# Xuye Qin
*Curriculum Vitae*

## PERSONAL DETAILS

| | |
|---|---|
| *Address* | No.969, Wenyi Road, Hangzhou 310000, China |
| *Phone* | (+86) 18616109836 |
| *Mail* | qin@qinxuye.me |

## OBJECTIVE

| | |
|---|---|
| *Position applied* | Software Engineer |
| *Research interests* | Distributed computing, data mining |

## EDUCATION BACKGROUND

**MS in Adaptive Distributed Computing Lab**          2011-2014
*Shanghai Jiao Tong University*

**BA of Software Engineering**          2005-2009
*Southeast University*

## WORK EXPERIENCE

**Research and Development Engineer**          2014-2015
*Alimama.inc*

**Senior Research and Development Engineer**          2015-present
*Aliyun.inc*

## SKILLS

| | |
|---|---|
| *Programming Languages* | Python, Java, C/C++, JavaScript, Go, C# |
| *Develop environments* | Linux, Emacs, Eclipse |
| *Open source* | Hadoop, Spark, MongoDB, Django, Tornado, OpenStack, Neo4j |

## COMMUNITY

*Blog*           http://qinxuye.me
*GitHub*        https://github.com/chineking
*BitBucket*    https://bitbucket.org/chineking

## PROJECT EXPERIENCE

### PyODPS DataFrame Framework          2015-present
*Python, ODPS, Work Project*

Pandas is a Python library which provides high-level API especially DataFrame API for the convenience of data processing. Meanwhile, ODPS is a big data processing service which is good at large-scale data processing. However, due to the limitation of the representation of SQL language, PyODPS DataFrame Framework aims at taking both the advantages of pandas and ODPS. Right now, all DataFrame operations are actually compiled into ODPS SQL.

### Digital Attribution API          2015
*Java, ODPS, Work Project*

A customer may receive various kinds of media touch points before a conversion. The attribution is to calculate how credits for conversions are assigned to those media touch points. Due to the complexity of attribution scenario including the variety of data sources, computing priorities, and dimensions, it's hardly possible to write every implementation of attribution for each scenario. Thus, we developed the attribution API for computation of each scenario according to the XML configurations which made by users. Besides, we provide the support of functions for the further filtering of data source, generating table view and so on. We utilize JavaCC for the lexical and syntax analysis to parse the configured functions.

### ODPS Python SDK          2014
*Python, Personal Project*

This Python implementation of ODPS SDK provides the operations to tables, resources and functions which built on ODPS platform. Tunnel API is also implemented to provide the ability of uploading data to or downloading data from the ODPS tables.

### Cola Distributed Crawler Framework          2013
*Python, Research Project*

Cola is a distributed crawler framework. A cola cluster consists of a single master and multiple workers. Users only need to write the url patterns and the parser function for each of them. Each worker shares a message queue node when running a job, and the target url will be put into the message queue. The message queue will also filter the duplicated target url.

### Sina Weibo User Recommendation System          2012
*Java, Hadoop, Lucene, Research Project*

The TF-IDF theory which used in search engine to measure the importance of a term in a web page was applied in our system. We scanned the content of users' statuses and get their highest TF-IDF term. The system recommended user who had the same highest TF-IDF term. I was in charge of data collection and implementation of this system.

### Cloud Backup Tool          2012
*Python, PyQt, Course Project*

Cloud backup tool supports multiple cloud storage. By setting backup folder, the files in the directory can sync with several cloud storage such as s3, google cloud storage etc, and encryption is also supported. A few coherence issues were involved in this project. I designed the architecture of the tool and implement cloud storage API.

## Personal Blog      2010

*Python, Django, Independent Project*

I developed the blog by utilizing Django framwork. The development included the parts of database design, backend, and frontend etc. jQuery and ajax were also used in this project. The blog supports the third-party authentication.

# INTERNSHIP

## One semester internship at Intel      2012

*Open-source Technology Center*

Virtualization on Intel Atom CPU to support quick switching between multiple mobile operating systems.

## Summer internship at Alibaba      2013

*Wireless search department*

Design and development of offline incremental indexing system.