

# POTSAWEE MANAKUL

[pm574@cam.ac.uk](mailto:pm574@cam.ac.uk) | [potsawee.github.io](https://potsawee.github.io) | [github.com/potsawee](https://github.com/potsawee) | [Google Scholar](#)

## EDUCATION

---

### University of Cambridge

Cambridge, UK

*Ph.D. Candidate in Information Engineering (Computer Science)*

*Oct 2019 – Feb 2024 (expected)*

- Work in the Machine Intelligence Lab under the supervision of Prof. Mark Gales
- Recipient of the Cambridge International & St John's College Scholarship (Full funding for PhD)
- Research Focus: Generative Language Models, Summarization, Question-Answering, NLP
- Publication at top-tier NLP/Speech conferences, including ACL, EMNLP, \*ACL, and InterSpeech

### University of Cambridge

Cambridge, UK

*B.A. and M.Eng. in Information and Computer Engineering*

*Oct 2015 – Jun 2019*

- **top 1%** in the first-year undergraduate examination
- **top 3%** in the second-year undergraduate examination
- **top 4%** in the third-year undergraduate examination
- **Honour with Distinction** in the fourth-year (Masters Program)  
→ Research project on deep learning for grammatical error detection and correction

## WORK EXPERIENCE

---

### SCB 10X

Bangkok, Thailand

*AI Researcher*

*Nov 2023 – Present*

- Develop state-of-the-art open-source Thai LLM – Typhoon-7B
- My role is under Venture Builder with the goal of establishing an AI startup

### Amazon

Berlin, Germany

*Applied Scientist Intern*

*Sep 2021 – Mar 2022*

- Project: Weakly Supervised Video Summarization for Target Video Classification Task
- A research paper at the Amazon Computer Vision Conference (ACVC) 2022
- Community Shopping Science Team with Hani Al-Shater and Nam Khanh Tran

### Machine Intelligence Lab, Engineering Department

Cambridge, UK

*Undergraduate Research Intern*

*Jun 2018 – Aug 2018*

- Developed deep learning models for grammatical error detection
- Co-authored a paper accepted at ICASSP 2019

### MediaTek

Cambourne, UK

*Software Engineering Intern*

*Jun 2017 – Sep 2017*

- Developed GUI tools for a 5G simulator using C#/.NET and SQLite

### Healthera

Cambridge, UK

*Software Engineering Intern*

*Jun 2016 – Aug 2016*

- Designed and worked on the front-end part of the company's website

## RECENT ML/NLP PROJECTS

---

### 1. SelfCheckGPT: <https://github.com/potsawee/selfcheckgpt>

- Developed methods to detect hallucinations in large language models such as GPT-3
- Interviewed by the Washington Post about LLM Hallucination – <https://wapo.st/3IMgiZw>
- Invited to give a talk at Deep Learning: Classics and Trends – <https://mlcollective.org/dlct/>
- >250 stars on GitHub

### 2. MQAG: <https://github.com/potsawee/mqag0>

- Developed automatic multiple-choice question generation and question answering systems
- Models are open-sourced on HuggingFace – <https://huggingface.co/potsawee>

### 3. LongSum: <https://github.com/potsawee/longsum0>

- Developed state-of-the-art automatic summary generation systems for long articles and podcasts
- Podcast summarization system won 1st place at Spotify Podcast Summarization Challenge 2020

## SELECTED RESEARCH PUBLICATIONS

---

- A. Liusie, **P. Manakul** and M.J.F. Gales, “LLM Comparative Assessment: Zero-shot NLG Evaluation through Pairwise Comparisons using Large Language Models”, in **EACL 2024** (main).
- **P. Manakul**, A. Liusie and M.J.F. Gales, “SelfCheckGPT: Zero-Resource Black-Box Hallucination Detection for Generative Large Language Models”, in **EMNLP 2023** (main).
- **P. Manakul**, A. Liusie and M.J.F. Gales, “MQAG: Multiple-choice Question Answering and Generation for Assessing Information Consistency in Summarization”, in **AACL 2023** (main). [Area Chair Award (Generation and Summarization)]
- **P. Manakul** and M.J.F. Gales, “Sparsity and Sentence Structure in Encoder-Decoder Attention of Summarization Systems”, in **EMNLP 2021** (main).
- **P. Manakul** and M.J.F. Gales, “Long-Span Summarization via Local Attention and Content Selection”, in **ACL 2021** (main).
- **P. Manakul** and M.J.F. Gales, “CUED.SPEECH at TREC 2020 Podcast Summarisation Track”, in Text REtrieval Conference (**TREC**) 2020. [Won 1st place in the Spotify Podcast Summarisation Track, out of 8 teams & 22 systems]
- **P. Manakul**, M.J.F. Gales, L. Wang, “Abstractive Spoken Document Summarization Using Hierarchical Model with Multi-Stage Attention Diversity Optimization”, in **InterSpeech 2020**. [Best Student Paper Finalist]
- Y. Lu, M. J. Gales, K. M. Knill, **P. Manakul**, L. Wang, Y. Wang, “Impact of ASR Performance on Spoken Grammatical Error Detection”, in **InterSpeech 2019**.

## PROFESSIONAL SERVICES

---

### Reviewer

- NeurIPS 2023, ICML 2023, InterSpeech 2023, ICLR 2024
- IEEE Transactions on Audio, Speech, and Language Processing 2022, 2023

### Undergraduate Teaching (Supervision at Cambridge)

- Third-year Undergraduate: 3F7 Information Theory and Coding
- Second-year Undergraduate: 2P7 Vector Calculus, Linear Algebra, Probability

### Lab Demonstration

- MPhil in Machine Learning and Machine Intelligence: MLMI13 Introduction to NLP
- First-year Undergraduate: Engineering Lego Lab

## SELECTED HONOURS & AWARDS

---

### Machine Learning Competition

- 1st place in the Medical Note Problem List Summarization at the BioNLP Workshop at ACL 2023
- 1st place in the Spotify Podcast Summarization Challenge at TREC 2020 Podcasts Track

### Scholarships and Awards

- Cambridge International & St John's College Scholarship (Full funding for PhD) — 2019
- Wright Prize (Academic Prize, St John's College) — 2019, 2018, 2017, 2016
- Winifred Georgina Holgate Pollard Memorial Prize (University Prize) — 2017
- Earle Prize (St John's College Year Group Prize) — 2017
- BP 1st Year Prize (sponsored by BP, based on the 1st year exam) — 2016
- Gaskell Prize (St John's College Year Group Prize) — 2016

### Olympiads

- Silver Medal from 45th International Physics Olympiad (IPhO) in Astana, Kazakhstan — 2014
- Gold Medal from 15th Asian Physics Olympiad (APhO) at National University of Singapore — 2014
- Bronze Medal from 14th Asian Physics Olympiad (APhO) in Bogor, Indonesia — 2013

*Updated: January 2024*