

Chester “Chet” Weger

chester.weger@gmail.com || 909-706-2026 || chet-weger.herokuapp.com

Experience

Senior Software Engineer, Data Services – [Motional](#) – Santa Monica, CA July 2020 – Present

- Developed enhancements for annotations data pipelines used to annotate lidar, video, and image data.
 - Converted image segmentation “masks” produced by [Scale](#) into internal “polygon” representation of image segmentation.
 - Added “Video” task to support behavior rule books team. Submitted pilot tasks to Scale.
 - Created new “Traffic Light” pipeline to support annotating traffic lights in house.
 - Reworked the image pipeline to remove a step annotating 2D-boxes, and instead inferred 2D-boxes from instance-segmentation polygon groups.
- Performed ad-hoc data-devops to correct systematic errors or fix operational errors.
 - Wrote and ran script to unrectify 133,434 2D-box annotations and 19,027 images to achieve data consistency across batches.
 - Answered questions on tooling and provided rapid hot-fixes and bug-fixes to unblock the data pipeline and annotations workflows.
- Created a *docker-compose* setup for running the annotations backend *REST API* composed of *MySQL*, *Nginx*, and *.NET* containers.
- Created integration tests for the annotations data pipelines. The integration tests, combined with many unit tests brought code coverage to over 60%, from under 5% when I joined.

Senior Data Engineer – *Pepsico Inc. via Pivotal Solutions* – Remote May 2020 – July 2020

- Ingested transactions from Amazon Seller Central API and loaded data in Snowflake using Python, SQL, & Airflow using Datavault methodology.
- Created data pipeline to load Walmart Online Grocery data into Snowflake.
- Created data pipeline to load Shipt out of stock records into Snowflake.

Senior Data Engineer – *Walt Disney Studio Technologies via Insight Global* – Burbank, CA Sept 2018 – April 2020

- Developed a suite of data pipelines processing marketing data from sources including Facebook Marketing API, Display & Video 360, Google Campaign Manager, and AdWords. This data provided reporting and analytics for a 40+ person Media and Analytics team spending millions on ads. Built these pipelines using *Python*, *Spark*, and *Airflow*.
- Collaborated with stakeholders to understand and apply data transformations, and troubleshoot and analyze data quality issues. Added unit tests to ensure code readability & maintainability. Automated reporting processes to save ~40 hours per week of manual excel-centric reporting.

Software Development Engineer – *The Rubicon Project* – Playa Vista, CA July 2014 – Sept 2018

- Wrote new Spark job for reporting on Rubicon’s “Prebid” traffic. Used Amazon EMR for rapid deployment and scaling.
- Designed & Developed Spark job for “First-Party” dataset persisted in Druid to sunset an older reporting product and add new functionality.
- Migrated several legacy Hadoop jobs processing *billions* of transaction logs per hour to Spark using Spark SQL and Dataset API, increasing performance and maintainability.
- Contributed to data-pipeline, API, and application layer for next-gen performance reporting product based on Druid capable of ingesting over a billion aggregated rows per month and allowing an arbitrary selection of over 40 dimensions and over 30 base and derived metrics.

Student – *HMC Clinic Program with The Rubicon Project – Respectful Cross Device Attribution* Sept 2013 – May 2014

- Worked with three other students, an academic advisor, and a liaison from *The Rubicon Project* to gain industry experience and explore novel approaches to tackling a big-data problem.

Software Engineering Intern – [Pagewoo/Nearwo](#) (now defunct) June 2013 – Aug 2013

- Developed a module to automatically generate information about a user’s business. Performed integration with the yelp business search API as well as building a scraping component to find, among other things, relevant pictures that would be useful in banner ad creation.

Technical Assistant – *Pomona College*; CS052: Foundations of Computer Science Aug 2012 – Dec 2012

- Responsibilities included grading homework using an automated script and holding office hours to answer questions on homework problems.

Skills

- Proficient writing readable and idiomatic code in a variety of languages including *Python*, *C#*, *Scala*, *JavaScript*, *Java*, *Bash*, and *SQL*.
- Pragmatic ability to balance the [benefits of functional programming](#), without [harming code readability and comprehension](#).
- Experience using Gitlab and Jenkins for testing and deployment. Used build tools including *setuptools*, *sbt*, *dotnet*, and *maven*. Proficient using Docker and docker-compose.
- Experience applying [Data Engineering best practices](#) such as making tasks idempotent, and separating ingestion from processing.
- Experience using *Druid* for OLAP reporting on extremely high-volume transaction logs. Experience with *data-modeling* for OLAP & OLTP apps. Experience applying *data warehouse* design and concepts.
- Experience using *Django*, *vert.x*, and other frameworks to develop *RESTful* web services following best practices.

Career Highlights

- Code Yellow (Motional) – reached 60% code coverage on a 22K line codebase which had less than 5% coverage when I joined.
- Developed pipelines for a 40+ person Media and Analytics team spending millions of dollars on programmatic ads on a variety of channels.
- Druid Performance Tuning – worked with [Imply.io](#) to implement optimizations for Druid querying and indexing. Benchmarked potential performance optimizations to evaluate cost effectiveness. Achieved over 40% reduction in average report query time.
- Won Internal Hackathon “Automation Prize” – used a Sankey chart to visualize the flow of money among marketplace entities.

Personal Projects

- [Interactive Tic-Tac-Toe](#): can you beat the AI?
- [Self Learning Meta Tic-Tac-Toe](#): watch the AI improve its state-evaluation constants using temporal difference learning.
- [Interactive Meta Tic-Tac-Toe](#): test your abilities against the AI. Adjust its search depth parameter to change its skill level. Adjust its evaluation constants to change its behavior.

Additional Courses

- [Functional Programming in Scala](#), – completed five part Coursera specialization on functional programming and Spark.

Education

Bachelor of Arts, *Computer Science, Minor in Mathematics, 2014, Pomona College*, Claremont, California