Sedrick Scott Keh
sedrickkeh  3Sedrick Keh

Email: sedrickkeh@gmail.com
Website: https://sedrickkeh.github.io

EDUCATION

- Carnegie Mellon University
  MS in Machine Learning, GPA: 4.08/4.33
  Pittsburgh, PA  Sept. 2021 – Dec. 2022

- Hong Kong University of Science and Technology
  BSc in Data Science; BSc in Mathematics – First Class Honors, GPA: 4.09/4.30
  Hong Kong  Sept. 2017 – Jun. 2021

PEER-REVIEWED PUBLICATIONS

[1] PINEAPPLE: Personifying INanimate Entities by Acquiring Parallel Personification data for Learning Enhanced generation
Sedrick Scott Keh, Kevin Lu, Varun Gangal, Steven Y. Feng, Harsh Jhamtani, Malihe Alikhani and Eduard Hovy.
International Conference on Computational Linguistics (COLING 2022). October 2022.

[2] EUREKA: EUphemism Recognition Enhanced through Knn-based methods and Augmentation
Sedrick Scott Keh*, Rohit K. Bharadwaj*, Emmy Liu, Simone Tedeschi, Varun Gangal and Roberto Navigli.
The Third Workshop on Figurative Language Processing (EMNLP 2022). December 2022.

[3] Exploring Euphemism Detection in Few-Shot and Zero-Shot Settings
Sedrick Scott Keh.
The Third Workshop on Figurative Language Processing (EMNLP 2022). December 2022.

[4] NewsPanda: Media Monitoring for Timely Conservation Action
Sedrick Scott Keh*, Zheyuan Ryan Shi*, David J. Patterson, Nirmal Bhagabati, Karun Dewan, Areeandra Gopala,
Pablo Izquierdo, Debojyoti Mallick, Ambika Sharma, Pooja Shrestha, and Fei Fang.
Conference on Innovative Applications of Artificial Intelligence (IAAI 2023). February 2023.

[5] PANCETTA: Phoneme Aware Neural Completion to Elicit Tongue Twisters Automatically
Sedrick Scott Keh, Steven Y. Feng, Varun Gangal, Malihe Alikhani and Eduard Hovy.
Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023). May 2023.

[6] Hashtag-Guided Low-Resource Tweet Classification
Shizhe Diao*, Sedrick Scott Keh*, Liangming Pan, Zhiliang Tian, Yan Song and Tong Zhang.
International World Wide Web Conference 2023 (WWW 2023). May 2023.

PREPRINTS AND PAPERS IN SUBMISSION

[7] Semi-Supervised Noisy Student Pre-training on EfficientNet Architectures for Plant Pathology Classification
Sedrick Scott Keh
arXiv:2012.00332. December 2020.

[8] Myers-Briggs Personality Indicator Classification and Language Generation Using Pre-trained Language Models
Sedrick Scott Keh, I-Tsun Cheng.
arXiv:1907.06333. July 2019.

[9] MagmaDNN: Towards High-Performance Deep Learning Using Magma
Daniel Nichols, Sedrick Scott Keh, Kam Fai Chan.
Technical Report. July 2019.
Research Experience

- Carnegie Mellon University – Language Technologies Institute  
  Sept. 2021 - Present
  Advised by Prof. Daniel Fried and Prof. Eduard Hovy
  - **Language Grounding:** Working on asking informative clarifying questions in a collaborative and visually-grounded uncertainty resolution task. Paper in progress.
  - **Creative Text Generation:** Using commonsense reasoning to generate figurative language.
    1. Proposed a model to generate diverse and meaningful personifications by incorporating commonsense concepts like animacy. First-author paper accepted at *COLING 2022* [1].
    2. Designed a system to automatically detect euphemisms in a sentence. System won 1st place at the Euphemism Detection Shared Task at the FigLang Workshop at EMNLP 2022. Also investigated euphemism detection in a few-shot and zero-shot setting. Two papers (1 first-author, 1 solo-author) accepted at *EMNLP 2022 FigLang Workshop* [2] [3].
  - **Controllable Text Generation:** Gathered a dataset of tongue twisters and proposed a model to generate tongue twisters by modeling linguistic difficulty. First-author paper accepted at EACL 2023 [5].

- Carnegie Mellon University – Institute for Software Research  
  Sept. 2021 - Present
  Advised by Prof. Fei Fang
  - **AI/NLP for Social Good:** Deployed a system to automatically extract and classify weekly news articles related to wildlife conservation. This is a collaboration with WWF offices in UK, India, and Nepal. First-author paper accepted to *IAAI 2023* [4].

- HKUST Statistics and Machine Learning Group  
  Feb. 2020 - Dec. 2021
  Advised by Prof. Tong Zhang
  - **Social Media Text Generation:** Created an intelligent social media captioning model which makes predictions by considering closely related posts and replies. Co-first-author paper under review [6].
  - **Story Generation:** Explored prompt-based methods for story completion tasks.

Awards and Scholarships

- **Rank 1 – EMNLP 2022 Figurative Language Workshop Shared Task**  
  2022
- **Scholarship – HKUST Full Undergraduate Scholarship ($560,000 HKD)**  
  2017-2021
- **Scholarship – S.S. Chern Class Mathematics Achievement Scholarship ($30,000 HKD)**  
  2018-2021
- **Finalist (Top 20; training pool for IMO)** – Philippine Mathematical Olympiad  
  2015, 2016, 2017
- **Finalist (Top 30; training pool for IOI)** – National Olympiad in Informatics  
  2017

Academic Service and Teaching

- **Reviewer:** ACL 2023, EMNLP 2022, FigLang Workshop at EMNLP 2022
- **Conference Volunteer:** EMNLP 2022 (Abu Dhabi, UAE)
- **Teaching Assistant:** 10-617 Intermediate Deep Learning (Graduate-level)  
  Fall 2022
- **Teaching Assistant:** 10-315 Advanced Introduction to Machine Learning  
  Spring 2022
  - Nominated for the 2022 Alan J. Perlis Graduate Student Teaching Award
- **Mentor:** CMU Undergraduate AI Mentorship Program

Industry Experience

- **Thinking Machines Data Science**  
  Machine Learning Research Intern (Geospatial Team)  
  Jun. 2021 - Aug. 2021
- **Senti AI Tech Labs**  
  NLP Research Intern  
  Jul. 2020 - Sept. 2020