

# Min-Hsien Weng

Email: [samminweng@gmail.com](mailto:samminweng@gmail.com)

ORCID: <https://orcid.org/0000-0003-2334-1863>

LINKEDIN PROFILE: <https://www.linkedin.com/in/min-hsien-weng-3b753821/>

## Areas of specialization

Software Verification; Programming Analysis; Compiler; Parallel Programming

Natural Language Processing; Large Language Model; Data Analytic and Visualization

## Skills

**Programming:** Java / C / Python / C# / JavaScript / PHP / HTML / CSS / OpenCL / Java Threading / Hadoop / Map Reduce / OpenMP / Shell Script / Rust / Dafny

**AI:** Deep learning (Large Language Model) / PyTorch / TensorFlow / Pandas / NLTK / Stanford CoreNLP / HDBSCAN / BERT (Word2Vec) / Matplotlib / D3.js / Google Chart

**IoT:** Arduino Uno / ESP32 Board / Sensors / MQTT / IFTTT / RFID / Kiosk

**Architect:** Compiler / C Memory Management / RFID Controller / Product Data Management

**System:** Linux / ANT Build / XAMPP / Android / Git / AWS Cloud / Docker

**DB:** SQLite / MySQL / Oracle SQL Server / MangoDB

**Academics:** Boogie / SAT Solvers / Theorem Proving / Latex / Haskell / Whiley / Lean

## Appointments held

- 2022-2023.7 Teaching Fellow, Waikato University Joint Institute at Zhejiang University City College (NZUWI), Hamilton, New Zealand
- 2019-2022 Postdoc, Waikato University, Hamilton, New Zealand
- 2014-2019 Doctoral Assistant, Waikato University, Hamilton, New Zealand
- 2002-2011 Associate Researcher, Industrial Technology Research Institute, Taiwan

## Education

- 2014-2019 PhD in CS, Waikato University, New Zealand
- 2012-2013 MSc in CS, Waikato University, New Zealand
- 2011-2012 PGDIP in CS, Waikato University, New Zealand
- 1999-2001 MASTER in IEEM, National Tsing Hua University, Taiwan
- 1995-1999 BACHELOR in IEM, National Yang Ming Chiao Tung University, Taiwan

## Membership

- 2023 Full Membership in Sigma Xi, the Scientific Honor Society
- 2022 Emerging Professional Member of Engineering NZ (Membership number: 2009229)
- 2023 Reviewer for the open access journal Urban Planning (ISSN: 2183-7635)

## Publications

### JOURNAL ARTICLES AND CONFERENCE PAPERS

- 2022 **MH Weng**, S Wu, M Dyer, "Identification and Visualization of Key Topics in Scientific Publications with Transformer-Based Language Models and Document Clustering Methods", *Applied Sciences* 12 (21), 11220
- 2021 **MH Weng**, R Malik, M Utting, "Automatic proofs of memory deallocation for a Whyley-to-C Compiler", *Formal Methods in System Design* 57 (3), 429-472
- 2021 **MH Weng**, S Wu, M Dyer, "AI Augmented Approach to Identify Shared Ideas from Large Format Public Consultation", *Sustainability* 13 (16), 9310
- 2021 M Dyer, S Wu, **MH Weng**, "Convergence of public participation, participatory design and NLP to co-develop circular economy", *Circular Economy and Sustainability* 1 (3), 917-934
- 2020 M Dyer, **MH Weng**, S Wu, T Garcia Ferrari, R Dyer, "Urban narrative: Computational linguistic interpretation of large format public participation for urban infrastructure", *Cogitatio* 5 (4), 20-32
- 2019 M Dyer, R Dyer, **MH Weng**, S Wu, T Grey, R Gleeson, TG Ferrari, "Framework for soft and hard city infrastructures", *Proceedings of the Institution of Civil Engineers-Urban Design and Planning* 172 (6), 219-227
- 2019 M Dyer, R Dyer, **MH Weng**, S Wu, T Grey, R Gleeson, TG Ferrari, "Urban narratives for city infrastructures", *WEC2019: World Engineers Convention 2019*, 1127
- 2019 K Mackness, M Dyer, R Dyer, A Hinze, T Garcia Ferrari, S Wu, R Wilkins, **MH Weng**, Urban narrative: Value based urban design", *2019 New Zealand Planning Institute Conference*, 1-24
- 2017 **MH Weng**, B Pfahringer, M Utting, "Static techniques for reducing memory usage in the C implementation of Whyley programs", *Proceedings of the Australasian Computer Science Week Multiconference*, 1-8
- 2016 **MH Weng**, M Utting, B Pfahringer, "Bound analysis for Whyley programs", *Electronic Notes in Theoretical Computer Science* 320, 53-67
- 2013 M Utting, **MH Weng**, JG Cleary, "The JStar language philosophy", *Parallel Computing* 40 (2), 35-50

### THESIS AND TECHNICAL REPORT

- 2019 **MH Weng**, "Efficient compilation of a verification-friendly programming language", *Thesis, Doctor of Philosophy (PhD)*, The University of Waikato
- 2019 Dyer, M., Dyer, R., Ferrari, T., **MH Weng**, Wilson, J., Wilkins, R., & Wu, S. "Data Collection, Data Analytics, Data Visualisations and Data Storytelling", *Report for Building Better Homes, Towns and Cities: Urban narratives (Contestable Research)*, 65pgs.
- 2013 **MH Weng**, "Automatic Parallelization of Data-Driven JStar Programs", *Thesis, Master of Science (MSc)*, The University of Waikato

### TALK

- 2016 Memory Optimization for C implementations of Whyley, SAPLING 2016, Australian National University, Canberra
- 2014 Bound Analysis for Whyley Programs, SAPLING 2014, Oracle Labs, Australia (Brisbane)

### PATTERNS

- 2012 Probability Time Division Multiplexing Polling Method and Wireless Identification Reader Controller Thereof, US US8233468 B2

2011 Ubiquitous Proxy Mobile Service Method and System And Computer Recordable Storage Medium For the Method, US US8037130 B2

#### CERTIFICATES

2023 Machine Learning Specialization, Coursera  
2022 Fundamentals of Deep Learning, NVIDIA  
2010 Cloudera Certified Hadoop Developer (CCHD), Cloudera

### Teaching

2022-23 Teaching Fellow at Waikato University Joint Institute at Zhejiang University City College (NZUWI)  
Taught the following Computer Science courses: COMPX201 Data structures and algorithms), COMPX222 (Web development), and COMPX318 (Internet of Things and Mobile computing).  
2020 Senior Tutor (6 months) at NZUWI Taught COMPX322 (Advanced Web Development) and COMPX202 (Mobile Computing)  
2014-2019 Doctoral assistant at Waikato university helps out a variety of undergraduate and postgraduate courses.

### Research Highlights

Some of my research projects are listed below.

**A large language model-based (GPT-3) topic analysis tool.**  
<https://github.com/samminweng/AIOnUrbanStudies>

**AI augmented approach** to identify shared ideas from large format public consultation (Urban Narratives). [https://github.com/samminweng/urban\\_narratives](https://github.com/samminweng/urban_narratives)

**A compiler** translates into efficient C code, and formally proves the memory safety of C code.  
<https://github.com/samminweng/WhileyOpenCL>

**Kaggle Notebook Expert** (top 0.2% of 320,000 participants from Sep 2023 - Present)

**Google - predict AI Model runtime** My solution leverages a GNN (Graph Neural Network) model, integrated with a modified BERT architecture, to train on the graph-based data structures generated by the compiler and predict the run-time of each AI model. [\[Bronze-medal solution\]](#)

**LLM Science Exam that uses LLMs to answer difficult science questions** My solution blends Retrieval Augmented Generation (RAG) with the generative power of 'Platypus2-70B', a large language model, to deliver accurate answers to a wide range of questions. RAG relevant Wikipedia texts as background knowledge, crafting the prompts that leverage Platypus2-70B's generative ability to predict the answer to the question. [\[Solution link\]](#)

**Pavilion of Dreams at 2010 Flora Expo** (the International Association of Horticultural Producers). I developed the software architecture that seamlessly integrates our custom RFID hardware, empowering visitors to unlock diverse interactions with exhibits in the building using their RFID bracelets. The exhibition drew over 100,000 visitors in just six months and it is also the first internationally recognized exposition that utilizes RFID technology in Taiwan and the seventh in Asia. Notably, two patents were filed during the project.