

Yuheng (Lyle) Li

 www.yuhengli.com

 [linkedin.com/in/yuhengl](https://www.linkedin.com/in/yuhengl)

 github.com/yuhengli

 yuhengli@cmu.edu

OBJECTIVE

Full-time Software Engineer

EDUCATION

M.S., Information Technology, **Carnegie Mellon University**, Pittsburgh, PA (GPA 3.95/4.0) 2015-2016
M.Eng., Data Mining, **University of Chinese Academy of Sciences**, China 2012-2015
B.S., Electronic Information Science and Technology, **Jilin University**, China (GPA 3.81/4.0) 2008-2012

TECHNICAL SKILLS

Language: Java, C#, Python, SQL, C/C++, JavaScript, HTML/CSS, JSP, Shell Scripting, Scala, R, MATLAB.

Tools: J2EE, .NET, Eclipse, Visual Studio, awk, MapReduce, AWS(EC2, EMR...), Azure, Spark, Kafka, Samza, JDBC, Oracle11g, HBase, MongoDB, MySQL, Hive, Impala, Cloudera, Tomcat, Undertow, Servlet, SVN/Git, Maven, Weka.

Knowledge: OOP/OOD, Algorithms, Software Architecture, Computer System, Cloud Computing, Data Mining, Machine Learning, NLP, Digital&Analog Electronics, ARM 7, FPGA Programming, Technology Law.

INTERNSHIP

Assistant Software Engineer, Chinese Academy of Science(CAS), Beijing China 2014-2015

- Responsible for developing GeoVis, a GoogleEarth-like Information Analysis System;
- 1 year+ software development and scientific research with 20+ people team, 10,000+ line codes experiences.

PROJECTS

Twitter Analytics Web Service, Pittsburgh, PA 02/2016 - now

- Used Hadoop on Amazon EMR to conduct ETL on 300 millions (1T) of raw tweet messages;
- Tuned and optimized HBase and MySQL performance to serve complex read queries;
- Built a RESTful web application running on high performance non-blocking Undertow server as frontend;
- Achieved 10,000+ rps using load balancer and backend cluster under \$0.85/hour limited budgets.

Social Networking Services with Heterogeneous Back-ends, Pittsburgh, PA 03/2016 - 04/2016

- Developed a social networking services with MySQL, MongoDB and HBase as backend, and undertow server as frontend using Javascript and PHP running on AWS EC2, EMR, ETL using Java, MapReduce;
- Achieved functions including realtime Timeline, graph based collaborative filtering recommendation, input auto-completion and text prediction based on n-gram language model;
- Optimized databased schema for metadata, graph model and JSON for realtime information retrieval;

Carnegie Financial Web Services for Mutual Fund Transaction, Pittsburgh, PA 01/2016 - 02/2016

- Developed a three-tier RESTful Web application with Java J2EE, JSP, Bootstrap, Tomcat and MySQL;
- Utilized Atomic Transactions to ensure the atomicity and consistency of fund transactions in DAO design;
- Implemented MVC design pattern for authentication control, role-based access control and fund operations;
- Analyzed and measured the performance and availability of the web service with Apache JMeter.

Twitter based Entity Linking System, Beijing, China 02/2015 - 06/2015

- Developed an entity linking system with C#,WPF; dataset from Twitter APIs & Wikimedia; ETL by Mapreduce;
- Achieved Twitter mention disambiguation by automatically linking to corresponding Wikipedia articles;
- Proposed an improved Entity Linking algorithm, which models multiple features(user preferences, popularity etc.) of wiki entities and levitated the In-KB accuracy by 5% (compared with KAURI) to 90.17%.

GeoVis Information Analysis System, Beijing China 08/2014 - 05/2015

- Developed a software plugin with C#, WPF, Oracle 11g and SVN based on a GoogleEarth Like 3D globe;
- Customized 50+ data-driven usercontrols for data visualization, like track simulation, tri-state treeview etc.;
- Enhanced the data visualization effect by implementing an adaptive K-means to cluster spatial and temporal data from different altitude scales, leading to real-time data rearrangement and visualization.