Hoyeol Kim

hoyeol.kim@proton.me | github.com/elibooklover | linkedin.com/in/hoyeol-kim-phd-82723663

Lead ML Data Linguist at AWS | iOS Developer

Experience

Amazon Web Services – AI Data, *Lead Machine Learning Data Linguist* | Texas, USA   *Jul 2022 - Present*
- 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 the data evaluation for RLHF, Responsible AI, and RAG for Amazon Titan (LLM).

Deepgram, *Multilingual NLP Analyst* | Contract | San Francisco, California, USA   *Apr 2022 - Jul 2022*
- 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* | College Station, Texas, USA   *Aug 2017 - May 2022*
- I taught a course per semester, such as "Topics in Digital Research" (`elibooklover.github.io/ENGL304-Spring2021`)

Gindo Solution, *Project Manager & iOS Developer* | Part-time | Arlington, Virginia, USA   *Nov 2017 - Oct 2018*
- I managed projects and guided UI/UX designs and functional developments for both Android and iOS apps.
- I developed several functions, such as notifications with badges, point systems, and payment functions, using UIKit and RESTful APIs.

Tulango *Marketing Lead* | Seoul, South Korea   *May 2016 - Nov 2017*
- I worked on language data for automatic tests, localized the app languages in English, Spanish, and Korean.
- I managed both online and offline app ads. Through online advertising, I gathered around 2,000 users in two weeks.

Education

Texas A&M University, *PhD in English with Computational Literary Studies* | College Station, Texas, USA | Supervised by Maura Ives, Amy Earhart, Shawna Ross, and Ruihong Huang (CS)   *Aug 2017 - May 2022*
- Dissertation: "Computational Approaches in the Humanities: from Sentiment Analysis to Deep Learning Colorization."
- Sentiment Analysis - Syuzhet and its Lexicons: `digitalhumanities.org/dhq/vol/16/2/000612/000612.html`
- EDA of Victorian400 Dataset with cGANs: `euppublishing.com/doi/abs/10.3366/ijhac.2021.0269?journalCode=ijhac`

Colorado State University Global, *Certificate in Computer Programming* | Aurora, Colorado, USA   *Aug 2023 - Apr 2024*
- My overall GPA is 4.0. I took the following courses: Programming I, II, and III (Java & C++), Data Structures and Algorithms, Software Engineering, and Platform-Based Development (Kotlin).

University of Exeter, *MA in English Literary Studies* | Exeter, UK   *Sep 2013 - Nov 2014*

Hankuk University of Foreign Studies, *BA in English Literature & Linguistics* | Seoul, South Korea   *Feb 2013*

Skills

Programming   I use Python for data management and analysis, and Swift for iOS development. During my PhD research, I utilized Python, R, SAS, PyTorch, TensorFlow, Django, Jekyll, and PHP. I also learned Java, Kotlin, and C++ through coursework.

Software   I am familiar with Bash, Unix, and Linux environments, in addition to Git version control.

Certificates   I hold a SAS Certified Base Programmer for SAS 9 certificate and a Japanese Language Proficiency Test N3 certificate.

Projects

SafeKey – Password Manager (iOS)   *Jan 2024 - Present*
iOS App, Personal Project: `apps.apple.com/us/app/safekey-password-manager/id6475980444`
- A secure password manager that supports syncing through iCloud. I developed this app, using SwiftUI, SwiftData, CloudKit, StoreKit, Local Authentication, UserDefaults, and KitTextField.

NeonLED – LED Banner (iOS)   *Nov 2023 - Present*
iOS App, Personal Project: `apps.apple.com/us/app/neonled-led-banner/id6474363843`
- I developed this app in SwiftUI and localized it for languages, such as Korean and Japanese.

Transliter (Python)   *Apr 2023 - Present*
Python Library, Personal Project: `pypi.org/project/transliter`
- Transliter transliterates multilingual text into Latin script.
- Transliter supports Korean, Japanese, Russian, Ukrainian, Bulgarian, Macedonian, Mongolian, Montenegrin, Serbian, and Tajiki.

AI Data Monthly/Weekly Business Reports (Python)   *Apr 2023 - Present*
Python Scripts, AWS AI Data
- All the data are extracted through API calls, processed through pandas, and exported as CSV files.

Victorian400 (ML Dataset)   *Dec 2019*
ML Dataset, Research Project: `kaggle.com/datasets/elibooklover/victorian400`
- I created, curated, and shared 400 image files for conditional generative adversarial networks to train with for a colorization project.

K-Town Rewards (iOS)   *Nov 2017 - Oct 2018*
iOS App: `appadvice.com/app/k-town-rewards/1374362677`
- I developed the app in UIKit, implementing features like badge notifications, point systems, and payment functions.