# Keith Wiley


**Contact**

12027 40th Ave NE
Seattle, WA 98125

505-615-4572

kwiley@keithwiley.com

http://keithwiley.com
https://github.com/kebwi


**Skills**

- *Programming*: Python, Java, C, C++ (*Familiarity*: Matlab/Octave)
- *Tools*: Eclipse, PyCharm, IPython Notebook/Jupyter, Xcode, Subversion (*Familiarity*: Git, Docker, Selenium, IntelliJ, Doxygen, CPPUnit)
- *MapReduce*: Hadoop/HDFS, Java MR, C++/Python Streaming MR, Hive-Python plugins, EMR
- *Cloud*: Tier 3, Hadoop-on-Azure, AWS/EMR/EC2/S3 (also direct-Hadoop-EC2 (nonEMR))
- *SQL/NoSQL*: Hive, Spark SQL (*Familiarity*: Cassandra, MongoDB, CouchDB, Riak, Redis, others)
- *Analytics*: scikit-learn, basic ML (decision trees, random forests), matplotlib (*Familiarity*: Splunk, Tableau)
- *APIs*: LinkedIn, Twitter, OpenLayers, general RESTful concepts
- *Web*: HTML, CSS, PHP, Perl/CGI, MySQL (*Familiarity*: Javascript, Open Street Map, OpenLayers)
- *Science*: Image (FFT, wavelet, coaddition) and Acoustic (spectrogram, octave-band) signal processing
- *Mobile Development*: Android, AndEngine

**Education**

| | | |
|---|---|---|
| Ph.D. Computer Science | University of New Mexico, Albuquerque | Jul 2006 |
| M.S. Computer Science | University of New Mexico, Albuquerque | Dec 2003 |
| B.A. Psychology | University of Maryland, College Park | Dec 1997 |

**Employment**

- *Atigeo, Aug 2013 – Sep 2016*

**Senior Software Engineer & Data Scientist/Engineer** – Developed Hive/MapReduce/Spark Python modules for ML & predictive analytics in Hadoop/Hive/Hue on AWS. Implemented a Python-based distributed random forest via Hive and Python streaming. Pipelined (ingest/clean/munge/transform) data for feature extraction toward downstream classification. ▶ Primary designer/developer of a scikit-learn based random forest & ensemble ML pipeline for cross-fold-validated predictive analytics, including insight via feature importance exposure. Performed statistical analysis and visualization of ML results via ROC curves & AUC. ▶ Onboarded multiple new hires to assist them in familiarizing with our vast and complex data pipeline and suite of databases and tables. ▶ Primary designer/developer of a pipeline that ingests/catalogs/stores/analyzes new datasets with final analytics/visualization. This project implemented an SOW whose completion was the keystone of a seven-figure contract.

- *Expedia via Slalom Consulting placement, Dec 2012 – Jul 2013*

**Big Data Engineer, Consultant** – Brief MongoDB project, then Hadoop/Hive on AWS, using EMR and nonEMR-Hadoop in EC2. Tasks: EC2-to-S3 data synch., Hive stand-up, AWS profiling. Accomplishments: Hadoop 2.0/YARN EC2 deployment. Amazon's own engineers were curious about my progress.

- *Slalom Consulting, Feb 2012 – Jul 2013*

**Big Data Engineer, Consultant** – *National Mobility* team (mgr. Jeff Rubingh), *National BI* team (mgr. Kevin Gregory), developing big data processing techniques. Focus: Hadoop MapReduce, Hive, Cloudera, Tier 3, Hadoop-on-Azure. Topics: CRM, NY MTA, Linked-In/Twitter APIs, some OpenLayers visualization.

- *University of Washington, Dept. of Astronomy, Feb 2010 – Jan 2012*

**Research Scientist IV** – LSST group (mgr. Andrew Connolly). Development of massively parallel image processing routines in Hadoop, namely image coaddition (multiple partially overlapping images are registered, stacked, and mosaiced together). Test dataset: SDSSDB (30TB, 4 million images), future applications to LSST (60PBs). Cluster (NSF CluE): 892 machines, 700TB storage, 3568 concurrent processes.

- *University of Washington, Applied Physics Lab, May 2007 – Feb 2010*

**Software Engineer IV** – Proj. 1: *Sonar Simulation Toolkit* (mgr. Robert Goddard), an eigenray model of underwater acoustics: Incorporation of external libraries, OO design, feature development, optimization/ performance-redesign, refactorization, unit-testing. Proj. 2: a real-time data-acquisition/FFT-processing system with low data-loss tolerances, rapid throughput, and amenability to future parallelism.

- *University of New Mexico, 1999 – 2007*

**Course Instructor** *(Jan 2007 – May 2007)* – CS241, Data structures/algorithms, taught in C.
**Graduate TAs and RAs** *(Sep 1999 – May 2006)* – taught 200–300-level C++ (6 semesters), various research.

- *The Institute for Genomic Research, Sep 1997 – Aug 1999*

**Software Developer** – C++ bioinformatics software development for DNA sequencing and closure analysis.

## Personal Projects

Sample only. Please see my website for a comprehensive listing and github for a few public disseminations.

*Image/Acoustic Signal Processing*

***Keith's Image Stacker:*** Multi-threaded (parallel) image stacking, Laplacian sharpening, wavelet denoising. Used by amateur astrophotographers, reviewed online and in *Astronomy* and *Sky & Telescope*.

***WildSpectra*** (collaboration: Dr. R. Haven Wiley, Biology dept, UNC-CH): Mac real-time spectrogram analyzer, used in Dr. Wiley's research lab and by researchers throughout the acoustic-biology community.

***Keith's iPod Photo Reader:*** Extracts images from iPod .ithmb image files. Implementation required reverse-engineering the image format from scratch.

*Data Analytics/ Dynamic Websites*

***Neuromorphic CM1K Emulator*** A Python emulator of General Vision's CM1K neuromorphic chip, including slides presenting modeling experiments. See personal website or github for more info.

***Movie Hurl*** (http://moviehurl.com): A Perl server-side website of user ratings of "shaky-cam" movies. Generic ratings: weighted average of anonymous submissions. Personalized predictions: correlated user-pair ratings across overlapping data subsets. See associated *New York Times* article under Publicity below.

*Android*

***Petri*** (game): Grow a cell culture in a Petri dish, fend off invasive cultures and phage outbreaks.

***WildSpectra Mobile:*** Real-time scrolling spectrograms (FFT and octave-band) on *Android* devices. Also: real-time waveform & FFT/octave spectrum, and post-recording editing/playback and file I/O.

***Shead Spreet:*** Spread sheet for *Android* devices with 300,000 installs, 8500 sales, and a 4.3/5 rating.

*Distributed Computing*

***Distributed Mandelbrot Set:*** Generates fractal images by farming job-segments to multiple computers. Networking coded from scratch using sockets. Automatic load-balancing ensures optimal performance.

*HCI*

***Druid*** (PhD thesis): Vector drawing program which permits interwoven surfaces (Celtic knots, Olympic rings, etc.) and which provides an isomorphic efficient user interface.

*Simulation*

Artificial life, evolutionary/genetic algorithms, cellular automata, robotics, flocking (please see my website).

*Web Design*

http://keithwiley.com, http://moviehurl.com, http://badlandswatches.com

## Positions, Publicity, Awards

- Movie Hurl (see Personal Projects above) was mentioned in a *New York Times* article at http://well.blogs.nytimes.com/2015/11/14/feeling-woozy-it-may-be-cyber-sickness, 2015.
- Numerous interviews & articles following my book's publication (see my website for links), 2014–2015.
- Advisor to the *Brain Preservatin Foundation*, 2014–present.
- Science Advisor to the *LifeBoat Foundation*, 2011–present.
- Proceedings chair for the *Computer Science at UNM Student Conference* committee, 2006.
- *Sky & Telescope* magazine. Software review: *Keith's Image Stacker* and *Keith's Astroimager*, Aug 2004.
- $1^{st}$ place in the *International Online ALife Contest*, *Cyberbotics Webots*, khepera robot sim., Jul 1999.

## Graduate Research (sample)

*Winter 2003–Summer 2006, Ph.D. thesis, UNM, C.S. Dept*
Design and implementation of *Druid* (see Personal Projects above).

*Spring 2001–Spring 2002, Autonomous Robotic Glider, UNM C.S. Dept/Sandia National Labs*
Use of genetic programming trees to evolve behavioral routines for autonomous robotic unpowered gliders.

## Publications

*Books*

*A Taxonomy and Metaphysics of Mind-Uploading.* Humanity+ Press and Alautun Press, 2014.

*Peer Reviewed*

The Fallacy of Favoring Gradual Replacement Mind Uploading Over Scan-and-Copy. *JoCS*, 2016.
Astronomy in the Cloud: Using MapReduce for Image Co-Addition. *PASP*, 2011.
Astronomical Image Processing with Hadoop. *ADASS*, 2010.
Parallel Distributed Image Coaddition with Hadoop. *Yahoo Hadoop Summit*, 2010.
Representation of Interwoven Surfaces in $2^1/_2$D Drawing. *IEEE CG&A*, 2006.
Representation of Interwoven Surfaces in $2^1/_2$D Drawing. *CHI*, 2006.

*Submitted or Under Review*

Mind Uploading and the Question of Life, the Universe, and Everything. 2015.
The Fermi Paradox, Self-Replicating Probes, and the Interstellar Transportation Bandwidth. 2011.

*Invited*

Long Exposure Webcams and Image Stacking Techniques. *The Art and Science of CCD Astronomy*, 2005.
Long Exposure Webcams and Image Stacking Techniques for the [...]. *Astronomy magazine*, 2003.
Pattern Evolver, An Evolutionary Algorithm that Solves [...]. *The Handbook of Genetic Algorithms*, 1999.

*Op-ed*

Mind Uploading and The Question of Life, the Universe, and Everything. *IEET*, 2015.
The Fallacy of Favoring Gradual Replacement Mind Uploading Over Scan-and-Copy. *IEET Magazine*, 2015.
'Interstellar' Might Depict AI Slavery. *H+ Magazine*, 2014.
Response to Susan Schneider's The Philosophy of 'Her'. *H+ Magazine*, 2014.
Implications of Computerized Intelligence on Interstellar Travel. *H+ Magazine*, 2011.