

# Danny To Eun Kim

NATURAL LANGUAGE PROCESSING | DIALOGUE SYSTEM | INFORMATION RETRIEVAL | MACHINE LEARNING

✉ to.kim.17@ucl.ac.uk | 🏠 kimdanny.github.io | 📷 kimdanny | 💼 danny-toeun-kim | 🐦 TEKnologyy | 🎓 To Eun Kim

## Education

### KAIST Graduate School of AI

PH.D. ARTIFICIAL INTELLIGENCE

- Incoming PhD student

Seoul, South Korea

Aug 2022 - Present

### University College London (UCL)

M.ENG. COMPUTER SCIENCE (BS/MS)

- UK First-Class Honours (3.8+/ 4.0 in US GPA scale)
- BSc Thesis: Exploring the potential of automating the process of clustering smell stories
- MEng Thesis: **Multi-Task Neural User Simulator for Task Oriented Dialogue System**

London, UK

Sep 2018 - Jun 2022

## Publications

### Attention-based Ingredient Phrase Parser

Z. SHI, P. NI, **T.E. KIM**, M. WANG, A. LIPANI (2022)

- [\[paper\]](#)

ESANN'22

July 2022

### A Multi-Task Based Neural Model to Simulate Users in Goal-Oriented Dialogue Systems

**T.E. KIM**, A. LIPANI (2022);

- Dialogue User Simulator that generates user-side utterance, predicts user satisfaction score and action [\[paper\]](#) [\[code\]](#)

SIGIR'22

July 2022

### Condit: 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 (2022)

- A technical report from Alexa Prize TaskBot challenge [\[paper\]](#)

Alexa Prize Proceedings

June 2022

## Research Experience

### Team Condit: Amazon 2021 Alexa Prize TaskBot Challenge

UCL **WEB INTELLIGENCE GROUP**

- Engineering Task-oriented Dialogue System by using Alexa Skills Kit, Evi, DynamoDB, Lambda, CodePipeline, and SageMaker.
- Constructed CI/CD Pipeline, Deployed multiple microservices with Flask and Docker, dealing with thousands of user conversations
- Research is focused on Information Retrieval, Knowledge Graph, Question-Answering, Ranking, and Dialogue State Tracking.
- Supervised by Professor **Aldo Lipani**

London, UK

Jun 2021 - Mar 2022

### Viral content Ranking and Generation

UCL **COMPUTATIONAL STATISTICS AND MACHINE LEARNING GROUP**

- Neural Learning-to-Rank model that ranks marketing articles in terms of future virality in social media
- Worked with Professor **Emine Yilmaz** and **Qiang Zhang**

London, UK

May 2021 - Mar 2022

### Bachelor Thesis: Olfactory NLP

UCL **HUMAN-COMPUTER INTERACTION CENTRE**

- Group Research Paper: *Exploring the potential of automating the process of clustering smell stories* [\[paper\]](#) [\[code\]](#)
- Literature Review: *An NLP Approach to Classify Smell Experiences* [\[paper\]](#)
- Supervised by Professor **Marianna Obrist**

London, UK

Oct 2020 - Mar 2021

## Employment

### Machine Learning Engineering Intern

**VECTOR AI**

- Document Understanding and automation for freight forwarding; Forward-deployed Engineering; MLOps

London, UK

Jun 2022 - Aug 2022

### Student Researcher

UCL **COMPUTATIONAL STATISTICS AND MACHINE LEARNING GROUP**

- Alexa Prize & Viral content ranking and generation

London, UK

May 2021 - Present

## Data Analyst Summer Intern

Yeongdeungpo-gu, Seoul, S.Korea

Jun 2020 - Aug 2020

### KB SECURITIES

- Rule-based Investment Funds Recommender System based on Network Analysis, Hierarchical Clustering and Association Rule Mining
- Piloted Stock Price Prediction System using LSTM Network and Natural Language News Data [\[Github Repo.\]](#)
- Piloted a Social Media Monitoring System using KB's NLP Engine and Automation (RPA) software
- Heavy use of SQL in Hadoop ecosystem (Map-reduce, Hive and Hue)

## Software (Dialogue System) Engineer & Industrial Exchange Network Partner

London, UK

ROYAL COLLEGE OF GENERAL PRACTITIONERS (RCGP)

Oct 2019 - Apr 2020

- Built a task-oriented chatbot that assists in checking software compliance regimes that British healthcare organisations have
- Built a RESTful API for user intents and entities detection using Dialogflow and React/Express/Node/MongoDB
- Made python scripts for Chatbot Integration Testing [\[Github Repo.\]](#)

## Selected Projects

### Unified API of Apache Kafka & Google Pub/Sub

PERSONAL PROJECT FOR ML MODEL SERVING

Nov 2021

- Made a Unified API of Apache Kafka and Google PubSub
- Simple ML model is served and does inferences in an event-driven way. [\[Github Repo.\]](#)

### Netflix-like Movie Analytics Web App

UCL

BEST COURSEWORK NOMINATED

Feb 2021 - Mar 2021

- Microservice architecture consisting of a NestJS REST API, PostgreSQL DB, and Redis cache - each component is containerized
- Designed a normalized ERD for Database design, and optimized queries, seeing more than 10% performance increase
- API can pre-cache anticipated requests, enabling extremely low latency infinite scrolling [\[Github Repo.\]](#)

### Credit Card Default prediction

UCL

BEST COURSEWORK NOMINATED

Dec 2020 - Jan 2021

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

### Propositional Semantic Tableau validator

UCL

COMPUTATIONAL LOGIC COURSEWORK

Dec 2019

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

### Relationship-Expert (Hackathon at FB London)

Facebook London

BUILDING CONVERSATIONAL AI FOR MESSENGER HACKATHON

Dec 7-8 2019

- Built a messenger-bot that checks if user's reply to one's lover is okay to send or not [\[Github Repo.\]](#)
- Fine-tuned a Transformer Language Model (BERT) using Pytorch for Sentiment Analysis
- Deployed the model with Facebook Messenger API (Node.js) and a Flask server

## Society/Service

### Sub-reviewer

CIKM 2022

CIKM 2022

July 2022

- Reviewed papers related to conversational AI submitted to the CIKM 2022

### UCL Artificial Intelligence Society

UCL, London, UK

DEVELOPER RELATIONS(19/20) → HEAD OF TUTORIALS TEAM(20/21) → ML OFFICER(21/22)

Sep 2019 - Present

- Main contributor/founder and lecturer of Machine Learning tutorials [\[Season 1\]](#) [\[Season 2\]](#) [\[YouTube playlist\]](#)
- Teaching Assistant in Advanced Python Tutorials (19/20)

### NLP paper reading group

virtual, S.Korea

SEASON 2 ORGANIZER [\[GITHUB REPO.\]](#)

May 2021 - Nov 2021

- Task-oriented Dialogue System Survey and Knowledge Graph construction [\[Slides\]](#) [\[Video\]](#)

## Professional Training

### Google Developers Korea Machine Learning Bootcamp

virtual (GDG Korea)

TRAINEE

Oct 2020 - Jan 2021

- 4-month-long ML Bootcamp hosted by Google Developers Korea. Education fees are all funded, selected in one of 8% of 3K+ applicants.
- Earned Deep Learning Certificate by weekly assignments, and attended weekly tech-talks and clinics

### KAIST Artificial Intelligence Bootcamp

Yangjae AI Hub, S.Korea

TRAINEE

Jun 2019 - Aug 2019

- Project 1: Kaggle Titanic - 81% Accuracy [\[Github Repo.\]](#)
- Project 2: Neural Style Transfer. Model comparison, Feature map visualisation and Hyperparameter Tuning [\[Github Repo.\]](#)

## Awards/Grants

---

2022	<b>Student Travel Grant</b>	conference registration fee	SIGIR '22
2021	<b>Centenary Award</b>	20/21 UCL best societies award (Head of Tutorials @ Artificial Intelligence Society)	UCL Student Union
2021	<b>CS Undergraduate Summer Research Internship Fund</b>		UCL CS Dept.

## Certifications

---

2020	<b>Deep Learning Specialization</b>	5 courses on Coursera - Prof. Andrew Ng	deeplearning.ai
2020	<b>The Fundamentals of Digital Marketing</b>	Google Adsense, Search console, SEO, etc.	Google D.G
2020	<b>Introduction to AI in the Data Center</b>	Data centre Architecture and GPU accelerated computing	NVIDIA
2020	<b>JavaScript Essential Training</b>		LinkedIn Learning
2019	<b>KAIST Artificial Intelligence Bootcamp</b>	Certificate of Completion	KAIST & Seoul Gvmt.

## Side Works

---

### Website Maintainer

UCL KOREAN SOCIETY

- User DB Management and Web-mail automation

London, UK

Jul 2020 - May 2021

### Korean Language Tutor

UCL KOREAN CULTURE SOCIETY

- Biweekly Korean lessons with K-Pop lyrics to UCL students

London, UK

Sep 2020 - May 2021

### Blogger

TEKNOLOGY: MY PERSONAL TECH BLOG

- Technical and random postings at <https://kimdanny.github.io>

2019 - Present

## Professional Skills

---

<b>Programming</b>	Python, Java, C, JS, Haskell, Arduino, R, $\text{\LaTeX}$
<b>ML / DL</b>	TensorFlow/Keras, PyTorch, OpenCV, Scikit-Learn
<b>Data Skill</b>	SQL, Hadoop, SAS, Numpy/Pandas/Matplotlib, Web-Scraping
<b>SW-Dev</b>	Apache Kafka, Node, Express, React, Flask, MongoDB, PostgreSQL,
<b>Platforms</b>	Linux, Git, Docker, AWS, Google Cloud Platform(GCP)