

MARK CHEN

(858)242-9619 | itisnanful@gmail.com | Github: <https://github.com/mAsCuLiNe>

EDUCATION

University of California San Diego, B.S. in Mathematics - Computer Science

Sep 2011 - Dec 2015

TECHNIQUE SKILLS

Language: Java (4yr), Python (4yr), C++/C, MATLAB

Database: PostgreSQL, MySQL, ActiveRecord, MongoDB

Cloud & System: AWS EC2/S3/RDS, Linux, RabbitMQ, Heroku, Google App Engine

Web development: HTML5/CSS3/JavaScript, Ruby on Rails, AngularJS, NodeJS, JQuery, Django, MeteorJS

Concepts: REST, SOAP, Database design, Design pattern, SQL optimization

PROJECTS

Empachat — Team Leader/Full Stack Developer (MeteorJS, MongoDB, Python) Feb 2016 - Mar 2016

A chat app that connects users with similar emotions and plays songs of similar emotions

Github: <https://github.com/BitTigerInst/Empachat> | Website: empachat.meteor.com

- Design app architecture with comprehensive resource discovery
- Mentor frontend and backend developers and help tackling hardest bugs

Distributed Data Crawler — Software Developer (Python, bash)

June 2015 - July 2015

*A Python-RabbitMQ crawling system with the ability to collect tens of terabytes in days with **minimal cost**.*

Github: <https://github.com/mAsCuLiNe/flexibloader>

- Designed & implemented the core of the master-slave mechanism, and 16 crawlers, using **design pattern**. Achieved minimal cost by memory optimization, machine failure retrieval, and deployment acceleration.
- Achievements: Scaled up to 1000 concurrent crawling jobs per second. Completed deploying 90 **AWS EC2 Linux** machines in 3 secs. Collected 40 million Flickr images (**8 TB**) in 70 hours with **\$100**. Collected and parsed data from all 1 million Chinese Wikipedia pages in 2 hours with **zero cost**.

Moorgoo — Frontend Developer (Ruby on Rails, html/css/javascript)

June 2014 - January 2015

An app for hosting videos that help students on UCSD courses and for making private tutor appointments

Website: <http://lovemoorgoo.herokuapp.com>

Implement UI of 72 videos categorized by 25 courses. Comprehensively tested core functions using Rspec.

WORKING EXPERIENCE

Full Stack Developer — UCSD Cognitive Science

February 2015 - present

- Worked with a group of data scientists on face similarity research. Self developed 4 data crowdsourcing tools with full stack web development (AngularJS, Python, html/css/js, Django, Linux, AWS, JQuery, MATLAB)
- Self developed and comprehensively tested a stimuli distributor. Distributed 900 Human Intelligence Tasks without missing any incomplete ones. Github: https://github.com/mAsCuLiNe/simi_server
- Data analysis and image processing using MATLAB and Python.

Data Engineer Intern — TuSimple Inc.

March 2015 - Sept 2015

- Worked with a group of data scientists on a recommendation system based on computer vision.
- Managed 20+ terabytes data from 30+ sources on Linux file system. Processed > 1TB data daily with Python.
- Developed web element detectors. Achieved 90% precision rate on web image caption detection.
- Implement data labeling tools by using python selenium and by generating in-browser highlights. Achieved image labeling rate of 1800 images per hour per person.
- Automated visualization of webpage-image data via selenium and SQL. 10x faster than visualizing manually.

COURSES

Operating System

Data Mining and Predictive Analytics

HCI Design

Server-side Web Application

Machine Learning

Numerical Analysis I & II