



Peter Lee

My research interest includes Data Infra, Data Science, Backend and IoT. I focus on big data and data infra since 2015. Contributor of TensorFlow, Rocket.Chat and Nvidia Jetson community and other open-source projects.



Taipei, Taiwan

peterlee0127@gmail.com [Website](#)

Experience

Vpon Taiwan - Data Engineer (Architect),

Oct 2018 - Present

- In data team: I design new system infra with a new stack. I also help our member to solve their daily problems.
- 1. Data migration: I help the data team to transfer existing projects from AWS to GCP. Example: Migrate the ETL pipeline from AWS to GCP. Moving our Gitlab from AWS to GCP (Gitlab CI/CD). Transfer data from Hive and S3 to GCP
- 2. Infrastructure: We use Google Cloud Platform and Amazon Web Service. I designed the architecture of our system. Like the Gitlab CI/CD, ETL flow, Data processing pipeline. For example GKE(Google Kubernetes Engine) with Istio, BigQuery, Cloud Function, and GCE.
- 3. CI/CD: Design CI/CD pipeline for our project. Our pipeline Build Docker Image, do Unit Test, then deploy to Container Registry. And these image will deploy on our Kubernetes cluster
- 4. Data tagging service: A high customize tagging service, it processes our data with many kinds of the tag. It provides a view of people interests, language, age, gender and so on.
- 5. ETL pipeline: Our daily cronjob to process many kinds of data. Like the data exchange with our partners. I deployed jobs on Apache Airflow.
- 6. Elastic Stack: Processing millions record data every day. It can validate data and watch our report in real-time.

Public Digital Innovation Space (PDIS) 唐鳳政委辦公室- Software Engineer,

Sep 2017 - Aug 2018

PDIS is a government organization in Taiwan, it also has another name: Executive Yuan - Ministry of Digital - Audrey Tang's Office. We use new technology to improve public participation and open government.

- 1: As a software developer: I use Line BOT API to connect our internal systems to make an easy-to-use app to help our colleagues saving their time. I also designed API interfaces to allow other internal systems to hook up with our bot system. Up to now, it already bridged across our meeting reserve system and electronic bulletin board system.
- 2: About productive and quality: I designed a software that can display real-time subtitle during streaming. To meet the schedule, I took only one week to make it from idea to prototype. This system now works perfectly in all of our video conferences.

3: As a quick learner and contributor: I involved in many open source projects. Like Sandstorm(A open source private cloud with the container), Rocket.Chat (A open source chat app which like Slack). We have fixed many issues and add lots of features, and giving back to the community.

4: Eager to learn everything: when I have free time, I like to think about how to make system more efficient. I still find more possibility to improve the current system. Like the subtitle display system, I would like to let some part of the system can automatically work. Let the computer can label the people with their name in real-time and display the information on the subtitle. On the road of development, I fixed various issues and bugs of Tensorflow. NVIDIA give me an award as the outstanding community developer in this April.

In 2019 and 2020, I still help PDIS to maintain our cloud services. Like our internal systems, website and translator.

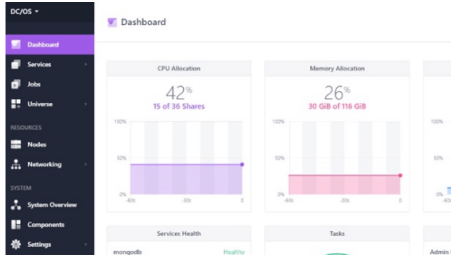
<https://pdis.nat.gov.tw/en/>

Nexdoor (91 App) - iOS Engineer, July 2013 - Sep 2014

Nexdoor is a famous App company in 2013, we build iOS/Android/Web for the enterprise. I was an iOS Engineer Intern when I was an undergraduate student. I developed some interesting project, like Geofence App, LBS Service framework and customize UI. In 2016, most of employees join new company 91 app.

91APP

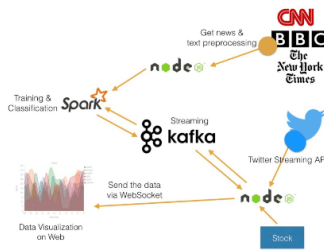
Publication



Implementation of Lambda Architecture: A Restaurant Recommender System over Apache Mesos

Advanced Information Networking and Applications (AINA) 2017 IEEE 31st International Conference

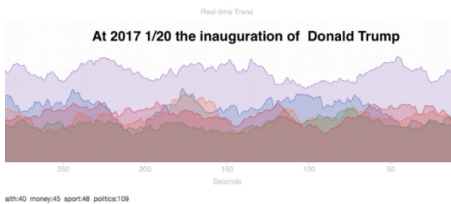
A Lambda Architecture at DC/OS, using Spark, Spark Streaming, Kafka, Hadoop HDFS. [Full Text](#)



Stock market analysis from Twitter and news based on streaming big data infrastructure

Awareness Science and Technology (iCAST) 2017 IEEE 8th International Conference

Real-time trend analysis by twitter, using Spark, Spark Streaming, Kafka, Classification Data and Visualize on Web. [Full Text](#)



Real-time Trend Analysis of Streaming Twitter and News Based on Big Data Infrastructure

電子情報通信学会-講演論文 信学技報, vol. 117, no. 184, SC2017-13, pp. 1-6, 2017年 8月.

SC2017-13 2017-08-18 (SWIM, SC) [Full Text](#)

Selected Awards

Nvidia Jetson developer community contributor

[Nvidia Jetson developer community contributor](#)

Porting TensorFlow for Nvidia Jetson.

TensorFlow Contributor

Porting TensorFlow to Nvidia Jetson. Repository [tensorflow-nvjetson](#). #26985 #20025, #19075, #17394.

Rocket.Chat Contributor

Update and patch Rocket.Chat for Sandstorm. Due to there is no official build since 2017. I started to fix issues and help new official build.

Selected Skills

Data

Kubernetes with Istio
Elasticsearch
Spark
Airflow
Kafka

Backend/Web

Node.js
MongoDB

General

DevOps
GCP/AWS
Gitlab CI/CD
Jenkins
Docker

Education

University Of Aizu, Fukushima, Japan.

2016 March - 2017 March Master Degree of Computer Science. Double Degree Program.

Tamkang University, Taipei, Taiwan.

2015 - 2017 Master Degree of Computer Science. Cloud Computing Lab.

2011 - 2014 Bachelor of Engineering (B.E.) Computer Science. Embedded System Lab.