

Andresa Rodrigues de Campos

Amazon
Applied Scientist
linkedin.com/in/andresa-campos

andesarodriguescampos@gmail.com
github.com/AndresaCampos
Phone: +1 (412) 304-5910

Education **Carnegie Mellon University, Pittsburgh, PA USA**

Ph.D., Physics, 2023 – Advisor: Scott Dodelson, Ph.D.

São Paulo State University, SP Brazil

M.S., Physics, 2018 – Advisor: Rogerio Rosenfeld, Ph.D.

Universidade Federal de São João del-Rei, MG Brazil

B.S., Physics (GPA: 9.23/10), Minor in Computational Physics, 2015

Dissertation *Dark Energy Science from 100 Million Galaxies: AI-Driven Analysis and Data-Intensive Techniques for Cosmological Discovery*

Research **Amazon**

Applied Scientist

November 2024 – present

- Designed and deployed a novel lossless prompt compression system for LLM-powered process mining, achieving 40% input token reduction and 30% latency improvement across 1.2M production events. Discovered that LLMs can learn compression dictionaries through in-context learning.
- Developed and deployed an ML classification system for property tax asset operations, improving accuracy from 64% to 95% and enabling automated classification for 92.6% of 180K+ annual asset additions.
- Built an AI-driven contract classification system using LLMs and chain-of-thought prompt engineering, achieving 88% accuracy across 48K contracts and increasing tax automation coverage from 35% to 95%+.
- Designed and deployed RAG pipelines for contract term extraction across multiple business domains, reducing manual review by 300+ hours/year with 91.6% accuracy.
- Designing a two-phase clustering algorithm combining cosine similarity with LLM validation for automated customer onboarding at scale. In progress.

Dark Energy Survey Collaboration

Analysis Team Lead

October 2019 – December 2023

- Developed a machine learning pipeline for supercomputer analysis, enhancing galaxy distance estimation from photometry in the largest galaxy catalog to date (100M galaxies), to improve understanding of dark energy.
- Employed self-organizing maps and Bayesian statistics to reduce galaxy distance estimation uncertainty by 66%, satisfying Department of Energy goals.
- Led a team of 5 scientists conducting hypothesis testing, consistently delivering high-quality results on schedule.

Carnegie Mellon University

Doctoral Researcher

September 2018 – December 2023

- Developed a statistical method for empirical model selection to address limitations in current modeling methods.
- Conducted comprehensive analysis identifying a 30% likelihood of inadequacy in a prevalent astrophysical model.
- Method was adopted by leading research collaborations (Hyper Suprime-Cam Subaru Telescope, Dark Energy Survey).

Teaching **Teaching Assistant – Carnegie Mellon University**

Physics II for Engineering and Physics Students, Spring 2019 & Fall 2018

Stars, Galaxies and the Universe, Fall 2018

Lecturer – Universidade Federal de São João del-Rei

Fundamentals of Physics I, Fall 2015

Introductory Experimental Physics, Fall 2015

Course Tutor – Universidade Federal de São João del-Rei

Calculus II, Fall 2012

Treatment and Representation of Exp. Measurements, Spring 2012 & Fall 2011

Publications Author or co-author of 79 peer-reviewed papers and preprints.

Total citations: 8,943; h-index: 41.

Profiles: ORCID | ADS Library | Google Scholar

First-author papers:

- **Andresa Rodrigues de Campos**, David Lee, Imry Kissos, Piyush Paritosh, “Loss-less Prompt Compression via Dictionary-Encoding and In-Context Learning: Enabling Cost-Effective LLM Analysis of Repetitive Data.” arXiv:2604.13066v1 (2026)
- **Campos, A.**; Yin, B.; Dodelson, S.; Amon, A.; Alarcon, A.; Sánchez, C.; Bernstein, G. M.; Giannini, G.; Myles, J.; Samuroff, S.; et al., “Enhancing weak lensing redshift distribution characterization by optimizing the Dark Energy Survey Self-Organizing Map Photo-z method.” arXiv:2408.00922 (2024).
- **Campos, Andresa**; Samuroff, Simon; Mandelbaum, Rachel, “An empirical approach to model selection: weak lensing and intrinsic alignments.” MNRAS, arXiv:2211.02800 (2023).

- Lemos, P.; Raveri, M.; **Campos, A.**; Park, Y.; Chang, C.; Weaverdyck, N.; Huterer, D.; Liddle, A. R.; Blazek, J.; Cawthon, R.; et al., “Assessing tension metrics with dark energy survey and Planck data.” MNRAS, arXiv:2012.09554 (2021) – Equal contributions from the first three authors.

Selected co-authored papers (by date, newest first):

- **DES Collaboration**; Abbott, T. M. C.; Adamow, M.; et al., “Dark Energy Survey Year 6 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing.” arXiv:2601.14559 (2026).
- **DES Collaboration**; Abbott, T. M. C.; Agüena, M.; et al., “Dark Energy Survey Year 6 Results: Cosmological Constraints from Cosmic Shear.” arXiv:2602.10065 (2026).
- Yin, B.; Amon, A.; **Campos, A.**; et al., “Dark Energy Survey Year 6 Results: Redshift Calibration of the Weak Lensing Source Galaxies.” arXiv:2510.23566 (2025).
- Samuroff, S.; **Campos, A.**; Porredon, A.; Blazek, J., “Joint constraints from cosmic shear, galaxy-galaxy lensing and galaxy clustering: internal tension as an indicator of intrinsic alignment modelling error.” Open J. Astrophys., arXiv:2402.15573 (2024).
- Samuroff, S.; Mandelbaum, R.; Blazek, J.; **Campos, A.**; MacCrann, N.; Zacharegkas, G.; Amon, A.; Prat, J.; Singh, S.; Elvin-Poole, J.; et al., “The Dark Energy Survey Year 3 and eBOSS: constraining galaxy intrinsic alignments across luminosity and colour space.” MNRAS, arXiv:2212.11319 (2023).
- Lemos, P.; Weaverdyck, N.; Rollins, R. P.; Muir, J.; Férté, A.; Liddle, A. R.; **Campos, A.**; Huterer, D.; Raveri, M.; Zuntz, J.; et al., “Robust sampling for weak lensing and clustering analyses with the Dark Energy Survey.” MNRAS, arXiv:2202.08233 (2023).
- Myles, J.; Gruen, D.; Amon, A.; Alarcon, A.; DeRose, J.; Everett, S.; Dodelson, S.; Bernstein, G. M.; **Campos, A.**; Harrison, I.; et al., “Mapping variations of redshift distributions with probability integral transforms.” MNRAS, arXiv:2210.03130 (2022).
- Secco, L. F.; Samuroff, S.; Krause, E.; Jain, B.; Blazek, J.; Raveri, M.; **Campos, A.**; Amon, A.; Chen, A.; Doux, C.; et al., “Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to modeling uncertainty.” Phys. Rev. D, arXiv:2105.13544 (2022).
- Breivik, K.; Connolly, A. J.; Ford, K. E. S.; Jurić, M.; Mandelbaum, R.; Miller, A. A.; Norman, D.; Olsen, K.; O’Mullane, W.; Price-Whelan, A.; et al., “From Data to Software to Science with the Rubin Observatory LSST.” arXiv:2208.02781 (2022).
- Myles, J.; Alarcon, A.; Amon, A.; Sánchez, C.; Everett, S.; DeRose, J.; McCullough, J.; Gruen, D.; Bernstein, G. M.; Troxel, M. A.; et al., “Dark Energy Survey Year 3 results: redshift calibration of the weak lensing source galaxies.” MNRAS, arXiv:2012.08566 (2021).
- Doux, C.; Baxter, E.; Lemos, P.; Chang, C.; Alarcon, A.; Amon, A.; **Campos, A.**; Choi, A.; Gatti, M.; Gruen, D.; et al., “Dark energy survey internal consistency tests of the joint cosmological probes analysis with posterior predictive distributions.” MNRAS, arXiv:2011.03410 (2021).

Invited Talks	From DES Y3 to Y6: SOMPZ’s Journey Towards Better Redshifts , Urbana-Champaign US, October 2023 Weak Lensing Photo-Z for DES Y6 , Cambridge UK, July 2022 Self-Organizing Maps , Duke NC, May 2022 SOMPZ for Lens and Source Galaxy Samples , Duke NC, May 2022 Empirical Model Selection , Duke NC, May 2022 Tension Metrics , Sussex UK, November 2019
Awards and	Best Poster Award , Amazon Internal Science Fair, 2025
Fellowships	Data Science for All Women Data science fellowship by Correlation One – Summer 2023 Dark Energy Survey Builder Two years of full-time equivalent work on infrastructure – Fall 2022 Dark Energy Survey Early Career Scientist Grant for participating in the Fall 2019 DES Collaboration Meeting Dark Energy Survey Observing Award Grant for observing for DES using the DECam at Cerro Tololo in Chile 7 nights in January 2019 – DES Year 6 CNPq Scholarship, São Paulo State University, SP, Brazil Graduate Research Fellowship, awarded in 2016 (2 years duration) FAPEMG Scholarship, UFSJ, MG, Brazil Undergraduate Research Fellowship, awarded in 2013 & 2014
Service to the Profession	Science Paper Reviewer, Amazon Internal Conferences Reviewed 12 papers across AMLC 2025, CSS 2026, and AMLC 2026, evaluating research for scientific rigor, novelty, and practical impact. McWilliams Software Development Series Co-organizer: Sessions with tutorials and talks related to software development and computing resources, 2021 – 2023
Students and Mentees	Boyan Yin : Undergraduate, Carnegie Mellon University Hannah Varekamp : Undergraduate, Carnegie Mellon University
Leadership	Dark Energy Survey Analysis Team Lead, Weak Lensing Redshift Distributions, 2020 – 2023 Analysis Team Co-Lead, Samplers and Tensions, 2018 – 2020
Collaborations	The Dark Energy Survey (DES) The LSST Dark Energy Science Collaboration (DESC)
Grants	Extracting Information from the Large Scale Structure of the Universe XSEDE Allocation: AST200006, 2021 (200,000 CPU hours) Co-lead and authored the proposal. PI: Scott Dodelson

Languages and Skills	Languages: English (fluent), Portuguese (native), Spanish (advanced) Programming Languages: Python (7+ yoe), Java, R, Bash, FORTRAN Machine Learning: PyTorch, Scikit-Learn, AutoGluon, TensorFlow, Keras NLP & LLMs: Prompt Engineering, RAG, Chain-of-Thought, In-Context Learning Data Analysis: Jupyter, NumPy, CuPy, SciPy, Pandas Data Visualization: Matplotlib, Seaborn, Plotly, Altair Markup Languages: L ^A T _E X, basic knowledge of HTML + CSS Databases: SQL, PostgreSQL High Performance Computing: Spark, PySpark, SLURM, MPI, mpi4py Tools: Git, GitHub, Linux
Outreach and Leadership	President – Tartan Salsa Student organization at Carnegie Mellon University promoting Latin American dances and culture. Led weekly classes averaging 80 attendees, grew membership by 70%, and secured over \$20,000 in funding over two years. 2021 – 2023
Relevant Coursework	Introduction to Machine Learning Neural Networks and Deep Learning in Science Modern Programming for Data Scientists Introduction to Large-Scale Computing in Science