ZHENHAO LI

(+44) (0) 7422538781 ⋄ zhenhao.li@epicgames.com ⋄ zhenhaoli.net ⋄ Github: Nickeilf

EDUCATION

Imperial College London, UK *October 2019 - Jan 2023; Jan 2025 - Aug 2025*
PhD in Natural Language Processing
Supervisors: Prof. Lucia Specia, Dr. Marek Rei
Thesis: Robustness in Natural Language Processing

Imperial College London, UK *September 2018 - September 2019*
MSc in Computing (Artificial Intelligence) *Degree: Distinction*
Supervisor: Prof. Lucia Specia

Nanjing University, China *September 2014 - June 2018*
BEng in Software Engineering
Supervisor: Prof. Jidong Ge

EMPLOYMENT

Lead Research Engineer Epic Games
AI for Gaming *Jan 2023 - Now*

· Continuously lead the development of the Epic Developer Assistant, including efforts such as Verse code generation, agentic text-to-game assistant, etc.

Research Engineer Epic Games
Content Moderation & Chatbot *Jan 2023 - Nov 2025*

· Research and development in content moderation, such as text toxicity detection, to support Epic Games ecosystem security.
· Built Epic Developer Assistant (EDA), a chatbot for game development in UEFN and Verse code generation.

Research Engineer Contex.ai (acquired by Epic Games)
Toxicity Detection *Dec 2021 - Jan 2023*

· Responsible for developing and improving the *text models* in toxicity detection. Built production-level single-task and multi-task text models, which outperform competitor products by over 0.15 F1 score.
· Implemented a novel *adversarial multi-task framework* to boost the performance on covert examples. Developed the initial framework of *conversational models*.

Research Assistant Imperial College London
Multimodal Machine Learning *June 2020 - Jan 2022*

· Built multimodal neural machine translation models robust to noise. Applied adversarial training with synthetic noisy data and improved performance by over 2 BLEU scores under noisy settings.

PUBLICATIONS

Huichi Zhou, Kin-Hei Lee, Zhonghao Zhan, Yue Chen, **Zhenhao Li**, Zhaoyang Wang, Hamed Haddadi, Emine Yilmaz. *Trustrag: Enhancing robustness and trustworthiness in rag*, 2025.

Zhenhao Li, Huichi Zhou, Marek Rei, Lucia Specia. *DiffuseDef: Improved Robustness to Adversarial Attacks via Iterative Denoising*, In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL), 2025.

Zhenhao Li, Marek Rei, Lucia Specia. *Multimodal Conversation Modelling for Topic Derailment Detection.* In Findings of the Association for Computational Linguistic: EMNLP, 2022.

Kaustubh D. Dhole, Varun Gangal, Sebastian Gehrmann, Aadesh Gupta, **Zhenhao Li**, Saad Mahamood, Abinaya Mahendiran, Simon Mille, Ashish Srivastava, Samson Tan, Tongshuang Wu, Jascha Sohl-Dickstein, Jinho D. Choi, Eduard Hovy, Ondrej Dusek, Sebastian Ruder, et al. *NL-Augmenter: A Framework for Task-Sensitive Natural Language Augmentation* arXiv preprint arXiv:2112.02721 (2021).

Genze Jiang, **Zhenhao Li**, Lucia Specia. *ICLs Submission to the WMT21 Critical Error Detection Shared Task.* In Proceedings of the Sixth Conference on Machine Translation (WMT), 2021.

Lucia Specia, Frdric Blain, Marina Fomicheva, Chrysoula Zerva, **Zhenhao Li**, Vishrav Chaudhary and Andr F. T. Martins. *Findings of the WMT 2021 Shared Task on Quality Estimation.* In Proceedings of the Sixth Conference on Machine Translation (WMT), 2021.

Zhenhao Li, Marek Rei, Lucia Specia. *Visual Cues and Error Correction for Translation Robustness.* In Findings of the Association for Computational Linguistics: EMNLP, 2021.

Khetam Al Sharou*, **Zhenhao Li***, Lucia Specia. *Towards a Better Understanding of Noise in Natural Language Processing.* In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP), 2021.

Lucia Specia, **Zhenhao Li**, Juan Pino, Vishrav Chaudhary, Francisco Guzmn, Graham Neubig, Nadir Durrani, Yonatan Belinkov, Philipp Koehn, Hassan Sajjad, Paul Michel and Xian Li. *Findings of the WMT 2020 Shared Task on Machine Translation Robustness.* In Proceedings of the Fifth Conference on Machine Translation (WMT), 2020.

Zhenhao Li, Marina Fomicheva, Lucia Specia. *"Exploring Model Consensus to Generate Translation Paraphrases"* In Proceedings of the Fourth Workshop on Neural Generation and Translation (WNGT), 2020.

Zhenhao Li, Lucia Specia. *"Improving Neural Machine Translation Robustness via Data Augmentation: Beyond Back-Translation"* In Proceedings of The 5th Workshop on Noisy User-generated Text (W-NUT), 2019.

Linxia Yao, Jidong Ge, Chuanyi Li, Yuan Yao, **Zhenhao Li**, Jin Zeng, Bin Luo, Victor Chang. *"Word Segmentation for Chinese Judicial Documents"* In Data Science, 5th International Conference of Pioneering Computer Scientists, Engineers and Educators, ICPCSEE, 2019.

OPEN SOURCE PROJECTS AND OTHER SERVICE

Core Contributor & Reviewer   GEM-benchmark/NL-Augmenter

· Acted as one of the framework organizers of NL-Augmenter – a repo for data augmentation in NLP
· Contributed to the core of data filtering and evaluation components, reviewed 30+ augmentation PRs

Workshop Co-organizer   Sixth Conference on Machine Translation (WMT21)
in Quality Estimation (QE) Shared Task   *Dec 2020 - Jul 2021*

· Acted as the co-organizer of WMT21 Quality Estimation Task 3
· Built a human annotation platform based on MT-EQuAl, which is used for the dataset collection

Workshop Co-organizer   Fifth Conference on Machine Translation (WMT20)
in Robustness Shared Task   *Apr 2020 - Dec 2020*

· Acted as the co-organizer of WMT20 Robustness Task and assisted in a new dataset
· Performed analysis on 23 system submissions from 9 teams, and gave a presentation on EMNLP 2020

Reviewer

· ACL (2025, 2023), EMNLP (2024, 2022), EACL (2025), COLING (2025), WMT (2025, 2022)

PATENT

Lucia Specia, Marina Fomicheva, Ozan Caglayan, **Zhenhao Li**, *"Multi-task neural network for toxicity detection"*, US Patent, Patent Number: US-20230419117-A1, 28 Dec 2023

Jidong Ge, Chuanyi Li, **Zhenhao Li**, Miaomiao Lei, Linxia Yao, Xiaoyu Zhou, Bin Luo, *"A Word Segmentation Method on Judicial Document Based on PageRank and Information Entropy."*, CN Patent, Application Number:201810534689.4, 30 May 2018

AWARDS

Second Place in Sentence-level Critical Error Detection of WMT21 QE shared task	Jul 2021
First Place in the English-Portuguese track of Duolingo STAPLE shared task	Apr 2020
Honorable Mention Prize in Mathematical Contest in Modeling	Apr 2017
Excellent Prize in the 12th Citi Financial Innovation Application Competition	Nov 2016
The People's Scholarship in China	2015, 2016, 2017

SKILL

Programming language: Python (expert), Java, Javascript, HTML, SQL
ML framework: Pytorch (expert), Huggingface (expert), Scikit-learn (expert), OpenNMT, Fairseq