Sugyeong Eo

KOREA UNIVERSITY

311 Aegineung Student Center, College of Informatics, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, Korea

(+82) 10-9398-3950 | djtnrud@korea.ac.kr

Research Interests

I am a Ph.D student in computer science and engineering at Korea University. I belong to NLP & AI Lab (Advisor: Prof. Heuiseok Lim). I am the founder and CSO of KU-NMT Group.

- Natural Language Processing
- Neural Machine Translation
- Quality Estimation
- Question Answer Generation
- Critical Error Detection
- Curriculum Learning

Education

Korea University *Seoul, South Korea*
Ph.D student in Computer Science and Engineering *Sep. 2020 - Present*
- **Advisor: Heuiseok Lim**
- **Natural Language Processing & Artificial Intelligence Lab**
- **CSO of KU-NMT Group.**

Hankuk University of Foreign Studies (HUFS) *Yongin & Seoul, South Korea*
B.A. in Linguistics & Cognitive Science(1st Major), Language & Technology(2nd Major), GPA: 3.9 / 4.5 *Mar. 2016 - Aug. 2020*

Academic Services

Program committe *South Korea*
Program committe *January.2022*
- NAACL 2022-Industry Track

Selected Publications

KU X Upstage's submission for the WMT22 Quality Estimation: Critical Error Detection Shared Task **(Ranked 1st on the leaderboard)**
  Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim
  Seventh Conference on Machine Translation, 2022

QUAK: A Synthetic Quality Estimation Dataset for Korean-English Neural Machine Translation
  Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Gyeongmin Kim, Jungseob Lee, Heui-Seok Lim
  Proceedings of the 29th International Conference on Computational Linguistics, 2022

Word-level Quality Estimation for Korean-English Neural Machine Translation
  Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim
  IEEE Access 10 (2022) pp. 44964–44973. IEEE, 2022

An Automatic Post Editing with Efficient and Simple Data Generation Method
  Hyeonseok Moon, Chanjun Park, Jaehyung Seo, **Sugyeong Eo**, Heuiseok Lim
  IEEE Access (2022). IEEE, 2022

Dense-to-Question and Sparse-to-Answer: Hybrid Retriever System for Industrial Frequently Asked Questions
  Jaehyung Seo, Taemin Lee, Hyeonseok Moon, Chanjun Park, **Sugyeong Eo**, Imatitikua D Aiyanyo, Kinam Park, Aram So, Sungmin Ahn, Jeongbae Park
  Mathematics 10.8 (2022) p. 1335. MDPI, 2022

Mimicking Infants' Bilingual Language Acquisition for Domain Specialized Neural Machine Translation
  Chanjun Park, Woo-Young Go, **Sugyeong Eo**, Hyeonseok Moon, Seolhwa Lee, Heuiseok Lim
  IEEE Access 10 (2022) pp. 38684–38693. IEEE, 2022

Return on Advertising Spend Prediction with Task Decomposition-Based LSTM Model
  Hyeonseok Moon, Taemin Lee, Jaehyung Seo, Chanjun Park, **Sugyeong Eo**, Imatitikua D Aiyanyo, Jeongbae Park, Aram So, Kyoungwha Ok, Kinam Park
  Mathematics 10.10 (2022) p. 1637. MDPI, 2022

SUGYEONG. EO · CURRICULUM VITAE 1

Empirical Analysis of Parallel Corpora and In-Depth Analysis Using LIWC
Chanjun Park, Midan Shim, **Sugyeong Eo**, Seolhwa Lee, Jaehyung Seo, Hyeonseok Moon, Heuiseok Lim
Applied Sciences 12.11 (2022) p. 5545. Multidisciplinary Digital Publishing Institute, 2022

Priming Ancient Korean Neural Machine Translation
Chanjun Park, Seolhwa Lee, Jaehyung Seo, Hyeonseok Moon, **Sugyeong Eo**, Heuiseok Lim
LREC 2022 (Proceedings of the Language Resources and Evaluation Conference), 2022, Marseille, France

A Self-Supervised Automatic Post-Editing Data Generation Tool
Hyeonseok Moon, Chanjun Park, **Sugyeong Eo**, Jaehyung Seo, SeungJun Lee, Heuiseok Lim
ICML 2022 - DataPerf workshop, 2022 (2022). 2022

BERTOEIC: Solving TOEIC Problems Using Simple and Efficient Data Augmentation Techniques with Pretrained Transformer Encoders
Jeongwoo Lee, Hyeonseok Moon, Chanjun Park, Jaehyung Seo, **Sugyeong Eo**, Heuiseok Lim
Applied Sciences 12.13 (2022) p. 6686. Multidisciplinary Digital Publishing Institute, 2022

A Dog Is Passing Over The Jet? A Text-Generation Dataset for Korean Commonsense Reasoning and Evaluation
Jaehyung Seo, Seounghoon Lee, Chanjun Park, Yoonna Jang, Hyeonseok Moon, **Sugyeong Eo**, Seonmin Koo, Heuiseok Lim
Findings of the Association for Computational Linguistics: NAACL 2022, 2022, Seattle, United States

PU-GEN: Enhancing generative commonsense reasoning for language models with human-centered knowledge
Jaehyung Seo, Dongsuk Oh, **Sugyeong Eo**, Chanjun Park, Kisu Yang, Hyeonseok Moon, Kinam Park, Heuiseok Lim
Knowledge-Based Systems (2022) p. 109861. Elsevier, 2022

Focus on FoCus: Is FoCus focused on Context, Knowledge and Persona?
SeungYoon Lee, Jungseob Lee, Chanjun Park, **Sugyeong Eo**, Hyeonseok Moon, Jaehyung Seo, Jeongbae Park, Heui-Seok Lim
Proceedings of the 1st Workshop on Customized Chat Grounding Persona and Knowledge, 2022

Plain Template Insertion: Korean-Prompt-Based Engineering for Few-Shot Learners
Jaehyung Seo, Hyeonseok Moon, Chanhee Lee, **Sugyeong Eo**, Chanjun Park, Jihoon Kim, Changwoo Chun, Heuiseok Lim
IEEE Access 10 (2022) pp. 107587–107597. 2022

Empirical Analysis of Noising Scheme based Synthetic Data Generation for Automatic Post-editing
Hyeonseok Moon, Chanjun Park, Seolhwa Lee, Jaehyung Seo, Jungseob Lee, **Sugyeong Eo**, Heui-Seok Lim
Proceedings of the Thirteenth Language Resources and Evaluation Conference, 2022

A New Tool for Efficiently Generating Quality Estimation Datasets
Sugyeong Eo, Chanjun Park, Jaehyung Seo, Hyeonseok Moon, Heuiseok Lim
NeurIPS 2021 - Data-centric AI (DCAI) workshop, 2021

Comparative Analysis of Current Approaches to Quality Estimation for Neural Machine Translation
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim, **(Equal Contribution (First Co-Author))**
Applied Sciences 11.14 (2021) p. 6584. Multidisciplinary Digital Publishing Institute, 2021

Dealing with the Paradox of Quality Estimation
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim, **(Equal Contribution (First Co-Author))**
MT Summit - Proceedings of the 4th Workshop on Technologies for MT of Low Resource Languages (LoResMT2021), 2021, Virtual

Should we find another model?: Improving Neural Machine Translation Performance with ONE-Piece Tokenization Method without Model Modification
Chanjun Park, **Sugyeong Eo**, Hyeonseok Moon, Heui-Seok Lim
NAACL 2021 - Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Papers, 2021

BTS: Back TranScription for Speech-to-Text Post-Processor using Text-to-Speech-to-Text
Chanjun Park, Jaehyung Seo, Seolhwa Lee, Chanhee Lee, Hyeonseok Moon, **Sugyeong Eo**, Heuiseok Lim
ACL 2021 - Proceedings of the 8th Workshop on Asian Translation (WAT2021), 2021, Online

An Empirical Study on Automatic Post Editing for Neural Machine Translation
Hyeonseok Moon, Chanjun Park, **Sugyeong Eo**, Jaehyung Seo, Heuiseok Lim, **(Equal Contribution (First Co-Author))**
IEEE Access 9 (2021) pp. 123754–123763. IEEE, 2021

Automatic Knowledge Augmentation for Generative Commonsense Reasoning
Jaehyung Seo, Chanjun Park, **Sugyeong Eo**, Hyeonseok Moon, Heuiseok Lim
NeurIPS 2021 - Data-centric AI (DCAI) workshop, 2021

How should human translation coexist with NMT? Efficient tool for building high quality parallel corpus
Chanjun Park, Seolhwa Lee, Hyeonseok Moon, **Sugyeong Eo**, Jaehyung Seo, Heuiseok Lim
NeurIPS 2021 - Data-centric AI (DCAI) workshop (2021). 2021

Honors & Awards

2022.12	**1st place in WMT 2022 QE Task 3, 2022**, WMT 2022 - EMNLP 2022	*Abu Dhabi*
2022.10	**Best Paper Award, 2022**, The 34th Annual Conference on Human & Cognitive Language Technology (HCLT2022)	*South Korea*
2021.10	**Best Paper Award, 2021**, The 33rd Annual Conference on Human & Cognitive Language Technology (HCLT2021)	*South Korea*
2022.11	**Ranked 4th, 2022**, CommonGen 1.1 Leaderboard	*No Location*

Selected Research Experience

(Project) Hyundai NGV - Question-answer pair generation for vehicle QA of natural language processing engine

Seoul, South Korea

RESEARCHER IN KOREA UNIV

December. 2021. - June. 2022.

- Research for question answer generation
- Supported by Hyundai NGV
- Skills: Question Answering, Language Modeling, Pytorch, Fairseq

(Project) NAVER PAPAGO - Parallel Corpus Filtering and Mining

Seoul, South Korea

RESEARCHER IN KOREA UNIV

December. 2021. - June. 2022.

- Research for Static and Dynamic Parallel Corpus Filtering and Mining
- Human-inspired curriculum learning based NMT
- Supported by NAVER PAPAGO
- Skills: Neural Machine Translation, Parallel corpus filtering, Pytorch, Fairseq

Data-centric Neural Machine Translation Research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

July. 2021. - December. 2021.

- Research for Data-centric Neural Machine Translation, Automatic Post Editing, Quality Estimation, Generative Commonsense Reasoning
- Develop Efficient tool for building high quality parallel corpus and A New Tool for Efficiently Generating Quality Estimation Datasets and A Self-Supervised Automatic Post-Editing Data Generation Tool
- Publication of 3 research paper in NeurIPS 2021 - Data-centric AI (DCAI) workshop: `https://eval.how/dcai2021/index.html`
- Publication of research paper in ICML 2022 - DataPerf workshop: `https://sites.google.com/view/dataperf2022`
- Skills: Neural Machine Translation, Web Application, Fairseq

Quality Estimation Research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

January. 2021. - Present

- Large Scale PLM based Quality Estimation development and research
- Low Resource QE and Pseudo QE data generation
- Publication of research paper in Applied Science (SCIE): `https://www.mdpi.com/2076-3417/11/14/6584`
- Publication of research paper in MT Summit 2021 - LoresMT Workshop: `https://aclanthology.org/2021.mtsummit-loresmt.1/`
- Publication of research paper in IEEE Access: `https://ieeexplore.ieee.org/document/9761258`
- Publication of research paper in COLING 2022: `https://aclanthology.org/2022.coling-1.460.pdf`
- 1st place in WMT 2022 QE Task 3: KU X Upstage's submission for the WMT22 Quality Estimation: Critical Error Detection Shared Task
- Skills: Neural Machine Translation, Pytorch, Quality Estimation, OpenKiwi, Huggingface, Fairseq

Automatic Post Editing Research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

January. 2021. - Present

- APE research based on Cross Lingual Pretrained Language Model
- Adpater based APE research
- Publication of research paper in IEEE Access (SCIE): `https://ieeexplore.ieee.org/document/9528385/`
- Research for Efficient and Simple Data Generation Method - Low Resource APE and Pseudo APE data generation
- Publication of research paper in IEEE Access (SCIE): `https://ieeexplore.ieee.org/document/9714400`
- Research for Synthetic Data Generation Using Noising Strategies for Automatic Post-editing
- Publication of research paper in LREC 2022: `http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.93.pdf`
- Skills: Neural Machine Translation, Pytorch, Automatic Post Editing, Huggingface, Fairseq

Korean Commonsense Generation and Reasoning Research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

June. 2021. - Present

- KommonGen: A Dataset for Korean Generative Commonsense Reasoning Evaluation
- Automatic Knowledge Augmentation for Generative Commonsense Reasoning
- Publication of research paper in NAACL 2022: `https://aclanthology.org/2022.findings-naacl.172/`
- Publication of research paper in Knowledge-Based Systems: `https://www.sciencedirect.com/science/article/abs/pii/S0950705122009546`
- Skills: Commonsense Reasoning, Pytorch, Huggingface

BTS: Speech recognition post-processor Research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

January. 2021. - Present

- Proposed BackTranScription(BTS). Back Transcription (BTS) has seen limited success without the role of phonetic transcription for building and creating vast datasets automatically.
- Publication of research paper in WAT 2021 (ACL Workshop): `https://aclanthology.org/2021.wat-1.10/`
- Skills: Neural Machine Translation, Pytorch, Huggingface, Fairseq

(Project) Domain Specialized Neural Machine Translation (Security)

Seoul, South Korea

RESEARCHER IN KOREA UNIV

May. 2020. - November.2020

- We propose Cross Communication Method (CCM), a new DS-NMT training methodology, inspired by cognitive scientific interpretation of cross language speech perception. This is the first study to reflect cognitive science theory in DS-NMT training methodology.
- Publication of research paper in IEEE Access: `https://ieeexplore.ieee.org/document/9751075`
- Skills: Neural Machine Translation, Pytorch, Python

Skills

Machine Learning & Deep Learning	Pytorch, Pytorch lightning, Keras, Tensorflow
Natural Language Processing	Overall NLU/NLG tasks, Huggingface, Fairseq
Product & Project & Data Manager	MS Office, Notion, Slack
Etc	Linux, Docker, Latex
Language Skills	Korean, English

Patents

2021.11 **DEVICE AND METHOD FOR GENERATING OF TRAINING DATA FOR QUALITY ESTIMATION IN MACHINE TRANSLATION,** Apply for a patent *South Korea*

2021.09 **DEVICE AND METHOD FOR GENERATING TRAINING DATA FOR AUTOMATIC POST EDITING,** Apply for a patent *South Korea*

Demo

2021.06 **BTS (BackTranScription) for STT Post-Processor,** `https://backtrans-317413.de.r.appspot.com/` *South Korea*

2021.10 **A Tool for generating High-Quality Parallel Corpus considering Human Translation Market,** `http://nlplab.iptime.org:9090/` *South Korea*

2021.10 **A New Tool for Efficiently Generating Quality Estimation Datasets,** `http://nlplab.iptime.org:9091/` *South Korea*

2021.11 **A Self-Supervised Automatic Post-Editing Data Generation Tool,** `http://nlplab.iptime.org:9092/` *South Korea*

Invited Talk & Teaching Experience

Hankuk University of Foreign Studies *South Korea*

PRESENTER *July. 2022*

- Basic practice of natural language processing for everyone
- Invited by Hankuk University of Foreign Studies, Interpreting and Translation Research Institute (ITRI)

All Publications

TOP CONFERENCE

KU X Upstage's submission for the WMT22 Quality Estimation: Critical Error Detection Shared Task (Ranked 1st on the leaderboard)
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim
Seventh Conference on Machine Translation, 2022

QUAK: A Synthetic Quality Estimation Dataset for Korean-English Neural Machine Translation
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Gyeongmin Kim, Jungseob Lee, Heui-Seok Lim
Proceedings of the 29th International Conference on Computational Linguistics, 2022

Empirical Analysis of Noising Scheme based Synthetic Data Generation for Automatic Post-editing
Hyeonseok Moon, Chanjun Park, Seolhwa Lee, Jaehyung Seo, Jungseob Lee, **Sugyeong Eo**, Heui-Seok Lim
Proceedings of the Thirteenth Language Resources and Evaluation Conference, 2022

Priming Ancient Korean Neural Machine Translation
Chanjun Park, Seolhwa Lee, Jaehyung Seo, Hyeonseok Moon, **Sugyeong Eo**, Heuiseok Lim
LREC 2022 (Proceedings of the Language Resources and Evaluation Conference), 2022, Marseille, France

A Self-Supervised Automatic Post-Editing Data Generation Tool
Hyeonseok Moon, Chanjun Park, **Sugyeong Eo**, Jaehyung Seo, SeungJun Lee, Heuiseok Lim
ICML 2022 - DataPerf workshop, 2022 (2022). 2022

A Dog Is Passing Over The Jet? A Text-Generation Dataset for Korean Commonsense Reasoning and Evaluation
Jaehyung Seo, Seounghoon Lee, Chanjun Park, Yoonna Jang, Hyeonseok Moon, **Sugyeong Eo**, Seonmin Koo, Heuiseok Lim
Findings of the Association for Computational Linguistics: NAACL 2022, 2022, Seattle, United States

Focus on FoCus: Is FoCus focused on Context, Knowledge and Persona?
SeungYoon Lee, Jungseob Lee, Chanjun Park, **Sugyeong Eo**, Hyeonseok Moon, Jaehyung Seo, Jeongbae Park, Heui-Seok Lim
Proceedings of the 1st Workshop on Customized Chat Grounding Persona and Knowledge, 2022

A New Tool for Efficiently Generating Quality Estimation Datasets
Sugyeong Eo, Chanjun Park, Jaehyung Seo, Hyeonseok Moon, Heuiseok Lim
NeurIPS 2021 - Data-centric AI (DCAI) workshop, 2021

Dealing with the Paradox of Quality Estimation
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim, **(Equal Contribution (First Co-Author))**
MT Summit - Proceedings of the 4th Workshop on Technologies for MT of Low Resource Languages (LoResMT2021), 2021, Virtual

Automatic Knowledge Augmentation for Generative Commonsense Reasoning
Jaehyung Seo, Chanjun Park, **Sugyeong Eo**, Hyeonseok Moon, Heuiseok Lim
NeurIPS 2021 - Data-centric AI (DCAI) workshop, 2021

How should human translation coexist with NMT? Efficient tool for building high quality parallel corpus
Chanjun Park, Seolhwa Lee, Hyeonseok Moon, **Sugyeong Eo**, Jaehyung Seo, Heuiseok Lim
NeurIPS 2021 - Data-centric AI (DCAI) workshop (2021). 2021

Should we find another model?: Improving Neural Machine Translation Performance with ONE-Piece Tokenization Method without Model Modification
Chanjun Park, **Sugyeong Eo**, Hyeonseok Moon, Heui-Seok Lim
NAACL 2021 - Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Papers, 2021

BTS: Back TranScription for Speech-to-Text Post-Processor using Text-to-Speech-to-Text
Chanjun Park, Jaehyung Seo, Seolhwa Lee, Chanhee Lee, Hyeonseok Moon, **Sugyeong Eo**, Heuiseok Lim
ACL 2021 - Proceedings of the 8th Workshop on Asian Translation (WAT2021), 2021, Online

INTERNATIONAL JOURNAL

Word-level Quality Estimation for Korean-English Neural Machine Translation
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim
IEEE Access 10 (2022) pp. 44964–44973. IEEE, 2022

An Automatic Post Editing with Efficient and Simple Data Generation Method
Hyeonseok Moon, Chanjun Park, Jaehyung Seo, **Sugyeong Eo**, Heuiseok Lim
IEEE Access (2022). IEEE, 2022

Dense-to-Question and Sparse-to-Answer: Hybrid Retriever System for Industrial Frequently Asked Questions
Jaehyung Seo, Taemin Lee, Hyeonseok Moon, Chanjun Park, **Sugyeong Eo**, Imatitikua D Aiyanyo, Kinam Park, Aram So, Sungmin Ahn, Jeongbae Park
Mathematics 10.8 (2022) p. 1335. MDPI, 2022

Mimicking Infants' Bilingual Language Acquisition for Domain Specialized Neural Machine Translation
Chanjun Park, Woo-Young Go, **Sugyeong Eo**, Hyeonseok Moon, Seolhwa Lee, Heuiseok Lim
IEEE Access 10 (2022) pp. 38684–38693. IEEE, 2022

Return on Advertising Spend Prediction with Task Decomposition-Based LSTM Model
Hyeonseok Moon, Taemin Lee, Jaehyung Seo, Chanjun Park, **Sugyeong Eo**, Imatitikua D Aiyanyo, Jeongbae Park, Aram So, Kyoungwha Ok, Kinam Park
Mathematics 10.10 (2022) p. 1637. MDPI, 2022

Empirical Analysis of Parallel Corpora and In-Depth Analysis Using LIWC
Chanjun Park, Midan Shim, **Sugyeong Eo**, Seolhwa Lee, Jaehyung Seo, Hyeonseok Moon, Heuiseok Lim
Applied Sciences 12.11 (2022) p. 5545. Multidisciplinary Digital Publishing Institute, 2022

BERTOEIC: Solving TOEIC Problems Using Simple and Efficient Data Augmentation Techniques with Pretrained Transformer Encoders
Jeongwoo Lee, Hyeonseok Moon, Chanjun Park, Jaehyung Seo, **Sugyeong Eo**, Heuiseok Lim
Applied Sciences 12.13 (2022) p. 6686. Multidisciplinary Digital Publishing Institute, 2022

PU-GEN: Enhancing generative commonsense reasoning for language models with human-centered knowledge
Jaehyung Seo, Dongsuk Oh, **Sugyeong Eo**, Chanjun Park, Kisu Yang, Hyeonseok Moon, Kinam Park, Heuiseok Lim
Knowledge-Based Systems (2022) p. 109861. Elsevier, 2022

Plain Template Insertion: Korean-Prompt-Based Engineering for Few-Shot Learners
Jaehyung Seo, Hyeonseok Moon, Chanhee Lee, **Sugyeong Eo**, Chanjun Park, Jihoon Kim, Changwoo Chun, Heuiseok Lim
IEEE Access 10 (2022) pp. 107587–107597. 2022

Comparative Analysis of Current Approaches to Quality Estimation for Neural Machine Translation
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, Heuiseok Lim, **(Equal Contribution (First Co-Author))**
Applied Sciences 11.14 (2021) p. 6584. Multidisciplinary Digital Publishing Institute, 2021

An Empirical Study on Automatic Post Editing for Neural Machine Translation
Hyeonseok Moon, Chanjun Park, **Sugyeong Eo**, Jaehyung Seo, Heuiseok Lim, **(Equal Contribution (First Co-Author))**
IEEE Access 9 (2021) pp. 123754–123763. IEEE, 2021

INTERNATIONAL CONFERENCE

Cross-lingual Zero-shot Transfer for Quality Estimation
Sugyeong Eo, Chanjun Park, Jaehyung Seo, Heonseok Moon, HeuiSeok Lim
International Conference on Interdisciplinary Research on Computer Science, Psychology, and Education (ICICPE 2021), 2021

A Study on Jeju Dialect ASR processor using BackTranScription
Chanjun Park, Jaehyung Seo, Seolhwa Lee, Heonseok Moon, **Sugyeong Eo**, Yoonna Jang, HeuiSeok Lim
International Conference on Interdisciplinary Research on Computer Science, Psychology, and Education (ICICPE 2021), 2021

Domain Specialized Automatic Post Editing System
Heonseok Moon, Chanjun Park, Jaehyung Seo, **Sugyeong Eo**, HeuiSeok Lim
International Conference on Interdisciplinary Research on Computer Science, Psychology, and Education (ICICPE 2021), 2021

Korean Generative Commonsense Reasoning
Jaehyung Seo, Chanjun Park, Heonseok Moon, **Sugyeong Eo**, HeuiSeok Lim

DOMESTIC JOURNAL

Policy-based performance comparison study of Real-time Simultaneous Translation
Jungseob Lee, Hyeonseok Moon, Jaehyung Park, **Sugyeong Eo**, Seungjun Lee, Seonmin Koo, Heuiseok Lim
Journal of the Korea Convergence Society 13.3 (2022) pp. 43–54. Korea Convergence Society, 2022

Research on Subword Tokenization of Korean Neural Machine Translation and Proposal for Tokenization Method to Separate Jong-sung from Syllables
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, HeuiSeok Lim
Journal of the Korea Convergence Society 12.3 (2021) pp. 1–7. Korea Convergence Society, 2021

Research on Recent Quality Estimation
Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, HeuiSeok Lim
Journal of the Korea Convergence Society 12.7 (2021) pp. 37–44. Korea Convergence Society, 2021

Study on Zero-shot based Quality Estimation
Sugyeong Eo, Chanjun Park, Jaehyung Seo, Hyeonseok Moon, Heuiseok Lim
Journal of the Korea Convergence Society 12.11 (2021) pp. 35–43. Korea Convergence Society, 2021

Filter-mBART Based Neural Machine Translation Using Parallel Corpus Filtering
Hyeonseok Moon, Chanjun Park, **Sugyeong Eo**, JeongBae Park, HeuiSeok Lim
Journal of the Korea Convergence Society 12.5 (2021) pp. 1–7. Korea Convergence Society, 2021

A study on performance improvement considering the balance between corpus in Neural Machine Translation
Chanjun Park, Kinam Park, **Sugyeong Eo**, Hyeonseok Moon, HeuiSeok Lim
Journal of the Korea Convergence Society 12.5 (2021) pp. 23–29. Korea Convergence Society, 2021

Recent Automatic Post Editing Research
Hyeonseok Moon, Chanjun Park, **Sugyeong Eo**, Jaehyung Seo, HeuiSeok Lim
Journal of Digital Convergence 19.7 (2021) pp. 199–208. Journal of Digital Convergence, 2021

The Verification of the Transfer Learning-based Automatic Post Editing Model
Hyeonseok Moon, Chanjun Park, **Sugyeong Eo**, Jaehyung Seo, Heuiseok Lim
Journal of the Korea Convergence Society 12.10 (2021) pp. 27–35. Korea Convergence Society, 2021

Study on Decoding Strategies in Neural Machine Translation
Jaehyung Seo, Chanjun Park, **Sugyeong Eo**, Hyeonseok Moon, Heuiseok Lim
Journal of the Korea Convergence Society 12.11 (2021) pp. 69–80. Korea Convergence Society, 2021

A Study on Verification of Back TranScription(BTS)-based Data Construction
Chanjun Park, Jaehyung Seo, Seolhwa Lee, Hyeonseok Moon, **Sugyeong Eo**, Heuiseok Lim
Journal of the Korea Convergence Society 12.11 (2021) pp. 109–117. Korea Convergence Society, 2021

DOMESTIC CONFERENCE

KoCED: English-Korean Critical Error Detection Dataset (Best Paper Award)
Sugyeong Eo, Suwon Choi, Seonmin Koo, Dahyun Jung, Chanjun Park, Jaehyung Seo, Hyeonseok Moon, Jeongbae Park, Heuiseok Lim
The 34rd Annual Conference on Human Cognitive Language Technology, 2022

Word-level Korean-English Quality Estimation
Sugyeong Eo, Chanjun Park, Jaehyung Seo, Heonseok Moon, HeuiSeok Lim
The 33rd Annual Conference on Human Cognitive Language Technology, 2021

Design Neural Machine Translation Model Combining External Symbolic Knowledge
Sugyeong Eo, Chanjun Park, HeuiSeok Lim
The 32st Annual Conference on Human Cognitive Language Technology, 2020

BackTranScription (BTS)-based Jeju Automatic Speech Recognition Post-processor Research
Chanjun Park, Jaehyung Seo, Seolhwa Lee, Heonseok Moon, **Sugyeong Eo**, Yoonna Jang, HeuiSeok Lim
The 33rd Annual Conference on Human Cognitive Language Technology, 2021

Verification of the Domain Specialized Automatic Post Editing Model
Heonseok Moon, Chanjun Park, Jaehyung Seo, **Sugyeong Eo**, HeuiSeok Lim
The 33rd Annual Conference on Human Cognitive Language Technology, 2021

KommonGen: A Dataset for Korean Generative Commonsense Reasoning Evaluation (Best Paper Award)
Jaehyung Seo, Chanjun Park, Heonseok Moon, **Sugyeong Eo**, Myunghoon Kang, Seounghoon Lee, HeuiSeok Lim
The 33rd Annual Conference on Human Cognitive Language Technology, 2021

BOOK CHAPTERS

Natural Languagae Processing Bible
Heuiseok Lim, **Korea University NLPAI Lab**
Human Science, `https://www.aladin.co.kr/shop/wproduct.aspx?partner=rss&ISBN=K412637214`, 2020