Hoyeol Kim

Lead Machine Learning Data Linguist at AWS | iOS Developer

GitHub: github.com/elibooklover
SafeKey:apps.apple.com/app/safekey-password
-manager/id6475980444
NeonLED:apps.apple.com/us/app/neonled-led-
banner/id6474363843
Transliter: pypi.org/project/transliter

WORK EXPERIENCE

Amazon Web Services — *Lead Machine Learning Data Linguist*

July 2022 - PRESENT, Texas, USA

As part of the AI Data team, I evaluate transcriptions and improve annotation, segmentation, tools for data QA, and guidelines for data/QA vendors. I write scripts to analyze and manage ML data, in addition to computing metrics, and write code for a weekly/monthly report tool. I work on data QA and model evaluation for the automatic generation of clinical notes (HealthScribe) and for toxicity detection. I work on RLHF, Responsible AI, and RAG for Amazon Titan (LLM).

Deepgram — *Multilingual NLP Analyst*

April 2022 - July 2022, San Francisco, California, USA

I evaluated and improved the performance and outputs of end-to-end deep learning ASR models. I evaluated multilingual data for variance in labeling, proposed better approaches for multilingual model architecture, and improved text cleaning pipelines.

Texas A&M University — *Graduate Instructor*

August 2017 - May 2022, College Station, Texas, USA

I taught digital humanities and English courses including "Topics in Digital Research" (https://elibooklover.github.io/ENGL304-Spring2021/)

Gindo Solution — *Project Manager & iOS Developer*

November 2017 - October 2018, Arlington, Virginia, USA

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

Tulango — *Marketing Lead*

May 2016 - November 2017, Seoul, Republic of Korea

I worked on language data for automatic tests, localized the app languages in English, Spanish, and Korean, and managed online and offline app ads.

Cyber Hankuk University of Foreign Studies — *Administrative Coordinator*

Oct 2014 - July 2015, Seoul, Republic of Korea

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

Republic of Korea Air Force — *Staff Sergeant*

April 2007 - June 2009, Republic of Korea

Field of Expert: Information System Management

EDUCATION

Texas A&M University — *PhD in English with Digital Humanities*

August 2017 - May 2022, College Station, Texas, USA

Dissertation: "Computational Approaches in the Humanities: from Sentiment Analysis to Deep Learning Colorization." Supervised by Maura Ives, Amy Earhart, Shawna Ross, and Ruihong Huang.

Colorado State University Global — *Certificate in Computer Programming*

August 2023 - April 2024, Aurora, Colorado, USA

Overall GPA is 4.0: Completed courses: Programming I, II & III (Java & C++), Data Structures and Algorithms. Currently in progress: Software Engineering, Platform Based Development.

University of Exeter — *MA in English Literary Studies*

September 2013 - November 2014, Exeter, UK

Hankuk University of Foreign Studies — *BA in English Literature & Linguistics*

March 2006 - February 2013, Seoul, Republic of Korea

SKILLS

Experienced in programming languages, including Python, Swift, Java, R, and SQL, with a focus on NLP and mobile development. Familiar with Bash, Unix, and Linux environments. Engineering communication skills and deep understanding of machine learning (PyTorch). Highly attentive to detail.

JOURNAL ARTICLES & CONFERENCES (SELECTED)

"Sentiment Analysis: Limits and Progress of the Syuzhet Package and Its Lexicons." *Digital Humanities Quarterly*, vol. 16, 2022. digitalhumanities.org/dhq/vol/16/2/000612/000612.html.

"*Victorian400*: Colorizing Victorian Illustrations." *IJHAC*, vol. 15, no. 1-2, Oct. 2021, pp. 186-202. DOI: 10.3366/ijhac.2021.0269.

"*The Automatic Generation of Contextual Questions and Answers for English Learners.*" with Minji Kang, Sungjin Ahn, Donghwan Kim, and Sung Won Han, *ACH2021*, Jul. 2021. DOI: 10.17613/pp84-yr61.

WORKSHOP

Instructor, "Deep Learning for Humanists" at Digital Humanities Summer Institute, University of Victoria, Canada, June 2022| June 2023.

LANGUAGES & CERTIFICATE

Korean (Proficient), English (Proficient), Japanese (N3), SAS Certified Base Programmer for SAS 9 (LQJZ1K12CBBEQ29V)