


# Jarrell WAGGONER, PhD

 | [malloc47.com](http://malloc47.com) | ✉ [jarrell.waggoner@gmail.com](mailto:jarrell.waggoner@gmail.com) | ☎ 847-261-4747 | { [github.com/](https://github.com), [linkedin.com/in/](https://www.linkedin.com/in/), @ } **malloc47**

Engineering leader with decades-long career spanning research science, early-stage startups, and Fortune 5 companies. Specializing in the intersection of data science and data engineering with extensive experience architecting MLOps tools and data platforms. Provides technical leadership across an org-level portfolio of teams/projects while staying hands-on developing backend systems, with particular emphasis on functional languages.

## Education

<b>Ph.D.</b>	COMPUTER SCIENCE & ENGINEERING	University of South Carolina	08/2013
<b>M.E.</b>	COMPUTER SCIENCE	University of South Carolina	05/2009

## Experience

**OPTUM, INC.** | Principal Software Engineer (Grade 30) Remote (NYC) 2021–Present  
Brought in through acquisition of Rally Health, continuing as architect of the former Rally data organization now housed within the rebranded Optum Digital entity.

- Serve as Domain Workgroup Representative responsible for curating the technologies used by the data and ML practices in Optum’s centralized technology governance model with influence across all of the Optum Tech organization.
- Lead an Architecture Advocates workgroup embedding architecture-focused engineers on individual teams, fostering a cross-cutting, big-picture view of our architecture across our data organization.
- Built a crawler to extract table-level data lineage from our **Airflow** DAGs, generating a 2000 node graph navigable through a custom-built **JavaScript** graph interface.

**RALLY HEALTH, INC.** | Software Architect Remote (NYC) 2019–2021  
(Acquired by Optum) | Principal Software Engineer Minneapolis, MN 2017–2019

Staff-level engineering leader responsible for overseeing technical design and engineering decisions across Rally’s entire data organization that reached a peak of 90+ engineers, analysts, managers, and data scientists.

- Brought on board while Rally was a small startup to architect a complete rewrite of the data platform for the whole company, moving from a fixed **Cloudera** cluster to a self-service platform using **Databricks** and **Redshift** fed by **Spark** ETLs written in **Scala** and scheduled with **Airflow** atop **Kubernetes**. Built consensus on new architecture and delivered working system within a year.
- Member of the Rally Engineering Technical Staff, responsible for making cross-cutting engineering decisions, evaluating potential acquisitions, managing the RFC process, and organizing technical interest groups.
- Heavily involved in defining team structure and hiring, conducting over 140 interviews for IC and management roles to scale the data organization from 4 data engineers to over 50 data engineers across 7 teams.
- Coordinated platform integration with 20 internal and external teams and vendors across an extensive range of projects including productionalized ML workflows, frontend/mobile event tracking, real time data processing, security/compliance/privacy, data ingestion APIs, data quality validation, and self-service internal product analytics.

**DRW HOLDINGS, LLC** | Software Engineer Chicago, IL 2016–2017

Part of the Trading Infrastructure team, developing the internal platform used by every trading desk at DRW. Built greenfield high-performance service-oriented systems using **Clojure** and **Java** while maintaining legacy applications in **Ruby** and **C#** among a catalog of over 50 microservices.

- Contributed to a **Ruby**-based reconciliation tool used to balance cash flows for high-volume trading.
- Extended a graph-based research workflow tool used for computing the value and settle price of options, futures, equities, and other financial instruments, written in **Clojure**.
- Developed and extended multiple UI frontends for internal tools using **React** and **Reagent**.

**GROUPON, INC.** | Senior Software Engineer (SDE IV) Chicago, IL 02/2016–09/2016  
| Software Engineer (SDE III) 08/2013–02/2016

Contributed to multiple teams solving cross-cutting data engineering and **MLOps** problems.

- **Flux team:** Responsible for a data pipeline management and machine learning platform used to run productionalized decision tree learning models to predict customer attrition, lifetime customer value, and merchant value. Spearheaded implementation of distributed systems for the feature store, job scheduling, and data catalog components used by all models on the platform. Fed from **Teradata**, written in **Clojure**, and backed by **Hive**.

- **Supply Intelligence team:** Tech lead overseeing the critical business automation of lead-to-salesperson assignment that previously required ~80 sales managers to conduct manually; led the effort to rearchitect this legacy system from an ad-hoc job scheduling platform written in **Ruby** and **Bash** to a multi-staged **Hadoop** pipeline written in **Clojure** allowing it to scale to 6M daily candidate assignments.
- **Project Genesis strike team:** Coordinated with product and business teams to build an ETL to inject 250K leads in **Salesforce** from scraped web data that increased the unassigned leads pool by 10X.

**TERRASTRIDE, INC.** | Technical Engineer

Columbia, SC 2012–2014

Founding engineer at startup creating the [huntstand.com](http://huntstand.com) web application. Written using **Python**, **Django**, and **Backbone.js**; deployed to **AWS**. Responsible for curating full technology stack and coordinating with 5 developers.

**UofSC COMPUTER VISION LAB** | Research Assistant

Columbia, SC 2011–2013

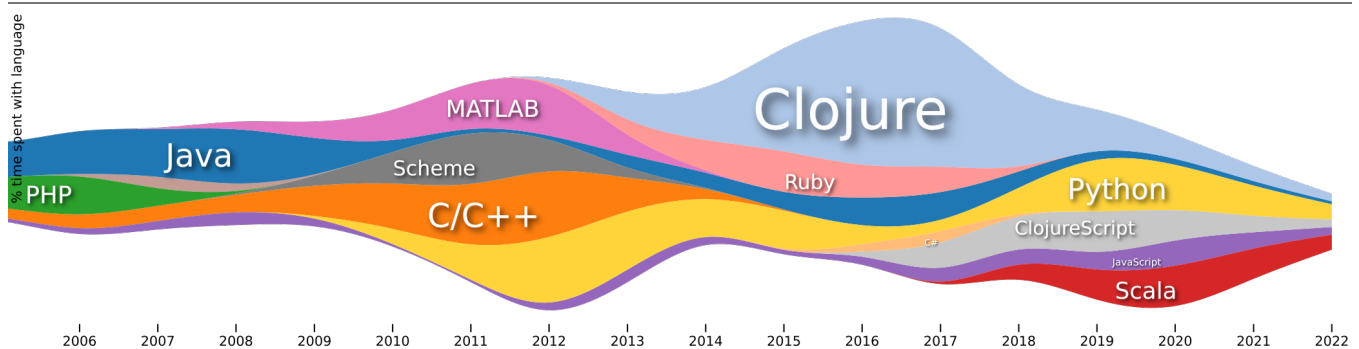
Conducted dissertation research on segmentation methods to extract important physical characteristics from image volumes of metallic and biologic materials, developed using **Python**, **NumPy**, **SciPy**, **OpenCV**, and **MATLAB**. Managed the lab computer network and organized weekly lab meetings. Created a **Django**-based web application for manual segmentation, an ML-trained classifier for assisted segmentation, and a fully automated energy-minimization segmentation approach, with large-scale evaluations on real and synthetic datasets. [[Dissertation](#)] [[Code](#)]

**DARPA MIND'S EYE PROGRAM** | Research Assistant

Columbia, SC 2010–2011

Built video event recognition systems for the **DARPA Mind's Eye Program**, collaborating with 10 students and faculty members across three institutions to create an AI system that describes events in a video clip as natural language sentences. Developed algorithms in **Scheme**, **Bash**, **MATLAB**, and **C** to process a corpus of 3480 videos extracted into over 1.5 million frames. Conducted distributed processing on the **Steele** cluster which was, at the time, among the top 500 most powerful supercomputing clusters. [Website: [0xab.com/research/video-events.html](http://0xab.com/research/video-events.html)] [[Code](#)]

## Languages



## Selected Publications and Presentations

- [1] Project Athena: Rally's next-generation data platform. *Optum/UHG/UHC Analytics Conference*. Eden Prairie, MN. September 2019.
- [2] Derrick. C. Spell, Ling-Yong Wang, Richard T. Shomer, Bahador Nooraei, **Jarrell Waggoner**, Xaio-Han T. Zeng, Jae Y. Chung, Kai-Chen Cheng, and Daniel Kirsche. QED: Groupon's ETL management and curated feature catalog system for machine learning. In *IEEE International Conference on Big Data*, pages 1639–1646, Dec 2016. [[Link](#)].
- [3] Rules engines: Logic as data structure. *Palmetto Open Source Software Conference*. Columbia, SC. April 2015. [[Slides](#)].
- [4] **Jarrell Waggoner**, Youjie Zhou, Jeff Simmons, Marc De Graef, and Song Wang. Topology-preserving multi-label image segmentation. In *IEEE Workshop on Applications of Computer Vision (WACV)*, pages 1084–1091, Waikoloa Beach, HI, 2015. [[PDF](#)].
- [5] **Jarrell Waggoner**, Youjie Zhou, Jeff Simmons, Marc De Graef, and Song Wang. Graph-cut based interactive segmentation of 3D materials-science images. *Machine Vision and Applications*, 25:1615–1629, 2014. [[PDF](#)].
- [6] **Jarrell Waggoner**, Jeff Simmons, Marc De Graef, and Song Wang. 3D materials image segmentation by 2D propagation: A graph-cut approach considering homomorphism. *IEEE Transactions on Image Processing*, 22, 2013. [[PDF](#)].
- [7] Andrei Barbu, Alexander Bridge, Zachary Burchill, Dan Coroian, Sven Dickinson, Sanja Fidler, Aaron Michaux, Sam Mussman, Siddharth Narayanaswamy, Dhaval Salvi, Lara Schmidt, Jiangnan Shangguan, Jeffrey Mark Siskind, **Jarrell Waggoner**, Song Wang, Jinlian Wei, Yifan Yin, and Zhiqi Zhang. Video in sentences out. In *Conference on Uncertainty in Artificial Intelligence*, pages 102–112, 2012. [[PDF](#)].
- [8] **Jarrell Waggoner**, Jeff Simmons, and Song Wang. Combining global labeling and local relabeling for metallic image segmentation. In *Proceedings of SPIE (Computational Imaging X)*, volume 8296, Burlingame, CA, 2012. [[PDF](#)].

## Interests

data visualization, geographic information systems, Linux, [music composition](#), functional programming