Draft version July 1, 2026
Typeset using LaTeX **RNAAS** style in AASTeX631

arXiv:2606.30855v1 [astro-ph.IM] 29 Jun 2026

Deep Learning for Astrophysics: An Open Textbook from the NASA Cosmic Origins AI/ML Science and Technology Interest Group

Yuan-Sen Ting,[1,2,3,*] Digvijay Wadekar,[4,*] Phill Cargile,[5] Carol Cuesta-Lazaro,[6,7,†] André Curtis-Trudel,[8] Gregory Green,[9] Ryan McClelland,[10] Daniel Muthukrishna,[11,5] Tri Nguyen,[12,13] Helen Qu,[14] Tomasz Rozanski,[15] Anna Scaife,[16] Jesse Thaler,[17,18] Licia Verde,[19] Francisco Villaescusa-Navarro,[14] John F. Wu,[20] Duo Xu,[21] Siyu Yao,[22] Alex Gagliano,[18,†] Siddharth Mishra-Sharma,[23,†] Andrew K. Saydjari,[24,‡] Georgios Valogiannis,[25,†] Peter Kurczynski,[10,§] and Swara Ravindranath[26,10,¶]

[1] *Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA*
[2] *Center for Cosmology and AstroParticle Physics (CCAPP), The Ohio State University, Columbus, OH 43210, USA*
[3] *Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[4] *Center for Gravitational Physics, Department of Physics, The University of Texas at Austin, Austin, TX 78712, USA*
[5] *Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
[6] *Institute for Advanced Study, Princeton, NJ 08540, USA*
[7] *Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA*
[8] *Department of Philosophy, University of Cincinnati, Cincinnati, OH 45221, USA*
[9] *School of Science, Westlake University, Hangzhou, Zhejiang 310030, China*
[10] *NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA*
[11] *MIT Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[12] *Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Northwestern University, Evanston, IL 60201, USA*
[13] *Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada*
[14] *Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA*
[15] *Research School of Astronomy and Astrophysics, The Australian National University, Canberra, ACT 2611, Australia*
[16] *Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, UK*
[17] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[18] *The NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI), Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[19] *ICREA and Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain*
[20] *Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[21] *Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada*
[22] *Department of Philosophy, School of Humanities, Shanghai Jiao Tong University, Shanghai, China*
[23] *Faculty of Computing and Data Sciences and Department of Physics, Boston University, Boston, MA 02215, USA*
[24] *Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*
[25] *Department of Astronomy and Astrophysics and Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[26] *Center for Research and Exploration in Space Science and Technology II (CRESST II), The Catholic University of America, Washington, DC 20064, USA*

ABSTRACT

The astronomical community's ability to use modern machine learning shapes the science return of upcoming facilities. Recent community assessments single out education as the principal barrier to adoption, because what limits uptake is the uneven understanding of these methods rather than their availability. The NASA Cosmic Origins Artificial Intelligence and Machine Learning Science and Technology Interest Group (AI/ML STIG) was formed to address this gap through short, domain-specific tutorials and community discussion. We present *Deep Learning for Astrophysics*, a freely available textbook at https://deeplearning4astro.com, curated from the group's lecture series. It collects 23 chapters across six parts from 17 lecturers, running from computational foundations and deep-learning architectures through generative modeling, simulation-based inference, and reinforcement learning to large-language-model agents, and closing with the practice of AI-laden science. Many chapters include executable notebooks. We also outline the group's plan for the coming year, centered on agentic research and on NASA's ASTRA mission-concept initiative.

THE GAP IS ADOPTION, NOT AVAILABILITY

The scientific case for machine learning in astronomy is established. Data from the Vera C. Rubin Observatory and the Nancy Grace Roman Space Telescope will strain traditional analysis, and learned methods already aid classification, inference, and instrument operation. Uptake remains the limiting problem. The NSF AI+MPS community paper (Ferguson et al. 2026) sets out strategic priorities for the mathematical and physical sciences, one of them to "foster education and workforce development in AI." It names an undertrained community, distrust of opaque methods, and resources written by and for computer scientists as barriers, and proposes modular, stackable training in response.

A review of deep learning in astrophysics (Ting 2026) makes a complementary point. The value of these methods rests on understanding why they work, on architectural choices that encode physical assumptions and on testability, rather than on predictive accuracy alone, and progress depends on teaching deep learning as an extension of statistical practice rather than a replacement for it. Understanding is the prerequisite for trust, and trust for adoption. This matters because AI can produce an appearance of understanding without the substance of it (Messeri & Crockett 2024). The community paper also flags a feedback loop: undertraining can lead to work with weak methodological grounding, which hardens the skepticism it provokes. Raising literacy can improve both the uptake and the quality of AI in the field.

Literacy matters, too, for those who never train a model. Astronomers referee papers, sit on panels, and edit journals, and as AI-generated work enters the literature, understanding the methods is what lets them separate scientific insight from plausible output. The field needs researchers able to judge the work, not only to produce it. Both assessments point to the same response: domain scientists, not only computer scientists, should build AI training tailored to their field, with domain-relevant examples and criteria for when a learned method is the right tool.

THE INTEREST GROUP

NASA's Cosmic Origins program supports the community studying how galaxies, stars, and cosmic structure form and evolve, with missions and concepts spanning the Hubble Space Telescope and the future Habitable Worlds Observatory. The AI/ML STIG is a Science and Technology Interest Group of its Program Analysis Group (COPAG), open to the national and international community.[1] Its premise is that training the community can increase the science return of NASA's missions, and that the training must be domain-specific and time-efficient.

The design responds to obstacles named in the community paper and the group's founding proposal. These include resistance to AI among established researchers, the difficulty of assessing interdisciplinary skill, faculty time constraints that make full courses impractical, fragmented resources that lack astronomical context, and a pace of change that outruns standard curricula, with LLM agents the current example. The response is a training sequence assembled from stackable short modules, each covering a self-contained competency. The aim is not to turn astronomers into machine-learning engineers but to help researchers connect astronomical questions with machine-learning methods. To that end, the group runs short tutorials on individual methods, with time for community questions and concerns about reliability. A leadership council of seven organizes the activities.[2] As a COPAG STIG, the group also reports community needs and technology gaps through COPAG channels.

THE TEXTBOOK

Over the past academic year the group ran an online lecture series.[3] We have curated these lectures into a single textbook, *Deep Learning for Astrophysics*, available online.[4] The lectures, notebooks, and slides were edited into a consistent chapter format and reviewed by the contributing lecturers. The executable notebooks and their outputs are preserved, so readers can run and modify the methods rather than only read about them (Figure 1). The exercises use real astronomical data, such as galaxy images, radio-galaxy morphologies, transient light curves, and stellar streams, rather than toy examples.

* Co-chair, NASA Cosmic Origins AI/ML STIG
† Leadership Council, NASA Cosmic Origins AI/ML STIG
‡ Hubble Fellow
Leadership Council, NASA Cosmic Origins AI/ML STIG
§ Chief Scientist, NASA Cosmic Origins Program Office
¶ Deputy Chief Scientist, NASA Cosmic Origins Program Office

[1] NASA AI/ML STIG home page at https://science.nasa.gov/astrophysics/programs/cosmic-origins/community/artificial-intelligence-machine-learning-science-technology-interest-group-ai-ml-stig/.

[2] The author list is ordered as co-chairs, then lecturers, then the leadership council, then the Cosmic Origins program officers, with each group alphabetical by surname. Yuan-Sen Ting is both a co-chair and a lecturer, and Carol Cuesta-Lazaro is both a lecturer and a member of the leadership council.

[3] Lecture series, recordings, and source materials at https://tingyuansen.github.io/NASA_AI_ML_STIG/.

[4] https://deeplearning4astro.com.

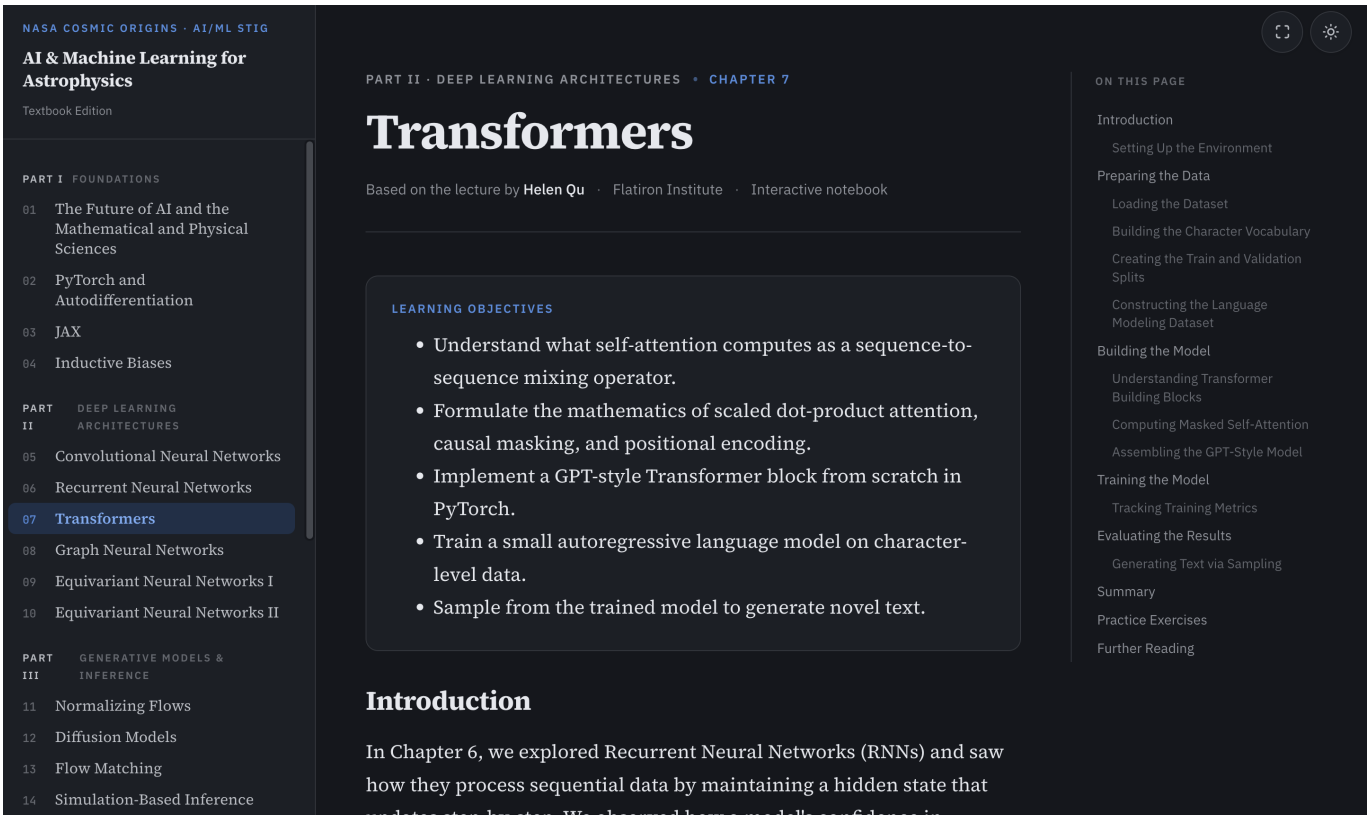

Figure 1. The web version of *Deep Learning for Astrophysics*. The textbook turns the AI/ML STIG lecture series into modular, domain-specific chapters that give astronomers practical entry points into modern machine-learning methods.

Each notebook chapter pairs the minimum theory a method needs with a worked astronomical example, so that its uses and its failure modes are learned together. Of the 23 chapters, seventeen are executable notebooks and six are concise readings, grouped into six parts. The intended reader is a graduate student or researcher who wants working knowledge rather than a survey. Because the parts are self-contained, the book can be read end to end or used as standalone modules in a course. The broader lecture series also included sessions on the AI funding landscape and on AI-enabled mission concepts; the textbook focuses on the technical and interpretive material most suited to chapter format.

Part I. Foundations

- The Future of AI and the Mathematical and Physical Sciences *(Jesse Thaler)*
- PyTorch and Autodifferentiation *(Phill Cargile)*
- JAX *(Phill Cargile)*
- Inductive Biases *(John F. Wu)*

Part II. Deep Learning Architectures

- Convolutional Neural Networks *(John F. Wu)*
- Recurrent Neural Networks *(Daniel Muthukrishna)*
- Transformers *(Helen Qu)*
- Graph Neural Networks *(Tri Nguyen)*
- Equivariant Neural Networks *(Anna Scaife)*

Part III. Generative Models & Inference

- Normalizing Flows *(Gregory Green)*
- Diffusion Models *(Duo Xu)*
- Flow Matching *(Tomasz Rozanski)*
- Simulation-Based Inference *(Tomasz Rozanski)*

Part IV. Reinforcement Learning

- Reinforcement Learning *(Carol Cuesta-Lazaro)*

Part V. Large Language Models & Agents

- LLM API Basics *(Yuan-Sen Ting)*
- RAG and Function Tools *(Yuan-Sen Ting)*
- Model Context Protocol *(Yuan-Sen Ting)*
- LLM as Agent *(Francisco Villaescusa-Navarro)*

Part VI. AI, Science & Society

- From Text to Spaceship *(Ryan McClelland)*
- AI and Scientific Publishing *(Licia Verde)*
- Understanding in AI-Laden Science *(Siyu Yao & André Curtis-Trudel)*

OUTLOOK

The group's focus for the coming year is agentic research, large-language-model systems that plan, execute, and revise analyses, and their use in scientific workflows. These systems are well suited to tasks where the goal can be stated and the result checked, rather than to open-ended discovery. This is where the book's closing question, what understanding and verification mean once AI enters the analysis, becomes operational. NASA's ASTRA initiative brings this question into mission-concept development. It invites the community to develop new astrophysics mission concepts,[5] and a concept that an agent helps shape requires human review against reasoning and requirements.

This emphasis on verifiability guides the other directions we expect to pursue. Generative design of instruments is one, where each candidate can be evaluated by simulation, as the textbook's chapter on generative design for spacecraft illustrates. Another is the constrained scheduling of facilities such as the Vera C. Rubin Observatory, the Nancy Grace Roman Space Telescope, and a future Habitable Worlds Observatory, where a proposed strategy can be scored before it is run. The work for the year ahead is to translate that standard into practice, building the tools and habits that let an agent's output be audited rather than accepted without verification.

These directions will shape the next lecture series. The textbook will continue to incorporate corrections and contributions from the community.

The textbook was produced by editing the original lecture materials into a consistent chapter format with substantial assistance from large language models (Anthropic's Claude Opus 4.8 and OpenAI's GPT-5.5), which were also used to copy-edit this note. In both cases the text, references, and affiliations were checked and verified by the authors. We thank the Cosmic Origins Program Office and the COPAG Executive Committee for their support, and the community members whose questions shaped the lectures.

REFERENCES

Ferguson, A., LaFleur, M., Ruthotto, L., et al. 2026, Machine Learning: Science and Technology, 7, 023001, doi: 10.1088/2632-2153/ae3e4e

Messeri, L., & Crockett, M. J. 2024, Nature, 627, 49, doi: 10.1038/s41586-024-07146-0

Ting, Y.-S. 2026, Annual Review of Astronomy and Astrophysics, 64, doi: 10.1146/annurev-astro-051024-021708

[5] The ASTRA Initiative, https://science.nasa.gov/astrophysics/programs/cosmic-origins/studies/astra-initiative/.