

# Ryan Reece, Ph.D.

Data scientist / machine learning scientist / physicist

ryan.reece@gmail.com

<http://rreece.github.io>

+1-817-888-7010

## OBJECTIVE

Looking for a challenging and rewarding opportunity in data science, with a focus on machine learning applications. Having completed a Ph.D. and postdoc in particle physics with the ATLAS experiment during the startup of the LHC and the discovery of the Higgs boson, I have [9+ years experience in data science](#), including techniques in data reduction, visualization, classification, statistical inference, and machine learning. I am passionate about using scientific techniques to solve important problems and about how technologies extend our reach.

## SKILLS

- **General:** data visualization, statistical analysis, data-driven modeling, anomaly detection, neural network classifiers, boosted decision trees, petabyte data reduction, object-oriented design, polymorphic interfaces, writing technical reports, working independently and in groups, presenting my ideas, graduate level physics and mathematics
- **Programming languages (fluent):** [C/C++/STL \(16+ years\)](#), [Python \(10+ years\)](#)
- **Data science software:** matplotlib, numpy, scipy, scikit-learn, pandas, ipython, [keras](#), [tensorflow](#), [BayesDB](#), ROOT, TMVA, HistFitter
- **General software:** Linux (Redhat/SLC/Ubuntu/Debian), bash, git, svn, SQL, UML, QT, Mathematica, Keynote/PowerPoint/Beamer
- **Markup languages:**  $\LaTeX$ , Markdown, (X)HTML with CSS, XML

## EXPERIENCE

**Postdoctoral Research Fellow** | July 2013 - August 2017

[Santa Cruz Institute for Particle Physics](#), The University of California, Santa Cruz, and  
The European Organization for Nuclear Research ([CERN](#)), Geneva, Switzerland

- I've spent 10 years, time as a postdoc and a Ph.D. student, as a member of the ATLAS experiment, a 3000+ person collaboration looking for new physics in high energy proton-proton collisions at the Large Hadron Collider (LHC).
- LHC data has a challenging rate ( $\sim 10$  PB/year) and requires distributed computing for the analysis. The ATLAS codebase has more than 10 million lines of C++ and almost as many lines of Python. It is not uncommon for an analysis project to depend directly on parts of the codebase that are order 100 thousand lines, or more.
- I have become an [expert in petabyte data reduction](#), high-throughput computing, world-wide grid computing, and [data visualization](#) as a primary user and supporter of our group's 200-CPU computing cluster, on which I have accumulated more than [350k CPU-hours](#).
- I have played leading roles in searches for signals of supersymmetry and other exotic decays. In two such projects, I was one of two "Editors" that serve as the [management leads of a team of 5-8 analyzers](#), eventually write the paper, and defend its approval—over a period of about a year.
- 2015-17, I moved to Geneva to be at CERN full-time. I supported the operations of the DAQ and detector monitoring systems of the SCT (a tracking sub-detector in ATLAS), taking on-call shifts for those systems.
- I have also served as a mentor for UCSC students, advising them in issues about analysis, computing, and living abroad.
- For most of 2017, I have been building [expertise in machine learning techniques](#), including deep learning frameworks and probabilistic databases. I have been doing research using Keras to build CNNs for particle classification, and in another project using BayesDB for anomaly detection.

**Graduate Researcher** | June 2006 - July 2013

[The University of Pennsylvania](#), Philadelphia, PA, and

The European Organization for Nuclear Research ([CERN](#)), Geneva, Switzerland

- I spent my first summers as a student with Penn (2006-08) at CERN, participating in the [integration and commissioning of custom electronics](#) for the Transition Radiation Tracker (TRT), the outermost sub-detector of the ATLAS tracker.
- 2009-12, throughout most of the running of the LHC I rotated with others the [on-call responsibility](#) for the TRT [data acquisition system \(DAQ\)](#).
- The focus of my Ph.D. research with the data from ATLAS has been on the identification of decays of tau leptons and their use in searches for new physics, basically a [pattern recognition](#) problem to identify a type of particle.
- 2009-10, I played leading roles in the development and commissioning of the cut-based tau identification used with the first ATLAS data.
- 2010-12, I helped develop a more advanced multivariate tau identification using [Boosted Decision Trees \(BDTs\)](#) which superseded the cut-based method as the ATLAS standard.
- I have a knack for developing data analysis frameworks. One of which, [pyframe](#), has been used by several analyses in ATLAS.

- The ATLAS and CMS experiments at the LHC [discovered the long-sought-after Higgs boson](#), evidence of which was announced on the 4th of July 2012.

## EDUCATION

- **Ph.D. Experimental Particle Physics**

The University of Pennsylvania (Philadelphia, PA), June 2006 - July 2013

thesis: "[A search for new physics in high-mass ditau events in the ATLAS detector](#)"

advisor: H.H. "Brig" Williams

- **B.S. Physics with Honors**

The University of Texas (Austin, TX), Aug 2003 - May 2006

thesis: "Late pulsing in the Hamamatsu R1408 PMT used in the Sudbury Neutrino Observatory"

advisor: Joshua R. Klein, GPA: 3.91/4, physics GPA: 3.94/4

## SHOWCASE

- More about my selected software projects here: <http://rreece.github.io/sw/>
- Download my thesis and publications here: <http://rreece.github.io/publications/>

Last updated: December 1, 2017