

# Wei-Hung Weng

Cambridge, MA, USA

☎ (+1) 857-400-4997 | ✉ ckbjimmy@mit.edu | 🏠 people.csail.mit.edu/ckbjimmy | 🌐 ckbjimmy

## Education

### Massachusetts Institute of Technology (MIT)

PHD IN COMPUTER SCIENCE (GPA: 4.8/5.0)

Cambridge, MA, USA

Sep 2017 - May 2022

### Harvard Medical School

MASTER OF MEDICAL SCIENCE (MMSc) IN BIOMEDICAL INFORMATICS (GPA: 4.0/4.0)

Boston, MA, USA

Sep 2015 - May 2017

### Chang Gung University

DOCTOR OF MEDICINE (MD) (GPA: 3.99/4.0)

Taoyuan, Taiwan

Sep 2004 - Jun 2011

## Research Experience

### Google Research

RESEARCH SCIENTIST

- Early scientific discovery research in machine learning for healthcare

Mountain View, CA

Feb 2022 - present

### Computer Science and Artificial Intelligence Laboratory (CSAIL), MIT

GRADUATE RESEARCH ASSISTANT IN CLINICAL DECISION MAKING GROUP (PROF. PETER SZOLOVITS)

- Learning representations for limited and heterogeneous medical data

Cambridge, MA

Sep 2017 - May 2022

### MIT Critical Data

INSTRUCTOR, MENTOR

- Lectured and mentored the course, conducted datathons and workshops for collaborative data science in medicine

Cambridge, MA

Jul 2017 - May 2022

### Google

RESEARCH INTERN

- Developing algorithms for mitigating negative interference in domain generalization (with Jonathan Deaton, Gamaleldin Elsayed, Jun-Aug 2021)
- Meta-learning with few-shot classification for extreme class imbalance in dermatology (with Yuan Liu, Vivek Natarajan, May-Aug 2020)
- Multimodal multitask representation learning for pathology metadata prediction (with Cameron Chen, Yuannan Cai, May-Aug 2019)

CA/MA

2019-2021

### Laboratory of Computer Science, Massachusetts General Hospital

POSTDOCTORAL RESEARCH FELLOW IN LABORATORY OF COMPUTER SCIENCE AND DEPARTMENT OF NEUROLOGY (DRS. HENRY

CHUEH, BRANDON WESTOVER)

- Machine learning, natural language processing and knowledge representation for clinical triage and undiagnosed disease identification

Boston, MA

Sep 2015 - May 2017

### Division of Clinical Informatics Solutions and Services, Philips Research North America

RESEARCH INTERN (DR. SANDEEP DALAL)

- Proposed the anatomy term extraction method for radiology reports using natural language processing and knowledge representation

Cambridge, MA

Jun 2016 - Aug 2016

### Molecular Imaging Center, National Taiwan University

RESEARCH ASSOCIATE (PROF. CHI-KUANG SUN)

- Developed the qualitative and deep learning methods to identify skin melanocytes in microscopic imaging
- Proposed a go-to market strategy of virtual biopsy microscopy for skin cancer detection

Taipei, Taiwan

May 2014 - Apr. 2015

### Kidney Research Center / Cancer Molecular Diagnostics Lab / Department of Psychology, Chang Gung Memorial Hospital

RESEARCH INTERN (PROFS. LEE-YUNG SHIH, YA-CHUNG TIAN, CHIN-YEN CHEN)

- Investigated cytokines effects on polyomavirus BK infection
- Clinical data analysis of concomitance of essential thrombocythemia and chronic myeloid leukemia
- Investigated the relationship between sleepiness scale and heart rate variability in medical students

Taoyuan, Taiwan

Jun 2008 - May 2011

### Cancer Immunology and Gene Therapy Lab, Johns Hopkins Medical Institute

RESEARCH INTERN (PROF. TZY-CHOU WU)

- Immunology/molecular cell biology techniques training and review writing

Baltimore, MD

Jul 2009 - Aug 2009

## Clinical Experience

### Chang Gung Memorial Hospital

RESIDENT PHYSICIAN IN PATHOLOGY AND GENERAL MEDICINE

- Supervised more than 10 medical clerks, interns and residents
- Selected as visiting scholar to Department of Pathology, University of Tokyo Hospital, Tokyo, Japan

Taoyuan, Taiwan

Aug 2012 - Apr 2014

### Republic of China Army

MEDICAL OFFICER OF HEALTH, SECONDARY LIEUTENANT

- Managed a medical clinic, and a shelter for soldiers with psychiatric diseases
- Volunteered in Mbanane Government Hospital in Swaziland, for critical ward care and rural outreaches

Taoyuan, Taiwan

Aug 2011 - Jul 2012

## Skills

Programming / Toolkit	<b>Domain</b>	Machine learning, Deep learning, Natural language processing, Computer vision, Graphic design, Medicine, Pathology
	<b>Languages</b>	Python, R, Linux, $\text{\LaTeX}$ , HTML5, Git, Tensorflow, PyTorch, Keras, Scikit-learn,
	<b>Medicine</b>	Bioconductor, cTAKES, MetaMap, UMLS, MIMIC, Photoshop, InDesign
		English, Mandarin, Taiwanese, Japanese

## Honors & Awards

2020	<b>Government Scholarship to Study Abroad</b> , Ministry of Education Republic of China (Taiwan)	Taipei, Taiwan
2019	<b>GSC Travel Grant</b> , MIT	Cambridge, MA
2017	<b>Rolf G. Locher Graduate Fellowship</b> , MIT	Cambridge, MA
2015	<b>Scholarship of Harvard University Alumni Association (HAA) in Taiwan</b> , HAA in Taiwan	Taipei, Taiwan
2015	<b>Wang Yuan-Chun Memorial Medical Contribution Scholarship</b> , HAA in Taiwan	Taipei, Taiwan
2015	<b>Semifinalist (TxNet - Education app for kidney transplant patients)</b> , BWH Startup Program	Boston, MA
2015	<b>First Prize Award (Fetal heart beat detector)</b> , HIT Biomedical Innovation Hackathon	Taipei, Taiwan
2011	<b>Objective Structured Clinical Examination (OSCE) Award</b> , Taiwan Association of Medical Education	Kaohsiung, Taiwan
2011	<b>Second Prize</b> , Medical Record Writing Competition	Kaohsiung, Taiwan
2011	<b>Great Honor</b> , Taiwan Medical Development Awards for Outstanding Writings	Taipei, Taiwan
2009	<b>Research Funding</b> , Medical Foundation in Memory of Dr. Deh-Ling Cheng	Kaohsiung, Taiwan
2006-08	<b>Presidential Award</b> , CGU Medicine (top 5% students)	Taoyuan, Taiwan
	<b>Conference Travel Award</b> , NeurIPS 2019, NeurIPS ML4H 2019, KDD 2019, NIPS ML4H 2017	

## Talks and Workshops

2021/05	<b>Invited speaker</b> , National Taiwan University, College of Medicine, Symposium of AI in Healthcare	Virtual
2020/02	<b>Invited speaker</b> , Advanced Biomedical Computation series in Brigham and Women's Hospital	Boston, USA
2020/02	<b>Invited mentor</b> , Society of Critical Care Medicine Datathon	Orlando, USA
2019/10	<b>Invited speaker</b> , Big Data in Bhutan	Thimphu, Bhutan
2019/09	<b>Invited speaker</b> , 2019 TMU-MIT Healthcare Datathon	Taipei, Taiwan
2019/08	<b>Invited mentor</b> , Aarhus Critical Care Datathon 2019	Aarhus, Denmark
2019/03	<b>Workshop moderator</b> , 2nd Big Data Machine Learning in Healthcare in Japan	Tokyo, Japan
2018/07	<b>Workshop moderator</b> , in Big Data for Health workshops and conference	Manila, Philippines
2018/07	<b>Workshop moderator</b> , ealthcare AI Datathon - Technical Workshop	Singapore
2018/07	<b>Invited mentor</b> , Healthcare AI Datathon	Singapore
2017/11	<b>Invited mentor</b> , Chinese PLA General Hospital-MIT Health Data Conference and Workshop	Beijing, China PRC
2017/07	<b>Invited speaker</b> , in UP-MIT-Stanford-AeHIN Big Data for Health Conference and Workshops for Asia-Pacific	Cebu, Philippines
2017/06	<b>Invited speaker</b> , in NUS-MIT Healthcare Analytics Datathon	Singapore
2016/10	<b>Invited panelist</b> , Deconstructing the Hype around Big Data and Precision Medicine	Taipei, Taiwan
2016/09	<b>Invited panelist, mentor and judge</b> , MIT-TMU Internet of Things Hackathon	Taipei, Taiwan

## Professional Activities

**Conference / Journal / Grant Reviewer**, AAAI (2021), ACL (2020), NeurIPS (2019, 2020), MICCAI (2020), NeurIPS Machine Learning for Health Workshop (2017, 2018), Machine Learning for Healthcare (2019, 2020), AMIA Joint Summits (2018, 2019), AMIA Annual Symposium (2018), Journal of the American Medical Informatics Association (journal), Journal of Biomedical Informatics (journal), PLoS One (journal), ACM Transactions on Computing for Healthcare (journal), IEEE Transactions on Pattern Analysis and Machine Intelligence (journal), IEEE Transactions on Medical Imaging (journal), IEEE Transactions on Industrial Informatics (journal), International Journal of Medical Informatics (journal), Poland National Science Center (grant proposal)

2018-2022 **Consultant**, HoloEducation (Startup: AI for medical education)

Cambridge, MA

2018-2021 **Consultant**, Vysioneer (Startup: machine learning for medical imaging)

Cambridge, MA

2018 **Organizing Committee and Moderator**, Boston Taiwanese Biotechnology Symposium 2018

Cambridge, MA

2018 **TBI Year-in-Review Workgroup Member**, AMIA Joint Summits on Translational Science (2017, 2018)

San Francisco, CA

2017 **Scientific Program Committee**, AMIA 2017 Annual Symposium

Washington, DC

## Academic Activities

2021 **Machine Learning for Healthcare**, Teaching assistant, 6.871, MIT

Cambridge, MA

2019 **Graduate Admissions Student Committee**, MIT EECS

Cambridge, MA

2018, 2019 **Collaborative Data Science in Medicine**, Instructor, HST.953, MIT

Cambridge, MA

2016 **Lab of Introduction of Biomedical Informatics**, Instructor, BMI701, Harvard Medical School

Boston, MA

## Publications

### SUMMARY

- Google scholar: <https://scholar.google.com/citations?user=CKE0A-AAAAAJ&hl=en>
- Total citation: 1542 / h-index: 14 / i10-index: 22 (Accessed May 13, 2022)

### CONFERENCES/WORKSHOPS

1. Chen RJ, Lu MY, **Weng WH**, Chen TY, Williamsom DFK, Manz T, Shady M, Mahmood F. Multimodal Co-Attention Transformer for Survival Prediction in Gigapixel Whole Slide Images. ICCV 2021.
2. **Weng WH**, Szolovits P, Weng TW. Robust Transfer Learning for Distributional Shift in Medicine. Under review.
3. **Weng WH**, Deaton J, Natarajan V, Elsayed GF, Liu Y. Addressing the Real-world Class Imbalance Problem in Dermatology. NeurIPS ML4H 2020.
4. McDermott MBA, Hsu TM, **Weng WH**, Ghassemi M, Szolovits P. ChexBERT: Approximating the CheXpert labeler for Speed, Differentiability, and Probabilistic Output. MLHC 2020.
5. Hu SY, Wang S, **Weng WH**, Wang J, Wang X, Ozturk A, Li Q, Kumar V, Samir AE. Self-Supervised Pretraining with DICOM metadata in Ultrasound Imaging. MLHC 2020.
6. Rawat BPS, **Weng WH**, Raghavan P, Szolovits P. Entity-Enriched Neural Models for Clinical Question Answering. ACL BioNLP 2020.
7. Hu SY, Wang S, **Weng WH**, Wang J, Wang X, Ozturk A, Li Q, Kumar V, Samir AE. Weakly Supervised Context Encoder using DICOM metadata in Ultrasound Imaging. ICLR AI4H 2020.
8. **Weng WH**, Chung YA, Tong S. Clinical Text Summarization with Syntax-Based Negation and Semantic Concept Identification. Preprint 2020.
9. Huang Z, Huang S, Chen L, **Weng WH**, Wang L, Cui X, Feng C, Zhang Z, Li T. Artificial Intelligence Assisted Early Warning System for Acute Kidney Injury Driven by Multi-Center ICU Database. medRxiv 2020.
10. **Weng WH**, Cai Y, Lin A, Tan F, Chen PHC. Multimodal Multitask Representation Learning for Metadata Prediction in Pathology. NeurIPS ML4H 2019 + Medical Imaging Meets NeurIPS 2019.
11. Ma R, Chen PHC, Li G, **Weng WH**, Lin A, Gadepalli K, Cai Y. Investigate BERT on Pathology Report Classification for Accelerating Human Annotation. NeurIPS ML4H 2019.
12. Mao H, Narayan A, Negi P, Wang H, Yang J, Wang H, Khani M, He S, Addanki R, Marcus R, Cangialosi F, **Weng WH**, Han S, Kraska T, Alizadeh M. Park: An Open Platform for Learning Augmented Computer Systems. ICML 2019 Reinforcement Learning for Real Life Workshop, NeurIPS 2019.
13. Hu SY, **Weng WH**, Lu SL, Cheng YH, Xiao F, Hsu FM, Lu JT. Multimodal Volume-Aware Detection and Segmentation for Brain Metastases Radiosurgery. MICCAI AIRT 2019.
14. **Weng WH**, Szolovits P. Representation Learning for Electronic Health Records. Preprint 2019.
15. **Weng WH**, Alsentzer E, Jin D, Min SY, Raghavan P, Szolovits P. Logical Form Information for Clinical Question Answering. KDD DSHealth 2019.

16. **Weng WH**, Chung YA, Szolovits P. Unsupervised Clinical Language Translation. KDD 2019. (oral) [press]
17. Chung YA, **Weng WH**, Tong S, Glass J. Towards Unsupervised Speech-to-Text Translation. ICASSP 2019. (oral)
18. Alsentzer E, Murphy JR, Boag W, **Weng WH**, Jin D, Naumann T, McDermott M. Publicly Available Clinical BERT Embeddings. NAACL-HLT 2019 Clinical NLP Workshop. (oral)
19. Liu G, Hsu TM, McDermott M, Boag W, **Weng WH**, Szolovits P, Ghassemi M. Clinically Accurate Chest X-Ray Report Generation. MLHC 2019.
20. Hsu TM, **Weng WH**, Boag W, McDermott M, Szolovits P. Unsupervised Multimodal Representation Learning across Medical Images and Reports. NeurIPS ML4H 2018.
21. Girkar U, Uchimito R, Lehman LH, Celi LA, Szolovits P, **Weng WH** (corresponding). Predicting Blood Pressure Response to Fluid Bolus Therapy Using Attention-Based Neural Networks. NeurIPS ML4H 2018.
22. **Weng WH**, Szolovits P. Mapping Unparalleled Clinical Professional and Consumer Languages with Embedding Alignment. 2018 KDD Machine Learning for Medicine and Healthcare. (oral)
23. Chung YA, **Weng WH**, Tong S, Glass J. Unsupervised Cross-Modal Alignment of Speech and Text Embedding Spaces. NIPS 2018. (spotlight) [press]
24. McDermott M, Vide S, Alenyà M, Tróconiz IF, Valencia JF, Borrat X, **Weng WH**, Szolovits P, Gambús PL. Continuous Prediction of Sedation Levels Based on Signal Inputs: Evaluation of Different Modeling Approaches Including Machine Learning. ISAP 2018.
25. **Weng WH**, Gao M, He Z, Yan S, Szolovits P. Representation and Reinforcement Learning for Personalized Glycemic Control in Septic Patients. NIPS Machine Learning for Health 2017.
26. **Weng WH**, Chung YA. Learning Image Representations using Deep Siamese CNNs for Content-Based Medical Image Retrieval. Medical Imaging meets NIPS 2017 / NIPS Machine Learning for Health 2017.
27. **Weng WH**, Waghlikar KB, Chueh HC. Computing Performance Analysis on Clinical Document-level Classification. AMIA 2017.
28. **Weng WH**, Waghlikar KB, Chueh HC. Classifiers for Identifying the Medical Specialty of Clinical Documents. 2017 AMIA CRI.
29. Sun CK, Wei ML, Su YH, **Weng WH**, Liao YH. Molecular imaging of melanin distribution in vivo and quantitative differential diagnosis of human pigmented lesions using label-free harmonic generation biopsy. SPIE Photonics West 2017.
30. **Weng WH**. Predictive Analytics Pipeline for Clinical Narrative Document Classification. Artificial Intelligence in Medicine 2016.
31. **Weng WH**. Medical Domain Classification with Sequential Label-Embedded Neural Concept Embedding Model. Artificial Intelligence in Medicine 2016.
32. **Weng WH**, Khatrri A, Waghlikar KB, Cohen AB, Chueh HC. Improving the Workflow of Curbside Consultation by Using Unstructured Clinical Notes - a Natural Language and Machine Learning-based Approach. AMIA 2016.
33. **Weng WH**, Waghlikar KB. Supervised Clinical Document Classification Pipeline. AMIA 2016 NLP WG Pre-Symposium.
34. **Weng WH**, Waghlikar KB. Classifying Clinical Documents into Medical Domains. AMIA 2016 NLP WG Pre-Symposium.
35. **Weng WH**, Liao YH, Tsai MR, Huang HY, Sun CK. Differentiating intratumoral melanocytes from Langerhans cells in non-melanocytic pigmented skin tumors in vivo by using third harmonic generation microscopy. Focus on Microscopy Conference 2016.
36. Huang HF, **Weng WH**, Hsu WH, Sun CK. Automated Detection of Noninvasive Imaging of Basal Cell Carcinoma by Convolutional Neural Network. NIPS 2015 Workshop on Machine Learning in Healthcare.
37. **Weng WH**, Tsai MR, Liao YH, Sun CK. Differentiating pigmented skin tumors by the tumor-associated melanocytes based on in vivo third harmonic generation microscopy. SPIE Photonics West 2015.
38. Lee SY, **Weng WH**, Sun CK. Super resolution brain imaging by using a two-photon fluorescence microscopy with harmonic modulation. SPIE Photonics West 2015.
39. Chou YH, Hung SY, Lee GG, **Weng WH**, Liao YH, et al. Nuclei location enhancement based on improved efficient ellipse hough transform for third harmonic generation microscopy imaging. IEEE ChinaSIP 2015.
40. Liao YH, **Weng WH**, Sun CK. Characterization of dendritic cells in pigmented skin tumors by harmonic generation microscopy. 23rd World Congress of Dermatology 2015.
41. Lee GG, Cai CS, Liao YH, **Weng WH**, Sun CK, et al. Quantitative Gabor feature analysis of collagen fibers in harmonically generated microscopy (HGM) imaging. OMICS International Global Summit and Expo on Multimedia and Applications ETMN Pre-conference Workshop 2015.
42. **Weng WH**, Tsai MR, Liao YH, Sun CK. Identifying melanocyte in pigmented skin lesions based on in vivo third harmonic generation microscopy. Biomedical Molecular Imaging 2014.
43. **Weng WH**, Liu WM, Tsai MR, Liao YH, Sun CK. In vivo quantification of melanin mass density in human by using third harmonic generation microscopy. Biomedical Molecular Imaging 2014.
44. **Weng WH**, Tian YC. Interleukin-1 beta Inhibits BK virus gene expression and replication in human renal proximal tubular epithelial cells. Taiwan Medical Development Awards for Outstanding Writings 2010.
45. **Weng WH**, Tian YC. IL-1 has suppressive effect on BKV replication. Summer Student Study for Infectious Disease 2009.

## JOURNALS

1. **Weng WH\***, Jin D\*, Sergeeva E\*, Chauhan G\*, Szolovits P. Explainable Deep Learning in Healthcare: A Methodological Survey. WIREs Mechanisms of Disease. 2022.
2. Jin D, Pan E, Oufattole N, **Weng WH**, Fang H, Szolovits P. What Disease does this Patient Have? A Large-scale Open Domain Question Answering Dataset from Medical Exams. Applied Sciences. 2021.

3. Wei ML, Liao YH, **Weng WH**, Shih YT, Sheen YS, Sun CK. A Study on Applying Slide-Free Label-Free Harmonic Generation Microscopy For Noninvasive Assessment of Melasma Treatments With Histopathological Parameters. *IEEE Journal of Selected Topics in Quantum Electronics*. 2021.
4. Amorim E, Mo SS, Palacios S, Ghassemi MM, **Weng WH**, Cash SS, Bianchi MT, Westover MB. Cost-effectiveness analysis of multimodal prognostication in cardiac arrest with EEG monitoring. *Neurology*. 2020.
5. Chen YC, Li X, Zhu H, **Weng WH**, Tan X, Chen Q, Sun Y, Wang X, Fan X. Monitoring Neuron Activities and Interactions with Laser Emissions. *ACS Photonics*. 2020.
6. Hombal V, Dalal S, **Weng WH**, Mankovich G, Fuller JT, Mabotuwana T, Gunn ML, Hall CS. Determining Follow-Up Imaging Study Using Radiology Reports. *Journal of Digital Imaging*. 2019.
7. Li X, Qin Y, Tan X, Chen YC, Chen Q, **Weng WH**, Wang X, Fan X. Ultrasound Modulated Droplet Lasers. *ACS Photonics*. 2019.
8. Chen YC, Li X, Zhu H, **Weng WH**, Tan X, Chen Q, Wang X, Fan X. Laser Recording of Subcellular Neuron Activities. *bioRxiv*. 2019 Jan 1:584938.
9. **Weng WH**, Waghlikar KB, McCray AT, Szolovits P, Chueh HC. Medical subdomain classification of clinical notes using machine learning-based natural language processing approach. *BMC Medical Informatics and Decision Making*. 17(1):155.
10. **Weng WH**, Liao YH, Tsai MR, Wei ML, Huang HY, Sun CK. Differentiating intratumoral melanocytes from Langerhans cells in nonmelanocytic pigmented skin tumors in vivo by label-free third-harmonic generation microscopy. *Journal of biomedical optics*. 2016 Jul;21(7):076009.
11. Liu CH, Tang WR, **Weng WH**, Lin YH, Chen CY. The process of coping with stress by Taiwanese medical interns: a qualitative study. *BMC medical education*. 2016 Dec;16(1):10.
12. Lin YH, Chen CY, Lin SH, Liu CH, **Weng WH**, Kuo TB, Yang CC. Gender differences in cardiac autonomic modulation during medical internship. *Psychophysiology*. 2013 Jun;50(6):521-7.
13. **Weng WH**, Shih LY. Occurrence of BCR-ABL1-positive chronic myeloid leukemia following essential thrombocythemia. *Acta haematologica*. 2011;126(4):220-3.
14. Hung CF, Monie A, **Weng WH**, Wu TC. DNA vaccines for cervical cancer. *American journal of translational research*. 2010;2(1):75.

## BOOK CHAPTER

1. **Weng WH**. Machine Learning for Clinical Predictive Analytics. in *Leveraging Data Science for Global Health*. Springer Nature. 2019.