

ETL Developer and Consultant

Adnan Siddiqi

(BS, Computer Engineering)

Email: kadnan@gmail.com

Home: <http://adnansiddiqi.me>

Blog: <http://blog.adnansiddiqi.me>

LinkedIn: <https://linkedin.com/in/kadnan>

Objective

A curious technologist having many years of experience looking for a remote opportunity working in an environment where I am given a chance to wear multiple hats and could utilize my skills to grow the business.

Skills

Languages: Python, PHP, Go, **Databases:-** MySQL, PostgreSQL, SQLite, MongoDB, Kafka, and ElasticSearch, **Frameworks:-** Laravel, Django, Flask, **DevOps:-** Python Fabric, Docker, K8s, AWS. **ETL:** Airflow, Bonobo, Luigi, Spark. **Others:** BioPython

Work Experience

ETL Developer

Tellus App, Inc - April 2017 - Present

- As an early member, I set up an AWS based infrastructure of data pipelines that fetch from multiple sources on a 24x7 basis.
- Used **Python Fabric** to manage and automate ~20 AWS instances.
- Writing scalable scrapers in **Python** to fetch hundreds of thousands of records from online data sources on a daily basis by using **requests** and **beautifulsoup** library. By introducing parallelism, I was able to download at least **300K records per day per site** with the help of using proxies.
- Using the **Selenium** library with **PhantomJS** and **ChromeDriver** to scrape data from dynamic websites that render data via Javascript.
- Setting up ETL/Data Pipelines for acquisition and transformation of data to make it available for other teams.

Self Employed/Freelance Consultant

December 2012 - Present

- Developed a Stellar Crypto Token based commodity trading exchange system in **Python** and **Laravel**.
- As a part of *eParts and More, Inc*, I wrote various automation tools and crawlers.
- As a part of *Inoutput*, I was responsible for developing project management and estimation tool in **PHP Laravel 5.2** and **MySQL**.
- As a freelance Contractor for *SmartConnect NV*, I wrote a **socket-based custom module in PHP** that provides a communication channel for the contact center system.

Project Lead

TradeKey Private Limited - January 2009 - December 2011

Senior Software Engineer

TradeKey Private Limited - January 2009 - December 2011

Side Projects

As a part of my journey exploring and learning new technologies, I got a chance to work on personal projects as well. Some are mentioned below:

- **GoCache**(<https://github.com/kadnan/GoCache>):- A simple Cache Server in **Go** based on the LRU algorithm which is also used on MemCache.
- **Fehrist**(<https://github.com/kadnan/Fehrist>):- A pure **Go** library for indexing different types of documents. Currently, it supports only CSV and JSON but flexible architecture gives you the liberty to add more documents.
- **ScrapeGen**(<https://github.com/kadnan/ScrapeGen>):- ScrapeGen is a simple tool written in **Python** that generates the code of a web scraper based on rules given in a file.