

Weizhe Lin

☎ (44) 7596497504 • ✉ wl356@cam.ac.uk • 🌐 LinWeizheDragon

Education

Trinity College, University of Cambridge <i>PhD in Engineering - Toyota Research Studentship</i> & Visiting PhD in Chemical Engineering and Biotechnology (Stranks Lab)	10/2021 – 11/2024 (expected)
Trinity College, University of Cambridge <i>BA & M.Eng Information Engineering - Full Oversea Scholarship & Cambridge Trust Scholarship</i>	10/2017 – 06/2021
The University of Hong Kong <i>B. Eng GPA 3.91/4.3 - HKU Entrance Scholarship</i>	08/2016 – 06/2017
Xiamen No.1 High School of Fujian, China <i>Ranked 68 out of 200,000 in Joint College Entrance Examination, Fujian Province, China</i>	09/2013 – 06/2016

Work Experience

Co-founder and CTO at To0space (Part-time Volunteering), Beijing	02/2023 – Now
<ul style="list-style-type: none">Lead the research department of To0space. The team consists of ~ 10 postgraduate students of Tsinghua University and Cambridge University. Helped to raise the first round of funding from MiraclePlus.Develop AIGC solutions for architectural design (https://www.to0space.com).Improved the user saving rate from 10% to 30%.	
Intern Applied Scientist II at Amazon Development Center, Cambridge	09/2022 – 12/2022
<ul style="list-style-type: none">Carry out research in Open-domain Table Question Answering systems. Developed Retrieval-augmented TableQA systems. Researched retrieval systems such as DPR/ColBERT.Supervisor: Adrià de Gispert Ramis (Principal Applied Scientist)	
Intern Researcher at Microsoft Software Technology Center, Beijing	06/2021 – 10/2021
<ul style="list-style-type: none">Worked in developing cutting-edge knowledge-graph-enhanced recommender systems that bring together structured & unstructured data using retrieval systems.Supervisor: Linjun Shou (Principal Applied Scientist Manager)	
Researcher at Computer Laboratory, University of Cambridge	06/2020 – 09/2020
<ul style="list-style-type: none">Improved multi-agent path planning with graph attention networksSupervisor: Dr. Amanda Prorok	
Researcher at Computer Laboratory, University of Cambridge	06/2019 – 09/2019
<ul style="list-style-type: none">Proposed a network-based novel multi-modal feature fusion framework to aid in prediction of psychological disorder.Developed a self-adaptor(fidgeting) detection system for automated detection of psychological distressSupervisor: Dr. Marwa Mahmoud	
Software Development Contractor for Bingo Century Investment Management Limited	04/2019 - Now
<ul style="list-style-type: none">Independently developed a real-time news monitoring system with front-end presentation and server-end analysis	
Cloud Engineering Intern in Informetis Europe Ltd.	08/2018 – 09/2018
<ul style="list-style-type: none">Developed a Django backend system which tracks power usage of IoT devices through their Power Supply monitors. The system is designed to help ML engineers develop more accurate algorithm for tracking various electrical appliancesGained skills in database management, involving the use of MySQL, Google Bigtable and Redis Caching.	

Projects

Knowledge-based visually-grounded language understanding	2021 - Current
<ul style="list-style-type: none">Ph.D. research supervised by Prof. Bill Byrne.Investigating efficient knowledge retrieval in visually-grounded language systems, e.g., knowledge-based visual question answering. Developed the RA-VQA framework, a Retrieval-augmented VQA system.	
Knowledge-aware multi-domain task-oriented dialogue systems (Master dissertation)	2020 - 2021
<ul style="list-style-type: none">Final year project supervised by Prof. Bill Byrne. Awarded Engineering Outstanding Dissertation Prize (Top 2).Utilising neural forms of graph networks in dialogue systems.	
Graph-based multi-robot path planning	2020 - 2021
<ul style="list-style-type: none">Supervised by Dr. Amanda Prorok.Imitation learning using Graph Neural Network to communicate between agents.Utilising graph attention neural network to leverage the performance of moving agents to their goals.	
COVID-19 diagnosis assist and CT denoising (AIXCOVNET Project Support Member)	2020 - 2022
<ul style="list-style-type: none">Working with Stranks Lab of Cavendish Laboratory, NHS(Addenbrooke's Hospital), Department of Radiology. Supervised by Sam Stranks. Performing CT denoising on datasets of COVID-19 and other commonly-seen lung diseases.	

Image reconstruction for hyperspectral microscopy using deep learning

2019 - Current

- Working together with Stranks Lab of Cavendish Laboratory. Collaborate with VISION Laboratory of Department of Physics. Supervised by Dr. Sam Stranks.
- Using machine-learning-based methods to denoise and reconstruct physics-informed images obtain by microscopy.
- Highly reduced the required laser exposure time for taking images for physics/material research.

Automatic fidgeting and self-adaptor detection for psychological distress from 2D videos

2019 - 2020

- Developed a fully automated system to detect the fidgeting behaviour (such as touching face by hand and rhythmic body motion). Supervised by Dr. Marwa Mahmoud.
- Performing classification based on multi-modal features extracted from interview videos.

LearnAh.uk - teach science the fun way using popular science videos

2018 - 2019

- Recommends popular videos using techniques of machine-learning based text analysis (Latent Semantic Indexing & Latent Dirichlet Allocation) based on teaching plans - [Link](#)
- UCL Institute of Education Knowledge Lab EDUCATE Graduate (with EU grant) + Y Combinator Startup School Graduate + Runner-up, Cambridge University Entrepreneur £2,000 competition (Social Enterprise)

Achievements

Publications and Presentations ([Link to Google Scholar Page](#))

- Weizhe Lin**, Jingbiao Mei, Jinghong Chen, Bill Byrne. PreFLMR: Scaling Up Fine-Grained Late-Interaction Multi-modal Retrievers. 2024. To appear at *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024)*. **Read**
- Jingbiao Mei, Jinghong Chen, **Weizhe Lin**, Bill Byrne, Marcus Tomalin. Improving hateful memes detection via learning hatefulness-aware embedding space through retrieval-guided contrastive learning. To appear at *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024)*. **Read**
- Jinghong Chen, **Weizhe Lin**, Bill Byrne. CONTROL-DAG: Efficient Controlled Decoding for Directed Acyclic Non-Autoregressive Text Generation. 2024. *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2024)*. **Read**
- Guangyu Yang, Jinghong Chen, **Weizhe Lin**, Bill Byrne. Direct Preference Optimization for Neural Machine Translation with Minimum Bayes Risk Decoding. 2024. *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2024)*. **Read**
- Weizhe Lin**, Jinghong Chen, Jingbiao Mei, Alexandru Coca, Bill Byrne. Finer-grained Late-interaction Multimodal Retrieval for Knowledge-based Visual Question Answering. 2023. *Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS)*. **Read**
- Kangyu Ji, **Weizhe Lin**, Yuqi Sun, Linsong Cui, Javad Shamsi, Yu-Hsien Chiang, Jiawei Chen, Elizabeth Tennyson, Linjie Dai, Qingbiao Li, Kyle Frohna, Miguel Anaya, Neil Greenham. Sam Stranks. Self-supervised deep learning for tracking degradation of perovskite LEDs with multispectral imaging. *Nature Machine Intelligence*. **Read**
- Weizhe Lin**, Rexhina Blloshmi, Bill Byrne, Adria de Gispert and Gonzalo Iglesias. An Inner Table Retriever for Robust Table Question Answering. 2023 **with Amazon**. *Proceedings of the 61th Annual Meeting of the Association for Computational Linguistics (ACL)*. **Read**
- Weizhe Lin**, Rexhina Blloshmi, Bill Byrne, Adria de Gispert and Gonzalo Iglesias. LI-RAGE: Late Interaction Retrieval Augmented Generation with Explicit Signals for Open-Domain Table Question Answering. 2023 **with Amazon**. *Proceedings of the 61th Annual Meeting of the Association for Computational Linguistics (ACL)*. **Read**
- Alexandru Coca, Bo-Hsiang Tseng, Jinghong Chen, **Weizhe Lin**, Weixuan Zhang, Tisha Anders and Bill Byrne. Grounding Description-Driven Dialogue State Trackers with Knowledge-Seeking Turns. 2023. *24th Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)* (**Best Long Paper Award**). **Read**
- Weizhe Lin**, Zhilin Wang, and Bill Byrne. FVQA 2.0: Introducing Adversarial Samples for Fact-based Visual Question Answering. 2023. *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Findings (EACL)*. **Read**
- Alexandru Coca, Bo-Hsiang Tseng, **Weizhe Lin**, Bill Byrne. More Robust Schema-Guided Dialogue State Tracking via Tree-Based Paraphrase Ranking. 2023. *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Findings (EACL)*. **Read**
- Weizhe Lin** and Bill Byrne. Retrieval Augmented Visual Question Answering with Outside Knowledge. 2022. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. **Read**
- Weizhe Lin**, Linjun Shou, Ming Gong, Jian Pei, Zhilin Wang, Bill Byrne and Daxin Jiang. Transformer-Empowered Content-Aware Collaborative Filtering. 2022 (**with Microsoft**). In *Proceedings of the RecSys 2022: Fourth Knowledge-aware and Conversational Recommender Systems Workshop (KaRS)*. **Read**
- Weizhe Lin**, Linjun Shou, Ming Gong, Jian Pei, Zhilin Wang, Bill Byrne and Daxin Jiang. Combining Unstructured Content and Knowledge Graphs into Recommendation Datasets. 2022 (**with Microsoft**). In *Proceedings of the RecSys 2022: Fourth Knowledge-aware and Conversational Recommender Systems Workshop (KaRS)*. **Read**
- Weizhe Lin**, Bo-Hsian Tseng and Bill Byrne. Knowledge-Aware Graph-Enhanced GPT2 for Dialogue State Tracking. 2021. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. **Read**

- **Weizhe Lin**, Indigo Orton, Qingbiao Li, Gabriela Pavarini, Marwa Mahmoud. Looking At The Body: Automatic Analysis of Body Gestures and Self-Adaptors in Psychological Distress. 2021. *IEEE Transactions on Affective Computing*. **Read**
- Zhilin Wang, **Weizhe Lin** and Xiaodong Wu. Learning similarity between movie characters and its potential implications on understanding human experiences. 2021. In *Proceedings of the 2021 NAACL Workshop WNU: 3rd Workshop on Narrative Understanding*. **Read**
- Qingbiao Li*, **Weizhe Lin*** (*equal contribution), Zhe Liu and Amanda Prorok. Message-Aware Graph Attention Networks for Large Scale Multi-Robot Path Planning. 2021. *IEEE Robotics and Automation Letters (RA-L)*. **Read**
- **Weizhe Lin**, Indigo Orton, Mingyu Liu, Marwa Mahmoud. Automatic Detection of Self-Adaptors for Psychological Distress. 2020. In *Proceedings of 2020 15th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2020)*. **<Oral session> Read**
- Ziheng Zhang*, **Weizhe Lin*** (*equal contribution), Mingyu Liu, Marwa Mahmoud. Multimodal Deep Learning Framework for Mental Disorder Recognition. 2020. In *Proceedings of 2020 15th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2020)*. **<Oral session> Read**
- Zhilin Wang, Elena Rastorgueva, **Weizhe Lin** and Xiaodong Wu. No you're not alone A better way to find people with similar experiences on Reddit. 2019. In *Proceedings of the 2019 EMNLP Workshop W-NUT: The 5th Workshop on Noisy User-generated Text*. **Read**
- **Weizhe Lin**, Xiaodong Wu, Zhilin Wang and Elena Rastorgueva. Author2Vec: A Novel Framework for Generating User Embedding. 2019. on *Arxiv*. **Read**
- Zhilin Wang, Xiaodong Wu, **Weizhe Lin** and Elena Rastorgueva. Detecting personal attributes through analyzing online forums. 2019. In *Cambridge Language Sciences Early Careers Researchers Symposium*. **Read**

Personal Awards

- Engineering Part IIB Project Prize (final year project **First Class Honour**) (2021)
- Runner-up in Integrated Design Project (Robot design challenge in Dept. of Engineering) (2019)
- Airbus Defense and Space Prize 2018 for Mars Lander Design and Programming Contest (Runner Up) (2018)
- 1st Year Structural Design Course Prize (Dept. of Engineering) (2017)
- Full Oversea Scholarship from Trinity College & Cambridge Trust Scholarship (2017)
- Centenary Prize for Top 3 Information, Electrical and Electronics Engineering Final Project of Hong Kong University (2017)
- Third place in Computer Science Projects of 15th China Future Scientist Award Program (2015)
- Silver Medal in Chinese Physics Olympiad (Fujian Division) (2015)

Websites and Society

Websites I built

- Website of Hercules Cambridge society: <https://www.herculescambridge.org.uk>
- Website of NLP project LearnAh <https://learnah.uk>

Director of Web Development of Hercules Cambridge (Design Collaboration) 07/2021 – 12/2022

- Lead the design of society website system. Leader of a small group of (7-8) engineers and designers.
- ([Link to society](#))

Data Analysis Mentor of Bridge for Enterprise (NGO) 10/2019 – 2021

- Supervise the development of Investor Selector Project (using NLP to filter investors for startups)
- ([Link to society](#))