

HOYEOL KIM

Machine Learning Data Linguist at AWS

hoyeolkim7@gmail.com

EDUCATION

Texas A&M University, College Station, USA

August 2017 - May 2022

Ph.D. in English with GR Certificate in Digital Humanities.

Dissertation: “Computational Approaches in the Humanities: from Sentiment Analysis to Deep Learning Colorization.”

Committee (English & Digital Humanities): Maura Ives (chair), Amy Earhart, Shawna Ross.

Committee (Computer Science & Engineering): Ruihong Huang.

University of Exeter, Exeter, United Kingdom

September 2013 - November 2014

M.A. in English Language and Literature.

Hankuk University of Foreign Studies, Seoul, Republic of Korea *March 2006 - February 2013*

B.A. Major in English Literature, Minor in English Linguistics.

COMPUTER SKILLS AND NATURAL LANGUAGES

Programming/Scripting Languages

Python, R, Swift, SAS, C++

Markup Languages

HTML, CSS, XML/TEI, L^AT_EX

Deep Learning Frameworks

PyTorch, TensorFlow

Web Frameworks/Platforms

Jekyll, WordPress, Django, Docker, AWS (S3 & EC2)

Databases

SQL, PostgreSQL, MySQL, SQLite

Operating Systems

macOS, Ubuntu, Windows, Raspbian

Natural Languages

Korean, English, Japanese (N3)

PEER-REVIEWED PUBLICATIONS

Journal Articles

“Sentiment Analysis: Limits and Progress of the Syuzhet Package and Its Lexicons.” *Digital Humanities Quarterly*, vol. 16, no. 2, 2022. digitalhumanities.org/dhq/vol/16/2/000612/000612.html.

“*Victorian400*: Colorizing Victorian Illustrations.” *International Journal of Humanities and Arts Computing*, vol. 15, no. 1-2, Oct. 2021, pp. 186-202. DOI: 10.3366/ijhac.2021.0269.

“Considering Kim’s Dual Identity in the Post-colonial Discourse.” *Journal of British and American Studies*, vol. 34, 2015, pp. 33-53.

Book Chapters

“The Challenges of Collaboration: Pursuing Computational Research in a Humanities Graduate Program.” *The Digital Futures of Graduate Study in the Humanities*, edited by Simon Appleford, Gabriel Hankins, and Anouk Lang, Part of the *Debates in the Digital Humanities* Series, U of Minnesota P (*Forthcoming*).

Tutorials

With Sungjin Ahn, Minji Kang, Donghwan Kim, and JunHyoung Park. “Deep Learning Tutorials with PyTorch.” <https://pseudo-lab.github.io/Tutorial-Book/>.

Encyclopedia Entries

The NAVER Encyclopedia project, led by Dr. Dongil Lee, was a commercial digital project that a number of scholars participated in. NAVER is one of the biggest tech companies in South Korea.

NAVER History Encyclopedia: the United Kingdom, edited by Dongil Lee, Nov. 2015.

- “Queen Victoria and 30 Scottish Kings and Queens.”

<http://terms.naver.com/entry.nhn?docId=3326637&cid=56790&categoryId=58124>

NAVER Literary Encyclopedia: Looking Closely at Literature, edited by Dongil Lee, Nov. 2013.

- “*Dubliners* by James Joyce.”

<http://terms.naver.com/entry.nhn?docId=2070515&cid=267&categoryId=1075>

- “*The Moonstone* by Wilkie Collins.”

<http://terms.naver.com/entry.nhn?docId=1996874&cid=267&categoryId=1075>

- “*The Return of the Native* by Thomas Hardy.”

<http://terms.naver.com/entry.nhn?docId=1994729&cid=267&categoryId=1075>

- “*New Grub Street* by George Gissing.”

<http://terms.naver.com/entry.nhn?docId=1994732&cid=267&categoryId=1075>

- “*The Time Machine* by Herbert George Wells.”

<http://terms.naver.com/entry.nhn?docId=1996900&cid=267&categoryId=1075>

EMPLOYMENT

Machine Learning Data Linguist

July 2022 - Present

Amazon Web Services, Texas, United States (Virtual)

Multilingual NLP Analyst (Contract)

April 2022 - July 2022

Deepgram, San Francisco, California, United States (Remote)

- I evaluated and improved the performance and outputs of end-to-end deep learning ASR models.

- I improved text cleaning pipelines, evaluated multilingual data for variance in labeling, and proposed better approaches for multilingual model architecture.

Graduate Assistant, Teaching (Instructor)

August 2017 - May 2022

Department of English, Texas A&M University, College Station, TX

Project Manager & iOS Developer (Part-time)

November 2017 - October 2018

Gindo Solution, Arlington, Virginia, United States (Remote)

- I managed projects and guided UI/UX designs and function developments for mobile apps. I also developed notification, point system, and payment functions.

Marketing Lead

May 2016 - November 2017

Tulango, Yonsei University, Seoul, Republic of Korea

- I worked on language data for automatic tests, localized the app languages in English, Spanish, and Korean, managed online and offline app ads, and developed public/private partnerships.

Administrative Coordinator

October 2014 - July 2015

Graduate School of TESOL, Cyber Hankuk University of Foreign Studies, Seoul, Republic of Korea

- I worked on admission and graduation processing, annual budgets, and course registration.

Staff Sergeant

April 2007 - June 2009

Republic of Korea Air Force Headquarters

- Field of Expert: Information Systems Management.

CONFERENCES

“The Application of Sentiment Analysis in Literary Studies.” KADH2021, Online, November 2021.

“The Challenges and Considerations of Computational Literary Analysis in Different Languages.” ACH2021, Online, July 2021.

With Minji Kang, Sungjin Ahn, Donghwan Kim, and Sung Won Han. “The Automatic Generation of Contextual Questions and Answers for English Learners” ACH2021, Online, July 2021.

<https://hcommons.org/deposits/item/hc:40637/>.

“Dickensian Sentiment and Sentiment Analysis of Victorian Novels.” Dickens Society Symposium 2021, Online, July 2021.

<https://www.victorianweb.org/authors/conferences/dickens2021.html>.

“Pursuing Computational Humanities Research as a Graduate Student in the Humanities.” Engaging Open Social Scholarship, Online, December 2020.

<https://inke.ca/engaging-open-social-scholarship/abstracts/#Kim-Hoyeol>.

With Maura Ives. “Colorization of Illustrations in Charles Dickens’ Novels Using Deep Learning.” DH2020, Ottawa, Canada, July 2020. https://dh2020.adho.org/wp-content/uploads/2020/07/183_ColorizationofIllustrationsinCharlesDickensNovelsUsingDeepLearning.html. (*Canceled due to COVID-19*)

“Victorian400 Dataset for Colorizing Victorian Illustrations.” Keystone DH2020, Philadelphia, USA, July 2020. (*Canceled due to COVID-19*)

“Limits and Progress of the Syuzhet Package and Its Lexicons.” ACH2019, Pittsburgh, USA, July 2019. <https://elibooklover.github.io/SentimentAnalysisSyuzhet/>.

“Digital Analysis of Dickensian Characters and the Language of Emotion.” Dickens Society Symposium 2018, Tübingen, Germany, July 2018.

TEACHING EXPERIENCE (SELECTED)

Instructor at Digital Humanities Summer Institute, University of Victoria

Deep Learning for Humanists (<https://dhsi.org/dhsi-2022-workshops/>) *June 2022, June 2023*

Graduate Assistant, Teaching (Instructor)

Department of English, Texas A&M University

ENGL 304: Topics in Digital Research *Spring 2021*

Course Website: <https://elibooklover.github.io/ENGL304-Spring2021/>

Tutorials: <https://elibooklover.github.io/Tutorials/>

ENGL 210: Technical Professional Writing *Spring 2022, Fall 2021*

ENGL 210: Technical Business Writing *Summer 2021, Spring 2021, Fall 2019, Summer 2019*

ENGL 203: Writing about Literature *Fall 2020, Spring 2020*

ENGL 104: Composition & Rhetoric *Spring 2019, Fall 2018*

Graduate Assistant (Grader), Department of English, Texas A&M University

ENGL 210: Technical Business Writing *Spring 2018*

ENGL 231: Survey of English I *Fall 2017*

SCHOLARLY DIGITAL PROJECTS

Victorian400 Dataset *2019 - 2021*

Colorizing Victorian Illustrations with the Victorian400 Dataset.

Ranked with a bronze medal in Kaggle and indexed in Google Dataset Search.

- GitHub: <https://github.com/elibooklover/Victorian400>
- Kaggle: <https://kaggle.com/elibooklover/Victorian400>

VictorianLit Dataset *2020 - 2021*

For Deep Learning-Based Sentiment Analysis of Victorian Literary Texts.

Indexed in Google Dataset Search.

- GitHub: <https://github.com/elibooklover/VictorianLit>

- Kaggle: <https://kaggle.com/elibooklover/VictorianLit>

Web Companion 2.0 for *Using Digital Humanities in the Classroom* (Peer-Reviewed)

With Shawna Ross, Claire Battershill, and Andrew Pilsch.

2020

CERTIFICATES

Digital Humanities Graduate Certificate at Texas A&M University

May 2022

Successfully completed a digital project along with four DH courses (STAT604, ENGL603/604/685).

Digital Editions, Start to Finish

May 2019

Issued by the Center of Digital Humanities Research, Texas A&M University.

SAS Certified Base Programmer for SAS 9 (92% Pass)

November 2018

Certificate Verification Number: LQJZ1K12CBBEQ29V

Certificate of Completion: Encoding and Analyzing Digital Editions

December 2017

Issued by Initiative for Digital Humanities, Media, and Culture, Texas A&M University.

FUNDING AND ACADEMIC AWARDS (SELECTED)

Graduate Assistantship, Department of English, Texas A&M University.

2017 - 2022

Graduate Merit Funding of \$1200, Department of English, Texas A&M University.

2021

Victorian400 Dataset: Professional Development Grant of \$161, Department of English, Texas A&M University.

2021

Victorian400 Dataset: Project Development Grant of \$563, Center of Digital Humanities Research, Texas A&M University.

2020

Victorian400 Dataset: Graduate Merit Funding of \$400, Department of English, Texas A&M University.

2020

SERVICE TO ACADEMIC ORGANIZATIONS

Collaborative Development Editor at *Digital Humanities Quarterly*

2022 - Present

Reviewer. DH2023. Graz, Austria.

2022 - 2023

Chair. "Texts and Images." Session 4. DHSI 2022. Online.

2022

Reviewer. DH2022. Tokyo, Japan.

2021 - 2022

Reviewer. DH Unbound 2022. Online.

2022

External Reviewer. *Digital Humanities Quarterly*.

2020 - 2022

- Reviewed five DH articles.

Chair. "Dickens in Translation." Panel Session. Dickens Society Symposium 2021. Online.

2021

Reviewer. ACH2021. Online.

2021

Invited Speaker for the New Graduate Students in English at Texas A&M University.

2020

Korean Translator. Around DH (<https://github.com/go-dh/arounddh2020>).

2020

Korean Translator. Multilingual DH (<http://multilingualdh.org/>).

2020

Reviewer. DH2020. Ottawa, Canada.

2019 - 2020