

(Danny) To Eun Kim

Information Retrieval | NLP | Dialogue Systems | Knowledge Graphs | Search and QA

✉ to.kim.17@ucl.ac.uk

🎓 [Google Scholar](#)

🐙 <https://kimdanny.github.io/>

Education	University College London (UCL) M.Eng. Computer Science (4-year integrated master's degree) Minor: Modern Applications of Engineering Mathematics Thesis: Multi-Task Neural User Simulator for Task Oriented Dialogue System	Sep 2018 - Aug 2022 First-Class Honours
Ongoing Work	Response ranking with conversational look-ahead (Targeting SIGIR '23 - 1st author) User satisfaction estimation with Hawkes Transformer (Targeting ACL '23 - 2nd author) Interactive Grounded Language Understanding (IGLU) - NeurIPS 2022 challenge Automated Clustering of Smell Stories (Targeting CHI '23 - 1st author)	
Publications	A Multi-Task Based Neural Model to Simulate Users in Goal-Oriented Dialogue Systems T.E. Kim , A. Lipani [paper] [poster] [code] <i>45th ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)</i> , 2022 Conditia: A State Machine Like Architecture for Multi-Modal Task Bots J. Ramos* , T.E. Kim* , Z. Shi , X. Fu , F. Ye , Y. Feng , A. Lipani [paper] <i>Alexa Prize TaskBot Challenge Proceedings</i> , 2022 Attention-based Ingredient Phrase Parser Z. Shi , P. Ni , M. Wang , T.E. Kim , A. Lipani [paper] [code] <i>European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)</i> , 2022	
Workshops & Presentations	NLP Interest Group at The Alan Turing Institute, <i>Can Dialogue Systems Have a Theory-of-Mind?</i> , 24th Nov., 2022. [slides] Knowledge Graphs Interest Group at The Alan Turing Institute (5th meet-up), <i>Personal Knowledge Graphs: Towards Better Conversational Assistants</i> , 27th Oct., 2022. [slides] Search-Oriented Conversational AI Workshop (SCAI '22), <i>A Multi-Task Based Neural Model to Simulate Users in Goal-Oriented Dialogue Systems</i> , 15th Jul., 2022. [recording] [slides]	
Research Experience	UCL Web Intelligence Group - Research & Teaching Assistant Conducting research with Prof. Emine Yilmaz and Dr. Aldo Lipani . <ul style="list-style-type: none">Highly contributed (first authored) to the <i>Alexa Prize TaskBot Challenge</i>, taking charge of engineering/DevOps challenges, such as AWS services integration, deployment of ranking/search/QA models, and making effective user feedback loops.Published three papers in top academic venues, and extending the research using novel ideas.Leading the <i>Viral Content Ranking and Generation</i> project. Aiming to construct a neural learning-to-rank model that ranks marketing articles in terms of future virality in social media.Participating in a NeurIPS 2022 challenge - Interactive Grounded Language Understanding in a Collaborative Environment (IGLU) with PhD students.Join in and present papers at regular reading group. Amazon Alexa Prize TaskBot Challenge - Team Conditia Co-led the team to the Quarter-Final stage. <ul style="list-style-type: none">Engineering a Task-oriented Dialogue System by using Alexa Skills Kit, Evi, DynamoDB, Lambda, CodePipeline, and SageMaker. Constructed and maintained CI/CD pipelines, deployed multiple micro-services with Flask and Docker, dealing with thousands of user conversations (details are addressed in the Master's thesis).Designed and implemented a novel state machine-like architecture for dialogue policy decision.Research were focused on Information Retrieval, Knowledge Graph, Question-Answering, Ranking, and Dialogue State Tracking.	May 2021 - Present May 2021 - Mar 2022

	UCL (Human-Computer) Interaction Centre - bachelor's group thesis Oct 2020 - Apr 2021 Project <i>Olfactory-NLP</i> with group members and Prof. Marianna Obrist . <ul style="list-style-type: none"> • Extracted smell-related phrases and clustered multiple passages into different olfactory groups. • Literature Review: <i>An NLP Approach to Classify Smell Experiences</i>. [paper] • <i>Exploring the potential of automating the process of clustering smell stories</i>. [paper] [code] • Final version will be submitted to HCI academic venue (CHI). 	
Industry Experience	Vector AI Machine Learning Engineer Intern <ul style="list-style-type: none"> • Cross-functional intern in both Core-ML and Forward-deployed engineering teams. • Tabular information extraction from bill of ladings and business registration documents based on OCR results for freight forwarding automation. • Use of modern dev tools, e.g., Gitlab, Jira, Sentry, Docker, Kubernetes, GCP. 	Jun 2022 - Jul 2022 London, UK
	KB Securities Data Analyst Intern <ul style="list-style-type: none"> • Rule-based Investment Funds Recommender System based on Network Analysis, Hierarchical Clustering and Association Rule Mining. • Stock Price Prediction System using LSTM Network and Natural Language News Data. [code] • Piloted a Social Media Monitoring System using KB's NLP Engine and Automation (RPA) software. Real-time sentiment monitoring in the comment sections. • Heavy use of SQL in the Hadoop ecosystem (Map-reduce, Hive, and Hue). 	Jun 2020 - Aug 2020 Seoul, South Korea
	Royal College of General Practitioners (RCGP) Software Engineer - UCL Industrial Exchange Network <ul style="list-style-type: none"> • Built a tree-based task-oriented chatbot that assists in checking software compliance regimes of British healthcare organizations. • Built a RESTful API for user intents and entities detection using Dialogflow and React, Express Node, and MongoDB. • Python scripts for chatbot integration testing. [code] 	Oct 2019 - Mar 2020 London, UK
Teaching Experience	UCL COMP0189 Applied Artificial Intelligence, <i>Teaching Assistant</i> UCL COMP0071 Software Engineering, <i>Teaching Assistant</i> UCL CEGE0004 Machine Learning for Data Science, <i>Teaching Assistant</i> UCL CEGE0096 Geospatial Programming, <i>Teaching Assistant</i> UCL Artificial Intelligence Society , <i>Lecturer</i> <ul style="list-style-type: none"> • Founder, maintainer, and lecturer of Machine Learning Tutorial Series • Season 20/21: [GitHub] [Lecture Recordings] Season 21/22: [GitHub] 	(Spring 2023) (Spring 2023) (Spring 2023) Fall 2022 2020 - 2022
Awards	2022 Automated Knowledge Graph Construction (AKBC) Conference – Registration fee waiver 2022 SIGIR Student Travel Grant – Registration fee waiver 2021 UCL Student Union best societies award – <i>Centenary Award (UCL AI Society)</i> 2021 UCL CS Undergraduate Summer Research Internship Fund – £3,200	
Community Involvement	The Alan Turing Institute NLP Interest Group, <i>member</i> The Alan Turing Institute Knowledge Graphs Interest Group, <i>member</i> Automated Knowledge Graph Construction (AKBC) Conference, <i>volunteer</i> ICLR 2023, <i>Reviewer</i> CIKM 2022, <i>Sub-reviewer</i> UCL Artificial Intelligence Society, <i>Head of Tutorials Team</i> NLP paper reading group (Jiphyeonjeon) , <i>Season 2,3 Organizer</i> <ul style="list-style-type: none"> • Season 2: Survey on Task-Oriented Dialog System and Knowledge Graph construction. [slides] [recording] (spoken in Korean) • Season 3: Recent Advances and Challenges in Task-oriented Dialogue Systems. [slides] [recording] (spoken in Korean) 	Nov 2022 - Present Sep 2022 - Present Nov 2022 Oct 2022 July 2022 Sep 2019 - Jun 2022 May 2021 - Nov 2022

UCL CS Transition Mentor, *Mentor - helping 1st year CS students*
UCL Korean Society, *IT Director - webmail automation*

Sep 2020 - Jun 2022
Jun 2020 - May 2021

Selected
Coursework &
Projects

COMP0087 Statistical Natural Language Processing

Spring 2022

Taught by Prof. Sebastian Riedel, received high mark (76.62%).

- *Fake reviews generations via large language models.* [\[paper\]](#)

Unified API of Apache Kafka & Google Pub/Sub

Nov 2021

- Simple ML model is served and does inferences in an event-driven way. [\[code\]](#)

COMP0022 Database and Information Management Systems

Spring 2021

Best group coursework nominated - Movie Analytics Web App.

- Microservice architecture consisting of a NestJS REST API, PostgreSQL DB, and Redis cache. Designed a normalized ERD for DB and optimized queries, seeing more than 10% speed increase. APIs can pre-cache anticipated requests, enabling extremely low latency infinite scrolling. [\[code\]](#)

BENG0095 Data Mining and Analysis

Winter 2020

Best group coursework nominated, received high mark (79.88%).

- Data visualization, model training and selection, hyperparameter tuning, and evaluation. [\[paper\]](#) [\[code\]](#)

Building Conversational AI for Messenger Hackathon

Dec 2019

Hackathon at Facebook London

- Built a Facebook's messenger-bot that checks if user's reply to one's lover is okay to send or not by fine-tuning a BERT model for sentiment analysis. [\[code\]](#)

Propositional Semantic Tableau Validator

Dec 2019

- Parse propositional logic, build a semantic tableau, and check its logical satisfiability (implemented in C) [\[code\]](#)

Professional
Training

Turing Commons: AI Ethics and Governance (The Alan Turing Institute) [\[website\]](#)

Nov 2022

Google Developers Korea Machine Learning Bootcamp

Oct 2020 - Jan 2021

Deep Learning Specialization (Prof. Andrew Ng on Coursera)

Dec 2020

UCLQ Summer School in Quantum Technology

Sep 2020

Introduction to AI in the Data Center (NVIDIA)

Mar 2020

KAIST Artificial Intelligence Bootcamp [\[Project 1\]](#) [\[Project 2\]](#)

Jun 2019 - Aug 2019

Major
Skills

Programming: Python, Java, C, L^AT_EX

ML/DL: PyTorch, TensorFlow/Keras, Scikit-Learn, OpenCV

Data skill: SQL, Hadoop, Numpy/Pandas/Matplotlib

SW Dev: Apache Kafka, Node, Express, React, Flask, MongoDB, PostgreSQL

Platforms: Linux, Git, Docker, AWS, Alexa Skills Kit, GCP

References

Prof. Emine Yilmaz

Professor and Turing Fellow of Computer Science in UCL, **Email:** emine.yilmaz@ucl.ac.uk

Dr. Aldo Lipani

Assistant Professor of Machine Learning in UCL, **Email:** aldo.lipani@ucl.ac.uk

Prof. Marianna Obrist

Professor of Multisensory Interfaces and Computer Science in UCL, **Email:** m.obrist@ucl.ac.uk