# William Wagner

75 N. Woodward Ave. Ubox 64338 Tallahassee, FL, US
(+1) 772-341-6961 | wcwwagner@gmail.com
http://williamwagner.me | https://github.com/wcwagner

## EDUCATION

**Florida State University**
Computer Science B.S. August 2014–PRESENT

- GPA 3.86/4.00
- Expected graduation: Spring 2018
- Honors: Dean's List (Fall 2014, 2016), President's List (Spring 2016)

## EXPERIENCE

**Parallel Architecture and Systems Lab,** Research Assistant, Sept. 2016 - PRESENT, FSU

- Built models that monitor and predict global disease trends using Wikipedia page view logs and official CDC disease data.
- Aggregated and cleaned terabytes of Wikipedia page view logs into weekly time series with Apache Spark.
- Technologies used: **Python**, Scikit-learn, Apache Spark, IPython, Git

**Bloomberg L.P.,** Incoming Software Engineer Intern, to begin May 2017, New York City

## PROJECTS

**Vestview,** web application — github: wcwagner/vestview

- Facilitates access to real-time and historical stock market prices, news, and graphs.
- Summarizes news articles by reducing content by up to 80%, using the TextRank algorithm.
- Technologies used: **Python**, **JavaScript**, Flask, SQL, jQuery, HTML/CSS, Git

**Flight-delay,** PySpark program — github: wcwagner/flight-delay

- Built a model that predicts flight delays with an accuracy of 80%, trained on approximately 15 million past flight records.
- Leveraged Apache Spark to preprocess the data and build a random forest classifier over a cluster of nodes.
- Technologies used: **Python**, PySpark, MLLib, Scikit-Learn, IPython, Git

**RateMyFSU,** chrome extension — github: wcwagner/RateMyFSU

- Helps students pick classes by displaying RateMyProfessor ratings in-line on FSU's class search page.
- Technologies used: **JavaScript**, jQuery, Chrome API, Git

**Sixdegrees,** C++ program — github: wcwagner/sixdegrees

- Calculates the degree of separation between two actors by finding the shortest path between them in a graph of over 150,000 movies.
- Implemented fast bi-directional breadth-first-search to efficiently find shortest path between two nodes.

## OPEN SOURCE CONTRIBUTIONS

**pandas,** Contributor — github: pydata/pandas

- Large python data analysis library.
- Wrote various bug fixes, along with unit tests to ensure correctness and compatibility in the code base.
- Added documentation to further explain panda's handling of mixed data types.

## TECHNICAL SKILLS

**Programming Languages:** C++(Proficient), Python(Proficient), JavaScript, SQL
**Frameworks/Software:** Bootstrap, Apache Spark, HTML/CSS, Git, Sphinx, Linux, PyCharm, Visual Studio