# Ryan Reece, Ph.D.

ryan.reece@gmail.com
http://rreece.github.io
https://www.linkedin.com/in/ryanreece
+1-817-888-7010

**Data scientist / machine learning scientist / physicist**
Palo Alto, CA

## EXPERIENCE

**Insight AI Fellow** | January 2018 - present
Insight Data Science, Palo Alto, CA

- 7 week fellowship: learned about data science and machine learning applications in a variety of business domains.
- Development of a cloud-based hyperparamter optimization platform: HYPR.AI, for automating the testing of many ML models in AWS/Paperspace containerized jobs.

**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 (~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.
- 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 to support the operations of the data acquisition system (DAQ) and detector monitoring systems of the SCT (a tracking sub-detector in ATLAS).
- 2017, I built more expertise in machine learning techniques, including deep learning frameworks and probabilistic databases. I did research into using Keras to build CNNs for particle classification, and another project using BayesDB and sklearn 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 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 of the cut-based tau identification used with the first ATLAS data.
- 2010-12, I helped develop advanced tau identification using Boosted Decision Trees (BDTs) which superseded the above.
- I have a knack for developing data analysis frameworks, 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 (Philadelpha, PA), June 2006 - July 2013
  thesis: "A search for new physics in high-mass ditau events in the ATLAS detector"
- **B.S. Physics with Honors**, The University of Texas (Austin, TX), August 2003 - May 2006
  thesis: "Late pulsing in the Hamamatsu R1408 PMT used in the Sudbury Neutrino Observatory"

## 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)
- **Programming languages (experienced):** javascript, SQL; **Markup languages:** LaTeX, Markdown, (x)html with css
- **Data science software:** matplotlib, numpy, scipy, scikit-learn, pandas, ipython, Keras, tensorflow, BayesDB, AWS, docker, ROOT, RooStats, TMVA
- **General software:** Linux (Redhat/SLC/Ubuntu/Debian), bash, git, svn, UML, QT, Mathematica, Keynote/PowerPoint