

Stephen Macke

PH.D. CANDIDATE · RESEARCH ASSISTANT

Thomas Siebel Center for Computer Science #2119B, 201 N. Goodwin Ave. Urbana, IL 61801

☎ (+1) 314-266-8333 | ✉ smacke@illinois.edu | 🏠 smacke.net | 📷 smacke | 📱 smacke

Education

University of Illinois at Urbana-Champaign

PH.D. IN COMPUTER SCIENCE

- ◇ Advisor: Aditya Parameswaran
- ◇ GPA: 4.0

Urbana, IL

2015–Present

Stanford University

M.S. IN COMPUTER SCIENCE

- ◇ GPA: 3.815

Stanford, CA

2013–2015

University of Tulsa

B.S. IN COMPUTER SCIENCE, APPLIED MATH MAJOR

- ◇ GPA: 4.0
- ◇ Summa Cum Laude, University Honors
- ◇ Departmental Honors in Computer Science, Mathematics

Tulsa, OK

2009–2013

Research Interests

- ◇ Interactive Data Analytics, Approximate Query Processing, Data Mining

Publications

- ◇ **Stephen Macke**, Yiming Zhang, Silu Huang, Aditya Parameswaran, “[Adaptive Sampling for Rapidly Matching Histograms](#)”, *Under Shepherding for 44th Int. Conf. on Very Large Data Bases (VLDB), Rio de Janeiro, Brazil, Aug. 2018*
- ◇ Doris Xin, **Stephen Macke**, Litian Ma, Jialin Liu, Rong Ma, Shuchen Song, Aditya Parameswaran, “[Helix: Holistic Optimization for Accelerating Iterative Machine Learning](#)”, *In preparation.*
- ◇ Doris Xin, Litian Ma, Jialin Liu, **Stephen Macke**, Shuchen Song, Aditya Parameswaran, “Accelerating Human-in-the-loop Machine Learning: Challenges and Opportunities”, *DEEM Workshop at SIGMOD Int’l Conf. on Management of Data, Houston, USA. June 2018*
- ◇ Doris Xin, Litian Ma, Jialin Liu, **Stephen Macke**, Aditya Parameswaran, “Accelerating Human-in-the-loop Machine Learning With Helix (Demo)”, *44th Int. Conf. on Very Large Data Bases (VLDB), Rio de Janeiro, Brazil, Aug. 2018*
- ◇ Tana Wattanawaroon, **Stephen Macke**, Aditya Parameswaran, “[Towards a Theory of Data-Diff: Optimal Synthesis of Succinct Data Modification Scripts](#)”, *ArXiv preprint.*
- ◇ Ahmed El-Kishky, Frank Xu, Aston Zhang, **Stephen Macke** and Jiawei Han, “Entropy-Based Subword Mining for Word Embeddings”, *SCLem Workshop at NAACL Conf. on Computational Linguistics, New Orleans, USA. June 2018*

Research and Industry Experience

University of Illinois at Urbana-Champaign

GRADUATE RESEARCHER

- ◇ Performing research toward the completion of a Ph.D. in the Data and Information Systems (DAIS) laboratory at UIUC, under the supervision of my advisor, Professor Aditya Parameswaran.

Urbana, IL

August 2015–Present

Alation, Inc.

SOFTWARE ENGINEERING INTERN

- ◇ Added several improvements to the Alation Catalogue’s database query search feature. Involved infrastructure improvements for keeping human- and machine-annotated metadata in the Catalogue (eventually) consistent with Elasticsearch indexes, as well as query search algorithmic enhancements to improve recall.

Redwood City, CA

Summer 2016

Palantir Technologies

SOFTWARE ENGINEERING INTERN

- ◇ Rewrote Palantir’s search API, and provided an implementation built on top of Elasticsearch.

Palo Alto, CA

Summer 2014

Palantir Technologies

SOFTWARE ENGINEERING INTERN

Palo Alto, CA

Summer 2013

- ◊ Wrote memory-efficient functionality for a server used to perform “live” reindexing of Lucene search indices. Also wrote integration tests which fuzzed the new code, ensuring correctness.

University of Illinois at Urbana-Champaign

NSF REU RESEARCHER

Urbana, IL

Summer 2012

- ◊ Performed research and development of data-parallel algorithms for stochastic, particle-resolved aerosol particle simulations (condensation and coagulation) under the supervision of Professor Matthew West.

Teaching Experience

Stanford University

COURSE ASSISTANT FOR CS149

Stanford, CA

Winter 2013-14 and 2014-15

- ◊ Course title: Parallel Computing
- ◊ Instructors: Alex Aiken and Kunle Olukotun
- ◊ Course material covers parallel architectures, threaded programming, cache coherence, and various programming models/frameworks (Java concurrency, functional, Hadoop/MapReduce, MPI, OpenMP, CUDA).
- ◊ In the latest offering I was responsible for managing AWS infrastructure and instances used for the course programming assignments, including large Hadoop clusters comprising over 80 AWS instances.

Stanford University

COURSE ASSISTANT FOR CS103

Stanford, CA

Fall 2013 and 2014

- ◊ Course title: Mathematical Foundations of Computing
- ◊ Instructor: Keith Schwarz
- ◊ Course material includes proof techniques (direct/contrapositive/contradiction, pigeonhole principle, induction), set theory / functions / relations, logic, automata, and complexity.

Selected Software Projects

JAVA DIRECT I/O (JAYDIO)

- ◊ Project provides direct I/O functionality for Java (bypassing FS cache)
- ◊ Hosted at <https://github.com/smacke/jaydio>
- ◊ Funded by Sourcegraph Open Source Fellowship (\$1000 stipend)

Honors & Awards

- 2015 **Awardee**, Andrew and Shana Laursen Fellowship
- 2015 **Awardee**, Diffenbaugh Graduate Fellowship
- 2014 **Awardee**, NSF Graduate Research Fellowship*
- 2014 **Awardee**, Sourcegraph Open Source Fellowship
- 2013 **1st Place**, DWR Governor’s Cup Business Plan Competition, Oklahoma division (\$22,000 prize)
- 2013 **Top-150 Score**, Putnam Competition (exact rank: 142.5, highest in Oklahoma/Arkansas region)*
- 2012 **Awardee**, Goldwater Scholarship*
- 2012 **World Finalist**, ACM ICPC World Finals in Warsaw, Poland
- 2009 **Awardee**, University of Tulsa Presidential Scholarship (covering all tuition and living expenses)

Skills

Programming Languages C/C++, Python, Java, Javascript

Machine Learning Tools Numpy, SciPy, Pandas, SciKit-Learn, CvxPy

High Performance Tools CUDA, OpenMP, MPI

Engineering Systems/Frameworks Git, ElasticSearch, PostgreSQL, Django, React+Redux

*Nationally competitive