Jarrell Waggoner, PhD

🗽 | malloc47.com | ✉ jarrell.waggoner@gmail.com | ☏ 847-261-4747 | { github.com/, linkedin.com/in/, @ } **malloc47**

Summary

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: 10 years of experience using Clojure to build data pipelines, microservices, and distributed systems.

Research Interests

computer vision, image segmentation, document image analysis, event recognition, image processing, artificial intelligence, pattern recognition & machine learning, data science, functional programming

Education

Doctor of Philosophy in Computer Science & Engineering — University of South Carolina — 08/2013
Advisor: Dr. Song Wang
Dissertation: "Multi-Label Segmentation Propagation for Materials Science Images Incorporating Topology and Interactivity"

Master of Engineering in Computer Science — University of South Carolina — 05/2009
magna cum laude

Industry 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. (Acquired by Optum) | Software Architect — Remote (NYC) — 2019—2021
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 web application. Written using **Python**, **Django**, and **Backbone.js**; deployed to **AWS**. Responsible for curating full technology stack and coordinating with 5 developers.

Palmetto Computer Labs | Project Manager — Columbia, SC — 2011—2013

Created and taught workshops on **git**, the **Linux** command line, **Android** development, and open source software for hundreds of students, developers, and government officials at IT-oLogy. Managed the Open IT Lab and associated projects. Assisted in planning POSSCON.

Elastic Vision Consulting | Contractor — Columbia, SC — 2011

Built a parser and generator for **XML** medical records formats (CCR and CCD) in a **Java** web application. Written using **JDOM**, **Xerces**, and **Hibernate**, and run on an **Axis2+Jetty6** driven server.

Research Experience

Research Assistant funded by AFOSR | Materials Volume Segmentation — 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]

Research Assistant funded by DARPA | Video Event Recognition 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] [Code]

NEH Fellow at the Center for Digital Humanities | Digital Collation 2009—2010
Developed the prototype for a *digital collation* application as part of the Sapheos / Paragon project to identify subtextual inconsistencies among multiple scanned copies of *The Faerie Queene* by Edmund Spenser. Created in **MATLAB** using **OpenCV** to process tens of thousands of book page images.
[Paper] [Narrative] [Code: github.com/malloc47/digital-collation]

Teaching Experience

GK-12 Fellow at Crayton Middle School | 8th Grade Science 2008—2009
Served in Crayton Middle School, coordinating with the classroom instructor to enhance the STEM curriculum and activities in an 8th grade science classroom. Subsequently coordinated and taught at the GK-12 Institute for Teachers, presenting the activities developed and delivered in the classroom.

Graduate Teaching Assistant at UofSC | Web Development 2007—2008, 2011
Supervised CSCE 145 labs, covering software development with Java, and taught CSCE 102, covering Javascript, HTML, and CSS. Taught CSCE 211 covering digital logic design.

Instructor for CSCE 204 at USCL | Introductory Programming Spring 2007
Hired as special faculty. Taught introductory Visual Basic for majors and non-majors. Selected textbooks, developed all course material, graded all assignments. Worked with Dr. Noni M. Bohonak

Camp Instructor for USCL Arts and Sciences Adventure Camp | 5th-8th Grade Students Fall 2006
Worked in collaboration with Dr. Dwayne Brown. One of two instructors teaching Math and Computer Science to grade school students.

Professional Tutor at USCL Academic Success Center | High School and College Students 2003—2007
Student and graduate tutor for college-level Mathematics, Computer Science, Physics, and English classes.

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] Youjie Zhou, Lili Ju, Yu Cao, **Jarrell Waggoner**, Yuewei Lin, Jeff Simmons, and Song Wang. Edge-weighted centroid voronoi tessellation with propagation of consistency constraint for 3D grain segmentation in microscopic superalloy images. In *CVPR Workshop on Perception Beyond the Visible Spectrum (PBVS)*, 2014. [PDF].

[5] Dhaval Salvi, **Jarrell Waggoner**, Andrew Temlyakov, and Song Wang. A graph-based algorithm for multi-target tracking with occlusion. In *IEEE Workshop on Applications of Computer Vision (WACV)*, 2013. [PDF].

[6] Dhaval Salvi, Jun Zhou, **Jarrell Waggoner**, and Song Wang. Handwritten text segmentation using average longest path algorithm. In *IEEE Workshop on Applications of Computer Vision (WACV)*, 2013. [PDF].

[7] Andrew Temlyakov, Pahal Dalal, **Jarrell Waggoner**, Dhaval Salvi, and Song Wang. Shape and image retrieval by organizing instances using population cues. In *IEEE Workshop on Applications of Computer Vision (WACV)*, 2013. [PDF].

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

[9] **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].

[10] **Jarrell Waggoner**, Youjie Zhou, Jeff Simmons, Ayman Salem, Marc De Graef, and Song Wang. Interactive grain image segmentation using graph cut algorithms. In *Proceedings of SPIE (Computational Imaging XI)*, Burlingame, CA, 2013. [PDF].

[11] 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].

[12] Andrei Barbu, Alexander Bridge, Dan Coroian, Sven Dickinson, Sam Mussman, Siddharth Narayanaswamy, Dhaval Salvi, Lara Schmidt, Jiangnan Shangguan, Jeffrey Mark Siskind, **Jarrell Waggoner**, Song Wang, Jinlian Wei, Yifan Yin, and Zhiqi Zhang. Large-scale automatic labeling of video events with verbs based on event-participant interaction. Technical report, 2012. [PDF].

[13] **Jarrell Waggoner**, Jeff Simmons, Marc De Graef, and Song Wang. Graph cut approaches for materials segmentation preserving shape, appearance, and topology. In *International Conference on 3D Materials Science*, pages 147–152, Seven Springs, PA, 2012. [PDF].

[14] **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].

[15] Zhiqi Zhang, Sanja Fidler, **Jarrell Waggoner**, Yu Cao, Sven Dickinson, Jeffrey Mark Siskind, and Song Wang. Superedge grouping for object localization by combining appearance and shape information. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3266–3273, Providence, RI, 2012. [PDF].

[16] Song Wang, **Jarrell Waggoner**, and Jeff Simmons. Graph-cut methods for grain boundary segmentation. *JOM Journal of the Minerals, Metals and Materials Society*, 63:49–51, 2011. [PDF].

[17] Andrew Temlyakov, Brent C. Munsell, **Jarrell Waggoner**, and Song Wang. Two perceptually motivated strategies for shape classification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2289–2296, 2010. [PDF].

[18] Zhiqi Zhang, Yu Cao, Dhaval Salvi, Kenton Oliver, **Jarrell Waggoner**, and Song Wang. Free-shape subwindow search for object localization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1086–1093, San Francisco, CA, 2010. [PDF].

Posters/Presentations

[P1] Project Athena: Rally's next-generation data platform. *Optum/UHG/UHC Analytics Conference.* Eden Prairie, MN. September 2019.

[P2] Rules engines: Logic as data structure. *Palmetto Open Source Software Conference.* Columbia, SC. April 2015. [Slides].

[P3] Python for computer vision. *All Things Open.* Raleigh, NC. October 2013. [Slides].

[P4] Interactive grain image segmentation using graph cut algorithms. *USC Graduate Student Day.* Columbia, SC. April 2013. [Slides].

[P5] Computer science: Research, industry, and entrepreneurship. *Careers in Science Lecture Series.* Lancaster, SC. March 2013. [Slides].

[P6] Extending Django. *Palmetto Open Source Software Conference.* Columbia, SC. March 2013. [Slides].

[P7] Interactive grain image segmentation using graph cut algorithms. *SPIE (Computational Imaging XI).* Burlingame, CA. February 2013. [Slides].

[P8] Android application development workshop. *Appathon Contest.* Columbia, SC. November 2012.

[P9] Homeomorphic multi-structure propagation for metallic image segmentation. *Gamecock Computing Research Symposium.* Columbia, SC. October 2012.

[P10] Open source and education. *SC Municipal Technology Association (SCMTA) Conference.* Charleston, SC. September 2012.

[P11] Open source and higher education. *SC Technical College System (SCTCS) Conference.* Columbia, SC. September 2012.

[P12] Introduction to android development. *Digital Humanities High Performance Computing (DHHPC) Workshop.* Columbia, SC. August 2012.

[P13] Combining global labeling and local relabeling for metallic image segmentation. *SPIE (Computational Imaging X).* Burlingame, CA. January 2012.

[P14] Open source and government. *SC Government Management Information Systems (SCGMIS) Software Developers Workshop.* Columbia, SC. January 2012.

[P15] Superpixel contour completion. *DARPA Mind's Eye PI Meeting.* Denver, CO. January 2011.

Honors/Awards

First Place	» UofSC	Gamecock Computing Research Symposium Poster Session	2012
First Place	» UofSC	Graduate Student Day Presentation	2012
Second Place	» UofSC	Graduate Student Day Presentation	2011
Honorable Mention	» UofSC	Graduate Student Day Presentation	2010
Inductee	» UofSC	Upsilon Pi Epsilon	2009
Recipient	» USCL	Clara P. Hammond Award	2004
Recipient	» USCL	Science and Mathematics Award	2004
Recipient	» USCL	Highest Academic Average Award	2004

Classes Taught

Open Source 101	» IT-oLogy	Open Source Software	2012—2013
Version Control 101	» IT-oLogy	git, github	2012—2013
Command Line 101	» IT-oLogy	Linux, Bash	2012—2013
CSCE 211	» UofSC	Digital Logic Design	Fall 2011
CSCE 102	» UofSC	HTML/CSS/JavaScript	Summer II 2008
CSCE 145 Lab	» UofSC	Algorithmic Design	Spring 2008
CSCE 145 Lab	» UofSC	Algorithmic Design	Fall 2007
CSCE 204	» USCL	Program Design and Development	Spring 2007
Math 241 & Math 242	» USCL	Maple	Spring 2007

Service

Judge	»	Rally Data Hackathon	2021
Mentor	»	Groupon internship program	2014
Book Reviewer	»	Practical Data Analysis, Packt Publishing	2013
Webmaster	»	Winter Vision Meetings	2013
Webmaster	»	Workshop on the Applications of Computer Vision	2013
Judge	»	Discovery Day — Undergraduate Research Presentations	2012
Reviewer	»	Pattern Recognition Letters	2011—2012
Reviewer	»	IEEE Transactions on Pattern Analysis and Machine Intelligence	2012
SysAdmin	»	Computer Vision Lab	2009—2012

Personal and Open Source Projects

COCKPIT | github.com/malloc47/cockpit
Single-page web application dashboard deplayed on a wall-mounted tablet that shows the local time, weather, and transit departure times. Written in **ClojureScript** using the **re-frame** framework and styled using the MUI **React** toolset.
[Blog Series: malloc47.com/building-a-personal-dashboard-in-clojurescript]

BEFUNGE.PY | github.com/malloc47/befunge.py
Complete Befunge interpreter written in **Python**. Implements the Befunge 93 specification, and is one of the closest Python equivalents to the **C** reference implementation.

TERM-DO | github.com/malloc47/term-do
An interactive terminal prompt that displays potential command completions as you type. A hybrid of gnome-do and Emacs's ido-mode. Works on many tested VT100 terminal types; built in **C++**. Includes client/server architecture implemented with **boost.interprocess** and a plugin system.

MATSCISEG | github.com/malloc47/matsciseg
Framework for propagated 3D volume segmentation, used in my dissertation work. Algorithms created in **Python** and **C++** and exposed as a web API using **Django**. Includes a web application that consumes the API created in **JavaScript**, and **jQuery**.

RATIO CONTOUR | github.com/malloc47/ratio-contour
Maintainer and contributor for the Ratio Contour project, a salient object detection and segmentation method used for computer vision applications. Developed in **C** and **MATLAB**.

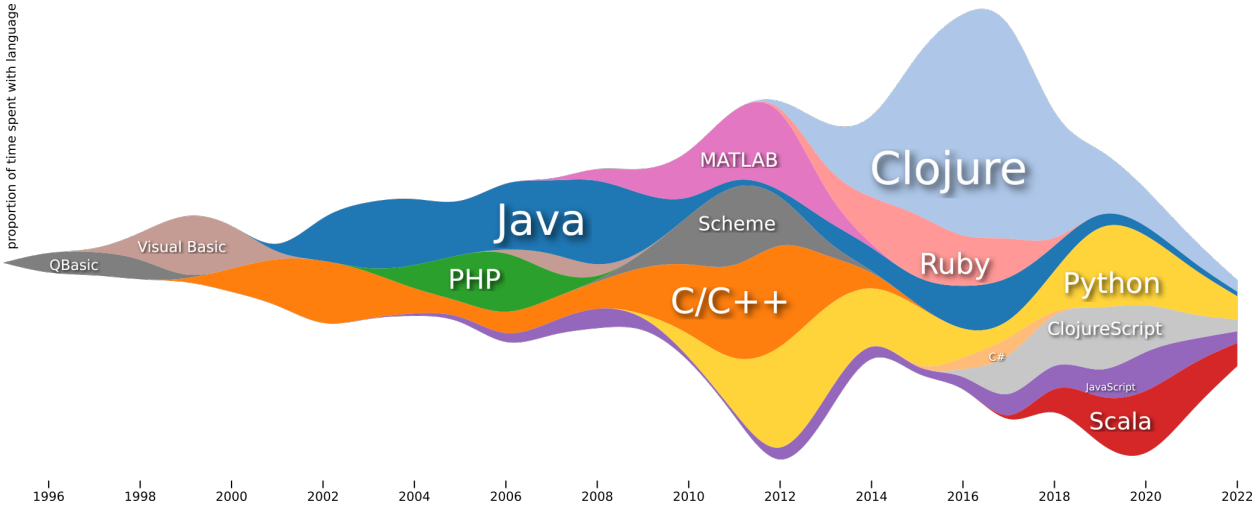
Digital Collation | github.com/malloc47/digital-collation

Research project to "collate" high-resolution documents by using image registration, accomplished using the **SIFT** feature detector and a thin plate spline warping technique, written in **MATLAB**.

Skills & Languages

- • • Clojure
- • • ClojureScript
- • • Python
- • • Bash
- • • Java
- • • C/C++
- • • JavaScript
- • Haskell
- • Emacs Lisp
- • LaTeX
- • • MATLAB
- • • Scala
- • • Scheme
- • • Ruby
- • • git
- • • Linux
- • • Django
- • React/re-frame
- • • PostgreSQL
- • • Airflow
- • • Hadoop ecosystem
- • • NumPy/SciPy
- • • OpenCV
- • • Spark

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



Interests

data visualization, geographic information systems, Linux, music composition