Dav J. W. Clark, Ph.D.

Data and Learning Scientist

917-544-8408 | davclark@gmail.com | 1201 W Mount Royal Ave #512, Baltimore, Maryland

https://www.linkedin.com/in/davclark | https://github.com/davclark

Summary

Team player who brings a pragmatic mix of tools, training, and community to bear on big challenges. Extensive and varied experience with technology including computational linguistics, statistics, machine learning, static and interactive visualization, physical computing. Demonstrated leadership in coalition building, teaching, and project incubation across academic, enterprise, consulting and start-up settings. Fast learner looking to join a great team.

Experience

Research Scientist, Kennedy Krieger Institute, Baltimore, MD	2016-2017
UC Berkeley, Berkeley, CA	
Fellow, Berkeley Institute for Data Science (BIDS)	2014-2016
Data Scientist, D-Lab	2013-2016
Instructor, Machine Learning (Masters in Data Science); Hacking Measurement	2014, 2015
Chief Scientist, Oroeco / Startup Chile, San Francisco, CA / Santiago, Chile	2012-2015
Partner, KeepOpen Web Design, Euless, TX (remote)	2004-2013
Principal Scientist, Entrieva (now LucidMedia), Reston, VA	2006-2007

Education

Ph.D., Psychology, UC Berkeley	2007-2013
Thesis: Climate change and conceptual change	
RCME Fellow (Full support / stipend for 2 years)	
MS, Cognitive Neuroscience, MIT	1999-2002
Thesis: Neurocognitive circuitry supporting neoword learning	
NSF Graduate Student Fellowship	
Jacob Javits Fellowship (declined)	
Triple Bachelor's degree, *U of MD, College Park*	1995-1999
BA w/ honors in Linguistics, BS in Computer Science, BS w/ high honors in Mathematics	
Thesis: Modeling language change with markov models	
Magna Cum Laude; Banneker/Key Scholarship (Full support / stipend for 4 years); Arts and Humanities Senior Scholar; ODK leadership honors fraternity	

Skills

Programming – **Data Science:** Python, R, Spark, SQL, MongoDB, HDF5; **Visualization:** ggplot2, Matplotlib, Bokeh, SVG (D3); **Front End:** Elm, HTML, CSS, JavaScript / React, Jekyll; **Back End:** Ruby on Rails, Express, Python, Tomcat

Ops – Amazon EC2 & MTurk, Docker, Ansible, Packer, Vagrant, Linux admin, system building, GPU

Science – **Data Collection:** brain imaging, physiology, computerized testing (desktop / web / mobile), surveys / crowd-sourcing, experiment design; **Statistics:** classical, Bayesian, non-parametric, time-series, machine learning

Writing – technical, grants, documentation, policy / MOU

Other Awards

UC Berkeley SPOT Award, outstanding service to D-Lab	2016
Harvard Business Plan Competition, runner up, ($4000 in-kind services)	2003
MIT Sloan Business Plan Competition, semi-finalist	2003

Community & Service

Data and Mental Health Working Groups, School Climate Collaborative, Baltimore, MD	2016-2017
EdX Data Czar, UC Berkeley	
Software Carpentry / Data Carpentry, SF Bay Area	2014-2016
Discussant, Numerous panels on data collection, sharing and ethics, SF Bay Area	2014-2016
Lead, BIDS Collaborative for graduate student data science projects, UC Berkeley	2015
Resident, ManyLabs (a Moore-funded science hackerspace), San Francisco, CA	2015
Presenter, Bloomberg Data for Good Exchange, New York, NY	2015
Chair, social science panel at SciPy 2014, Austin, TX	2014
Training Assistant, Data Science for Social Good, Chicago, IL	2013
President, MIT Graduate Student Volunteer Corps, Cambridge, MA	2001-2002

Sample Open Source Projects

marketflow – efficient ingest of large financial datasets	http://marketflow.readthedocs.io/en/latest/
rpy2 – interoperate with R from Python	http://rpy2.readthedocs.io/
NiPype – data processing graphs for brain imaging analysis	http://nipype.readthedocs.io/
Hacking Measurement – incubating grad student projects	http://hackingmeasurement.berkeley.edu/
MTurk Admin – longitudinal experiments on Amazon MTurk	https://github.com/davclark/mturk_admin
Taichi Site – static site using Elm for video-based practice	https://github.com/davclark/taichi-site
Hotaru – controlling an LPD8806 LED strip using React.js	https://github.com/davclark/hotaru-material-ui