

To Eun Kim

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

✉ to.kim.17@ucl.ac.uk

 <https://kimdanny.github.io/>

Education	University College London (UCL) M.Eng. Computer Science (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)

T.E. Kim, A. Lipani. *A Multi-Task Based Neural Model to Simulate Users in Goal-Oriented Dialogue Systems*. [Short Paper]. International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022 (SIGIR '22). [\[paper\]](#) [\[poster\]](#) [\[code\]](#)

J. Ramos*, **T.E. Kim***, Z. Shi, X. Fu, F. Ye, Y. Feng, A. Lipani. *Conditia: A State Machine Like Architecture for Multi-Modal Task Bots*. Alexa Prize TaskBot Challenge Proceedings, 2022. [\[paper\]](#)

Z. Shi, P. Ni, M. Wang, **T.E. Kim**, A. Lipani. *Attention-based Ingredient Phrase Parser*. European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2022 (ESANN '22). [\[paper\]](#) [\[code\]](#)

Under Review **T.E. Kim**, S. Mitroi, R. Kumar, N. Ranade, K. Selvaratnam, L. Xu, E. Maggioni, M. Obrist.
Exploring the potential of automating the process of clustering smell stories. International Journal of Human-Computer Studies (IJHCS '23).

Workshops & Presentations *Recent Advances in Applications of Knowledge Graphs*. To be presented at the 5th Meet-up of Knowledge Graphs Interest Group at Alan Turing Institute, 27th Oct., 2022.

A Multi-Task Based Neural Model to Simulate Users in Goal-Oriented Dialogue Systems. Presented at Search-Oriented Conversational AI Workshop, 2022 (SCAI '22). [\[recording\]](#) [\[slides\]](#)

Research Experience	UCL Web Intelligence Group - Graduate Research Assistant Conducting and assisting research with Prof. Emine Yilmaz and Dr. Aldo Lipani.	May 2021 - Present
---------------------	---	--------------------

- Highly contributed (first authored) to the *Alexa Prize TaskBot Challenge*, 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 *Viral Content Ranking and Generation* 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 regular reading group.

Amazon Alexa Prize TaskBot Challenge - Team Conditia	May 2021 - Mar 2022
Co-led the team to the Quarter-final stage.	

- 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 was focused on Information Retrieval, Knowledge Graph, Question-Answering, Ranking, and Dialogue State Tracking.

UCL (Human-Computer) Interaction Centre - bachelor's group thesis Oct 2020 - Apr 2021
Project *Olfactory-NLP* with group members and Prof. Marianna Obrist.

- Extracted smell-related phrases and clustered multiple passages into different olfactory groups.

- Literature Review: *An NLP Approach to Classify Smell Experiences*. [\[paper\]](#)
- *Exploring the potential of automating the process of clustering smell stories*. [\[paper\]](#) [\[code\]](#)
- Edited version is recently submitted and under review in the International Journal of Human-Computer Studies.

Industry Experience	Vector AI	Jun 2022 - Jul 2022
	Machine Learning Engineer Intern	London, UK
	<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. 	
	KB Securities	Jun 2020 - Aug 2020
	Data Analyst Intern	Seoul, South Korea
	<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). 	
	Royal College of General Practitioners (RCGP)	Oct 2019 - Mar 2020
	Software Engineer - UCL Industrial Exchange Network	London, UK
	<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] 	
Teaching Experience	UCL COMP0189 Applied Artificial Intelligence, <i>Teaching Assistant</i>	(Spring 2023)
	UCL COMP0071 Software Engineering, <i>Teaching Assistant</i>	(Spring 2023)
	UCL CEGE0004 Machine Learning for Data Science, <i>Teaching Assistant</i>	(Spring 2023)
	UCL CEGE0096 Geospatial Programming, <i>Teaching Assistant</i>	Fall 2022
Awards	2022 SIGIR '22 Student Travel Grant – Free registration fee	
	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	CIKM 2022, <i>Sub-reviewer</i>	July 2022
	UCL Artificial Intelligence Society , <i>Head of Tutorials Team</i>	Sep 2019 - Jun 2022
	<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] 	
	NLP paper reading group (Jiphyeonjeon) , <i>Season 2,3 Organizer</i>	May 2021 - Nov 2022
	<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) 	
	UCL CS Transition Mentor, <i>Mentor - helping 1st year CS students</i>	Sep 2020 - Jun 2022
	UCL Korean Society, <i>IT Director - webmail automation</i>	Jun 2020 - May 2021
	UCL Korean Culture Society, <i>Korean Language Tutor - biweekly lessons</i>	Sep 2020 - Mar 2021
	Stanford CS224n study group [slides]	Winter 2020
Selected Coursework & Projects	COMP0087 Statistical Natural Language Processing	Spring 2022
	Taught by Prof. Sebastian Riedel, received high mark (76.62%). <ul style="list-style-type: none"> • <i>Fake reviews generations via large language models</i>. [paper] 	

	Unified API of Apache Kafka & Google Pub/Sub <ul style="list-style-type: none"> Simple ML model is served and does inferences in an event-driven way. [code] 	Nov 2021
	COMP0022 Database and Information Management Systems Best group coursework nominated - Movie Analytics Web App. <ul style="list-style-type: none"> 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] 	Spring 2021
	BENG0095 Data Mining and Analysis Best group coursework nominated, received high mark (79.88%). <ul style="list-style-type: none"> Data visualization, model training and selection, hyperparameter tuning, and evaluation. [paper] [code] 	Winter 2020
	Building Conversational AI for Messenger Hackathon Hackathon at FaceBook London <ul style="list-style-type: none"> 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] 	Dec 2019
	Propositional Semantic Tableau Validator <ul style="list-style-type: none"> Parse propositional logic, build a semantic tableau, and check its logical satisfiability (implemented in C) [code] 	Dec 2019
Professional Training	Google Developers Korea Machine Learning Bootcamp Deep Learning Specialization (Prof. Andrew Ng on Coursera) UCLQ Summer School in Quantum Technology Introduction to AI in the Data Center (NVIDIA) KAIST Artificial Intelligence Bootcamp [Project 1] [Project 2]	Oct 2020 - Jan 2021 Dec 2020 Sep 2020 Mar 2020 Jun 2019 - Aug 2019
Major Skills	Programming: Python, Java, C, \LaTeX 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, GCP	
References	Prof. Emine Yilmaz Professor and Turing Fellow at UCL, Dept. of Computer Science, Email: emine.yilmaz@ucl.ac.uk Dr. Aldo Lipani Assistant Professor in Machine Learning at UCL, Email: aldo.lipani@ucl.ac.uk Prof. Marianna Obrist Professor of Multisensory Interfaces at UCL, Dept. of Computer Science, Email: m.obrist@ucl.ac.uk	