WeiJia Jin

Links to code samples on Github by project

1. [CallistoReinvented: Web application for intelligent job search and recommendation (2014)](#) -
Ruby on Rails, Python, SQL

Summary: Build web application using MVC framework, agile and test-driven development. Designed automated tagging system for job listings using NLP techniques. Relevance ranking for search results and job recommendation system are individually customized for users based on their profile information and browsing history.

Documentations: [Requirements and Specifications](#), [Design Document](#)

Major code contributions:

- Parse HTML data, automate the tagging of job listings using NLP techniques, update server with new job listings: [htmlparser.py](#)   [jobtagger.py](#)   [post_jobs.py](#)
- Backend model and controllers: [jobposting.rb](#)   [company.rb](#)   [jobposting_controller.rb](#)   [companies_controller.rb](#)
- Testing using Rspec: [jobposting_spec.rb](#)   [company_spec.rb](#)   [jobpostings_controller_spec.rb](#)   [companies_controller_spec.rb](#)

2. [New Algorithm for Generating Dorling Cartograms (2014)](#) - **Python (pandas, json), D3.js**

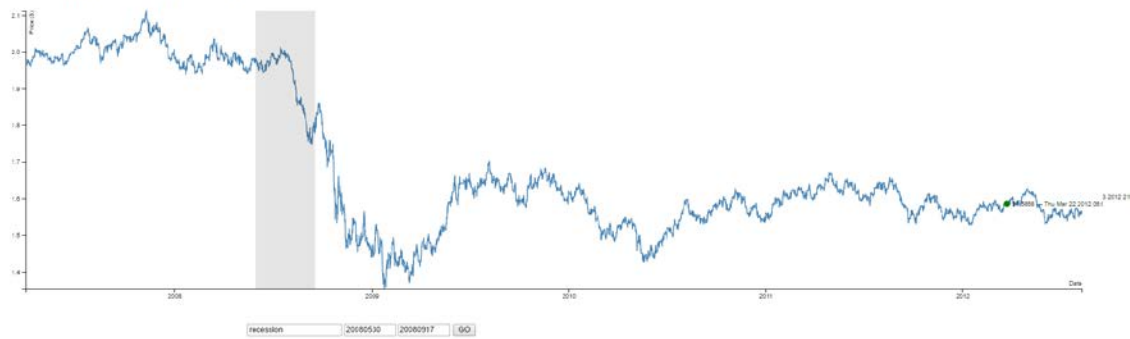Paper: http://goo.gl/NTo0Os

Designed important constraints and optimizations for the automated drawing of Dorling cartograms. Data wrangling of Yelp Challenge Dataset and web data to aggregate and generate new features. Created web visualization with animated cartogram drawing process.

3. [SimpleDB: Basic database management system (2013)](#) - **Java**

Summary: Implemented core modules for data access and storage, query processing operators, buffer pool management, query optimization, transactions, locking, and concurrent queries.

4. [Predictive Model of Foreign Exchange Rate with News Matching (2014)](#) - **R, D3.js, HTML**

Summary: Created two-month forecast of US-SGD exchange rates using time series analysis methods. Built web visualization in D3 with click and drag function for convenient query range selection, sends query request to NYTimes article search API and displays the results.

5.   General coding exercises - **Python**