

# NAVEEN NATARAJAN

Email: naveen.natz@gmail.com

Website: <http://www.nnatz.com/>

Phone: 443-562-8451

---

## PROFESSIONAL EXPERIENCE

### Dow Jones

July 2013 - Present

#### Position: Application Developer

- Develop and update the existing customer self-service site that is used by all the WSJ & Barron's customers.
- Developed new version of the existing customer self-service site using Django.
- Built a fault-tolerant abstraction layer that would handle request from all the customer service systems.
- Built a coupon manager system using Django.
- Developed an encryption and decryption algorithm in Java for storing documents in Nuxeo.
- Built many proof of concepts using Python, NodeJS and Java.

### Children's Hospital of Philadelphia (CHOP)

June - Aug 2012

#### Position: Database Developer Intern

- Twelve weeks of summer internship in the Database Services Team as a Database Administrator.
- Worked on Oracle, Microsoft SQL Server and Epic Cache Databases.

## TECHNICAL SKILLS

- **Languages:** Python, Java, JavaScript
- **Databases:** Relational(MySQL, PostgreSQL, Oracle), NoSQL(Hive,Cassandra)
- **Frameworks:** Hadoop, Django, Amazon Web Services, GreenPlum(MPP)
- **Technologies:** Git, Sqoop, Nginx, Jenkins, Bootstrap, Query Optimization, ETL, Web Robots
- **Operating Systems:** Linux, Windows

## EDUCATION

### Johns Hopkins University, Baltimore, MD

Aug 2011 - May 2013

- Master of Science & Engineering in Computer Science (MSE)

### Anna University, India

Aug 2007 - Apr 2011

- Bachelor of Technology in Information Technology (B.Tech)

## PROJECTS

### Analyzing Large Data Sets Using Hadoop

#### Technology Used: Hadoop, Hive, Python, Java, AWS

Used hadoop and hive to analyze large data sets like Aadhar, NYC Subway data, etc. Built map-reduce programs in both Python and Java. I ran the map-reduce program against the data sets on a four node Amazon Elastic Map Reduce(EMR) cluster.

### Coupon Management System

#### Technology Used: Python, Django, Mysql, Nginx, AWS

Developed a coupon management system web portal for the circulation team using Django, which is used to manage corporate coupon codes for WSJ and Barron's subscriptions. Prior to this the circulation team was using excel sheet to manually add and modify the coupons.

### Abstraction Layer

#### Technology Used: Python, Django, Postgres

Designed and developed an abstraction layer that handles all the calls made to the customer service systems. It is designed to be fault tolerant and also tolerant with respect to changes in contract between systems. This simplified the maintenance and development of the customer service systems.

### Comparison of Iterative Processing of Large Data Sets in Hadoop and HaLoop

#### Technology Used: Java, Hadoop and HaLoop frameworks

This was an Academic project. The Hadoop platform lacks built-in support for iterative programs, which arise naturally in many applications including data mining, web ranking, graph analysis etc. HaLoop extends MapReduce with programming support for iterative applications. This involved writing chained map reduce programs. I evaluated both the frameworks on real queries and real datasets such as Twitter tweets, mutual friends and web ranking on Amazon EMR.

### Molecular Dynamics Database Project (Mddb)

#### Technology Used: Python, Greenplum/PostgreSQL, ETL process, Query Processing & Optimization

This was my Master's degree research project. Built an adaptive control system used for scientific simulations. This involved implementing ETL process in which included extracting data from different file formats, cleaning the data, transforming and loading it into a Greenplum database. Designed a schema for the database and ways to efficiently load and manage the data in the database, which is in the order of forty terabytes. Designed queries to query the data efficiently and analyze them and also worked on the optimization of these queries.