

# Matthew R. Walker

matt.r.walker@gmail.com

8150 Raritan St San Antonio, TX 78254

## EDUCATION

### University of Texas at Austin

MS, Computer Sciences, May 2005

Advisor: Dr. Inderjit Dhillon

### Rice University

BS, Electrical and Computer Engineering, May 2003

Magna cum laude

## EXPERIENCE

### Etsy – Large-Scale Analytics

Dec 2009 – present

Created EMR-based analytics platform with cascading.jruby query language, extensive custom operator library written in Scala, S3 storage, and EC2 automation and monitoring. CEP-like visit query operators allowed fine-grained attribution of desired actions to site features and behaviors. Enabled analyses ranging from A/B/A testing, to search experiments, to member lifecycle, to general business intelligence. Scaled system to over 300 jobs, powering analytics and popular applications on the site like the Facebook gift recommender as well as internal Hive and dashboards.

### Adtuitive – Online Advertising Research

May 2009 – Dec 2009

First hire brought on to research domain-specific page-to-product matching algorithms. Built on-demand EC2 sandbox for search index analysis and algorithm experimentation. Researched IDTokenSets and custom algorithm built on Needleman-Wunsch for product name extraction. Applied engineering skills to create affiliate tracking system, provide searchable product tags, and build web ETL pipeline in cascading.jruby.

### Pulse Meridian – Data-Driven Business Optimization

Jun 2007 – Aug 2008

Founded startup with graduate school colleagues. Developed multilinear and multivariate polynomial models to optimize semiconductor yield for Pintail Technologies. Provided CNET data mining consulting services to lift click rates in their online advertisement platform. Built several web applications, one currently used to measure the efficacy of lithotripsy treatments by a urology practice.

### Pervasive Software – Dataflow Engineering

May 2007 – May 2009

Converted dataflow framework into pure Java library, seeing project through to general availability. Designed composition language and embedded library in JRuby to further improve developer productivity. Implemented, benchmarked, and optimized several key operators used for delimited text I/O,  $k$ -means co-clustering, support vector machine classification, and Monte Carlo simulation.

### Pervasive Software – Hyper-Parallel Dataflow

Nov 2006 – May 2007

Conducted original research in concurrent programming using a Java framework for dataflow applications. Implemented and benchmarked a variety of novel dataflow formulations of algorithms including decision tree training, rule-based name standardization, and  $k$ -means clustering.

### Scott & White Hospital – Database Optimization

Oct 2006 – present

Developed, modified, and maintained library of stored procedures, views, and Crystal Reports in pathology department's PowerPath system. Collaborated with physicians to generate reports for auditing, administration, and research.

### Texas A&M University – Secure Distributed Applications

Jun 2005 – Nov 2006

Designed and implemented distributed applications spanning authentication, identity management, web apps, provisioning, and bulk mail. Coordinated subsystem deployment, testing, interface features, and notification content.

### University of Texas at Austin – Research

Sep 2003 – Dec 2004

Conducted original research in data mining and optimization focusing on constrained  $k$ -means clustering. Solved optimization problems in C++ using MOSEK, while addressing noisy constraints.

**MIT Lincoln Laboratory – Satellite Tracking***Jun 2002 – Aug 2002*

Upgraded real-time ionospheric monitoring system from FORTRAN and C to object-oriented C++. Measured and corrected for ionospheric activity's effect on satellite tracking using GPS signal bias.

**Rice University – Parallel Disk Scheduling***Jun 2001 – Aug 2001*

Simulated the PC-OPT parallel disk scheduling algorithm. Developed software testing environment and collected data to identify the optimal prefetch schedule.

**Rice University – Device Drivers***Jun 2000 – Aug 2000*

Implemented device drivers and firmware for VME cards used as high-speed data selectors in the Muon Detector at CERN.

**TEACHING**

Computer fluency teaching assistant, fall 2004

Artificial intelligence lab assistant, spring 2003

Object-oriented programming lab assistant, spring 2002

Digital logic design lab assistant, spring 2001

**SKILLS**

**Programming Languages:** Ruby, Scala, Java, Bash, SQL

**Technologies:** EMR (Hadoop and Hive), S3, EC2, git, sbt, PostgreSQL, MySQL, Lucene

**Communication:** Two years teaching experience including lectures and labs. Technical presentations ranging from tutorials, to design overviews, to paper presentations. See also publications.

**PUBLICATIONS**

Nena Marin, Srivatsava Daruru, Joydeep Ghosh, **Matt Walker**. Pervasive Parallelism in Data Mining: Dataflow Solution to Co-Clustering Large and Sparse Netflix Data. To appear in *The Fourteenth International Conference on Knowledge Discovery and Data Mining (KDD)*, July 2009.

**Matt Walker**, Kevin Irwin. Four Paths to Java Parallelism. *Java Developer's Journal*, Vol 13, Issue 9, September 2008.

Jim Falgout, **Matt Walker**. It's a Multi-Core World: Let the Data Flow. *Java Developer's Journal*, Vol 12, Issue 8, August, 2007.

**M. Walker**. GRIMS/RTSEQ Object-Orientation Report. *MIT Lincoln Laboratory, Project Report SS-7*, November 7, 2003.

**HONORS**

Eagle Scout

Phi Beta Kappa

Eta Kappa Nu

Donald D. Harrington Fellowship at UT, 2003–2004

Dean's Excellence Award from UT Natural Sciences, 2003

NSF Honorable Mention, 2003 and 2004

Hertz Foundation Finalist, 2003

Herbert Allen Outstanding Junior Award from Rice Engineering Alumni, 2002

Junior and Senior Merit Awards in ECE from Rice Engineering Alumni, 2002 and 2003

Dean's List and President's Honor Roll at Rice University, 1999–2003

Ford Motor Company/Golden Key Undergraduate Scholar Award

Louis J. Walsh Scholarship in Engineering at Rice University

National Merit Scholar

Advanced Placement National Scholar

Abell-Hanger Foundation Scholarship from Texas Interscholastic League Foundation