

# Michiharu Yamashita

+1 (814) 321 8176 | michiharu@psu.edu | mickeymst.github.io/

## EDUCATION

---

### Pennsylvania State University

University Park, PA

*Ph.D. in Information Science and Technology (Advisor: Prof. Dongwon Lee)*

*Aug 2020 – Present*

- Thesis: Understanding and Modeling User-oriented Career Recommender Systems for the Future of Work

### Tokyo Institute of Technology

Tokyo, Japan

*M.Eng. in Computational Intelligence and Systems Science (Advisor: Prof. Kazuo Yano)*

*Apr 2015 – Mar 2017*

- **Salutatorian (2nd place in the department)**
- Thesis: Predicting Optimal Work Environment from Wearable Sensors

### University of Tsukuba

Tsukuba, Japan

*B.S. in Management Science and Engineering (Advisor: Prof. Ushio Sumita)*

*Apr 2011 – Mar 2015*

- Thesis: Network Analysis and Visualization for Mobile Applications' Competitiveness

## RESEARCH EXPERIENCE

---

### PIKE Research Group at Penn State

Aug 2020 – Present

*Research Assistant (Advisor: Prof. Dongwon Lee)*

*University Park, PA*

- Research Topics: Graph Neural Networks, NLP, Recommendation System, Computational Jobs Marketplace
- Developing user-oriented job recommendation systems and robust and secure models for jobs marketplace.

### Megagon Labs

Apr 2017 – Mar 2018

*Research Engineer*

*Tokyo, Japan*

- Research Topics: NLP, Entity Matching, Entity Extraction
- Developed NLP-specific models and embedded the modules into the company system.

### Hitachi Central Research Laboratory / Tokyo Institute of Technology

Apr 2015 – Mar 2017

*Research Assistant (Advisor: Prof. Kazuo Yano)*

*Tokyo, Japan*

- Research Topic: Network Science, Wearable Sensors, People Analytics
- Developed machine learning models for wearable sensors' data and visualized the sensor data into graph.

### Sumita Research Group at University of Tsukuba

Jan 2014 – Mar 2015

*Research Assistant (Advisor: Prof. Ushio Sumita)*

*Tokyo, Japan*

- Research Topic: Network Analysis, Visualization, Operations Research, Ranking Algorithm
- Developed a competitive score algorithm for app installing with graph embeddings.

## EMPLOYMENT EXPERIENCE

---

### Indeed

May 2022 – Aug 2022

*Summer Intern*

*Austin, TX (Remote)*

- Developing robust job recommendation models to deliver high quality matches to job seekers and employers.
- Doing research on NLP and deep learning techniques in computational jobs marketplace.

### Freelance Machine Learning Consultant

Jun 2019 – Jul 2020

*Machine Learning Engineer*

*Tokyo, Japan*

- Conducted machine learning projects with a big tech company, a tech startup, and an education startup.
- Developed job mobility prediction models, machine learning models, query optimization tools, etc.

### Recruit Holdings

Apr 2017 – Jan 2019

*Machine Learning Engineer*

*Tokyo, Japan*

- Developed recommendation systems, multi-view click prediction models, and GIS-based applications.
- Developed a pedestrian congestion visualization algorithm using GIS data and OpenStreetMap.
- Developed the geo-topic model to obtain the user interest from POI.

### Ohma

Mar 2016 – Mar 2017

*Software Engineer (Advisor: Prof. Yutaka Matsuo)*

*Tokyo, Japan*

- Developed multiple ML-related systems: network visualization, entity extraction, face recognition, search engine.

### BigGorilla (Open-source Components for Data Integration)

Apr 2017 – Mar 2018

Megagon Labs

Tokyo, Japan

- Developed an entity matching and entity extraction module and NLP related frameworks.
- Applied NLP modules into the companies and promoted open-source components BigGorilla.

### Spysee2 (People Search Engine)

Mar 2016 – Mar 2017

Ohma, Inc.

Tokyo, Japan

- Developed a network visualization search engine SPYSEE2 which had 1M+ visits per month.
- Developed entity extraction, entity linking, and face recognition systems from unstructured and noisy web data.
- Crawled millions of web pages efficiently using programs on AWS.

### Mobile App Competitiveness Visualization

Apr 2014 – Mar 2015

Fuller, Inc.

Tsukuba, Japan

- Developed a ranking algorithm for mobile app competitive analysis.
- Developed a network visualization using app downloading flow.

## PUBLICATIONS

---

1. Jingyi Xie\*, **Michiharu Yamashita\***, Zekun Cai\*, and Aiping Xiong. A user study on the feasibility of topic-aware misinformation warning on social media. In *Proceedings of the Human Factors and Ergonomics Society (HFES)*, 2022
2. **Michiharu Yamashita**, Jia Tracy Shen, Hamoon Ekhtiari, Thanh Tran, and Dongwon Lee. James: Job title mapping with multi-aspect embeddings and reasoning. *arXiv preprint arXiv:2202.10739*, 2022
3. **Michiharu Yamashita**, Yunqi Li, Thanh Tran, Yongfeng Zhang, and Dongwon Lee. Looking Further into the Future: Career Pathway Prediction. In *ACM WSDM 2022 Workshop on Computational Jobs Marketplace*, 2022
4. Jia Tracy Shen, **Michiharu Yamashita**, Ethan Prihar, Neil Heffernan, Xintao Wu, Ben Graff, and Dongwon Lee. MathBERT: A Pre-trained Language Model for General NLP Tasks in Mathematics Education. In *NeurIPS 2021 Workshop on Math AI for Education*, 2021 (**Best Paper Award**)
5. Jia Tracy Shen, **Michiharu Yamashita**, Ethan Prihar, Neil Heffernan, Xintao Wu, Sean McGrew, and Dongwon Lee. Classifying Math Knowledge Components via Task-Adaptive Pre-Trained BERT. In *International Conference on Artificial Intelligence in Education (AIED)*, pages 408–419. Springer, 2021
6. **Michiharu Yamashita**, Shota Katsumata, and Yusuke Fukasawa. Discovery of User Preferences from Big Geospatial Data Using Topic Models. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 4387–4392. IEEE, 2018
7. **Michiharu Yamashita**, Hideki Awashima, and Hidekazu Oiwa. A Comparison of Entity Matching Methods between English and Japanese Katakana. In *Proceedings of the Fifteenth Workshop on Computational Research in Phonetics, Phonology, and Morphology at EMNLP*, pages 84–92, 2018
8. Kent Kawai, **Michiharu Yamashita**, and Yutaka Matsuo. Face Recognition System Based on Convolutional Neural Network Robust to Occlusion and Low Quality Images. In *The 31st Annual Conference of the Japanese Society for Artificial Intelligence*, pages 3M21–3M21. JSAI, 2017
9. **Michiharu Yamashita**, Nobuo Sato, and Kazuo Yano. Enhancing Collective Happiness by Controlling Room Temperature Using Big Data from Wearable Sensors. In *The 2016 IEICE General Conference*, volume 115, pages 31–34. IEICE, 2016

## HONORS AND AWARDS

---

2020 Funds for Excellence in Graduate Recruitment Scholarships from Penn State  
2018 Recruit Holdings The Best Freshman Award 2018  
2018 Recruit Holdings R&D MVP Award 2017  
2017 Full Repayment Exemption of Graduate Student Loan \$20,000 for Excellent Achievement  
2017 Salutatorian at Tokyo Institute of Technology  
2017 The Second Best Master Thesis Award from Tokyo Institute of Technology  
2015-2017 Full Tuition Exemption from Tokyo Institute of Technology  
2011-2015 Full Tuition Exemption from University of Tsukuba

## TECHNICAL SKILLS

---

**Languages:** Python, Java, C/C++, SQL, Swift, R, JavaScript, HTML/CSS

**Frameworks:** Tensorflow, Keras, Flask, Elasticsearch

**Developer Tools:** Docker, Google Cloud Platform, AWS