Donghyeok Choi

dhchoi.net

EDUCATION

Ph.D. Graduate School of Culture Technology, KAIST, 2024

Thesis Title: Bureaucratic Success in the Joseon Dynasty: A Computational Study of Intergenerational Mobility Patterns (Advisor: Juyong Park)

M.E. Graduate School of Culture Technology, KAIST, 2018

B.E./B.A. Computer Science and Engineering & History (double major), Sungkyunkwan University, 2016

PUBLICATIONS

2026 **Donghyeok Choi**, Javier Cha, Juyong M. Park†, "Total Success Index and the Longitudinal Dynamics of Bureaucratic Stratification in Joseon Korea", *Physica A: Statistical Mechanics and its Applications*, 688, 131353. DOI: 10.1016/j.physa.2026.131353 (Elsevier; SCIE; IF 3.3; Scimago Q2)

2026 Yan Hon Michael Chung, **Donghyeok Choi**, "Fine-tuning Vision-Language Models as OCR Systems for Low-Resource Languages: A Case Study of Manchu", *Computational Humanities Research*, 2, e20. DOI: 10.1017/chr.2026.10042 (Cambridge University Press)

2026 **Donghyeok Choi**, ChangHwan Kim†, "Meritocracy with Multiple Rules: How Special Civil Service Examinations Reproduced Privilege in Joseon Korea", *Social Science History*. DOI: 10.1017/ssh.2026.10135 *(first author; corresponding author: ChangHwan Kim)* (Cambridge University Press; SSCI, A&HCI, Scopus; Scimago Q1 (History))

2026 **Donghyeok Choi**, "Lifespans, Exam-Passing Age, and Careers of Joseon Yangban: A Prosopographic Analysis of Linked Biographical Records", *The Review of Korean Studies*, 29(1). DOI: 10.25024/review.2026.29.1.002 (The Academy of Korean Studies; Scopus, ESCI, KCI)

2024 **Donghyeok Choi**, "Bureaucratic Success in the Joseon Dynasty: A Computational Study of Intergenerational Mobility Patterns". *PhD Dissertation, Graduate School of Culture Technology, KAIST*

APPOINTMENTS

Full-time

2025– Postdoctoral Fellow, Department of History, Hong Kong Baptist University

2024–25 Postdoctoral Fellow, Department of History, The University of Hong Kong

Part-time

2024–26 Lecturer, Digital History in Global Asia Program, Department of History, Lingnan University

2023–24 Lecturer, College of Liberal Arts, Korea University

2022–23 Lecturer, College of Liberal Arts, Sungkyunkwan University

TEACHING EXPERIENCE

2025–26 Digital Data Management in Asian Studies, Lingnan University

2025–26 Capstone Advisor, Digital History in Global Asia Program (4 students), Lingnan University

2024 Introduction to Digital History in Global Asia, Lingnan University

2024 Artificial Intelligence Software Developer Training Program, Saltlux

2023 Introduction to Digital Humanities II, Korea University

2023 Seocho AI College Full Stack Developer Course, KAIST IT Academy

2022 Korean Artificial Intelligence Data Science, Sungkyunkwan University

TECHNICAL SKILLS

AI & Machine Learning Fine-tunes vision-language models for OCR and large language models for named-entity recognition and record linkage, then integrates them into human-in-the-loop research pipelines for training-data creation, retrieval, and evaluation; builds agent infrastructure for LLM-driven software and data workflows, including tool integration via the Model Context Protocol (MCP), persistent memory, and multi-agent orchestration.

Research Computing & Infrastructure Designs and operates research computing infrastructure, with hands-on experience across hardware selection, procurement, installation, and administration of an NVIDIA H200 system at HKBU; performs distributed model training on GPU clusters using data and model parallelism, and deploys containerized training and inference workloads across on-premises and cloud environments.

Software & Data Engineering Builds full-stack research software and data platforms in Python and JavaScript/TypeScript, drawing on years of research and freelance work across web development, data processing, analysis, and visualization; specializes in database architecture and schema design for complex humanities data, using PostgreSQL as the default and MongoDB or Neo4j when the data model calls for them.

CURRENT RESEARCH PROJECTS

2026– **Bandit: A Human-in-the-Loop OCR Platform for East Asian Historical Documents**
PI and Lead Developer | Self-initiated
Develops an end-to-end OCR platform that converts difficult East Asian historical documents into verified, reusable data. It combines fine-tuned vision-language models, self-managed GPU infrastructure, and an expert correction interface that feeds human review back into training and evaluation.

2025– **Yaudang: A Multilingual OCR Platform for Hong Kong Colonial Archives**
PI and Lead Developer | Self-initiated
Builds a multilingual research platform for Hong Kong's colonial archives, unifying English, traditional Chinese, and Japanese sources across heterogeneous page layouts. It couples vision-language OCR with expert review to turn fragmented collections into a structured corpus and reusable infrastructure for computational Hong Kong history.

2025– **Playing Heaven: Remapping Early Modern Neo-Confucian Worlds with AI**
Lead Developer | Schmidt Sciences, Humanities and AI Virtual Institute (HAVI), HKD 5,300,000 (PI: Javier Cha, University of Hong Kong)
Architects the project's research software and AI workflows, translating ambitious questions about early modern Neo-Confucian textual worlds into a cross-institutional research program that keeps historical inquiry, data, models, and interpretation in one coherent system.

2025– **Nabi / Nabi X: AI-Powered Digital Humanities Education Platform**
Lead Developer | HKU Teaching Development Grants (TDG and TDLEG), HKD 800,000
Leads the development of an AI-powered teaching platform that turns digital humanities methods into guided, classroom-ready workflows. Designed the system from pedagogical concept to production deployment with support from two HKU teaching grants, creating reusable infrastructure for teaching across institutions.

Other Research Experience

2023–25 Complex Network Science-Based Investigation of the Korean Legal Code for Law and Development in the Post AI Era, National Research Foundation of Korea (NRF)

2021–22 Imagining Post-AI Culture Technology, KAIST

2020–21 AI Data Construction Project: Emotional Recognition and Video Summary, National Information Society Agency (NIA)

2021 AI Data Construction Project: Movie and Image Editing Data, National Information Society Agency (NIA)

2016–19 Social Network Effects on the Dynamics of Groups and Organizations, National Research Foundation of Korea (NRF)

PEER REVIEW

2027 HKADH 2027, Hong Kong

2026 Digital Scholarship in the Humanities

2026 Journal of Open Humanities Data (2 reviews)

2025 Digital Humanities 2025, Lisbon

2024 Digital Humanities 2024, Washington, DC

2023 Digital Humanities 2023, Graz

2022 Digital Humanities 2022, Tokyo

CONFERENCE PRESENTATIONS

2027 "From Merit to Lineage: The Changing Basis of Bureaucratic Career Attainment in Joseon Korea, 1393–1894", Association for Asian Studies 2027 Annual Conference, Boston, United States, Mar 2027 (forthcoming). **Conference Presentation**

2027 "Computational Reading with Sparse Autoencoders: Tracing Changes in Semantic Features", HKADH 2027 (Model/Making: Generating Digital Humanities), Hong Kong, 2027 (accepted). **Conference Presentation**

2026 "Bandit: A Human-in-the-Loop OCR System for East Asian Historical Documents with Vision-Language Model Fine-tuning", JADH (Japanese Association for Digital Humanities) 2026, Kyushu University, Fukuoka, Japan, Sep 2026. **Conference Presentation**

2026 "Playing Heaven: Remapping Early Modern Neo-Confucian Worlds with AI", Digital Humanities 2026 (ADHO), Mini-conference, Daejeon, Republic of Korea, Jul 2026. **Conference Presentation**

2026 "Contours of Discourse: Visualizing Seminal Works in East Asian Digital Humanities as Word Embeddings", Association for Asian Studies 2026 Annual Conference, Vancouver, Canada, Mar 2026. **Poster Presentation**

2025 "A Prosopographical Analysis of Yangban Lifespans and Career Trajectories in Joseon Dynasty", Digital Humanities and Social Sciences Korea Conference, University of Pennsylvania, Apr 2025. **Conference Presentation**

2025 "Mapping Power Shifts: Topological Analysis of Communication Networks During the Gyeyu Revolt in Joseon Korea", HKADH, Hong Kong, Jan 2025. **Conference Presentation**

2023 "Unraveling the Joseon Dynasty: A Quantitative Approach to Historical Dynamics", CUDAN, Tallinn, Estonia, Dec 2023. **Conference Presentation**

2023 "Historical Dynamics of the Kingdom of Chosŏn's Governance: Patterns of Meritocratic Bureaucracy and Consequences of Systemic Corruption", Netsci, Vienna, Austria, Jul 2023. **Conference Presentation**

2022 "Historical Data Engineering: Preprocessing the Annals of the Chosŏn Dynasty and the Examination Rosters", Advanced Topics in Digital Korean Studies, Seoul, Republic of Korea, May 2022. **Conference Presentation**

2020 "Entry and Rise In Bureaucracy: Patterns of Career Advancement in Joseon Dynasty", Digital Humanities, Virtual Conference, Jul 2020. **Poster Presentation**

2018 "Lifelong Career Success: Joseon Dynasty's Bureaucrats", COMPLEX NETWORKS Cambridge, United Kingdom, Dec 2018. **Conference Presentation**

2017 "Measuring Success from Historical Data: The Annals of the Joseon Dynasty", THE 4TH SATELLITE ON QUANTIFYING SUCCESS, Indianapolis, United States, Jun 2017. **Conference Presentation**

INVITED TALKS

2026 "Bandit: A Human-in-the-Loop OCR System for East Asian Historical Documents", Korean Association for Digital Humanities (KADH) 2026 Conference, Online, Nov 2026 (forthcoming). **Invited Talk** (host: Jae-Yon Lee)

2026 "Manchu OCR, AI, and the Future of Qing Archival Research", Institute of History and Philology, Academia Sinica, Taipei, Taiwan, Aug 2026. **Invited Talk** *Workshop on Manchu Archival Document Reading and Multiethnic History Research* (host: Ling-Wei Kung)

2026 "From Reading to Discovery: AI-Assisted Workflows for East Asian Historical Texts", Hanyang University, Seoul, Republic of Korea, Jul 2026. **Invited Talk** *Global Digital History Workshop* (host: Soo-Hyun Mun)

2026 "From Reading to Discovery: AI-Assisted Workflows for East Asian Historical Texts", CrossAsia, Staatsbibliothek zu Berlin, Jun 2026. **Invited Talk** *CrossAsia DH Lunchtalks*

2026 "From Reading to Discovery: AI-Assisted Workflows for East Asian Historical Texts", KAIST, Apr 2026. **Online Guest Lecture** *Introduction to Digital History; Introduction to Digital Humanities* (host: Donghyun Woo)

2025 "Fine-tuning Large-Language Models for Humanities Research", Hong Kong University of Science and Technology, Hong Kong, Oct 2025. **Invited Workshop** *Digital Humanities Open Workshop Series* (host: Yan Hon Michael Chung)

2025 "Quantifying Bureaucratic Success in Joseon Korea: A Digital Analysis of Gwageo Passers and Their Career Trajectories with Implications for Ming-Qing Keju Studies", China Studies Forum, Shanghai, China, Oct 2025. **Invited Talk** *Panel on Digital Intelligence and Chinese Studies, organized by Renmin University of China* (host: Jing Zhang)

2025 "Quantifying Historical Success: Digital Prosopography and the Bureaucratic Success Index", Hong Kong Baptist University, Hong Kong, Jun 2025. **Invited Workshop** *CHRP Digital Humanities Workshop Series 4; organized by Academy of Chinese, History, Religion and Philosophy*

2024 "Data Visualization", The University of Hong Kong, Hong Kong, Nov 2024. **Guest Lecture** *HUDT1002* (host: Jenny Kwok)

2024 "Knowledge Graph and Network Analysis Tools", Seoul National University, Apr 2024. **Guest Lecture** *Seminar in Application of Cognitive Science (Digital Humanities Theory and Practice)* (host: Eunsoo Lee)

AWARDS AND HONORS

Recognition for "The Land of Extinct: How Does the Local City Collapse"

Contributed to geographical data visualization, highlighting population dynamics and urban concentration trends in Korea. This work parallels my research on success concentration trends during the late Joseon Dynasty.

2023 UNESCO and Asia-Pacific Broadcasting Union's T4P Media Award, Grand Prize in the Digital Category (Link)

2022 SIGMA Data Journalism Award Shortlist (Link)

2021 Korea Data Journalism Award for Data Visualization of the Year (Link)

Further details and project overview available at Project Website.

Updated October 2026