# John F. Wu — CV

Email: jowu@stsci.edu

Website: jwuphysics.github.io

ORCID: 0000-0002-5077-881X

Space Telescope Science Institute

3700 San Martin Drive

Baltimore, MD 21218

**EDUCATION** Ph.D. in Physics and Astronomy Piscataway, NJ 2013 - 2019 Rutgers, The State University of New Jersey B.Sc. in Physics/Astrophysics, with MCS Honors Pittsburgh, PA Carnegie Mellon University 2009 - 2013 Appointments Assistant Astronomer Baltimore, MD 2022 - Present Space Telescope Science Institute Associate Research Scientist Baltimore, MD Center for Astrophysical Sciences, Johns Hopkins University 2022 - Present Department of Computer Science, Johns Hopkins University 2024 - Present Postdoctoral Researcher Baltimore, MD Space Telescope Science Institute 2020 - 2021 Center for Astrophysical Sciences, Johns Hopkins University 2019 - 2020Graduate Research Assistant Piscataway, NJ Rutgers, The State University of New Jersey 2013 - 2019 Undergraduate Research Assistant Pittsburgh, PA McWilliams Center for Cosmology, Carnegie Mellon University 2012 - 2013 Research Intern Pittsburgh, PA Carnegie Mellon University CyLab 2011 Professional Memberships 2015 - Present American Astronomical Society 2021 - Present International Astronomical Union Grants and Awards Maryland Academy of Sciences, Outstanding Young Scientist award, \$5,000 2024 NSF, Explore ACCESS GPU allocation, equivalent to \$1,386 2023 2022 STScI, Director's Discretionary Research Funding, \$68,110 Google, GCP Research Credits Program, \$5,000 2019 Rutgers, Robert A. Schommer Prize, \$500 2018 USAID, Research and Innovation Fellowship, \$11,636 2016 Rutgers, Special Study Award, \$1,350 2014 Rutgers, Claud Lovelace Graduate Fellowship & Excellence Fellowship Supplement, \$1,000 2013 CMU, Senior Leadership Recognition 2013

# Advising and Mentorship

Co-advisor, Sophia Rivera (JHU/Undergrad)	2024 - Present
Co-advisor, Nikhil Garuda (Arizona/Undergrad)	2024 - Present
Primary supervisor, Kiera McCormick (Loyola/STScI SASP + JSALT Intern)	2024
Primary supervisor, Christine Ye (Stanford/JHU JSALT Intern)	2024
Primary supervisor, Charles O'Neill (ANU/JHU JSALT Intern)	2024
Primary supervisor, Alina Hyk (Oregon State/JHU JSALT Intern)	2024
Co-advisor, Austin Larson (JHU/Masters in CS)	2024
Primary advisor, Mikaeel Yunus (JHU/PhD student in Physics and Astronomy)	2023 - Present
Primary advisor, Harish Krishnakumar (Princeton/Undergrad)	2022 - Present
Primary advisor, Phani Velicheti (Arizona/STScI Space Astronomy Summer Program)	2022
Primary advisor, Ziting Guo (Yale-NUS/Undergraduate capstone project)	2021 - 2022
Mentor, Kamonte Johnson (Frostburg State University/CollegeBound Foundation)	2020 - 2022
Co-advisor, Antoine Washington (Rutgers University/Undergraduate senior thesis)	2017 - 2020
Co-advisor, Marcell Howard (Case Western Reserve University/REU)	2018
Co-advisor, Manuel Perez III (University of Redlands/REU)	2017
Seminars and Talks (†Invited)	
†Seminar, Rutgers University Astronomy Seminar	2024
†Talk, HWO Generative AI Task Group	2024
†Talk, ESA Space Science and Machine Learning	2024
†Talk, NOIRLab, Rare Gems in Big Data	2024
†Talk, LSST-DA Informatics & Statistics Science Collaboration meeting	2024
Talk, Aspen Center for Physics, Diffuse Cosmic Backgrounds []	2024
†Colloquium, Swarthmore Physics & Astronomy, Colloquium	2024
†Seminar, Roman Virtual Lecture Series	2024
Poster, ICML, Machine Learning for Astrophysics	2023
Talk, STScI, Roman Science Inspired by Emerging JWST Results	2023
Poster, Flatiron CCA, Cosmic Connections Symposium	2023
†Talk, KITP, Galaxy Formation and Evolution in the Data Science Era	2023
†Seminar, University of Helsinki Astrophysics Seminar (virtual)	2022
†Seminar, Carnegie Observatories Lunch Talks	2022
†Seminar, UMBC Astrophysics Seminar	2022
†Special Talk, $STScI$	2021
Talk, JHU/STScI HotSci Series	2021
†Seminar, Université de Montréal, Astrophysics Seminar (virtual)	2021
$\dagger Seminar,\ University\ of\ Toronto,\ Statistics\ and\ Machine\ Learning\ Journal\ Club\ (virtual)$	2021
†Seminar, Western Sydney University, Machine Learning in Astronomy	2021
Seminar, Space Telescope Science Institute, Galaxies Journal Club	2021
†Seminar, Fermilab, Cosmic Physics Center Seminar (virtual)	2021
Poster, NeurIPS, Machine Learning for the Physical Sciences workshop (virtual)	2020
†Talk, NCSA – Accelerated Artificial Intelligence for Big-Data Experiments (virtual)	2020
Seminar NOIRLah Flash Seminar (virtual)	2020

+Comings Wayne State University Partials / Astro / Neelson Physics Comings (winters)	อกอก
†Seminar, Wayne State University, Particle/Astro/Nuclear Physics Seminar (virtual) †Talk, The ISM in the Era of Big Data (AAS 236, virtual)	2020 2020
Talk, JHU Astro Coffee	2020 2020
•	2020 2020
†Talk, Astronomers Turned Data Scientists Meeting (AAS 235)	2020 2020
Poster, AAS 235th Meeting	
Seminar, STScI, Science Coffee Seminar	2019
Seminar, JHU, CAS Wine & Cheese Seminar	2019
Deep learning workshop, MIAPP – Galaxy Evolution in a New Era of HI Surveys	2019
Talk, ESO — Nine Billion Years of Gas Evolution	2019
†Seminar, Rutgers Statistics, Foundations of Probability Seminar	2019
Dissertation talk, AAS 233rd Meeting	2019
Seminar, Princeton, Galread Seminar	2018
Seminar, Princeton, Data Science/COMPASS Seminar	2018
Seminar, University of Cape Town, Lunch Seminar	2018
Poster, École Normale Supérieure – Galaxy Evolution Across Time	2017
Talk, Princeton-Rutgers Extragalactic Science Day	2016
Talk, AAS 227th Meeting	2016
Seminar, Australian Astronomical Observatory, Colloquium	2015
SERVICE  Journal Reviewer for $AJ$ , $ApJ$ , $A&A$ , $JOSS$ , $MNRAS$ , and $PNAS$	
Machine Learning conference/workshop paper reviewer for $NeurIPS$ and $ICML$	
Member, NOIRLab Data Science Advisory Sub-committee	2022 - Present
Member, STScI Honors Committee	2024 – Present
Member, JHU/STScI Joint Colloquium Committee	2023 - 2024
Panel Moderator, NeurIPS, Machine Learning and the Physical Sciences	2023
Reviewer, NASA Astrophysics Data Analysis Program (ADAP)	2023
Reviewer, NASA Postdoctoral Program (NPP)	2023
Member, NRAO Science Review Panel (SRP)	2022 - 2023
Founder/Organizer, STScI Machine Learning Reading Group (MLRG)	2022 - 2023
Guest Editor, Annual Reviews of Astronomy and Astrophysics (ARA&A Vol. 63)	2023
Program Coordinator & Diversity Lead, KITP Program - galevo23	2023
Coordinator, STScI Science Staff Retreat	2022
Leveler, JWST Cycle 1 Panel	2021
Co-organizer, Low Density Universe (LDU) Meetings	2020 - 2021
STScI Liason, JHU Physics and Astronomy Postdocs + Research Scientists	2020
Session Chair, AAS 236 ISM-BIG meeting-in-meeting	2020
Co-organizer, JHU CAS Astro Coffee	2020
Co-organizer, Rutgers Gaia DR2 Hackathon	2018
Co-organizer, Rutgers SPS/RAS Astro Hack Sessions	2018
Webmaster, Rutgers Physics GSO and SSPAR	2014 - 2017
Time Allocation Committee Member, SALT 2015-2 Rutgers TAC	2014 - 2017
Local Organizing Committee Member, 2015 PHISCC Workshop	2015 2015
Organizer, Student Seminars in Physics and Astronomy at Rutgers (SSPAR)	2014 - 2015
Vice President, Rutgers Physics Graduate Student Organization (GSO)	2014 - 2015

# TEACHING AND OUTREACH

Guest Lecturer, NAACL and CLSP Summer School	2024
Podcast Guest Speaker, Where What If Becomes What's Next	2024
Guest Lecturer, Swarthmore College, Introduction to Radio Astronomy	2024
Ingenuity Speaker Series, Ingenuity Project (Baltimore Polytechnic Institute)	2023
Guest Lecturer, LSST Data Science Fellowship Program (DSFP)	2023
Speaker, Linda Hall Library - "How Do I Become an Astronomer?"	2023
Speaker, Astronomy on Tap Baltimore	2023
Podcast Guest Speaker, Times Higher Education (THE) Campus Podcast	2022
Guest Speaker, Marymount School of New York, Independent Science Research	2021
Teaching Assistant, STScI ML Office Hours	2021
Guest Lecturer, Rutgers Byrne Seminar: The Poetry of Astronomy	2016, 2019
Certificate, Seminar In Graduate Mentoring in Astronomy and Physics (SIGMA-P)	2018
Plenary Talk, Friends of Rutgers Astronomy	2017
Leadership Team, Parsons Community Outreach	2015 - 2016
Volunteer, Parsons Community Outreach	2013 - 2016
Teaching Assistant, Rutgers 343: Observational Radio Astronomy	2015
Public Talk, Rutgers Astronomical Society	2014
Certificate, Developing Educational Leaders among TAs in Physics (DELTA-P)	2013
Workshops and Other Experience (*Lead)	
*JSALT: Frederick Jelinek Memorial Summer Workshop  Evaluating Large Language Models for Research Astronomy	Baltimore, MD June – Aug 2024
Aspen Winter Meeting	Aspen, CC
Diffuse cosmic backgrounds and the low surface brightness universe	Mar 2024
Galaxy Formation Meeting (Biosphere 2) Wide-Field Spectroscopy meets Galaxy Formation Theory	Tucson, AZ  Mar 2023
*Kavli Institute of Theoretical Physics Building a Physical Understanding of Galaxy Evolution with Data-driven Astronomy	Santa Barbara, CA y Jan – Mar 2023
Pascal Institute The Self-Organized Star Formation Process	Paris, France Sept 2019
MIAPP Topical Workshop  Nine Billion Years of Gas Evolution	Munich, Germany  July 2019
USAID Research & Innovation Fellowship  Improving the LADUMA Pipeline Using MeerKAT Early Science Data  Cape	e Town, South Africa Sept – Nov 2016
SKA Pathfinders HI Science Coordination Committee 2015 PHISCC Workshop: HI Surveys Get Real	Piscataway, NJ Mar 2015
Vatican Observatory Summer School  Galaxies, Near and Far, Young and Old	astel Gandolfo, Italy  June 2014
NRAO Synthesis Imaging Workshop  14th Synthesis Imaging Workshop	Socorro, NM May 2014

### PROFESSIONAL COLLABORATIONS

4MOST Wide Area Vista Extragalactic Survey (WAVES/ORCHIDSS): Member

ALMA Lensing Cluster Survey (ALCS): Member

Dark Energy Spectroscopic Instrument (DESI): External Collaborator (LOWZ Program)

DECam Local Volume Exploration (DELVE): WIDE Survey WG Member

Deep Skies Lab: Contributor

HWO AI/ML WG: Data Processing; Gen AI; Mission Ops Task Forces

LADUMA: Pipeline & Calibration WG, Source-finding WG, and Ancillary Data WG Member

LSST-DA Galaxies Science Collaboration: Member

LSST-DA Informatics & Statistics Science Collaboration: Member

UniverseTBD: Member

### Telescope Observing Proposals

Anglo-Australian Telescope (AAT)	
CoI, N0331 (5 nights), N0334 (4 nights)	2015, 2017
Atacama Large Millimeter/submillimeter Array (ALMA)	
PI, one proposal (9.4 hrs – partially observed in Cycle 7)	Cycles 7, 8
CoI, four proposals (120.9 hrs)	Cycles 2, 6, 7, 8
CTIO Blanco/DECam	
CoI, 2023B-646244 (54 nights)	2023B
Gemini South/Flamingo2	
CoI, Fast Turnaround (5.8 hrs)	2021B
Green Bank Telescope	
CoI, 24B-384 (60 hrs)	2024B
Southern African Large Telescope (SALT)	
PI, 2016-1-SCI-040 (3.9 hrs), 2015-2-SCI-052 (3.9 hrs), DDT (1.6 hrs)	2015-1 - 2016-1
CoI, 2017-1-MLT-014 (11.3 hrs), 2016-2-SCI-051 (20.4 hrs)	2016-2 — 2017-1
Very Large Array (VLA)	
CoI, 19A-433 (10 hrs)	2019A

# John F. Wu — Publication List

Names of students for whom I am the primary supervisor are underlined. For an up-to-date list of my publications, please see my ADS Library or my ORCID.

#### JOURNAL ARTICLES - FIRST AUTHOR AND MAJOR CONTRIBUTIONS

- 15. Learning Galaxy Astrophysics from Interpretable Sparse Feature Networks Wu, J. F., ApJL, submitted.
- 14. How the Galaxy-Halo Connection Depends on Large-Scale Environment Wu, J. F., Jespersen, C., Wechsler, R. H., 2024, ApJ, in press. arXiv:2402.07995.
- 13. Deep Learning Cosmic Ray Transport from Density Maps of Simulated, Turbulent Gas Bustard, C., Wu, J. F., 2024, MLS&T, 5, 1, 015028.
- Target Selection and Sample Characterization for the DESI LOW-Z Secondary Target Program
   Darragh-Ford, E., Wu, J. F., Mao, Y.-Y., Wechsler, R. H., et al. 2023, ApJ, 954, 149.
- Identification of Galaxy-Galaxy Strong Lens Candidates in the DECam Local Volume Exploration Survey Using Machine Learning
   Zaborowski, E., Drlica-Wagner, A., Ashmead, F., Wu, J. F., et al., 2023, ApJ, 954, 68.
- 10. Quantifying Roman WFI Dark Images with the Wavelet Scattering Transform Velicheti, P. D., Wu, J. F., Petric, A. O., 2023, PASP, 135, 1050.
- Identification of galaxy shreds in large photometric catalogs using Convolutional Neural Networks
   Teodoro, E. M., Peek, J. E. G., Wu, J. F., 2023, AJ, 165, 123.
- 8. A Machine Learning Approach to Enhancing eROSITA Observations Soltis, J., Ntampaka, M., Wu, J. F., et al., 2022, ApJ, 940, 60.
- 7. Extending the SAGA Survey (xSAGA). I. Satellite Radial Profiles as a Function of Host-galaxy Properties
  - Wu, J. F., Peek, J. E. G., Tollerud, E. J., et al, 2022, ApJ, 927, 121.
- 6. Predicting the Spectrum of UGC 2885, Rubin's Galaxy with Machine Learning Holwerda, B. W., Wu, J. F., Keel, W. C., Young, J., et al., 2021, ApJ, 914, 142.
- Connecting Optical Morphology, Environment, and HI Mass Fraction for Low-Redshift Galaxies Using Deep Learning
   Wu, J. F., 2020, ApJ, 900, 148.
- 4. The Star-Forming Interstellar Medium of Lyman Break Galaxy Analogs Wu, J. F., Baker, A. J., Heckman, T. M., et al., 2019, ApJ, 887, 251.
- 3. Using convolutional neural networks to predict galaxy metallicity from three-colour images Wu, J. F., Boada, S., 2019, MNRAS, 484, 4683.
- 2. Herschel and ALMA Observations of Massive SZE-selected Clusters Wu, J. F., Aguirre, P., Baker, A. J., et al., 2018, ApJ, 853, 195.

1. Galaxy Candidates at  $z \sim 10$  in Archival Data from the Brightest of Reionizing Galaxies (BORG[z8]) Survey

Bernard, S. R., Carrasco, D., Trenti, M., Oesch, P. A., Wu, J. F., et al., 2016, ApJ, 827, 76.

#### PEER-REVIEWED MACHINE LEARNING PAPERS

NOTE: MACHINE LEARNING CONFERENCE PAPERS ARE RIGOROUSLY REFEREED BY MULTIPLE REVIEWERS AND ARE CONSIDERED EQUIVALENT TO JOURNAL PUBLICATIONS.

- The Multimodal Universe: Enabling Large-Scale Machine Learning with 70TBs of Astronomical Scientific Data
   Angeloudi, E., Angeloudi, J., Bowles, M., et al., 2024, NeurIPS, accepted.
- Towards Interpretable Scientific Foundation Models: Sparse Autoencoders for Disentangling Dense Embeddings of Scientific Concepts
   O'Neill, C., Ye, C., Iyer, K., Wu, J. F., 2024, NeurIPS: FM4Science workshop, accepted.
- 9. Sparse autoencoders for dense text embeddings reveal hierarchical feature sub-structure O'Neill, C., Ye, C., Wu, J. F., Iyer, K., 2024, NeurIPS: Sci4DL workshop, accepted.
- 8. Estimating Dark Matter Halo Masses in Simulated Galaxy Clusters with Graph Neural Networks
  Garuda, N., Wu, J. F., Nelson, D., Pillepich, A., 2024, NeurIPS: ML4PS workshop.
- Conditional Diffusion Models for Generating Images of SDSS-Like Galaxies
   Yunus, M., Wu, J. F., Heckman, T. M., Holwerda, B. W., 2024, ICML: AI4Science workshop.
- 6. Steering semantic search with interpretable features from sparse autoencoders O'Neill, C., Ye, C., **Wu**, **J. F.**, Iyer, K., 2024, NeurIPS: MINT workshop, accepted.
- 5. Predicting dark matter halo masses from simulated galaxy images and environments Larson, A., Wu, J. F., Jones, C., 2024, ICML: AI4Science workshop.
- 4. Disentangling Dense Embeddings with Sparse Autoencoders O'Neill, C., Ye, C., Iyer, K., Wu, J. F., 2024, TMLR, submitted.
- 3. Learning the galaxy-environment connection with graph neural networks Wu, J. F., Jespersen, C., 2023, ICML: ML4astro workshop, 1, arXiv:2306.12327.
- 2. Identifying AGN host galaxies with convolutional neural networks
  Guo, Z., Wu, J. F., Sharon, C. E., 2022, NeurIPS: ML4PS workshop, 63, arXiv:2212.07881.
- 1. Predicting galaxy spectra from images with hybrid convolutional neural networks Wu, J. F., Peek, J. E. G., 2020, NeurIPS: ML4PS workshop, 3, arXiv:2009.12318.

#### Journal Articles - Surveys and Collaborations

13. ALMA Lensing Cluster Survey: Physical characterization of near-infrared-dark intrinsically faint ALMA sources at z=2-4 Tsujita, A., Kohno, K., Huang, S., et al. 2024, arXiv::2406.09890

- 12. PHANGS-ML: Dissecting Multiphase Gas and Dust in Nearby Galaxies Using Machine Learning
  - Baron, D., Sandstrom, K. M., Rosolowsky, E., et al. 2024, 968, 24.
- 11. The cold interstellar medium of a normal sub- $L^*$  galaxy at the end of reionization Valentino, F., Fujimoto, S., Giménez-Arteaga, C., et al. 2024, A&A, 485, 138.
- Katachi: Decoding the Imprints of Past Star Formation on Present Day Morphology in Galaxies with Interpretable CNNs Alfonzo, J. P., Iyer, K. G., Akiyama, M., et al. 2024, ApJ, 967, 152.
- 9. The SAGA Survey. V. Modeling Satellite Systems around Milky Way-mass Galaxies with Updated UniverseMachine
  Wang, Y., Nadler, E. O., Mao, Y.-Y., et al. 2024, ApJ, accepted, arXiv:2404.14500
- 8. The SAGA Survey. IV. The Star Formation Properties of 101 Satellite Systems around Milky Way-mass Galaxies
  Geha, M., Mao, Y.-Y., Wechsler, R. H., et al. 2024, ApJ, accepted, arXiv:2404.14499
- 7. The SAGA Survey. III. A Census of 101 Satellite Systems around Milky Way-mass Galaxies Mao, Y.-Y., Geha, M., Wechsler, R. H., et al. 2024, ApJ, accepted, arXiv:2404.14498
- JWST constraints on the UV luminosity density at cosmic dawn: implications for 21-cm cosmology Hassan, S., Lovell, C. C., Madau, P., et al. 2023, ApJL, 958, 3.
- A variable active galactic nucleus at z = 2.06 triply-imaged by the galaxy cluster MACS J0035.4-2015
   Furtak, L., Mainali, R., Zitrin, A., et al. 2023, MNRAS, 522, 5142.
- 4. The DECam Local Volume Exploration Survey Data Release 2
  Drlica-Wagner, A., Ferguson, P. S., Adamów, M., et al. 2022, ApJS, 261, 38.
- 3. LADUMA: The First Untargeted Detection of an OH Megamaser at z > 0.5 Glowacki, M., Collier, J. D., Kazemi-Moridani, A., et al., 2022, ApJL, 931, 7.
- 2. The DECam Local Volume Exploration Survey: Overview and First Data Release Drlica-Wagner, A., Carlin, J. L., Nidever, D. L., et al., 2021, ApJS, 256, 2.
- ALMA Lensing Cluster Survey: an ALMA galaxy signposting a MUSE galaxy group at z=4.3 behind "El Gordo"
   Caputi, K. I., Caminha, G. B., Fujimoto, S., et al., 2021, ApJ, 908, 146.

#### STSCI TECHNICAL REPORTS

- 2. The Roman Data Monitoring Tool Wu, J. F., Otor, O. J., Schultz, W., 2024, Roman Technical Report, under review.
- Simulating Cosmic Rays for the Roman Wide Field Instrument
   Wu, J. F., Sanchez, J., Casertano, S., Desjardins, T., 2023, Roman Technical Report,
   Roman-STScI-000502.

1. Effective use of machine learning to empower your research Wu, J. F., 2022, Times Higher Education – Campus, Feature Article.

## UNREFEREED CONFERENCE PAPERS AND WHITE PAPERS

- 8. Designing an Evaluation Framework for Large Language Models in Astronomy Research Wu, J. F., Hyk, A., McCormick, K., Ye, C., et al. 2024, ICML: AI4Science workshop, submitted, arXiv:2405.20389.
- Constructing Impactful Machine Learning Research for Astronomy: Best Practices for Researchers and Reviewers Huppenkothen, D., Ntampaka, M., Ho, M., Fouesneau, M., et al. 2023, arXiv:2310.12528.
- 6. NANCY: Next-generation All-sky Near-infrared Community surveY
  Han, J. J.; Dey, A.; Price-Whelan, A. M. et al., 2023, Roman Core Community Survey
  White Papers.
- 5. Optical, Radio Continuum and HI Deep Spectroscopic Survey (ORCHIDSS) Duncan, K., Baker, A., Best, P., et al., The Messenger, 190, 25.
- 4. Roman Ultra Deep Field Koekemoer, A. M. et al. 2021, Roman Early-Definition Astrophysics Survey Opportunity.
- 3. Obscured AGN Hiding High Growth at the Cosmic Noon Petric, A. et al. 2021, Roman Early-Definition Astrophysics Survey Opportunity.
- 2. Herschel And ALMA Observations Of The ISM In Massive High-Redshift Galaxy Clusters Wu, J. F. et al. 2017, Galaxy Evolution Across Time, 51.
- 1. LADUMA: Looking at the Distant Universe with the MeerKAT Array
  Blyth, S. et al. 2016, Proceedings of MeerKAT Science: On the Pathway to the SKA, 4.