement of model prediction on document-level named entity recognition". Bioinformatics 2023. (Within the top 10% of JCR, IF 6.937) [\[Github\]](#) [\[Paper\]](#)

[C10] Mujeen Sung*, **Minbyul Jeong***, Yonghwa Choi, Donghyeon Kim, Jinhyuk Lee, Jaewoo Kang, "BERN2 an advanced neural biomedical named entity recognition and normalization tool". Bioinformatics 2022. (33,351,060 API calls from 118 countries across 2,919 organizations, Within the top 10% of JCR, IF 6.937) [\[Github\]](#) [\[Paper\]](#) [\[Homepage\]](#)

[C9] Seongjun Yun, **Minbyul Jeong**, Sungdong Yoo, Seunghun Lee, SeanS. Yi, Raehyun Kim, Jaewoo Kang, Hyunwoo J. Kim, "Graph Transformer Networks Learning Meta-path Graphs to Improve GNNs". Neural Networks 2022. [\[Github\]](#) [\[Paper\]](#)

[C8] Wonjin Yoon, Jaehyo Yoo, Sumin Seo, Mujeen Sung, **Minbyul Jeong**, Gangwoo Kim, Jaewoo Kang, "Data-Centric and ModelCentric Approaches for Biomedical Question Answering". In: Proceedings of CLEF 2021. [\[Github\]](#) [\[Paper\]](#)

[C7] Meijun Liu, Yi Bu, Chongyan Chen, Jian Xu, Daifeng Li, Yan Leng, Richard Barry Freeman, Eric Meyer, Wonjin Yoon, Mujeen Sung, **Minbyul Jeong**, Jinhyuk Lee, Jaewoo Kang, Chao Min, Min Song, Yujia Zhai, Ying Ding, "Pandemics are catalysts of scientific novelty: Evidence from COVID-19". Journal of the Association for Information Science and Technology (JASIS&T) 2022. [\[Paper\]](#)

[C6] Minbyul Jeong, Mujeen Sung, Gangwoo Kim, Donghyeon Kim, Wonjin Yoon, Jaehyo Yoo, Jaewoo Kang, "Transferability of Natural Language Inference to Biomedical Question Answering". In: Proceedings of CLEF 2020. [[Github](#)] [[Paper](#)]

[C5] Jinhyuk Lee, Sean S. Yi, **Minbyul Jeong**, Mujeen Sung, Wonjin Yoon, Yonghwa Choi, Miyoung Ko, Jaewoo Kang, "Answering Questions on COVID-19 in Real-Time". In: Proceedings of 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. [[Github](#)] [[Paper](#)]

[C4] Jian Xu, Sunkyu Kim, Min Song, **Minbyul Jeong**, Donghyeon Kim, Jaewoo Kang, Justing F. Rousseau, Xin Li, Weijia Xu, Vette I. Torvik, Yi Bu, Chongyan Chen, Islam Akef Ebeid, Daifeng Li, Ying Ding "Building a PubMed knowledge graph". Scientific Data 2020. [[Paper](#)]

[C3] Seongjun Yun, **Minbyul Jeong**, Raehyun Kim, Jaewoo Kang, Hyunwoo J. Kim, "Graph Transformer Networks", In: Proceedings of NeurIPS 2019. (AI Top-tier Conference). [[Github](#)] [[Paper](#)]

[C2] Wonjin Yoon, Jinhyuk Lee, Donghyeon Kim, **Minbyul Jeong**, Jaewoo Kang, "Pre-trained Language Model for Biomedical Question Answering". In: Proceedings of ECML PKDD 2019. [[Github](#)] [[Paper](#)]

[C1] Donghyeon Kim, Jinhyuk Lee, Chan Ho So, Hwisang Jeon, **Minbyul Jeong**, Yonghwa Choi, Wonjin Yoon, Mujeen Sung, Jaewoo Kang, "*BERN: A neural named entity recognition and multi-type normalization tool for biomedical text mining*". IEEE Access 2019. [[Github](#)] [[Paper](#)]

Reviewer

Conferences:

- Conference on Neural Information Processing Systems (2024, 2025, 2026)
- the Association for Computational Linguistics (2025, 2026)
- International Conference on Learning Representations (2024, 2025, 2026)
- Empirical Methods in Natural Language Processing (2024, 2025, 2026)
- North American Chapter of the Association for Computational Linguistics (2024, 2025)

Journals:

- Experimental & Theoretical Artificial Intelligence (2024)
- Neural Networks (2023)
- Knowledge-Based Systems (2023)
- Applied Intelligence (2022, 2023)
- Data Technologies and Applications (2022)