

EUJEONG CHOI

L I N G U I S T | A I D A T A B U I L D I N G



☎ +82 10-3595-0608
✉ eujeonglesleychoi@gmail.com
📍 Seoul, South Korea
🌐 <https://bethegiver.github.io/>

PROFILE

I am an enthusiast of **Data-Centric AI**.

During the last 5 years, I invested in creating robust and trustworthy **NLP Datasets** with a reliable **Data Pipeline**. In this process, I have taken part in both academic and industrial applications, and am especially interested in **optimizing data production without compromising its quality**.

SKILLS

- English & Korean Bilingual
- ML Data Scheme Generation
- Annotation Guideline & Instruction
- In-house Annotator Management
- Fluent in: Python, Spreadsheet
- Experienced in: Github, Label Studios

AWARDS

TOXIC COMMENT PREDICTION & COLLECTION SERVICE

Korean Language Information Processing System Competition
2020

EDUCATION

BACHELOR OF LINGUISTICS & COMPUTER SCIENCE (LB&C)

Korea University
2017-2022

EXPERIENCE

DATA MANAGER

Jul 2021 - Present

Upstage (<https://upstage.ai/>)

- Built two main NLP Datasets (with in-house annotators)
 - Involved Domain: Movie, Beauty Review & Query
- Devised a Data Schema applicable to the model & Managed annotators
- Took part in automating & accelerating the Data Production Pipeline

AI LANGUAGE TECH LABS INTERN

Apr 2020 - Oct 2020

SK Telecom (<https://www.skt.ai/kr/>)

- Organized the training data for Chatbot training and validation

ASSISTANT RESEARCHER AT KOREA UNI. Mar 2019 - Jun 2020

PUBLICATIONS

HCI PROJECT : HUMAN-LIKE CHATBOT AGENTS

[1] Evaluating How Users Game and Display Conversation with Human-Like Agents
CODI@Coling 2022

[2] Assessing How Users Display Self-Disclosure and Authenticity in Conversation with Human-Like Agents: A Case Study of Luda Lee
AACL-IJCNLP 2022