

Personal Statement

- I am a problem solver with research and project experience in media, telco, health and financial services. My computational mindset, prototyping skills and product experience enable me to identify data-driven business opportunities and deliver impact-oriented solutions in both start-up and enterprise environments.

Education

Melbourne University

PhD in Computer Science (Natural Language Processing)

2010 – 2014

Melbourne, Australia

Dalian University of Technology

Bachelor and Master in Software Engineering

2003 – 2009

Dalian, China

Experience

Accenture

Data Engineering Manager

Sep 2018 – Present

Sydney, Australia

- I provided for tech architecture design, cloud integrations and solution price estimations.
- I led a diverse team for 5G management system design and implementation for a major telco company.

Kaplan

NLP Specialist

May 2017 – Aug 2018

Sydney, Australia

- I optimised a text-based compliance system with more than 100x speed up. I also introduced a cloud-based serverless design which significantly boosted the system throughput and scalability.
- I applied active learning and other pragmatic methods to improve machine learning data quality.
- I co-developed an in-house data anonymisation system for personal and identifiable information.

Hugo AI

Data Scientist

Mar 2016 – Mar 2017

Sydney, Australia

- I led the design and development for an automatic user profile summarisation system. The work features data crawling, cleaning, reconciliation and natural language generation from various social media sources. The deliverables facilitate online recruiters for more effective and efficient talent acquisitions.

IBM Research

Research Scientist

Apr 2014 – Mar 2016

Melbourne, Australia

- I designed and developed a number of social media analytics, including a text-based [geotagger](#), [Social Media Analytics in Real Time](#) and a biomedical [named entity recogniser](#). These analytics aim to enhance IBM Watson in dealing with domain specific data. The deliverables have been utilised by clients from fashion, sports and insurance.

Melbourne University/NICTA

PhD and Research Assistant

Jul 2010 – Mar 2014

Melbourne, Australia

- I applied NLP/ML methods to improve the utility of social media for various downstream applications.
- I was an early contributor for "[Early detection and mitigation of disastrous events with broadband-enabled social interaction technologies](#)".

Microsoft Research Asia

Fulltime Research Intern

Jan 2007 – Aug 2007, Sep 2009 – Jun 2010

Beijing, China

- › I co-developed an English writing assistant which has been productised as [Bing Search Dictionary](#) in China. My work includes bilingual text mining, short sentence categorisation, incremental bilingual text crawling. This product had won [2010 Wall Street Journal Asian Innovation Award](#).

Professional Skills

Frontend HTML, CSS, Javascript (jQuery, Vue and React)

Programming Python, R, Bash, Java, C#, C/C++

Modelling Jupyter, pandas, scikit-learn, spaCy, word2vec, fasttext, Keras (TensorFlow), AllenNLP (PyTorch)

Backend MySQL, Postgress, Redis, Kafka, Memcached, MongoDB

Cloud Certified AWS architect (W1LNNXZ212QE11WZ), familiar with VPC, EC2, S3, Lambda, CloudWatch, etc.

Container/DevOps Docker, docker-compose, git/bitbucket, JIRA/Trello, Jenkins/Bamboo, Ansible

General Skills

Language English (professional), Mandarin (native), Japanese (learner)

Presentation I am actively involved in professional presentations including publishing technical articles, giving invited talks in universities, demonstrating showcases for industrial partners and clients.

Management I am capable of mapping ideas into actionable product roadmaps; I like to lead a small agile team with competent professional and communication skills to execute the plan and get it done on time.

More Information

My Research I worked on [Improving the Utility of Social Media with Natural Language Processing](#). In particular, I studied Twitter text normalisation and geolocation prediction in depth. The details are too long, so I present my tweet length summary: *U wanto undrstnd TWEETS lik dis? Make them readable like this; and if there are too many to read, select those from a region.*

My Academic Achievement

- › I published a few top-tier conference and journal papers (e.g. ACL, EMNLP, COLING and JAIR). My H10-index is 15 and the most cited paper is over 500 (by 2020). See [Google Scholar](#) for a full list of publications.
- › I had co-organised two workshops on social media analytics ([W-NUT 2015](#) and [W-NUT 2016](#)).
- › I had one issued patent (US9569551B2) in United States.
- › I had one software copyright (2015SR225534) in China.

My Project I like coding some side projects for fun. They are like Unix utilities dedicated to solve a particular information need. The most recent one is [panda-lens.com](#). It is a pure data-driven real-time information consumption platform.

My Hobby Cook gourmet food daily; Play badminton weekly; Travel 3 times a year; Had one detective novel (written in Mandarin)