

Didier Deshommes

Executive Summary

Summary: Back-end Python engineer with experience in creating and implementing scalable solutions in RESTful API design, web-crawling, data processing/analytics pipelines, database schemas, analytics and search.

Experience

Spring 2009 -
February 2016

Systems and architecture Engineer, *Parse.ly*, New York, NY.

- Designed and created an API for our publishers to access their analytics and recommendations data. The API powers multiple live dashboards and widgets. It serves on average 500 requests per second and peaks up to 1000-1500 requests per second. The documentation for the API is here: <https://www.parsely.com/docs/api/overview/endpoint.html>. Technologies used: Tornado, Redis, MongoDB, ElasticSearch, Cassandra.
- Designed and created crawling system that processes millions of URLs per month. Crawlers collect informative metadata from publishers' page for display on the analytics dashboard. Technologies used: Scrapy, MongoDB, Redis.
- Implemented several features in our real-time and batch analytics pipeline that processes hundreds of millions of events per day. Technologies used: MapReduce, Pig, Spark, Storm, ElasticSearch, Cassandra.
- Designed and implemented flexible auto-complete search for our dashboard analytics users. Users could complete either the beginning, middle or end of several words. Users were able to search millions of titles, authors and sections in a unified way in less than 200ms.
- Automated configuration and deployment for machines that were used to do API, crawling or search work using Chef and Vagrant.
- Designed and developed named entity recognition system using Solr, the Wikipedia index, Wikipedia traffic data, and NLTK. With this feature, users were able to see what topics were trending on the web based on the traffic they got from Wikipedia traffic.
- Wrote and launch the back-end prototype to the Parse.ly reader, an RSS feed reader that displays articles from the web based on a user's interests, using Django and Solr.

Winter 2007 -
Spring 2008

Senior Python contractor, *Wordstream*, Boston, MA.

- Designed and implemented a keyword-matching algorithm that accepts documents as its input and produces a list of suggested hyperlinks to be added by a user. The suggested keywords are based on the user's profile data.

2007 - 2008

Software Engineer, *Intelligent Information Systems*, Durham, NC.

- Designed and implemented a website so clients could access and browse property inspection information
- Implemented and improved a procedure to port and test programs written in VB6 to VB.NET. Converted over 100 programs written in VB6 code to VB.NET using a combination of the MS conversion tool and Python.

Education

Fall 2006 -
Winter 2008

MS Applied Mathematics, *North Carolina State University*, *Raleigh, NC*.

Fall 2001 -
Winter 2006

BS Computer Science, *North Carolina State University*, *Raleigh, NC*.

Fall 2001 -
Winter 2006

BS Applied Mathematics, *North Carolina State University*, *Raleigh, NC*.

Extra

- Presenter, PyData: "Wikipedia Indexing and Analysis" - see <https://vimeo.com/53091620>
- Contributor to open source projects - see <https://github.com/dfdeshom>