

Byungjun Kim (김병준)

Assistant Professor of Cultural Informatics, Graduate School of Korean Studies,
The Academy of Korean Studies

bjkim@byungjunkim.com · bjkim@aks.ac.kr · +82-31-730-8161

byungjunkim.com · Lab: digitalhumanities.kr · [Google Scholar](https://scholar.google.com/citations?user=bjkim) · [ORCID 0000-0001-9925-9343](https://orcid.org/0000-0001-9925-9343)

RESEARCH PROFILE

My research lies in digital humanities and computational social science: I analyze large-scale bibliographic and textual data to study how scholarly and cultural knowledge is produced and communicated.

Keywords: digital humanities · computational social science · cultural analytics · bibliometrics and science of science · natural language processing and text mining · scholarly communication and open access · Korean datasets and language models

EDUCATION

Mar. 2018 – Aug. 2022	Ph.D. in Data Science , Sungkyunkwan University (Dept. of Interaction Science) Dissertation: <i>Mapping the Knowledge Structure of Korean Humanities: Bibliographic Data Analysis of Humanities Journal Articles in the Korea Citation Index, 2004–2019</i> . Advisor: Jang Hyun Kim.
Mar. 2015 – Feb. 2017	M.S. in Culture Technology , KAIST (Graduate School of Culture Technology) Thesis: <i>The Shift of Critical Language: The Shifting Aspects of Concept Words in Literature-Magazine Criticism, 1995–2015</i> . Advisors: Wonjae Lee and Bong Gwan Jun.
Mar. 2008 – Feb. 2015	B.A. in Korean Language and Literature , Sungkyunkwan University (dual major: Film, Television and Multimedia)

ACADEMIC & RESEARCH APPOINTMENTS

Sep. 2024 – Present	Assistant Professor , Cultural Informatics, Graduate School of Korean Studies, The Academy of Korean Studies
Sep. 2024 – Present	Co-Director , Center for Digital Humanities (digitalhumanities.kr)
Jan. 2023 – Aug. 2024	Research Assistant Professor , Center for Digital Humanities & Computational Social Sciences, KAIST
Mar. 2024 – May 2024	Research Fellow , Trier Center for Digital Humanities, Universität Trier, Germany (CLS INFRA)
Nov. 2021 – Dec. 2022	Research Associate , Center for Digital Humanities & Computational Social Sciences, KAIST
Jan. 2018 – Nov. 2021	Graduate Research Assistant , Data Science & Social Analytics Lab (DSSAL), Sungkyunkwan University
Jun. 2017 – Dec. 2018	Manager , Data Research Team, Bigvalue (real estate data analysis)

PEER-REVIEWED INTERNATIONAL JOURNAL ARTICLES (SSCI · A&HCI · SCIE · SCOPUS)

Underline indicates corresponding authorship throughout this CV.

1. Kim, G., Park, S., Ji, H., Lee, H., Lee, B., Jeong, C., & **Kim, B.** (2026). KNoTE: A TEI encoded XML dataset of modern Korean novels. *Journal of Open Humanities Data*, 12, 102. doi.org/10.5334/johd.590
2. Lim, I., Ji, H., & **Kim, B.** (2026). KPoEM: A human-annotated dataset for emotion classification and RAG-based poetry generation in Korean modern poetry. *The Review of Korean Studies*, 29(1), 161–206. [10.25024/review.2026.29.1.006](https://doi.org/10.25024/review.2026.29.1.006)
3. **Kim, B.**, Jee, J., & Woo, D. (2026). Mapping the semantic transformations of major powers in Cold War East Asia: A word embedding approach using Chinese and Korean newspapers (1946–95). *Digital Scholarship in the Humanities*, 41(2), 778–793. doi.org/10.1093/llc/fqag003
4. Hong, Y., **Kim, B.**, Jeon, J., & Kim, L. (2025). Has higher education become more interdisciplinary? A longitudinal analysis of syllabi using natural language processing. *Humanities and Social Sciences Communications*, 12(1), 1841. doi.org/10.1057/s41599-025-06126-7
5. Lee, S., **Kim, B.**, & Jun, B. G. (2024). Automatic detection of four-panel cartoon in large-scale Korean digitized newspapers using deep learning. *Journal of Open Humanities Data*, 10, 36. doi.org/10.5334/johd.205
6. Jee, J., **Kim, B.**, & Jun, B. G. (2024). The role of English Wikipedia in mediating East Asian historical disputes: The case of Balhae. *Asian Journal of Communication*, 34(4), 399–418. doi.org/10.1080/01292986.2024.2342822
7. **Kim, B.**, Yang, S., & Kim, H. (2024). Voices of transitions: Korea's online news media and user comments on the energy transition. *Energy Policy*, 187, 114020. doi.org/10.1016/j.enpol.2024.114020
8. Song, J., **Kim, B.**, Kim, M., & Iverson, P. (2023). The Korean Speech Recognition Sentences: A large corpus for evaluating semantic context and language experience in speech perception. *Journal of Speech, Language, and Hearing Research*, 66(9), 3399–3412. doi.org/10.1044/2023_JSLHR-23-00137
9. Kim, Y., **Kim, B.**, Park, M., Nam, W., & Kim, J. H. (2023). A soft power challenge, or an opportunity? A big data analysis on Chinese soft power during the COVID-19 pandemic. *Foreign Policy Analysis*, 19(3), orad011. doi.org/10.1093/fpa/orad011
10. Lee, K. R., Kim, J. H., Jang, J., Yoon, J., Nan, D., Kim, Y., & **Kim, B.** (2023). News big data analysis of international start-up innovation discourses through topic modelling and network analysis: Comparing East Asia and North America. *Asian Journal of Technology Innovation*, 31(3), 581–603. doi.org/10.1080/19761597.2022.2134154
11. Lee, K. R., **Kim, B.**, Nan, D., & Kim, J. H. (2021). Structural topic model analysis of mask-wearing issue using international news big data. *International Journal of Environmental Research and Public Health*, 18(12), 6432. doi.org/10.3390/ijerph18126432
12. **Kim, B.**, Yoo, M., Park, K. C., Lee, K. R., & Kim, J. H. (2021). A value of civic voices for smart city: A big data analysis of civic queries posed by Seoul citizens. *Cities*, 108, 102941. [10.1016/j.cities.2020.102941](https://doi.org/10.1016/j.cities.2020.102941)

REFEREED INTERNATIONAL CONFERENCE PROCEEDINGS (AI/NLP)

1. van Cranenburgh, A., Yang, X., Alvanita, Di Domenico, C. N., Ferragud, M., Graciotti, A., Ion, A. G., **Kim, B.**, Park, S., Visser Solissa, N., Zhou, X., & Pianzola, F. (2026). GOLEMcoref: A multilingual coreference

dataset of fiction. Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 472–480. <https://aclanthology.org/2026.acl-short.39/>

PEER-REVIEWED DOMESTIC JOURNAL ARTICLES (KCI)

1. 구슬, 김가연, **김병준**. 「LLM 상용화 이후 한국 인문사회과학 학술 초록의 어휘 변화: 2004~2024년 KCI 초록 데이터 기반 분석」. *언어와 정보*, 30(1): 1–24, 2026. doi.org/10.29403/LI.30.1.1
2. 최주찬, **김병준**, 김경민, 허예슬. 「소설의 시대 31년, 그 백 번째 물음 — 데이터로 읽는 『현대소설연구』 1~99호」. *현대소설연구*, 100, 2025. doi.org/10.20483/JKFR.2025.12.100.9
3. **김병준**, 노대원. 「생성형 AI는 인문학 연구를 어떻게 바꿀까?」. *영주어문*, 59, 2025. doi.org/10.30774/yjll.2025.2.59.203
4. 최진석, 최새현, **김병준**, 허예슬, 최주찬, 황호덕. 「두 박완서, 제도와 젠더 혹은 디지털 작가 연구의 시작 — 2010년대의 『여성문학연구』를 중심으로」. *개념과 소통*, 34, 2024. doi.org/10.15797/concom.2024..34.004
5. 서재현, 하나, 김민우, 한지민, **김병준**, 박소정. 「멀리서 읽는 “저희” — ELECTRA 기반 의미 분류 모델과 N-gram을 활용한 두 가지 ‘저희’의 통시적 변화 양상 연구」. *대동문화연구*, 128, 2024. doi.org/10.18219/ddmh..128.202412.513
6. (KCI 미등재) **김병준**. 「근대 국한문혼용체 자료 서브워드 기반 형태소 분석기의 설계와 적용」. *디지털인문학*, 1(2), 2024. doi.org/10.23287/KJDH.2024.1.2.5
7. (우수등재) 장우리, **김병준**, 전봉관. 「한국 최초 신문연재만화 『명탕구리』의 서사 구조와 문화사적 의미 연구」. *인문콘텐츠*, 74, 2024. doi.org/10.18658/humancon.2024.09.231
8. 이윤호, **김병준**. 「한국 학술장 내 웹툰 관련 연구의 확장 and 분화: KCI 논문 서지 데이터 기반 구조적 토픽 모델링 분석, 2003~2023」. *애니메이션연구*, 20(3), 2024. doi.org/10.51467/ASKO.2024.09.20.3.80
9. (KCI 미등재) **김병준**. 「디지털인문학의 미래, 스마트 빅데이터: 독일 트리어 디지털인문학 센터 CLS 연구 펠로우 후기」. *Digital+Humanities*, 2024. doi.org/10.5281/zenodo.13293696
10. 허예슬, **김병준**, 최주찬, 최진석. 「푸코의 초상: 한국 현대문학 학술장의 푸코 인용 양상 변화, 2008~2021」. *사회와 이론*, 46, 2023. doi.org/10.17209/st.2023.11.46.305
11. **김병준**, 전봉관. 「민족, 국민, 국가 — 시계열 워드 임베딩을 활용한 조선일보 기사의 민족 담론 의미 변동 추적 (1920~40)」. *현대소설연구*, 90, 2023. doi.org/10.20483/JKFR.2023.06.90.5
12. (우수등재) 전준, **김병준**, 김재홍, 김란우. 「전산사회과학 연구과정의 블랙박스 열기: 아카데미 데이터베이스를 활용한 비교사회학 연구를 중심으로」. *한국사회학*, 57(2): 131–157, 2023. doi.org/10.21562/kjs.2023.05.57.2.131
13. 최진석, **김병준**, 허예슬, 최주찬, 황호덕. 「김윤식과 우리 시대, 인용의 인구사회학적 시작: 현대문학연구자의 성별 및 세대별 김윤식 저술 인용 양상 연구(2004~2019)」. *국제어문*, 96, 2023. doi.org/10.31147/IALL.96.12
14. Yang, J., **Kim, B.**, & Lee, W. (2022). Media coverage of the conflicts over the 4th Industrial Revolution in the Republic of Korea from 2016 to 2020: A text-mining approach. *Asian Journal of Innovation and Policy*, 11(2), 202–226. <https://doi.org/10.7545/ajip.2022.11.2.202>
15. 서재현, **김병준**, 김민우, 박소정. 「멀리서 읽는 “우리” — Word2Vec, N-gram을 이용한 근대 소설 텍스트 분석」. *대동문화연구*, 115, 2021. doi.org/10.18219/ddmh..115.202109.349
16. **김병준**, 천정환. 「박사학위 논문(2000~2019) 데이터를 통해 본 한국 현대문학 연구의 변화와 전망」. *상허학보*, 60, 2020. doi.org/10.22936/sh.60..202010.012
17. 전성규, **김병준**. 「디지털인문학 방법론을 통한 『서북학회월보』와 『태극학보』의 담론적 상관관계 연구」. *개념과 소통*, 23, 2019. doi.org/10.15797/concom.2019..23.005

18. **김병준**, 전봉관, 이원재. 「비평 언어의 변동: 문예지 비평 텍스트에 나타난 개념단어의 변동 양상, 1995~2015」. *현대문학의 연구*, 61, 2017. doi.org/10.35419/kmlit.2017..61.002
19. 이원재, **김병준**, 전봉관. 「작가-비평가 관계와 비평가의 구조적 위치가 소설 단행본 판매량 증감에 미치는 영향: 2010~2015」. *한국현대문학연구*, 48, 2016. [KCI: ART002106457](https://kci.go.kr/ART002106457)
20. 전봉관, **김병준**, 이원재. 「문예지를 매개로 한 한국 소설가들의 사회적 지형: 1994~2014」. *현대소설연구*, 61, 2016. [KCI: ART002092223](https://kci.go.kr/ART002092223)

OTHER SCHOLARLY PUBLICATIONS (ROUNDTABLES, REPORTS & ESSAYS)

1. **김병준**, 유인혁, 김화선, 최성민, 허민석, 배하은, 김완중. 「분업과 연대, 학술 공공성의 새로운 형식들: 한국현대문학 편집인의 열린 대화 (좌담)」. *한국학*, 48(1): 107-131, 2025. doi.org/10.25024/ksq.48.1.202503.107
2. **김병준**. 「학위논문 서지 데이터 분석을 통해 본 한국 현대문학 연구의 변화와 전망」. *상허학보*, 74, 2025. doi.org/10.22936/sh.74..202506.016
3. 최진석, **김병준**, 서희경, 이용희, 이해령, 홍덕구. 「〈특집〉 제1회 한국현대문학자 대회 — 좌담: 학회란 무엇인가」. *사이월 SAJ*, 2024. [KCI: ART003087807](https://kci.go.kr/ART003087807)
4. **김병준**. 「계량서지학 기반 개념사 연구의 가능성」. *개념과 소통*, 28, 2021. [KCI: ART002803884](https://kci.go.kr/ART002803884)

BOOKS & EDITED VOLUMES

1. 이하람, 김희진, 강수진, 김용수, **김병준**, 박언주, 설연지, 윤미선, 장성현, 전보미, 정서현. 『디지털 인문학의 길잡이: 입문』 (가제). 서울대학교출판문화원, 2026 (근간).
2. 고부열, **김병준**, 남지만, 박철우, 신원철, 안광호, 전현주, 정진한, 최형근, 허인영, 허철. 『동·서 디지털 인문학, 구축·활용·번역』. 동문연, 2026. ISBN 9791124374047. [안양대학교 HK+ 동서교류문헌총서 7]
3. 양소은, 최석영, **김병준**. 『생성형 AI 활용 격차와 AI 리터러시』. 한국언론진흥재단, 2025.
4. **김병준**. 「디지털 인문학: 한국 학술 데이터의 새로운 지평」. 『한국학의 현재와 미래』, 이영준, 김동규, 오영진 엮음. 민음사, 2025.
5. 전봉관, 장우리, 이서준, **김병준**. 『멍텅구리: 한국 최초 신문 연재 네컷만화로 100년 전 날것의 식민지 조선을 보다』. 더숲, 2024. ISBN 9791194273080.
6. 류인태, 박지은, 권기성, 김바로, **김병준**, 김지선, 박진호, 양승목, 이민철, 이재연, 장문석, 지영원, 한희연. 『디지털로 읽고 데이터로 쓰다: 디지털 한국어문학의 모색』. 박문사, 2023. ISBN 9791192365268.
7. 이재열, 하상응, 임동균, 이원재, **김병준**, 조은아, 강정한, 이호영, 한준. 『플랫폼 사회가 온다: 20대와 50대의 단절과 전승』. 한울아카데미, 2021. ISBN 9788946080744.
8. 김장현, 김윤환, 이겨레, **김병준**. 『데이터기반 디지털 사회과학: 교육과 연구사례』. 성균관대학교출판부, 2019. ISBN 9791155503850.
9. 박건철, 김장현, **김병준**, 이겨레. 『도시데이터표준분석모델: 민원분석편』. 서울디지털재단, 2019. ISBN 9791196601584.

DATASETS

1. Kim, G., Park, S., Ji, H., Lee, H., Lee, B., Jeong, C., Lee, J.-y., Jo, G., Lim, I., Ismayilov, O., and **Kim, B.** (2026). KNoTE (Korean Novel TEI-Encoded) dataset, v4 [Data set]. Zenodo. doi.org/10.5281/zenodo.19813475

2. Lim, I., Ji, H., Koo, S., Jung, S.-y., Yun, J., and **Kim, B.** (2025). KPoEM (Korean Poetry Emotion Mapping) dataset, v4 [Data set]. Zenodo. doi.org/10.5281/zenodo.15598092

GRANTS & FUNDED PROJECTS

- **(PI)** The Gobaek (Confession) Benchmark: An Evaluation Dataset for Self-Rationalization of Colonial Male Intellectual Personas. Research-Education Linkage Program, The Academy of Korean Studies. KRW 16,000,000. Sep. 2026 – Feb. 2027.
- **(PI)** Knowledge Graphs of Literary Crossings: Modeling European-Korean Literary Translation History Using Linked Open Data. Korea-Germany R&D Network Program, National Research Foundation of Korea. KRW 20,000,000. Mar. 2026 – Feb. 2027.
- **(PI)** TEI/XML Construction Methodology for Korean Modern Literature in Humanities Data Design Education. Research-Education Linkage Program, The Academy of Korean Studies. KRW 16,000,000. Sep. 2025 – Feb. 2026.
- **(PI)** Development of Advanced Natural Language Processing and Large Language Model-Based Digital Korean Studies and Education Methodology. Research-Education Linkage Program, The Academy of Korean Studies. KRW 16,000,000. Mar. 2025 – Aug. 2025.
- **(Co-PI)** Exploring the Co-evolutionary Dynamics of Artificial Intelligence and Society: A Computational Social Science Approach. Global Humanities and Social Sciences Convergence Research Program, National Research Foundation of Korea. KRW 660,000,000. Jun. 2024 – May 2027.
- **(PI)** Integrative Approaches in Computational Literary Studies and Global Knowledge Structures. CLS INFRA TNA Fellowship, European Union Horizon 2020. EUR 7,240. Mar. 2024 – May 2024.
- **(PI)** Building an Academic Corpus and Language Model in the Humanities and Social Sciences to Advance Korean AI. Creative-Challenge Research Program, National Research Foundation of Korea. KRW 140,000,000. Jun. 2023 – May 2025.

WORKING PAPERS (UNDER REVIEW / IN PROGRESS)

- "Who Carries the Theory?: A Computational Analysis of Conceptual Travel between Korean and English Literary Studies in South Korea." Invited article for a special issue of *Journal of Cultural Analytics* (Scopus). Minor revision. Corresponding author.
- "Peripheralizing Social Science Academia: A Relational Landscape of the Global and Local Social Science." Under review at *American Journal of Sociology* (SSCI). Co-first author.
- "An LLM Construction Study of the Veritable Records of the Joseon Dynasty (조선왕조실록 LLM 구축 연구, working title)." In preparation for *International Journal of Korean History* (Scopus). Co-author.

SELECTED CONFERENCE PRESENTATIONS

- **Kim, B.**, Ji, H., Kim, G., Jeong, C., Lee, H., & Park, S. (2026, July). Building a Korean literature TEI/XML database: From data acquisition to AI-enhanced reading [Workshop]. ADHO Digital Humanities Conference (DH2026), Daejeon, South Korea. doi.org/10.5281/zenodo.21830474
- 김가연, 박선영, **김병준**. 「2008~2024 KCI 데이터로 본 한국문학/한국영문학 담론의 시계열적 비교 분석」. 제3회 한국현대문학자대회, Feb. 2026. doi.org/10.5281/zenodo.18738333

- Lim, I., & **Kim, B.** (2026, January). Co-Reading: A human-AI co-reader media system for poetry-emotion-color [Paper presentation]. Hong Kong Association for Digital Humanities Conference (HKADH2026), Hong Kong. <https://doi.org/10.5281/zenodo.18753718>
- **Kim, B.**, & Kim, Y. (2025, July). Digital intellectual history of modern Korean literary studies: Bibliometric analysis of KCI and OpenAlex data sets [Short Paper]. ADHO Digital Humanities Conference (DH2025), Lisbon, Portugal. <https://doi.org/10.5281/zenodo.16404466>
- Lim, I., Ji, H., & **Kim, B.** (2025, April). Decoding the poetic language of emotion in Korean modern poetry: Insights from a human-labeled dataset and AI modeling [Paper presentation]. Digital Humanities and Social Sciences Korea Conference, Philadelphia, PA, United States.
- **김병준.** 「AI 대전환, 오픈엑세스 학술 생태계에 가져올 변화」. 2024 Open Access Conference, Seoul, Oct. 2024 (oral).
- **Kim, B.**, & Jeon, J. (2024, May). Varieties of Korean studies, now and then: A digital intellectual history of Korean studies, 2000–2023 [Paper presentation]. Keimyung International Conference on Korean Studies (KICKS), Daegu, South Korea.
- **Kim, B.** (2024, May). OpenAlex database for computational literary studies and digital intellectual history [Invited talk]. TCDH Research Colloquium, Trier, Germany.
- **Kim, B.**, & Jun, B. G. (2023, June). Quantitative methodology on the success factors of Squid Game based on worldwide audience responses [Paper presentation]. 31st AKSE Conference, Copenhagen, Denmark.

Additional conference presentations are listed at byungjunkim.com/talks and on [Google Scholar](https://scholar.google.com/citations?user=byungjunkim).

TEACHING EXPERIENCE

- Exploring Humanities Data Design, Cultural Informatics, The Academy of Korean Studies (Fall 2025).
- Introduction to Digital Humanities: Foundations of Methodology, Cultural Informatics, The Academy of Korean Studies (Fall 2025).
- Methodology of Digital Korean Language and Literature Research 2, Graduate School, Dept. of Korean Language and Literature, Sungkyunkwan University (Spring 2025).
- Methodology of Digital Content Compilation in Foreign Languages, Cultural Informatics, The Academy of Korean Studies (Spring 2025).
- Research on Humanities Information Compilation 1, Cultural Informatics, The Academy of Korean Studies (Spring 2025).
- Computer Programming for Humanities Knowledge, Cultural Informatics, The Academy of Korean Studies (Fall 2024).
- Social Network Analysis for Digital Humanities, Hallym University (2024).
- Introduction to Digital Humanities II: Computational Literary Studies, Korea University (Fall 2023; selected as a Best Course).
- Literature as Data: The Potential of Quantitative Text Analysis, Winter Institute in Digital Humanities, Korea (Jan. 2023).
- Digital Humanities and Natural Language Processing courses, Sungkyunkwan University (2020–2021).

- Digital Humanities courses, Hankuk University of Foreign Studies (Spring 2021); AI and VR liberal arts courses, Baekseok University (2020–2021).

TECHNICAL SKILLS & METHODS

- **Programming and data analysis:** Python, R, SQL; reproducible, version-controlled data workflows.
- **NLP and machine learning:** topic modeling (STM, BERTopic), word and dynamic embeddings, text classification (ELECTRA/BERT), deep learning, LLM/RAG pipelines.
- **Bibliometrics and network analysis:** KCI, OpenAlex, citation and reference data; Gephi, networkx, igraph, VOSviewer, CiteSpace.
- **Knowledge organization and data infrastructure:** corpus and dataset construction, metadata and TEI/XML, Linked Open Data (RDF/SPARQL), RDBMS, NoSQL, web scraping and crawling.
- **Visualization:** statistical graphics, network visualization, data storytelling for the humanities and social sciences.

PROFESSIONAL SERVICE

- Local Organizer and Technical Support Committee Chair, ADHO Digital Humanities Conference 2026 (DH2026), Daejeon, 2025 – 2026.
- Co-President, Knowledge Sharing Alliance (지식공유연대; open-access advocacy coalition of Korean scholarly societies), 2019 – Present.
- Board of Directors and Editorial Board, Korean Association for Digital Humanities (KADH), 2022 – Present.
- Operating Committee Member, Korean Modern Literature Scholars' Conference, 2023 – Present.

AWARDS & HONORS

- 2021 SKKU Graduate Student Paper Award, Sungkyunkwan University (Oct. 2021).
- Best Paper Award, 2021 National Graduate Student Conference, Korean Society for Journalism and Communication Studies (Aug. 2021).
- BK21+ Scholar, National Research Foundation of Korea (2019–2020); Graduate Scholarship, Sungkyunkwan University (2018–2020); National Scholar, KAIST (2015–2017).

REFERENCES

Dr. Jang Hyun Kim, Professor of Data Science, Sungkyunkwan University (doctoral advisor).
alohakim@skku.edu

Dr. Wonjae Lee, Professor of Sociology, KAIST (master's advisor). wnjlee@kaist.ac.kr

Dr. Bong Gwan Jun, Professor, School of Digital Humanities and Computational Social Sciences, KAIST (master's advisor). junbg@kaist.ac.kr