Zachary Hafen-Saavedra

PhD, Theoretical and Computational Astrophysics

z.hafen.saavedra@gmail.com || zhafen.github.io ||

Experience

McCue Prize Postdoctoral Fellow July 2020 - Present
University of California, Irvine Irvine, CA

- Developed a custom-built C++ backend wrapped in a user-friendly Python frontend to process and retrieve >200,000 scientific abstracts from the NASA astrophysics data system
- Employed natural language processing of scientific text to extract actionable quantitative trends
- Interfaced with modern APIs to select 10s of GB of data from >2 TB of remote data
- Led a cross-discipline, international collaboration to enable an analysis requiring expertise from observers, data analysts, and large- and small-scale simulators
- Led testing of a Bayesian parameter estimation procedure to identify viable datasets for value extraction
- Extended statistical software to enable resampling of multi-dimensional distributions
- Modernized forward-modeling software to draw on open source atomic spectra data
- Developed an analysis-management tool to automate parameter exploration
- Organized a meeting of key galaxy-shape experts to identify target measurements and simulations

GK-12 Graduate Fellow June 2014 - July 2020
Northwestern University Evanston, IL

- Created quantitative visualizations, educational explanatory visualizations, and award-winning artistic visualizations to communicate core messages
- Processed tens of TB of >20-dimensional data using high-performance-computing resources to reduce to <100 GB of highly-interpretable data
- Employed modern software development best practices (unit testing, version control, etc.) to maintain a broad suite of essential software
- Utilized and modified a C simulation code to generate >100,000-CPU-hour simulations of entire galaxies
- Developed public code for time-dependent analysis of fluid simulations to isolate driving behavior
- Developed software to analyze mapped data, including jagged and sparse array handling, multiple save formats, numpy and pandas utilities, and easy utilization of item attributes
- Utilized and modified viz software to create interactive 3D visualizations of complex interactions between merging and orbiting galaxies
- Led one of Chicago's first data-science education initiatives for high-school students to bring opportunities to >100 students primarily from underrepresented-in-STEM backgrounds
- Collaborated with a multidisciplinary range of scientists to publish (to date) 35 papers, 6 as a lead author

Skills

Techniques: data analysis (time-series, sparse, big data, data cleaning), natural language processing, code testing (unit, integration, functional, CI), simulations, frequentist/Bayesian statistics, multiscale/hierarchical modeling, analysis pipeline development, forward modeling/mock data, technical writing
Soft skills: technical leadership and management, public speaking, mentoring (technical, professional)
Tools: Python (numpy, pandas, scikit, matplotlib), C/C++, Unix/Bash, parallel computing, git, nltk

Education

Northwestern University May 2020
PhD, MS, Physics and Astronomy Evanston, IL
Computational and Theoretical Astrophysics specialization

University of Northern Colorado May 2014
BS, Physics, emphasis: Math Greeley, CO