

# Zhaohao Zeng

📍 0524, Bldg 63, Waseda Univ, Tokyo, Japan  
☎ +81 070-4015-6227  
✉ zhaohao.zeng93@gmail.com  
🔗 Github.com/matthew-z

## OBJECTIVE

Software Engineer at Airbnb, New Graduate (China)

## EMPLOYMENT

2016 – PRESENT

### Big Data Engineer

Silicon Studio Corporation, Tokyo, Japan

Part time. Participating in Game Data Science research, Integrating Data Science R/Python scripts into the Spark pipeline, and implementing related API.

## EDUCATION

2015 – 2017(EXPECT)

**Master, Computer Science**, GPA 3.7/4.0,  
Waseda University, Japan  
Supervisor: Prof.Tetsuya Sakai

2013 – 2015

**Bachelor, Electronics**, with 2:1 Honours,  
The University of Liverpool, U.K.

2011 – 2013

**Electronics Science and Technology**,  
Xi'an Jiaotong - Liverpool University, P.R.China  
*Transferred to Liverpool after Year-1*

## SKILLS

Python, R, SQL, Scala, C++, LaTeX,  
Apache Spark, Django, Linux, Emacs, Vim

Chinese, English(GRE 322/340), Japanese

## SELECTED PROJECTS

MAY 2016 – PRESENT

Data Crawling & Dialogue Evaluation

### *Evaluation of HelpDesk System on Weibo*

Built a helpdesk dialogue test collection by crawling dialogues on Weibo with **Python**, and designed a suite of language-independent evaluation measures for quantifying the success of each dialogue.

**Publication:** *Evaluating Helpdesk Dialogues: Initial Considerations from An Information Access Perspective*. IPSJ-SIGNL, to appear, Osaka, Japan, Oct 2016.

AUG 2016 – SEP 2016

Web Development & Open Source Contribution

### *Google Person Finder*

Person Finder is a Google-owned open-source project based on **Python**, **Django** and **AppEngine**. I improved its searching, indexing and query preprocessing modules to let it handle Chinese and Japanese better. My commits have been merged into this product.

OCT 2015 – APR 2016

Machine Learning & Big Data

### *Cross-Device Targeting in Big Data*

Proposed a framework using unsupervised learning algorithms to target different devices for serving targeted advertising to consumers across multiple digital devices, with **Apache Spark** and **Scala**. A big dataset was provided by our partner Wider Planet, Inc.

JAN 2016 – FEB 2016

Machine Learning

### *Kaggle: Airbnb New User Bookings*

Applied a solution based on Gradient Boosting to predict Airbnb new users first booking destination with **R**, and got a nearly 0.8 nDCG@5 score.

JAN 2016 – FEB 2016

Data Mining & Visualisation

### *Network Analysis of Contributors on Github*

Analysed and visualised the relationships between different programming languages and popular repositories on Github with **Google BigQuery** and **Apache Spark**.

SEP 2014 – MAY 2015

NP Problem Algorithm

### *Automatic Timetable Scheduling Algorithm*

Proposed a framework based on genetic algorithm and closed pattern mining to solve an NP problem: university timetable scheduling problem, with C++.

## HONOURS

2015 Monbukagakusho Honours Scholarship  
The Ministry of Education(MEXT),  
Japanese Government.

## MOOC CERTIFICATES

STANFORD UNIVERSITY

Machine Learning

Algorithms: Design and Analysis

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Text Retrieval and Search Engines

Text Mining and Analytics

Pattern Discovery in Data Mining

UNIVERSITY OF WASHINGTON

Data Manipulation at Scale: Systems and Algorithms

Practical Predictive Analytics: Models and Methods