Jarrell Waggoner

Biographical

Work Address	3000 K St NW #350 Washington, DC 20007
Phone	847-261-4747
email	jarrell.waggoner@gmail.com

Online

Website	www.malloc47.com
Twitter	@malloc47
github	github.com/malloc47
LinkedIn	linkedin.com/in/malloc47

Interests computer vision, image processing, artificial intelligence, pattern recognition & machine learning, data science, data engineering, functional programming, web development, GIS, Clojure

Education

Aug. 2013	**Ph.D.**	Computer Science & Engineering	**University of South Carolina**
May 2009	**M.E.**	Computer Science	**University of South Carolina**

Experience

2017—Present

Architect at Rally Health, Inc.

Seasoned engineering leader responsible for overseeing technical design and engineering decisions across Rally's entire data organization of 50+ engineers, analysts, and data scientists.

- Architected 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 while growing the team from only myself to an independent platform team of 10 engineers.
- Member of the Rally Engineering Technical Staff, responsible for making cross-cutting engineering decisions, evaluating potential acquisitions, signing off on major company-wide architectural changes, and organizing technical interest groups.
- Heavily involved in defining team structure and hiring, conducting over 100 interviews for IC and management roles to scale the data organization from 4 engineers to over 50 engineers.
- Fostered good design practices across the data organization by auditing design documents, leading architecture meetings for multiple teams to grow design aptitude, and spearheading the standardization of best-practices on the team including code review, documented on-call workflow, CI/CD , monitoring/alerting, idempotent ETL workflows, and reproducible EDW layout.
- Coordinated with over 16 internal and external teams across an extensive range of projects including productionalized ML workflows, frontend/mobile event tracking, real time data processing, data anonymization, security/compliance/privacy requirements, data ingestion APIs, data quality validation, and self-service internal product analytics.

2016—2017

Software Engineer at DRW Holdings, LLC

Member 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 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**.

2013—2016

Senior Software Engineer at Groupon, Inc.

Contributed to three engineering teams: The Flux team building Data Science pipelines, the Project Genesis strike team integrating scraped web data into **Salesforce**, and served as Tech Lead of the Supply Intelligence team creating internal sales tools to optimize Groupon's supply funnel.

- Built a **PostgreSQL**-backed high-performance caching and write management system in **Clojure** around the **Salesforce** API that hits 10K req/min.
- Managed critical business automation of the sales lead assignment process that previously required an estimated 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** to handle over 6M accounts.
- Coordinated with product and business teams to ETL 250K leads in **Salesforce** from scraped web data.

— Built out an ETL management and machine learning platform using **Python**, **Clojure**, **Hive**, and **Spark** to run mission-critical Decision Tree Learning models to predict customer attrition, lifetime customer value, and merchant value.

— Mentored interns and junior developers, established best practices, and led multiple major technical initiatives on a team of 5 developers.

2012—2014 **Technical Engineer** at TerraStride, Inc.

Software developer in an agile startup environment creating the 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.

2011—2013 **Research Assistant** at USC Computer Vision Lab

Dissertation research on computer vision models and algorithms for materials science image segmentation in **Python**, **NumPy**, **SciPy**, **OpenCV**, and **MATLAB**. Created a web interface using **Django**, **JavaScript**, and **jQuery**. Conducted large-scale analysis using a 98-core high-performance computing system.

2010—2011 **Research Assistant** for the DARPA Mind's Eye Program

Researched video event recognition for the DARPA Mind's Eye program. Collaborated with 10 students and faculty members across three institutions. Developed algorithms in **Scheme**, **Bash**, **MATLAB**, and **C** to process a corpus of 3480 videos extracted into over 1.5 million frames. Distributed processing over 7 HPC machines. 0xab.com/research/video-events.html , github.com/malloc47/video-in-sentences-out

2009—2010 **NEH Fellow** at the USC Center for Digital Humanities (Sapheos/Paragon Project)

Developed the prototype for a *digital collation* application to identify sub-textual inconsistencies among multiple copies of *The Faerie Queene* by Edmund Spenser. Created in **MATLAB** using **VLFeat** and **OpenCV** to process tens of thousands of book page images. github.com/malloc47/digital-collation

Skills & Languages

◦ • • Bash	◦ • • GNU/Linux	◦ • • Java	• Nix/NixOS	• • Scala
◦ • • Clojure	◦ • • Hadoop ecosystem	• • JavaScript	◦ • • PostgreSQL	◦ • • Scheme
◦ • • git	• Haskell	• • LaTeX	◦ • • Python	◦ • • Spark

• Small-scale or personal projects • • Used in production ◦ • • Used in large-scale production systems

Selected Publications

[1] 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].

[2] **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].

[3] **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].

[4] **Jarrell Waggoner**. *Multi-Label Segmentation Propagation for Materials Science Images Incorporating Topology and Interactivity*. Dissertation, University of South Carolina, 2013. [PDF].

[5] **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].

[6] 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].

[7] **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].

Recent Talks

[1] Project athena: Rally's next-generation data platform. *Optum/UHG/UHC Analytics Conference.* Eden Prairie, MN. September 2019.

[2] Rules engines: Logic as data structure. *Palmetto Open Source Software Conference.* Columbia, SC. April 2015. [Slides].

[3] Python for computer vision. *All Things Open.* Raleigh, NC. October 2013. [Slides].

[4] Extending django. *Palmetto Open Source Software Conference.* Columbia, SC. March 2013. [Slides].

Activities

open source software, GIS visualization, Linux, music composition