Check for updates

Received: 5 July 2023 | Accepted: 6 October 2023

DOI: 10.1002/aaai.12159

SPECIAL TOPIC ARTICLE

AI Institute in Dynamic Systems: Developing machine learning and AI tools for scientific discovery, engineering design, and data-driven control

J. Nathan Kutz[1] | **Steven L. Brunton**[1] | **Krithika Manohar**[1] | **Hod Lipson**[2] | **Na Li**[3]

[1]University of Washington, Seattle, Washington, USA

[2]Columbia University, New York, New York, USA

[3]Harvard University, Cambridge, Massachusetts, USA

Correspondence
J. Nathan Kutz, University of Washington, Seattle, WA 98195, USA.
Email: kutz@uw.edu

Funding information
United States National Science Foundation, Directorate of Engineering, Grant/Award Number: 2112085

Abstract

The mission of the AI Institute in Dynamic Systems is to develop the next generation of advanced machine learning (ML) and AI tools for controlling complex physical systems by discovering physically interpretable and physics-constrained data-driven models through optimal sensor selection and placement. The research effort is anchored by a common task framework (CTF) that evaluates the performance of ML algorithms, architectures, and optimization schemes for the diverse tasks required in engineering applications. The aim is to push beyond the boundaries of modern techniques by closing the loop between data collection, control, and modeling, creating a unique and cross-disciplinary architecture for learning physically interpretable and physics constrained models of complex dynamic systems from time series data. The CTF further supports sustainable and open-source challenge datasets, which are foundational for developing interpretable, ethical, and inclusive tools to solve problems fundamental to human safety, society, and the environment.

The emergence of machine learning (ML) and AI algorithms is transforming every scientific and engineering discipline. The widespread adoption of ML/AI algorithms represents the second computational revolution. As in the early 1990s, the widespread availability of computers allowed for significant advancements in science and engineering by allowing for numerical proxies for modeling physical systems. Similarly, ML/AI is allowing for the modeling and control of physical systems directly from the quantity and quality of data acquired from emerging sensor technologies. The goal of the AI Institute in Dynamic Systems is to develop the next generation of advanced ML tools for controlling complex physical systems by discovering physically interpretable and physics-constrained data-driven models through optimal sensor selection and placement. The work is anchored by a common task framework (CTF), which is a suite of challenge data sets with a diverse set of objectives relevant to engineering design and control. Thus the institute evaluates the performance of ML algorithms and allows us to push beyond the boundaries of modern techniques by closing the loop between data collection, control, and modeling, thus creating a unique and cross-disciplinary architecture for learning physically interpretable and physics constrained models of complex

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2024 The Authors. Association for the Advancement of Artificial Intelligence.

dynamic systems from time series data. The CTF further supports sustainable and open-source challenge datasets, which are foundational for developing interpretable, ethical, and inclusive tools to solve problems fundamental to human safety, society, and the environment.

As such, the AI Institute develops the fundamental mathematical and computational tools required to empower AI for real-time sensing (Manohar et al. 2018), learning (Chen et al. 2022; Fasel et al. 2022), decision making (Hu et al. 2023), and prediction capabilities that lead the way towards safe, reliable, efficient, explainable, and ethical data-enabled engineering and science systems. The institute is organized around three key thrusts that are used in synergistic partnership for solving today's grand-challenge problems. Thrust 1 is centered around sensing and data-acquisition, which is required of all AI agents engineered to perform in real-world applications. This thrust also builds fundamental optimization tools, which underpin all ML and AI learning algorithms (Manohar et al. 2018). Thrust 1 enables Thrust 2, which aims to build a diversity of dynamic models directly from sensor data (Chen et al. 2022; Fasel et al. 2022; Gao and Kutz 2022). With the development of dynamic models, data-driven control protocols are developed in Thrust 3 (Hu et al. 2023). The understanding of models and controls then re-integrates with sensing where better sensing strategies can be constructed with knowledge from Thrusts 2 and 3. This creates a virtuous cycle of data-driven integration strategies for broad application across science and engineering. Critically, all thrusts integrate around the CTF, which allow the institute to test, evaluate, and compare diverse algorithms for executing Thrusts 1–3 in application areas, which are relevant to the science and engineering community. The centrality of the CTF to the institute allows the broader community access to data, algorithms, and innovations. As such, it creates a nexus point for scientific activity.

Based at the University of Washington in Seattle, the AI Institute team is geographically centered across the greater Pacific Northwest region, while maintaining key connections to east coast technology hubs. This includes in the Northwest the University of Nevada Reno, the University of Hawaii Manoa, Montana State University, Portland State University, the University of Alaska, and Boise State University. In addition, we are partnered with Harvard University and Columbia University. The geographic connections help facilitate active collaboration among team members through a shared sense of purpose, while remaining connected to other hubs. A long term goal for the institute is to serve as a nexus point for facilitating the rapid transfer of intellectual ideas and cutting edge educational opportunities in AI for dynamic systems from established technology hubs.

More broadly, the institute is developing the three critical efforts necessary for driving forward AI for dynamic systems: (i) the mathematical foundations of AI, (ii) grand challenge applications for AI, and (iii) a transformational educational and workforce development infrastructure for AI Engineering. Physics-informed ML is emerging as a leading paradigm for bringing together AI and engineering dynamic systems, providing new capabilities in real-time sensing, learning, decision making, and predictions that are safe, reliable, efficient, ethical, and imbued with uncertainty quantification (UQ). The AI Institute team is developing a general and flexible ML framework to rapidly learn new physics, enforce known physical constraints, and discover them directly. Methods are continuously evaluated through demonstrations of real-world grand challenge applications, so that the methods are physically motivated and catalyze advances across multiple applied domains. To facilitate reproducibility and robustness across disciplines, the institute has developed the CTF that leverages SageBionetworks, a nonprofit company dedicated to hosting challenge data sets, for evaluating ML/AI algorithms on challenge data sets in physics and engineering, thus providing a broad service to the engineering ML community. All methods and datasets are made available to the public through actively maintained open source software packages.

Research: The foundations of physics-informed ML are rooted in four key disciplines: (i) control theory, (ii) probability and statistics, (iii) optimization, and (iv) dynamical systems (modeling). The integration of all four of these disciplines is critical for the development of ML algorithms that can be leveraged by engineered systems, as illustrated in the integration of methods in Figure 1. Establishing rigorous mathematical connections between these disciplines is a driving inspiration for the institute efforts in reframing the foundations of ML for the engineering sciences. A key step in moving the field forward is to address the current lack of educational infrastructure around training, which integrates these four core background areas, especially as they are applied to pressing challenges in engineering dynamic systems. The institute trains students and postdocs to have deep expertise in both the core mathematical architectures as well as domain applications through the development of new open-source and multimodal undergraduate and graduate curricula.

Given the institute objectives, efforts are organized around three key thrusts that are used in synergistic partnership. Figure 2 shows how each thrust is critical for all other thrusts. Thrust 1 is built around sensing and data-acquisition, which is required of all AI agents engineered to perform in real-world applications. This thrust also builds fundamental optimization tools, which underpin all ML and AI learning algorithms. Thrust 1 enables Thrust 2,

23719621, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aaai.12159, Wiley Online Library on [30/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

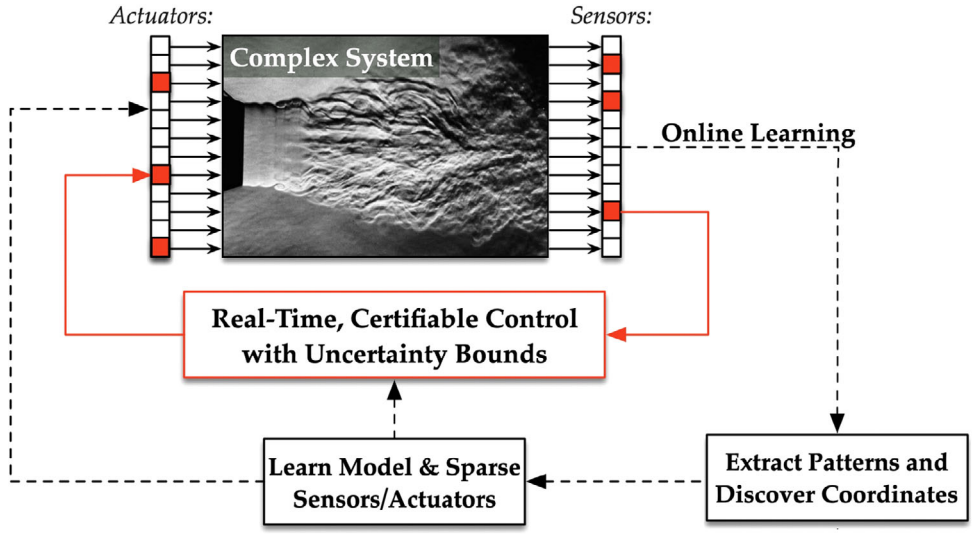


FIGURE 1 Overview of methodological innovations proposed for physics-informed machine learning for dynamic systems by the AI Institute in Dynamic Systems.

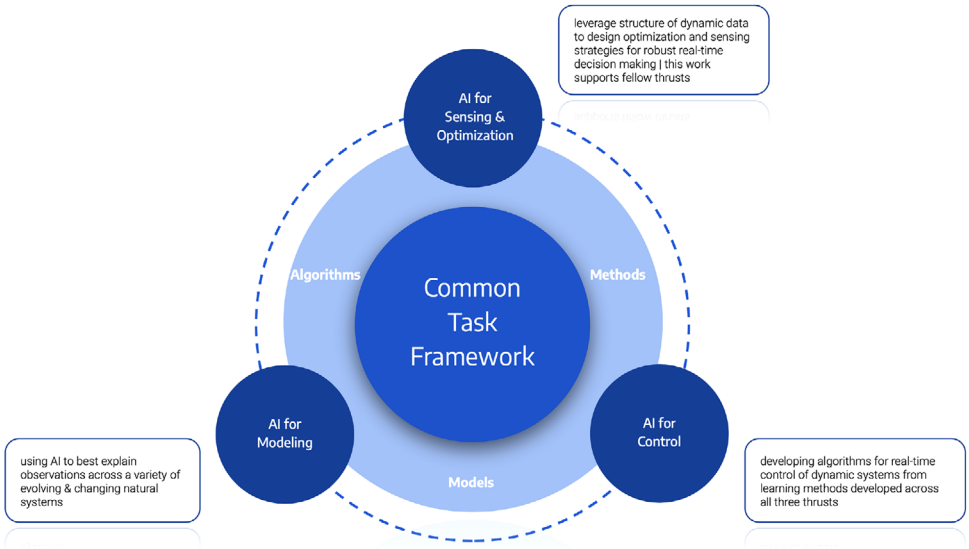


FIGURE 2 Integration and partnering of thrusts. Each thrust is integral and critical to each other thrust in the overall sending, modeling, control paradigm for AI integration.

which aims to build a diversity of dynamic models directly from sensor data. With the development of dynamic models, data-driven control protocols are developed in Thrust 3. The understanding of models and controls then reintegrates with sensing where better sensing strategies can be constructed with knowledge from Thrusts 2 and 3. This creates a virtuous cycle of data-driven integration strategies for broad application across science and engineering. Critically, all thrusts integrate around the CTF, which allow us to test, evaluate, and compare diverse algorithms for executing Thrusts 1–3 in application areas which are relevant to the science and engineering community.

Thrust 1: AI for sensing and optimization: Much of the power of modern AI derives from our ability to fit highly expressive models to large sets of high-dimensional data, by numerical optimization. Physics and dynamics-informed AI poses additional challenges beyond the standard issues of scale and dimensionality: the goal is to identify accurate physical models that support downstream control and decision making with guaranteed performance. This goes well beyond the standard demands of ML/AI, where the optimization goal is often just to accurately label a given training set. Physical data are structured: individual signals are often sparse in an

23719621, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aaai.12159, Wiley Online Library on [30/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

appropriate basis, batches of signals are often low-rank, while data generated by systems with a few important underlying degrees of freedom typically concentrate near low-dimensional manifolds (Champion et al. 2019; Chen et al. 2022). From a computational perspective, correctly inferring these low-dimensional models in noisy, incomplete, or otherwise unreliable data typically requires us to solve a high-dimensional nonconvex optimization problem, to global optimality. This requires new perspectives and computational schemes to address the emerging challenges posed for optimization in the context of complex dynamical systems. Thus, the institute has leveraged the structure of dynamic data to design guaranteed optimization and sensing strategies for safe, efficient, robust real-time decision making (Hu et al. 2023; Manohar et al. 2018).

Thrust 2: AI for physics-informed models: A significant strength of the institute team is in learning physically interpretable models of dynamical systems from off-line and/or on-line streaming data (Baddoo et al. 2023; Brunton, Proctor, and Kutz 2016; Champion et al. 2019; Chen et al. 2022; Gao and Kutz 2022; Schmidt and Lipson 2009). Physics informed learning is of growing importance for scientific and engineering problems. Physics-informed simply refers to the ability to constrain the learning process by physical and/or engineering principles. For instance, conservation of mass, momentum, or energy can be imposed in the learning process. In the parlance of ML, the imposed constraints are referred to as regularizers. Thus, physics informed learning focuses on adding regularization to the learning process to impose or enforce physical priors. Known physics (e.g., invariances, symmetries, conservation laws, constraints, etc.) may be incorporated at various stages of the learning process. For example, rotational invariance is often incorporated by augmenting the training data with rotated copies, and translational invariance is often captured using convolutional neural network architectures. In kernel-based techniques, such as Gaussian process regression and support vector machines, symmetries can be imposed by means of rotation-invariant, translation-invariant, and symmetric covariance kernels. Additional physics and prior knowledge may be incorporated as additional loss functions or constraints in the optimization problem. Thus, the institute leverages data and AI to automatically discover the underlying mechanisms and mathematical representations that best explain observations across a variety of evolving/changing natural systems. These mathematical formulations are being used to gain insights, make predictions, and discover governing equations and coordinate systems in a variety of application areas already (Baddoo et al. 2023; Champion et al. 2019; Chen et al. 2022; Fasel et al. 2022; Gao and Kutz 2022).

Thrust 3: AI for data-driven control: The remarkable successes of ML can be also leveraged towards the control of modern complex dynamical systems. Reinforcement learning (RL) is a class of ML that addresses the problem of learning to control physical systems by explicitly considering their inherent dynamical structure and feedback loop. To date, the successes of RL have been limited to very structured or simulated environments, and its successes in real-world systems are few. RL has significant challenges involving: (i) Scalability: How to develop scalable RL methods for large-size network multiagent (MARL) dynamical systems? (ii) Robustness: How to maintain the performance (efficiency and safety) of the learned policies even when there is a model class mismatch? (iii) Safety: How to guarantee RL maintains stability and stays in the safe constraint while still learning efficiently? The institute is leading the developing of critically enabling mathematical and computational architectures to overcome these challenges which are present in a diverse number of applications involving complex dynamical systems. Thus, the institute is innovating algorithms for real-time control of dynamic systems for their safe, reliable, and efficient operation leveraging principles from model-based and model-free control, optimization, and learning methods developed across all the three thrusts of modeling, control, and sensing/optimization. Institute partners are already establishing the theoretical foundations of policy optimization for learning control policies, thus revolutionizing how modern RL control can be embedded in engineering systems (Hu et al. 2023).

Education in AI: The institute has a significant and focused effort on open-source educational courses and algorithms in AI for science and engineering, including the development of open source packages for engineering AI applications (de Silva et al. 2021; Kaptanoglu et al. 2022; Pan et al. 2023; Van Breugel et al. 2022). This serves as the foundation for nurturing and developing the next generation of talent in the sciences. Through these initiatives, we aim to support the growth and learning of individuals interested in this field, ensuring a continuous pipeline of skilled professionals, and advance innovations broadly in academia and industry. To efficiently disseminate the knowledge and tools that the institute is developing, the institute will provide the mathematical and software foundations necessary for diverse students to contribute to new AI for dynamic systems research. The CTF helps promote an active and broad engagement, with code and data openly available to student, postdocs, and professionals in the field overall. More broadly, to reach a broad and diverse community, our education plan integrates active learning classroom experiences with state-of-the-art online lectures, bootcamps, and workshops to teach students from K-12 through graduate school both fundamentals and

23719621, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aaai.12159, Wiley Online Library on [30/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

practical skills. All members of the institute are involved with the educational and workforce development efforts, providing educational materials that leverage each AI Institute members strengths. This institute effort enriches the broader community along with providing the best and most complete open-source educational materials in the engineering and physical sciences for graduate students, undergraduate students, and postdocs alike. Indeed, the overarching goal of the institute is to design and implement a comprehensive education and training plan that integrates ML and artificial intelligence seamlessly into undergraduate and graduate engineering curriculum. This involves identifying existing and desired skills and objectives and developing modular curriculum content to include in existing and new courses. This content is actively deployed within the institute and made available more broadly for other institutions and industry partners.

In addition to academic research efforts and partnerships, the institute is committed to building bridges to industry and national laboratories whose futures depend on the successful integration of ML/AI algorithms in the design, execution, and control of their emerging technologies. The institute has a large number of existing partnerships with traditional engineering companies to modern tech companies, balancing engineering design with state-of-the-art ML. Partnerships also exist with national laboratories who have developed significant efforts around physics-informed ML with applications to national interest.

Summary: From autonomy to turbulence, the design of computational algorithms capable of reliably controlling today's modern, high-dimensional complex systems remain a grand challenge scientific endeavor. Such systems are characterized by the manifestation of physical phenomena that can span multiple spatial and temporal scales and involve coupling between multiple types of physics, which are often unknown and unmeasured. Indeed, our current mathematical approaches to actuate such systems are insufficient to establish reliable and assured control paradigms. Emerging ML and artificial intelligence (AI) methods are transforming many technological landscapes due to their ability to flexibly and adaptively integrate measurement data into the modeling framework. In regards to engineered dynamic systems, ML and AI strategies are now being engineered by the AI Institute in Dynamic Systems to form the fundamental underpinnings of *physics-informed ML* strategies, which account for the constraints imposed by the underlying physics of the system, while providing greater flexibility to represent more complex physical phenomena. The institute's development of physics-informed ML holds tremendous promise for a new modeling paradigm to control modern dynamic systems. This integration also is offering transformative impact in education and training by bridging the gap between traditional engineering disciplines and modern ML methods.

ACKNOWLEDGMENTS

This material is based upon work supported by the National AI Research Institutes program supported by NSF and DHS under Award No. 2112085. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict.

ORCID

J. Nathan Kutz https://orcid.org/0000-0002-6004-2275

REFERENCES

Baddoo, P. J., B. Herrmann, B. J. McKeon, J. Nathan Kutz, and S. L. Brunton. 2023. "Physics-Informed Dynamic Mode Decomposition." *Proceedings of the Royal Society A* 479(2271): 20220576.

Brunton, S. L., J. L. Proctor, and J. N. Kutz 2016. "Discovering Governing Equations from Data by Sparse Identification of Nonlinear Dynamical Systems." *Proceedings of the National Academy of Sciences* 113(15): 3932–37.

Champion, K., B. Lusch, J. N. Kutz, and S. L. Brunton. 2019. "Data-Driven Discovery of Coordinates and Governing Equations." *Proceedings of the National Academy of Sciences* 116(45): 22445–51.

Chen, B., K. Huang, S. Raghupathi, I. Chandratreya, Q. Du, and H. Lipson. 2022. "Automated Discovery of Fundamental Variables Hidden in Experimental Data." *Nature Computational Science* 2(7): 433–42.

de Silva, B. M., K. Manohar, E. Clark, B. W. Brunton, J. N. Kutz, and S. L. Brunton. 2021. "PySensors: A Python Package for Sparse Sensor Placement." *Journal of Open Source Software* 6(58): 2828.

Fasel, U., J. N. Kutz, B. W. Brunton, and S. L. Brunton. 2022. "Ensemble-SINDy: Robust Sparse Model Discovery in The Low-Data, High-Noise Limit, with Active Learning and Control." *Proceedings of the Royal Society A* 478(2260): 20210904.

Gao, L., and J. N. Kutz. 2022. "Bayesian Autoencoders for Data-Driven Discovery of Coordinates, Governing Equations and Fundamental Constants." *Proceedings of the Royal Society A*. arXiv preprint arXiv:2211.10575.

Hu, B., K. Zhang, N. Li, M. Mesbahi, M. Fazel, and T. Başar. 2023. "Toward a Theoretical Foundation of Policy Optimization for Learning Control Policies." *Annual Review of Control, Robotics, and Autonomous Systems* 6: 123–58.

Kaptanoglu, A. A., B. M. de Silva, U. Fasel, K. Kaheman, A. J. Goldschmidt, J. Callaham, C. B. Delahunt, et al. 2022. "PySINDy: A Comprehensive Python Package for Robust Sparse System Identification." *Journal of Open Source Software* 7(69): 3994.

Manohar, K., B. W. Brunton, J. N. Kutz, and S. L. Brunton. 2018. "Data-Driven Sparse Sensor Placement for Reconstruction: Demonstrating The Benefits of Exploiting Known Patterns." *IEEE Control Systems Magazine* 38(3): 63–86.

23719621, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aaai.12159, Wiley Online Library on [30/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Pan, S., E. Kaiser, B. M. de Silva, J. N. Kutz, and S. L. Brunton. 2023. "PyKoopman: A Python Package for Data-Driven Approximation of the Koopman Operator." arXiv preprint arXiv:2306.12962.
Schmidt, M., and H. Lipson. 2009. "Distilling Free-Form Natural Laws from Experimental Data." *Science* 324(5923): 81–85.
Van Breugel, F., Y. Liu, B. W. Brunton, and J. N. Kutz, 2022. "PyNumDiff: A Python Package for Numerical Differentiation of Noisy Time-Series Data." *Journal of Open Source Software* 7(71): 4078.

How to cite this article: Kutz, J. N., S. L. Brunton, K. Manohar, H. Lipson, and N. Li. 2024. "AI Institute in Dynamic Systems: Developing machine learning and AI tools for scientific discovery, engineering design, and data-driven control." *AI Magazine* 45: 48–53. https://doi.org/10.1002/aaai.12159

AUTHOR BIOGRAPHIES

J. Nathan Kutz is the Robert Bolles and Yasuko Endo Professor of Applied Mathematics and Electrical and Computer Engineering at the University of Washington. He is also Director of the AI Institute in Dynamic Systems. His research interests include dynamical systems and machine learning with applications to diverse scientific disciplines that include optics, neuroscience, and fluid dynamics.

Steven L. Brunton is the James Morrison Professor of Mechanical Engineering at the University of Washington. He is also an Associate Director of the AI Institute in Dynamic Systems. His research interests include machine learning, fluid dynamics, dynamical systems, and control.

Krithika Manohar is an Assistant Professor of Mechanical Engineering at the University of Washington. Her research interests include sensing and sensor placement, model reduction, and dynamical systems.

Hod Lipson is the Director of Columbia University's Creative Machines Lab. Lipson's work focuses on evolutionary robotics, design automation, rapid prototyping, artificial life, and creating machines that can demonstrate some aspects of human creativity.

Na Li is the Winokur Family Professor of Electrical Engineering and Applied Mathematics in the School of Engineering and Applied Sciences (SEAS) at Harvard University. Her research lies in control, learning, and optimization of networked systems, including theory development, algorithm design, and applications to real-world cyber-physical systems, such as energy systems, buildings, multirobots, and bio-medical systems.

23719621, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aaai.12159, Wiley Online Library on [30/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License