

# Zhaohao Zeng

Tel: (+81) 070-4015-6227 Address: 05-24, Bldg 63, Waseda Univ, Tokyo  
E-mail: [zhaohao.zeng93@gmail.com](mailto:zhaohao.zeng93@gmail.com) GitHub: [www.github.com/matthew-z](http://www.github.com/matthew-z)

## Experience

---

Oct 2016 - Present	Silicon Studio Corporation <b>Big Data Engineer</b> Part time(onsite, 28 hours per week). I am participating in Game Data Science research, and integrating Data Science and Machine Learning R/Python scripts into the Spark pipeline, and implementing related API and visualisation tools.	Tokyo, Japan
--------------------	---	--------------

## Education

---

Sep 2015 - Present	Waseda University <b>Master of Engineering in Computer Science</b> Supervisor: Prof. Tetsuya Sakai	Tokyo, Japan
Sep 2013 - Jun 2015	The University of Liverpool <b>Bachelor of Engineering in Electronics</b>	Liverpool, UK
Sep 2011 - Jun 2013	Xi'an Jiaotong - Liverpool University <b>Bachelor of Engineering in Electronics Science and Technology</b> Transferred to Liverpool after completing Year-0 & Year-1.	Suzhou, China

## Publication

---

Sakai, T., **Zeng, Z.** and Luo, C.:Evaluating Helpdesk Dialogues: Initial Considerations from An Information Access Perspective. IPSJ - Special Interest Group of Natural Language (IPSJ-SIGNL 228) , Osaka, Japan, Sep 2016.

## Honors

---

2015	<b>Monbukagakusho Honors Scholarship</b> The Ministry of Education(MEXT), Japanese Government.
2016	<b>Partial Tuition-waiver Scholarship</b> Waseda University
2013	<b>Partial Tuition-waiver Scholarship</b> The University of Liverpool

## Selected Projects

---

- 2016 **Evaluation Measure of Helpdesk System** Data Crawling & Dialogue Evaluation  
Built a helpdesk dialogue test collection by crawling dialogues on Weibo with Python. Built an annotation platform to annotate the dialogues with subjective labels and nugget labels for evaluation via crowdsourcing. Designed a suite of language-independent evaluation measures for quantifying the success of each dialogue.
- 2016 **Google Person Finder** Web Development & Open Source Contribution  
Contributed to a Google-owned open-source project called Person Finder, which is based on Python, Django and AppEngine. Improved Person Finder's searching, indexing, and query pre-processing modules to help it handle Chinese and Japanese better.
- 2015 **Cross-Device Targeting in Big Data** Machine Learning & Big Data  
Proposed a framework using unsupervised learning algorithms to target user across multiple digital devices, with Apache Spark and Scala. Helped online advertising publisher better track user and recommend products. Utilized a big dataset which was provided by our partner Wider Planet, Inc.
- 2016 **Kaggle: Airbnb New User Bookings** Machine Learning  
Attended a data science contest on Kaggle to predict Airbnb new user's first booking destination with R. Handled the incomplete data provided by Airbnb into dummy encoding in this research, and applied a solution based on Gradient Boosting. Received a nearly 0.8 nDCG@5 score with the test data.
- 2015 **Network Analysis of Contributors on Github** Machine Learning & Visualisation  
Utilized contributors' activities log on Github to analyze the relationships between different programming languages and popular repositories. Visualized the correlation between different repositories, e.g. jQuery has a great influence on other front-end tools, on Github with a graph. Conducted the research using Google Cloud Platform and Apache Spark to handle the large dataset.
- 2014 **Algorithm Automatic Timetable Scheduling Algorithm** NP problem  
Developed a genetic algorithm based solution to solve an NP problem: university timetable scheduling problem with C++, Python and Objective-C. Applied parameters tuning and closed pattern mining algorithms to accelerate the processing.

## MOOC Certificates

---

University of Washington	<a href="#">Practical Predictive Analytics</a>
University of Washington	<a href="#">Data Manipulation at Scale</a>
Stanford University	<a href="#">Machine Learning</a>
Stanford University	<a href="#">Algorithms: Design and Analysis</a>
John Hopkins University	<a href="#">The Data Scientists Toolbox</a>
The University of Tokyo	<a href="#">From the Big Bang to Dark Energy</a>
University of Pennsylvania	<a href="#">Probability</a>
University of Illinois at Urbana-Champaign	<a href="#">Pattern Discovery in Data Mining</a>
University of Illinois at Urbana-Champaign	<a href="#">Text Retrieval and Search Engines</a>
University of Illinois at Urbana-Champaign	<a href="#">Text Mining and Analytics</a>