

# Chia-Hsuan Chang

shane.chang.tw@gmail.com • <https://ponshane.github.io>

## Research interest

Natural Language Processing, Health Informatics, Generative Artificial intelligence, Data Science, AI Bias and Fairness

## Education

- 2016 – 2021    **National Sun Yat-Sen University** – Kaohsiung, Taiwan  
PhD in Information Systems  
Dissertation: [A Study on Text Analysis Across Languages](#)  
Advisor: San-Yih Hwang. *GPA: 4.1/4.3*
- 2015 – 2016    **National Sun Yat-Sen University** – Kaohsiung, Taiwan  
MBA in Information Systems  
Advisor: San-Yih Hwang. *GPA: 4.1/4.3*
- 2011 – 2014    **National Sun Yat-Sen University** – Kaohsiung, Taiwan  
BA in Information Systems  
*GPA: 3.75/4.3*

## Research experience

- Jun. 2023 – Present    **Remote Research Volunteer**  
Supervisor: Grace Lu-Yao (Thomas Jefferson University SKCC)  
→ Using electronic health records to improve treatment options for patients who have prostate cancer and lung cancer
- Feb. 2023 – Present    **Postdoctoral Researcher**  
Mentor: Christopher C. Yang (Drexel University CCI)  
→ Investigating potential applications of large language models in the healthcare domain  
→ Studying fairness issue of predictive models deployed in healthcare service
- Oct. 2022 – Jan. 2023    **Postdoctoral Researcher**  
Mentor: San-Yih Hwang (National Sun Yat-Sen University)  
→ Developed a prompt method to detect adverse drug effects from social media  
→ Proposed a topic identification algorithm for multilingual corpus using large language models

Feb. 2020 – **Visiting Ph.D. Student**

Feb. 2021 Mentor: Christopher C. Yang (Drexel University CCI)

→ Proposed algorithms to bridge the language gap in consumer health vocabulary and consumer health search

→ Developed topic identification methods to reveal topic trends from large scale COVID-19 tweets

### Peer-reviewed publication

2024 Classifying Cancer Stage with Open-Source Clinical Large Language Models

**Chang, C. H.**, Lucas, M. M., Lu-Yao G., Yang, C. C.

*IEEE International Conference on Healthcare Informatics – Accepted*

2024 Constructing Cross-lingual Consumer Health Vocabulary with Word-Embedding from Comparable User Generated Content

**Chang, C. H.**, Wang, L., Yang, C. C.

*IEEE International Conference on Healthcare Informatics – Accepted*

2024 An ExplainableFair Framework for Predictive Modeling in Healthcare

Lucas, M. M., Wang, X., **Chang, C. H.**, Yang, C. C.

*IEEE International Conference on Healthcare Informatics – Accepted*

2023 Examining Public Awareness of Ageist Terms on Twitter: Content Analysis

Schramm, E., Yang, C. C., **Chang, C. H.**, Mulhorn, K., Yoshinaga S., Huh-Yoo, J.

*JMIR Aging* (indexing: ESCI)

2023 Resampling for Mitigating Bias in Predictive Model for Substance Use Disorder Treatment Completion

Lucas, M. M., **Chang, C. H.**, Yang, C. C.

*IEEE International Conference on Healthcare Informatics – Houston, USA*

2023 Prompting for Few-shot Adverse Drug Reaction Recognition from Online Review

**Chang, C. H.**, Chang, F. Y., Hwang, S. Y., and Yang, C. C.

*IEEE International Conference on Healthcare Informatics – Houston, USA*

2023 On Bridging Consumer Health Search Across Languages Using Cross-lingual Word Space

**Chang, C. H.**, & Yang, C. C.

*Electronic Commerce Research and Applications* (indexing: SCIE, SSCI)

2021 Topics and Sentiments of Public Concerns Regarding Covid-19 Vaccines: Social Media Trend Analysis

Monselise, M., **Chang, C. H.**, Ferreira, G., Yang, R., & Yang, C. C.

*Journal of Medical Internet Research* (indexing: SCIE)

- 2021 A Word Embedding-based Approach to Cross-lingual Topic Modeling  
**Chang, C. H.**, & Hwang, S. Y.  
*Knowledge and Information Systems* (indexing: SCIE)
- 2021 Learning Bilingual Sentiment Lexicon for Online Reviews  
**Chang, C. H.**, Hwang, S. Y., & Wu, M. L.  
*Electronic Commerce Research and Applications* (indexing: SCIE, SSCI)
- 2021 What Are People Concerned About During the Pandemic? Detecting Evolving Topics about COVID-19 from Twitter  
**Chang, C. H.**, Monselise, M., & Yang, C. C.  
*Journal of Healthcare Informatics Research* (indexing: ESCI)
- 2020 Employing Word Mover's Distance for Cross-lingual Plagiarized Text Detection  
Chia-Ming Chang, **Chang, C. H.**, and Hwang, S. Y.  
*Annual Meeting of the Association for Information Science and Technology – Online*
- 2019 An Approach to Cross-lingual Sentiment Lexicon Construction  
**Chang, C. H.**, Wu M. L., and Hwang, S. Y.  
*IEEE International Congress on Big Data (BigDataCongress) – Milan, Italy*
- 2018 Incorporating Word Embedding into Cross-lingual Topic Modeling  
**Chang, C. H.**, Hsu, T. X., and Hwang, S. Y.  
*IEEE International Congress on Big Data (BigData Congress) – San Francisco, USA*

## Professional service

- 2023, 2024 **Reviewer.** IEEE Transactions on Knowledge and Data Engineering (TKDE)
- 2023, 2024 **Reviewer.** Knowledge and Information Systems (KAIS)
- 2023, 2024 **Reviewer.** Journal of Healthcare Informatics Research (JHIR)
- 2023 **Reviewer.** International Journal of Data Science and Analytics (JDSA)

## Award and honor

- 2021 Exceptional Doctoral Dissertation (National Sun Yat-Sen University)
- 2021 Honorary Member (The Phi Tau Phi Scholastic Honor Society of the R.O.C.)
- 2020 Scholarship (Ministry of Science and Technology, Taiwan)  
*NT 900,000 for one year, Graduate Students Study Abroad Program*
- 2016 – 2020 Industry-Academy Cooperation Scholarship (Ministry of Education, Taiwan)  
*NT 1,200,000 for four years*
- 2016 – 2018 PhD Two-Year Scholarship (National Sun Yat-sen University)  
*NT 684,000 for two years*

2015 – 2016      Five-Year BA-MA Program Scholarship (National Sun Yat-sen University)  
*NT 120,000 for one year*

## Teaching experience

- Summer 2019 &  
Summer 2021      **Instructor, Fresh Undergraduate Training for Text Mining Research (NSYSU TAR Lab)**  
→ Organized the [curriculum](#) for fresh undergraduates to get familiar with necessary mathematical skills for machine learning, including linear algebra, multivariate calculus, and optimization approach  
→ Guided the paper discussions on the developments of NLP, such as topic modeling, word embedding, and transformer
- Autumn 2017 &  
Autumn 2018      **Teaching assistant, MIS 532: Process Management in Service Computing (National Sun Yat-sen University)**  
→ Taught students to implement microservices (SOAP/WSDL and RESTful API)  
→ Taught students to deploy microservices using docker and docker-compose
- Spring 2016 &  
Spring 2017      **Teaching assistant, MIS 205 & MIS 512: Database Management (National Sun Yat-sen University)**  
→ Prepared the hand-on materials for helping students learn the CRUD operations in relation database (i.e., OracleDB) and document-oriented database (i.e., MongoDB)
- Summer 2016      **Instructor, Fresh Undergraduate Training for Text Mining Research (NSYSU TAR Lab)**  
→ Organized the [curriculum](#) for fresh undergraduates to learn fundamental knowledge about conducting researches on text mining  
→ Taught introduction to clustering algorithms, MongoDB operations, and packages for performing NLP

## Technical skill

### Programming languages

Python, R, Java, Bash, SQL, HTML/CSS/JS,  $\LaTeX$

### Software

Git, Linux, Docker,  
Spacy, LangChain, HuggingFace Transformers, PyTorch, Keras,  
MongoDB, Chroma DB, MySQL

### Medical Language Processing

Universal Medical Language System (UMLS), ICD-10

**Languages**

Chinese (native), English (fluent)