

# Pedro Rodriguez | Research Scientist

✉ me@pedro.ai • www.pedro.ai • Github: EntilZha, Twitter: @EntilZhaPR

## Education

### University of Maryland at College Park

*Ph.D. in Computer Science*

2015 - 2021

Advisor: Jordan Boyd-Graber, Thesis: Evaluating Machine Intelligence with Question Answering

### University of California at Berkeley, Computer Science

*Bachelor of Arts in Computer Science with coursework in Physics and Astrophysics*

2010 - 2014

## Research Experience

### Meta Reality Labs

*Research Scientist*

### Natural Language Processing, Multimodal

*April 2021 - Present*

- Research in improving metrics, datasets, and evaluation methodologies in multimodal retrieval.
- Research in using Item Response Theory to interpret NLP benchmark data.
- Contributed a task to Dynabench.org and am a co-organizer of the DADC workshop.

### University of Maryland at College Park

*Research Assistant to Jordan Boyd-Graber, PhD Candidate*

### Machine Learning, Deep Learning, NLP

*August 2015 - April 2021*

- Research experience in natural language processing, question answering, and evaluation methodologies.
- Created an AI for a trivia game (Quiz Bowl) that beat the best players in the world. <https://youtu.be/bYFqMINXayc>

### Google Research - Zurich

*Research Intern*

### Question Answering, Fact Verification, NLP

*Winter 2019/2020*

- Investigated feasibility of hard mining in neural retrieval models for fact verification (e.g., FEVER).
- Discovered that false negatives (an artifact of how the dataset was built) makes hard mining fail for FEVER model training.

### Facebook Research

*Research Intern*

### Conversational NLP

*Summer 2019*

- Conversational NLP research project on indirectly measuring "interestingness" of Wikipedia facts.
- Published internship project as an EMNLP paper, released open dataset, and open sourced code.

### Microsoft Research - Information and Data Sciences

*Research Intern*

### Machine Learning, Deep Learning, NLP

*Summer 2018*

- Used search logs to identify topical chains that kept users engaged.
- Built a conversational NLP system using topic chains to guide topic exploration.

**UC Berkeley AMPLab:** Undergraduate Research Assistant, Fall 2014: Optimized Gibbs LDA on Spark

**Boise State University Geophysics:** Cryosphere Scientist, Field Researcher, May 2013-October 2014

**UC Berkeley Center for Time Domain Informatics:** Undergraduate Researcher in Astrophysics, 2012

## Industry Experience

### Riot Games

*Data Science Intern*

### Machine Learning, NLP, Bayesian Modeling

*Summer 2017*

- Improved NLP model effectiveness for automatically sorting and/or answering 12,000 player support tickets per day.
- Created data labeling process for player support tickets to increase effectiveness and enable incremental future improvement.
- Used Machine Learning ablation studies, Bayesian modeling, and Monte Carlo simulations to understand the behavior of the new sportsmanlike play system called Honor 2.0. This also included querying over 56 million records with Spark SQL.

### Oracle Data Cloud/Datalogix

*Data Science Research Intern*

### Machine Learning, Data Engineering, DevOps

*Summer 2016*

- Architected and implemented ETL pipeline with Apache Spark/Airflow which ingests 200GB/day of advertising data.
- Improved throughput from AWS S3 to Apache Spark by 40-60x which resulted in substantial cost savings.

### Zillow Group/Trulia

*Data Scientist*

### Machine Learning, Crowdsourcing, DevOps

*January 2015 - July 2015*

- Designed and iteratively improved crowdsourcing experiments to produce rank order for machine learning generated photo albums. In doing so wrote an Amazon Mechanical Turk library which managed approximately 20,000 tasks.
- Architected and implemented infrastructure for internal data science APIs used in 100% of visits to trulia.com.

**Avalanche Safety Educator:** Volunteer for National Ski Patrol, Professionally in South America, 2010-2017

## Publications

---

- **Pedro Rodriguez** and Jordan Boyd-Graber. Evaluation Paradigms in Question Answering. EMNLP 2021.
- **Pedro Rodriguez**, Joe Barrow, Alexander Hoyle, John P. Lalor, Robin Jia, and Jordan Boyd-Graber. Evaluation Examples are not Equally Informative: How should that change NLP Leaderboards? ACL 2021.
- **Pedro Rodriguez**, Paul Crook, Seungwhan Moon, and Zhiguang Wang. Information Seeking in the Spirit of Learning: a Dataset for Conversational Curiosity. EMNLP 2020.
- **Pedro Rodriguez**, Shi Feng, Mohit Iyyer, He He, and Jordan Boyd-Graber. Quizbowl: The Case for Incremental Question Answering. ArXiv 2019.
- Tristan Thrush, Kushal Tirumala, Anmol Gupta, Max Bartolo, **Pedro Rodriguez**, Tariq Kane, William Gaviria Rojas, Peter Mottson, Adina Williams, Douwe Kiela. Dynatask: A Platform for Creating Dynamic AI Benchmark Tasks. Under review.
- Max Bartolo, Hannah Rose Kirk, Tristan Thrush, Katerina Margatina, **Pedro Rodriguez**, Mimansa Jaiswal, Pontus Stenetorp, Robin Jia, and Douwe Kiela. The First Workshop on Dynamic Adversarial Data Collection (DADC). To occur at NAACL 2022.
- Naveen Raman, **Pedro Rodriguez**, and Jordan Boyd-Graber. What more can Entity Linking do for Question Answering?. NeurIPS 2020 Workshop on Human and Model in the Loop Evaluation and Training Strategies 2020.
- Eric Wallace, **Pedro Rodriguez**, Shi Feng, and Jordan Boyd-Graber. Trick Me If You Can: Adversarial Writing of Trivia Challenge Questions. TACL 2019.
- Denis Peskov, Joe Barrow, **Pedro Rodriguez**, Graham Neubig, and Jordan Boyd-Graber. Mitigating Noisy Inputs for Question Answering. Interspeech 2019.
- Shi Feng, Eric Wallace, Alvin Grissom II, Mohit Iyyer, **Pedro Rodriguez**, and Jordan Boyd-Graber. Pathologies of Neural Models Make Interpretations Difficult. EMNLP 2018
- Jordan Boyd-Graber, Shi Feng, and **Pedro Rodriguez**. Human-Computer Question Answering: The Case for Quizbowl. The NIPS '17 Competition: Building Intelligent Systems, 2018
- Santiago Rodriguez and **Pedro Rodriguez**. Classification of snow anisotropic surfaces for mountainous terrain with mixed vegetative cover using optical and thermal spectra from Landsat-8. ISSW 2018.
- Santiago Rodriguez, **Pedro Rodriguez**, and Bryan Biggs. Ski Touring Bitacora - An innovative approach for trip planning, recording observations, and risk management while traveling in Avalanche terrain. ISSW 2018.
- **Pedro Rodriguez**, Santiago Rodriguez, and Julian Carielo. The Stability Wheel: An Intuitive and Didactic Decision-Making Framework. ISSW 2014.
- Santiago Rodriguez, Hans Peter Marshall, and **Pedro Rodriguez**. Applications of Low Cost and Low Power FMCW Radar in the Characterization of Dry Snow. ISSW 2014.
- **Pedro Rodriguez**, Joshua S. Bloom, and Joseph W. Richards. Cross Matching MACC and ROSAT: Characterization of Matching and Resulting Class Trends. AAS Abstracts 2013.

## Service

---

**ACL 2021:** Sub-committee Co-Chair for Diversity and Inclusion

**Reviewer For:** ACL ARR 2022, ACL ARR 2021, EMNLP 2021, ACL 2021, NAACL 2021, EMNLP 2020, ACL 2019

**Colorado Data Science Team:** Founding Member and former captain.

## Skills

---

**Tools:** Python Ecosystem (PyTorch, pandas...), SQL, Apache Spark, Scala/Java, HTML/CSS/JS, Tex, Rust

**Open Source:** Maintained open source projects for 7 years, help users fix issues and review/merge pull requests.

## Open Source Code

---

I have always been a strong advocate for open source software and release software whenever possible.

**Py-IRT:** Item Response Theory library written in Python and Pyro.

**ACL 2021 IRT Paper:** Source code for ACL 2021 IRT paper using Py-IRT.

**EMNLP 2020 Curiosity Paper:** Source code and dataset for EMNLP 2020 Curiosity paper.

**Dataset Viewer:** Dataset viewer for datasets I've published.

**FEVER Fact Verification System:** System developed during Google Research internship.

**Qanta.org:** An NSF funded project for advancing AI question answering by learning to play quiz bowl.

**PyFunctional:** Python library for functional pipelines, 7 years old, Over 2K Github stars.

**wikidata-rust:** Fast Wikidata dump parser written in Rust.

**Apache Spark:** Contributed pull requests that added SparkSQL primitives.

Last Updated: February 19, 2022