

Education

- 2013.9—**M.S. candidate in Computer Science**, *Courant Institute of Mathematics, New York University, NY*,
Present Expected graduation on Dec. 2015, GPA:3.56.
Courses: Advanced Database Systems, Data Mining, Big Data, Production Quality Software, Open Source Tools
- 2014.9—12 **Online program: ETL for Data Warehouse**, *University of California, Irvine Extension, CA*, GPA:4.0.
- 2008—2012 **B.A. in Management Information Systems**, *Beijing Institute of Petrochemical Technology, China*.

Work Experience

- 2014.1—**Application Developer Intern**, Information Tech Services, New York University, NY.
Present Developed a batch process program that extracts projects management and HR data from multiple data sources, transforms data, and loads data by Java, SQL and JDBC into Oracle 11g database on daily process
Designed and developed a web application to manage NYUs employees, departments and work groups map information with Java EE, Servlet, JSP, Javascript, jQuery, AJAX, and XML
- 2012.7—**Software Developer Intern**, Beijing Huaxin Haitong Tech Co. Ltd, China.
- 2012.10 Accomplished business and information systems requirements analysis and design via communicating with clients
Completed database design, modeling and SQL queries in SQL Server, and systems development by ASP.NET

Project Experience

- 2014.12 **Yelp Restaurants Data Analysis**, *Python, Yelp API, Google Map API*.
Extracted and merged 16k+ restaurants data in NYC from Yelp and Google Map APIs by ETL method
Retrieved 150 millions reviews of restaurants and 15k+ users attribute from Yelp by writing a web spider program
Analyzed the distribution and correlation of restaurants, plus used NLTK to find keywords in users' reviews
- 2014.12 **Quroa-like Question-Answer-Vote Platform**, *Python, Google AppEngine, DataStore, Jinja2, Github*.
Designed and developed a web application independently that allows users to create, modify, view and vote questions and answers, mark and search questions by tags, and upload images
URL: kl1885-ost-14fall.appspot.com, Github: github.com/liukun1016/OST_Project_14Fall
- 2014.11 **Distributed Database with Multi-version Concurrency and Recovery**, *Java, Eclipse*.
Implemented databases concurrency by a multi-version consistency algorithm, deadlock avoidance by Wait-die protocol and Two Phase Locking, and system recovery from failures by available replication algorithm
Realized this database with 10 sites and 20 data items unevenly distributed, and passed 95% of the 20 test cases
- 2014.11 **ETL Designing for Northwind Datawarehouse**, *SQL Server, Visual Studio, SSIS/SSDT*.
Created data profile and source/target map, and intermediate Extract, Clean, Error, Update and Insert tables
Designed multiple ETL work flows, e.g. customer, product, order and order line data for Northwind DW
- 2012.5 **Cloth Material Info Management Systems**, *C#, SQL Server, Visual Studio*.
Analyzed use case and work flow, e.g. vendors, inventory, warehouse, and etc. with abstract factory design model
Finished database design and modeling in MS SQL Server and system development with C# independently
Honored as one of the two Undergraduate Graduate Excellence Project & Thesis of the entire department

Tehcnical Skills

- Programming **Proficient: SQL, Java Core, Python, Unix Shell script; Basic: C++, Java EE, Javascript, jQuery, AJAX, HTML, CSS, XML.**
- Tools & IDE **MySQL, Oracle 11g, SQL Server, Eclipse, JDeveloper, Visual Studio, Visio, Sublime, Github.**

Publication

- 2014.3 **Kun Liu**, Teresa Piliouras, Yiwei Jia, Pui Lam Yu, Qian Zeng, Jeanne Lauer, Maigh Attre and Julie Ortiz, *Others may see someone eating a cookie - I see the start of a protein kinase cascade reaction*, IEEE ISEC Conference 2014, DOL:10.1109/ISECon.2014.6891042.