



## A synergistic future for AI and ecology

Barbara A. Han<sup>a,1</sup> , Kush R. Varshney<sup>b,1</sup> , Shannon LaDeau<sup>a</sup> , Ajit Subramaniam<sup>c</sup> , Kathleen C. Weathers<sup>a</sup> , and Jacob Zwart<sup>d</sup>

Edited by Mary Power, University of California, Berkeley, CA; received March 21, 2023; accepted July 25, 2023

Research in both ecology and AI strives for predictive understanding of complex systems, where nonlinearities arise from multidimensional interactions and feedbacks across multiple scales. After a century of independent, asynchronous advances in computational and ecological research, we foresee a critical need for intentional synergy to meet current societal challenges against the backdrop of global change. These challenges include understanding the unpredictability of systems-level phenomena and resilience dynamics on a rapidly changing planet. Here, we spotlight both the promise and the urgency of a convergence research paradigm between ecology and AI. Ecological systems are a challenge to fully and holistically model, even using the most prominent AI technique today: deep neural networks. Moreover, ecological systems have emergent and resilient behaviors that may inspire new, robust AI architectures and methodologies. We share examples of how challenges in ecological systems modeling would benefit from advances in AI techniques that are themselves inspired by the systems they seek to model. Both fields have inspired each other, albeit indirectly, in an evolution toward this convergence. We emphasize the need for more purposeful synergy to accelerate the understanding of ecological resilience whilst building the resilience currently lacking in modern AI systems, which have been shown to fail at times because of poor generalization in different contexts. Persistent epistemic barriers would benefit from attention in both disciplines. The implications of a successful convergence go beyond advancing ecological disciplines or achieving an artificial general intelligence—they are critical for both persisting and thriving in an uncertain future.

complex systems | ecosystems | interpretable AI | deep learning | prediction

Ecological understanding is paramount for confronting multiple linked phenomena including the increased frequency of disease outbreaks, exponential losses of global biodiversity, and profound impacts of climate change. These crises share a commonality: they emerge from perturbations of complex systems whose high dimensionality underpins nonlinear dynamics that are difficult to predict. Advances in AI have the potential to transform our understanding of ecological systems (1). At the same time, ecological systems are themselves an impetus for the advancement of AI. Challenges that are commonplace in multiscale, context-dependent, and imperfectly observed ecological systems offer a panoply of problems through which AI moves closer to realizing its full potential. Predicting and purposefully managing the outcomes of perturbations to natural complex systems is a great

challenge of our time, one that demands bold synergistic convergence of AI and ecological science to achieve a more holistic understanding that informs action—creating system wisdom toward a resilient future (Fig. 1).

Just as humans learn patterns from data to build intelligent thinking about a system, ML underpins modern AI, the proximate goal of which is to execute tasks or make decisions in a particular domain. Applying AI tools to ecological domains has vastly expanded our abilities to quantify phenomena that were previously unquantifiable or difficult to observe (2), and offers the possibility of faster and more accurate predictions about ecological systems. Recent examples include studying the interactions between organisms and their environments through camera and acoustic data (3–5); distilling earth systems satellite data into meaningful ecological functions (e.g., productivity) (6); analyzing animal behavior through deep learning and pose estimation (7); bioinformatics for prediction and validation of whether new viruses are capable of human infection (8); and which animal species are most likely to harbor them (9).

These are examples of "AI for ecology"—the application of existing AI tools to ecological problems. The other direction, however, of ecological science inspiring new paradigms in AI is just as important. An ultimate goal in AI research is to achieve artificial *general* intelligence (AGI) that can extrapolate and reason about other domains and systems similar to human intelligence. It is likely that AGI will need a combination of data-driven ML and new ways of representing and reasoning from diverse knowledge types to meet the challenge of trustworthy predictions about no-analog futures, such as those we anticipate on a rapidly changing planet. Meeting these challenges may be helped by a fundamental shift in how AI and ecological research propel each other.

Here, we identify a convergence on the horizon between ecological research, which has traditionally lagged behind advances in AI and computational science and AI research (Fig. 2). This convergence seeks new thought paradigms to support intelligent extrapolation to unobserved (or

Author affiliations: <sup>a</sup>Cary Institute of Ecosystem Studies, Millbrook, NY 12545; <sup>b</sup>IBM Research - Thomas J. Watson Research Center, Yorktown Heights, NY 10598; <sup>c</sup>Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964; and <sup>d</sup>U.S. Geological Survey, Water Resources Mission Area, Integrated Information Dissemination Division, San Francisco, CA 94116

Author contributions: B.A.H., K.R.V., S.L., A.S., K.C.W., and J.Z. wrote the paper.

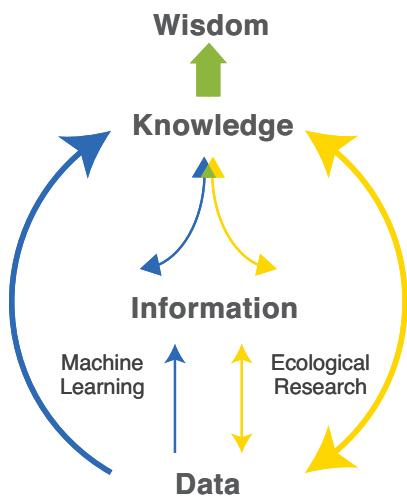
The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2023 The Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

<sup>1</sup>To whom correspondence may be addressed. Email: hanb@caryinstitute.org or kvarshn@us.ibm.com.

Published September 11, 2023.



**Fig. 1.** A diagram illustrating interconnections between data, information, knowledge, and wisdom. **Data** reflect raw observations or measurements (e.g., satellite data such as sea surface temperature (SST) at a specific location), while synthesizing these measurements into a meaningful form constitutes **information** (e.g., providing spatial or temporal context to SST measurements in the form of time series maps). **Knowledge** adds context, providing similar examples or comparisons to other knowledge systems (e.g., some marine organisms will experience heat stress). Finally, **wisdom** considers all of these elements as well as societal or cultural values to assess potential actions (e.g., carbon emission limits to mitigate detrimental climate warming effects). Arrows represent how machine learning (ML, blue) and ecological systems research (yellow) achieve connections within this framework. ML can transform data into information but can also bypass the information step to direct inference conveyed as knowledge. In contrast, the bidirectional arrows for ecology represent iterative feedback to the process of data collection in order to achieve knowledge through statistical modeling and hypothesis refinement. Intentional and synergistic advances in ML-AI and ecological systems science may spur greater understanding, prediction, and protection of complex systems function if we can align strengths from each discipline to explicitly identify biases and manage uncertainty and different ways of knowing, especially at information and knowledge levels.

unobservable) systems and futures. Our argument extends well beyond the "AI for X" paradigm of simply applying AI to myriad domains (represented by "X"). A shift toward coproduced, convergent research has the potential to fuel the next generation of AI advances and ecological understanding.

This convergence would be helped by the amalgamation of two cultures of scientific understanding (22). In ecology, a main objective is to understand complex systems spanning the physical and the biological. This understanding is often captured in mathematical models, which reflect our hypotheses about the mechanisms that act in concert to produce observable outcomes. But ecology neither succumbs to the "unreasonable effectiveness of mathematics" (like purely physical phenomena) (23) nor to the "unreasonable effectiveness of data" (like certain biological phenomena) (24). Simple mathematical equations cannot fully capture the essence of an ecological system. Likewise, complicated ML models do not generalize well enough to capture nonlinearities from unexpected perturbations—ML models derive algorithmic understanding from data that are often about one particular system. Furthermore, because the goal of ecological research is not just prediction but phenomenological and mechanistic understanding of complex systems, ecologists use a variety of modeling techniques at different scales that together coherently capture emergent properties in a holistic system-level way that is conducive to further study (25). This approach contrasts with AI research and development, where conflicts among multiple models are not undesired because the

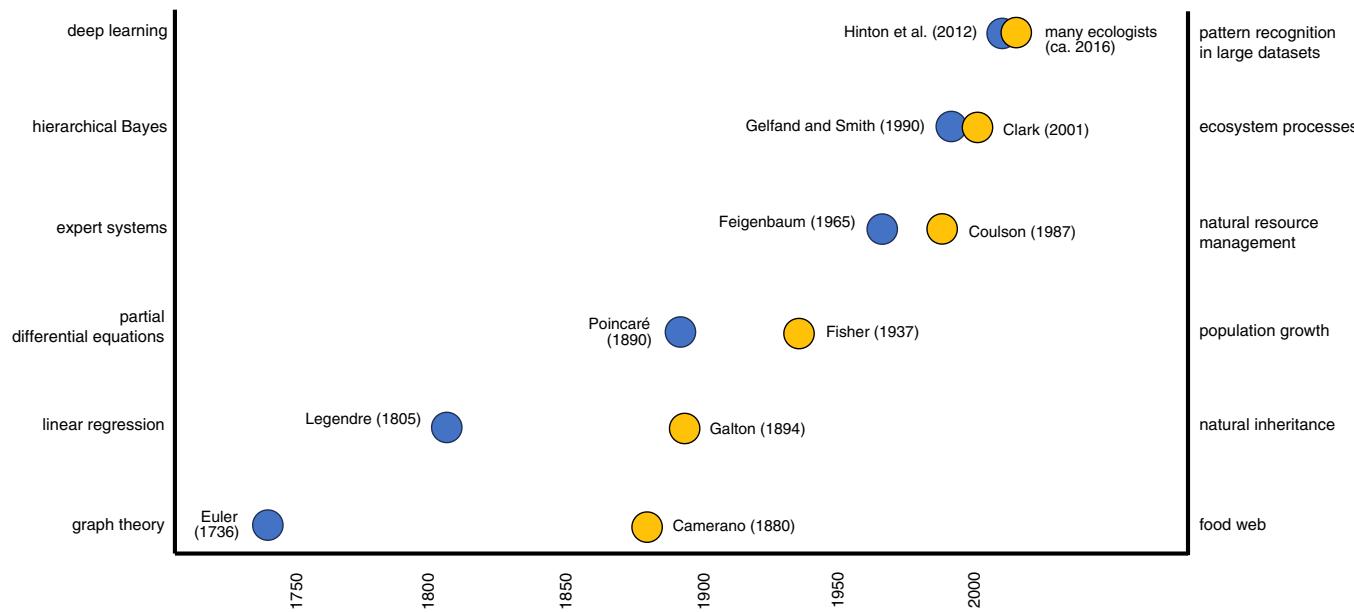
models are only being used for prediction (which one performs best), not for explanation (why does one perform better than another) (26). For example, large language models in AI show emergent behaviors that are not present in smaller language models (27), but current techniques in AI do not explain them. The modus operandi of ecological systems research, which prioritizes understanding the mechanisms generating the data we observe, can push AI researchers to focus on methods permitting deeper understanding of the causes underlying such observed phenomena—what quantitative changes in a system lead to qualitative changes in system behavior (27).

A critical emergent behavior of ecological systems—an intelligence if you will—is their spectacular resilience to perturbation. This quality further points to ecology as an inspiration for AI, which has only experienced mixed success with imbuing resilience into the existing brittle intelligence of neuroinspired AI (28). Thus, it is not only the different and combined ways of modeling ecological systems that can help push AI research, but the ecological systems themselves that can inspire robust multiscale architectures for AI. [This echoes the synergistic relationship between quantum chemistry and quantum computing (29).]

By coproducing an interwoven research path, AI and ecological understanding are well positioned to mutually advance beyond where each discipline could arrive independently. A future where AI research development is synergistic with ecological research could advance the quest to understand complex ecological systems at temporal and spatial scales relevant to societal needs. Below, we provide an account of the current state of the art in AI research (Section 1). We briefly trace the history of ecological systems modeling up to the present use of deep neural networks (Section 2) and expand upon the opportunity for convergent research in AI and ecology. We then conjecture examples of this convergent research paradigm, exploring how the study of ecology can advance AI (Section 3), how AI is advancing ecology (Section 4), and opportunities for synergistic research to accelerate mutual discovery and advancement (Section 5). We conclude by identifying some distinct and common biases that should be addressed by AI and ecology disciplines and highlight some shared opportunities for more responsible development and deployment of AI (Section 6).

## 1. AI: The Current State of the Art

In the last 10 y, deep neural networks (also known as deep learning) have become synonymous with AI due to their powerful modeling abilities. Familiar successes include more accurate clinical diagnoses from radiological imaging and increasingly rapid analysis and decision-making used for self-driving car technology. In December 2022, ChatGPT, a language model based on deep learning, showed the rapidly expanding potential of deep learning models. Deep neural networks are a subfield of ML. ML derives its intelligence from the patterns in data, whether tabular data, time series, images, or text. The goal is to generalize the patterns from these data to new unseen data points. ML models have many forms; some are simpler, like decision trees and linear models, and others are more complex. For example, artificial neural networks are inspired by the wiring of neurons in brains and



**Fig. 2.** The arc of computational methods in ecological modeling (yellow) has lagged behind the development of the computational methods themselves (blue) but is close to converging. The figure is meant to be evocative, not comprehensive. It is biased toward Western scientific traditions and data points may be contested. Figure citations (10–21), with Lederberg (20) as a reference for Feigenbaum 1965, and Krizhevsky et al. (10) to accompany Hinton et al. (19).

utilize a structure of layers that compute weighted dot-products fed through nonlinear functions. Deep neural networks have many layers and are trained on large datasets.

Despite the recent prominence of ML and particularly deep learning within AI, several other approaches to AI have developed in parallel and may overcome the limitations of deep learning to model complex systems and enable more resilient intelligence. One example is symbolic AI, which involves logical reasoning over knowledge graphs. A knowledge graph is different from statistical data in that it explicitly captures concepts and their semantic relationships. A graph might have "animals", "cows", "plants", and "grasses" as vertices with edges between them indicating that cows "are" animals, grasses "are" plants, and cows "eat" grasses. An AI system can then reason broader generalizations, such as that some animals eat some plants. Knowledge graphs are one type of representation used in symbolic AI. Other representations are ontologies, logical rule sets, probabilistic dependency graphs, differential equations, and analytic equations; each has reasoning algorithms built on top of them (30). This kind of knowledge and reasoning constitutes expert systems.

The present state of the art in AI is based on "foundation models" trained on unimaginably large datasets used as base models for many different tasks (31). A process of fine-tuning on a small domain-specific dataset specializes the foundation model for the task. Moreover, foundation models are being used within generative models, which are able to create new data (32), for example, to generate the sequences and molecular structures of new viral variants through which to investigate hypotheses about infection risk to humans and other animals.

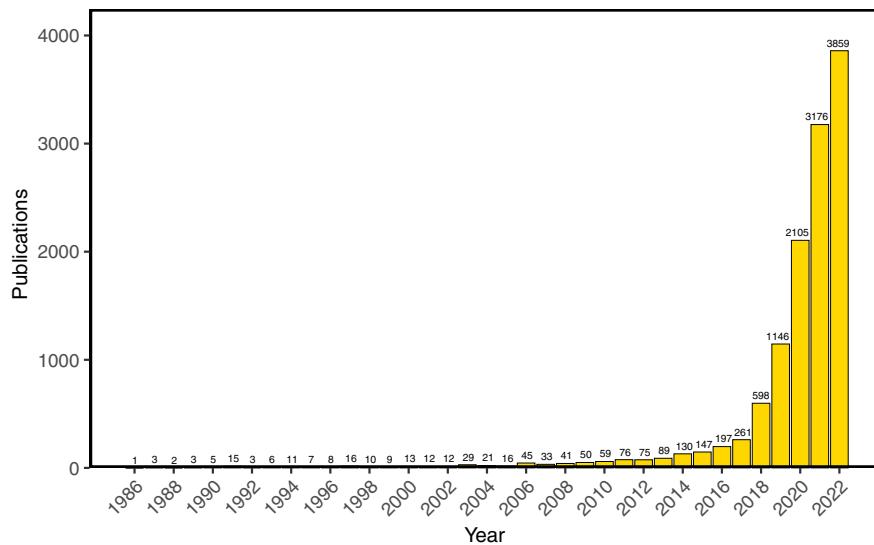
In another vein, neurosymbolic AI is combining the best qualities of deep neural networks and knowledge-based symbolic methods to push beyond each method's unique limitations (33). Neurosymbolic AI systems are considered to be broader than more narrowly abled deep learning methods and are a step farther on the path toward AGI. They can help

enable challenging knowledge discovery tasks (generating new hypotheses) and deal with heterogeneous and complex data of different modalities, scales, qualities, and quantities—all of which are common in ecological inquiry of complex systems. Like Bayesian statistical approaches, neurosymbolic AI can incorporate diverse modes of expert knowledge in its reasoning that are not always presented as data. It may also be more adaptable and robust than other AI approaches and may give explainable outputs that lend mechanistic insight, which is a guiding principle of ecological research (34).

## 2. Ecological Systems Modeling

Prediction in ecological systems epitomizes the difficulty in modeling complex systems more generally, which are defined by nonlinear dynamics that arise from feedbacks and dependencies that span multiple scales of time, space, and social dimensions. Because the science of ecological systems sits at the intersection of multiple mature subdisciplines, many physical and biological principles can guide our understanding about these systems—for example, the physics of hydrology, biogeochemistry, and landscape ecology (35), or the principles of fitness in population dynamics. Extracting information from observations about ecological systems ideally would consider the stochasticity and context dependence that are inherent in ecological systems. Ecologists have met this challenge with a plethora of modeling approaches, some focusing on system components (e.g., questions such as, What are the biophysical controls on the distribution and abundance of a limited resource?) and others focusing on modeling the bigger picture (e.g., How will ecological communities function differently in a rapidly changing landscape?). Improving systems-level prediction in ecology may also be fruitful for the development of novel AI, and these innovations may happen more quickly than they have in the past.

Adoption of new computational methods has generally exhibited a multiyear lag in ecological modeling (Fig. 2).



**Fig. 3.** The number of papers per year for the query on Web of Science: ((TS=(“artificial intelligence” OR “machine learning”))) AND WC=(Ecology OR “environmental sciences”). Certain data included herein are derived from Clarivate Web of Science (40). © Copyright Clarivate 2023. All rights reserved.

Examples include graph theory (e.g., for modeling food webs), linear regression (e.g., for modeling heredity), partial differential equations (e.g., for modeling population growth), expert systems (e.g., for managing environmental decisions), hierarchical Bayesian methods (e.g., for estimating tree fecundity), and deep learning (e.g., for assessing biodiversity). The historical lag between the invention of a computational method and its integration into ecological modeling appears to be closing (36) (Fig. 2), and the use of AI in ecological modeling has greatly increased (Fig. 3). ML-based prediction has started supplementing theory-driven prediction in areas such as hydrology (37), zoonotic disease ecology (38), and forest ecology (39). However, the scope of application is still largely limited to pattern-recognition and prediction, and AI remains underutilized as a tool, for example, for synthesizing big data in ecology or for identifying and prioritizing novel hypotheses about ecological function for further investigations.

Analogous to brain-inspired breakthroughs in deep learning (41), the self-organizing properties and processes of biology are likely to hold hidden inspirations for AI system design that will be revealed by a closer coupling of AI research and ecological problems. Extending such biological inspiration offers a new lens highlighting opportunities for AI innovations that borrow from empirically observed intelligent decision-making from simple organisms, like slime molds, that defy our current notions of “intelligence” (42). Similarly, evolutionary principles that are fundamental to biology and ecology have inspired AI research. Evolutionary computing is an ecology-inspired branch of AI that applies genetic algorithms to direct the evolution of systems for applied goals, whose advances stem from *in silico* studies, such as directed evolution in microbes (43). Symbolic regression is another form of evolutionary computing that is currently being advanced through applications in ecology (44, 45), leading to human-interpretable functional equation models of complex ecological systems that are compositions of more primitive equations. Biodiversity measurements in ecology, which often serve as a surrogate for ecosystem complexity, are

inspiring AI researchers to develop new ways to measure unwanted bias in training data (46).

We are bullish about the synergistic advances that can be made between AI and ecological research—where ecological theory has potential to advance AI research frontiers; where existing AI methods are infused into the ecological modeling of complex systems; and where coproduced AI and ecological research hold promise for critical mutual advances.

### 3. Ecology for AI: Resilience Theory

In ecology, resilience is the capacity of a system to resist or recover from perturbation. Ecological systems are resilient because the ecological function, or the roles that are played by members of the system, are both redundant and context-dependent. Understanding the resilience of our ecological systems is one of the most critical questions facing modern science. Progress in our ability to measure and make predictions about system resilience will determine the degree to which we are able to prepare for and respond to the reverberating impacts of global climate change and land use affecting the biological processes and higher-order interactions that underpin ecosystem resilience (47). Resilience theory in ecology (48) may offer clues to AI researchers for building more robust and adaptable systems involving feedback loops, redundant pathways, and satisficing behaviors that identify which fundamental principles of a system would ideally be quantified and captured in order to reproduce resilience (49). These AI systems could themselves be used to model and investigate ecological resilience. Out-of-distribution generalization and resilience to distribution shift is an active area of AI research (50). Future AI technologies that could model such context-dependent behavior may benefit from complex and nonlinear interactions with robustness built-in. Purposeful joint advances between AI research and ecology have the potential to extend general systems theory, where novel AI that is inspired and constrained by the complexities of an ecological system serve as a throughway to other domains,

such as psychology or economics where prediction is similarly difficult because of complexity spanning multiple interacting scales.

#### 4. AI for Ecology: Knowledge-Guided ML

Extant deep learning algorithms are data hungry and can produce predictions that are inconsistent with reality due to their architectures not including prior knowledge of the phenomena they are modeling. The growing field of knowledge-guided ML (KGML) is one way to advance both AI and ecology, especially under data-sparse conditions, which are still common in many areas of ecology (51–54). KGML seeks to inject scientific knowledge into the underpinning architecture of ML algorithms to assist resulting models in making more physically consistent predictions (55). This idea is consistent with the incorporation of prior knowledge in Bayesian statistical approaches, which have been leveraged for similar purposes in ecological studies but are often limited by both demands for data and computational costs (e.g., refs. 56 and 57). Examples of injecting knowledge guidance into ML models include customizing loss functions to obey physical laws (53, 58), using existing ML architectures [e.g., long short-term memory (LSTM) (59) or developing new ones to better represent reality [e.g., mass conserving-LSTM (60), recurrent graph networks (61), using process-based model outputs as inputs to ML models or as pretraining datasets (62), using partial differential equations representing systems in ML models (63), or embedding neural networks into hierarchical models (64). Going forward, ecological modeling may inspire more advanced architectures that include prior ontological knowledge in combination with hierarchies, physical and biological laws, and differential equations, along with decentralized and emergent training paradigms.

#### 5. Accelerated Discovery through Synergies between AI and Ecology

AI systems are beginning to expand beyond pattern recognition to hypothesis generation and discovery, in part by revealing missing links among variables in high-dimensional networks that represent complex systems. These missing links represent unanticipated interactions or dependencies in system components spanning multiple scales. The diversity and breadth of the states and processes in ecological systems offer exciting potential for advancing AI capacity to identify such missing links and generate novel hypotheses.

One opportunity for synergy between AI and ecology is in a problem known as *mode collapse* in AI research, where algorithms are unable to capture the full diversity of a multimodal distribution because modeling is necessarily focused on a small number of observed modes. Ecology and AI have alternately approached this long-standing problem—diffusion models were a starting point in ecology (65) that are now showing some success in addressing this problem in generative AI (66); but ecological modeling has advanced greatly since then, with telegraph models, reaction-diffusion models, and population cycling models making continued advances to dealing with mode collapse in ecological systems. Here, coproduced research holds mutual benefits for both AI and ecology—a relevant example is the potential for AI-generated

hypotheses about multimodal distributions to elucidate what is driving the bimodal peaks in the spillover transmission of Ebola virus from wildlife hosts (67). Such generative AI would benefit from the inclusion of advanced ecological modeling techniques.

Unlike many AI systems, ecological systems are understood through theory-based rules. For example, our understanding of biogeochemistry is founded on rules in physics and chemistry, geology, hydrology, and biology. Our understanding of how populations change over time is founded on evolutionary rules underpinning the concept of fitness, which is fundamental for predator-prey interactions, competition among organisms, and the structure of food webs. However, although the theory-based approach provides understanding of mechanisms (68), it may not be adequate to meet the ecological crises we face. In AI research, foundation modeling adopts a similar philosophy—that there are rules underpinning patterns and prediction—but in contrast to ecology, foundation models learn rules algorithmically by leveraging vast amounts of available data rather than through decades of applying the scientific method. Reinforcement learning is another active frontier of AI research where such rules are explicitly incorporated to generate hypotheses about how systems will evolve and stabilize over time. Thus, despite the complexity and seemingly chaotic unpredictability of ecological systems, governing rules, whether they are theory-derived or data-derived, offer anchor points from which we may observe, explain, and predict emergent properties of complex ecological systems. This approach holds particular promise for incorporating the social dimensions of complex systems, whose effects on ecological systems have historically been under-recognized and remains a transdisciplinary frontier.

#### 6. Responsible AI and Ecology

The increasing dominance of industry in AI research is leading to continued improvements of products and platforms that may provide useful methods, whereas research on ethical AI and other societal considerations is comparatively neglected (69, 70). Nevertheless, calls for safe, ethical and responsible AI as well as research into methods for mitigating biases are increasing (71–73). Indigenous, feminist, decolonial, and other critical perspectives provide a grounding for such AI research (74–76), but those perspectives remain on the fringe of AI and face epistemic barriers that would need to be overcome to be incorporated into mainstream AI research (77).

Similarly, the history of ecology is rooted in colonialism; ecological protectionism was often used to justify control of the environment; and exclusionary practices continue to generate inference that is beneficial for only a privileged subset of society (78). Ecologists are more recently working to develop practices for linking diverse knowledge to better understand socioecological systems. For example, indigenous knowledge (a.k.a. traditional ecological knowledge), obtained with informed consent and with an understanding of clear reciprocal benefits to the community, has strengthened research directions in ecology, extending to conservation, responsible management, stewardship, and the relational ethics between people and nature (79). As an example, the nuanced understanding of Arctic snow and sea-ice conditions imparted by indigenous knowledge was used to guide unmanned aerial vehicles and

satellite data collection. This information was used to better understand and manage ringed seals, a culturally important species in a changing climate (80).

Nuances of social and cultural structures and their integration are essential for gaining knowledge and wisdom about complex systems, as well as for responsibly affecting their future. For example, systemic racism has profoundly affected the social, ecological, and evolutionary characteristics of urban environments and is important to consider for ensuring social justice and strengthening the resilience of these systems to climate change (81). It is critical that both AI and ecology continue to expand epistemic boundaries, recognizing different ways of knowing as scientifically valid and respecting indigenous data sovereignty. In order to do this successfully, both could learn from qualitative research methodologies from social science (82) and specifically to capture "the context inextricably tied to the Indigenous epistemology, while maintaining standardized temporal and representational parameters that will make it consistent with other datasets for integration and analysis" (the hierarchical picture of Fig. 1 is contested under various epistemologies) (83). Additionally, following the CARE principles [Collective Benefit, Authority to Control, Responsibility, and Ethics, (84)] for indigenous data can help ensure that these data are useful for indigenous peoples and remain under their control, while also advancing knowledge within AI and ecology.

## 7. Looking toward Convergence

The research enterprise has been moving along disciplinary trajectories, fast approaching an important convergence between AI and the science of ecological systems. Convergent research in biomedicine underscores the unrealized potential

of AI for achieving what were previously characterized as "moonshots"—cures for infectious diseases that have yet to emerge (85) and for noninfectious diseases stemming from multiple sets of interacting factors (86).

Accelerating such convergent breakthroughs would be helped by investment on multiple fronts: confronting and ameliorating biases and limitations in the data—and ways of knowing—that currently exist, considering transdisciplinary thinking and practices to bridge philosophical and ethical differences in what constitutes knowledge (87), and building trust while exploring new disciplinary languages and perspectives. Investing in such intentional convergence has the potential to yield transformative perspectives and solutions that are as unimaginable and disruptive as recent breakthroughs in chatbots and generative deep learning. In an age where rapid environmental changes pose existential risks, strategic synergies between ecological system science and AI can help propel us to better understand and potentially restore the resilience of the ecological systems upon which we depend.

**Data, Materials, and Software Availability.** Literature search outputs for Fig. 3 data have been deposited in Figshare (<https://doi.org/10.25390/caryinstitute.22312177>) (40).

**ACKNOWLEDGMENTS.** We acknowledge the partial support from the NSF (DBI Grant 2234580 and DEB Grant 2200158), Cary Institute Science Innovation Fund, and Lamont-Doherty Earth Observatory Climate and Life Fellowship. We thank Amy Schuler and Adrian Castellanos for performing the literature search and creating Fig. 3, respectively; Leslie Tumblety for assistance with Fig. 1; and Nicole Herman-Mercer and Elmera Azadpour for feedback on inclusion of indigenous knowledge in ecological and AI research. We thank Alison Appling and two anonymous peers for reviewing an earlier version of this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government.

1. G. L. W. Perry, R. Seidl, A. M. Bellvé, W. Rammer, An outlook for deep learning in ecosystem science. *Ecosystems* **25**, 1700–1718 (2022). 10.1007/s10021-022-00789-y.
2. M. Nauta, D. Bucur, C. Seifert, Causal discovery with attention-based convolutional neural networks. *Mach. Learn. Knowl. Extr.* **1**, 312–340 (2019).
3. D. Stowell, M. D. Wood, H. Pamula, Y. Stylianou, H. Glotin, Automatic acoustic detection of birds through deep learning: The first Bird Audio Detection challenge. *Methods Ecol. Evol.* **10**, 368–380 (2018).
4. K. Katja *et al.*, FathomNet: A global image database for enabling artificial intelligence in the ocean. *Sci. Rep.* **12**, 15914 (2022).
5. R. C. Whytock *et al.*, Robust ecological analysis of camera trap data labelled by a machine learning model. *Methods Ecol. Evol.* **12**, 1080–1092 (2021).
6. M. A. Syariz, C.-H. Lin, M. Van Nguyen, L. M. Jaelani, A. C. Blanco, WaterNet: A convolutional neural network for Chlorophyll-a concentration retrieval. *Remote Sens.* **12**, 1966 (2020).
7. M. W. Mathis, A. Mathis, Deep learning tools for the measurement of animal behavior in neuroscience. *Curr. Opin. Neurobiol.* **60**, 1–11 (2020).
8. K. Wadhawan *et al.*, "Towards interpreting zoonotic potential of betacoronavirus sequences with attention" in *ICLR 2021* (Workshop, Machine Learning for Preventing and Combating Pandemics, 2021).
9. I. R. Fischhoff, A. A. Castellanos, J. P. G. L. M. Rodrigues, A. Varsani, B. A. Han, Predicting the zoonotic capacity of mammals to transmit SARS-CoV-2. *Proc. Biol. Sci.* **288**, 20211651 (2021).
10. A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet classification with deep convolutional neural networks" in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. Burges, L. Bottou, K. Q. Weinberger, Eds. (Curran Associates Inc., 2012).
11. A. E. Gelfand, A. F. M. Smith, Sampling-based approaches to calculating marginal densities. *J. Am. Stat. Assoc.* **85**, 398–409 (1990).
12. L. Euler, De integratore aequationum differentialium. *Novi Comment. Acad. Sci. Petropolitanae* **8**, 3–63 (1763).
13. L. Camerano, *Dell'equilibrio dei Viveni Merce la Reciproca Distrizione* (Laboratorio del R, Museo Zoologico di Torino, 1880).
14. F. Galton, *Natural Inheritance* (Macmillan and Co., New York, NY, 1894).
15. R. A. Fisher, The wave of advance of advantageous genes. *Ann. Eugen.* **7**, 355–369 (1937).
16. A. M. Legendre, *Mémoire sur les Opérations Trigonométriques: Dont les Résultats Dépendent de la Figure de la Terre* (F. Didot, 1805).
17. H. Poincaré, Sur le problème des trois corps et les équations de la dynamique. *Acta Math.* **13**, 1–270 (1890). 10.1007/bf02392506.
18. J. S. Clark *et al.*, Ecological forecasts: An emerging imperative. *Science* **293**, 657–660 (2001).
19. G. Hinton *et al.*, Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Process. Mag.* **29**, 82–97 (2012).
20. J. Lederberg, "How DENDRAL was conceived and born" in *A History of Medical Informatics* (Association for Computing Machinery, 1990), pp. 14–44.
21. R. N. Coulson, L. J. Folse, D. K. Loh, Artificial intelligence and natural resource management. *Science* **237**, 262–267 (1987).
22. L. Breiman, Statistical modeling: The two cultures (with comments and a rejoinder by the author). *SSO Schweiz. Monatsschr. Zahnheilkd.* **16**, 199–231 (2001).
23. E. P. Wigner, "The unreasonable effectiveness of mathematics in the natural sciences" in *Mathematics and Science* (World Scientific, 1990), pp. 291–306.
24. A. Halevy, P. Norvig, F. Pereira, The unreasonable effectiveness of data. *IEEE Intell. Syst.* **24**, 8–12 (2009).
25. O. Gimenez *et al.*, Statistical ecology comes of age. *Biol. Lett.* **10**, 20140698 (2014).
26. A. Birhane, D. J. T. Sumpter, The games we play: Critical complexity improves machine learning. *arXiv* [Preprint] (2022). <https://doi.org/10.48550/arXiv.2205.08922> (Accessed 30 June 2023).
27. J. Wei *et al.*, Emergent abilities of large language models. *arXiv* [Preprint] (2022). <https://doi.org/10.48550/arXiv.2206.07682> (Accessed 30 June 2023).
28. G. Indiveri, T. K. Horiuchi, Frontiers in neuromorphic engineering. *Front. Neurosci.* **5**, 118 (2011).
29. Y. Cao *et al.*, Quantum chemistry in the age of quantum computing. *Chem. Rev.* **119**, 10856–10915 (2019).
30. Z. Cui, T. Gao, K. Talamadupula, Q. Ji, Knowledge-augmented deep learning and its applications: A survey. *arXiv* [Preprint] (2022). <https://doi.org/10.48550/arXiv.2212.00017> (Accessed 30 June 2023).
31. R. Bommasani *et al.*, On the opportunities and risks of foundation models. *arXiv* [Preprint] (2021). <https://doi.org/10.48550/arXiv.2108.07258> (Accessed 30 June 2023).
32. L. Floridi, M. Chiratti, GPT-3: Its nature, scope, limits, and consequences. *Minds Mach.* **30**, 681–694 (2020).
33. M. K. Sarker, L. Zhou, A. Eberhart, P. Hitzler, Neuro-symbolic artificial intelligence: Current trends. *arXiv* [Preprint] (2021). <https://doi.org/10.48550/arXiv.2105.05330> (Accessed 30 June 2023).
34. A. Tocchetti *et al.*, Robustness: A human-centered perspective on technological challenges and opportunities. *arXiv* [Preprint] (2022). <https://doi.org/10.48550/arXiv.2210.08906> (Accessed 30 June 2023).
35. K. C. Weathers *et al.*, Frontiers in ecosystem ecology from a community perspective: The future is boundless and bright. *Ecosystems* **19**, 753–770 (2016).

36. E. J. Rykiel, Artificial intelligence and expert systems in ecology and natural resource management. *Ecol. Model.* **46**, 3–8 (1989).

37. G. S. Nearing *et al.*, What role does hydrological science play in the age of machine learning? *Water Resour. Res.* **57**, e2020WR028091 (2021).

38. B. A. Han, J. P. Schmidt, S. E. Bowden, J. M. Drake, Rodent reservoirs of future zoonotic diseases. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7039–7044 (2015).

39. Z. Liu *et al.*, Application of machine-learning methods in forest ecology: Recent progress and future challenges. *Environ. Rev.* **26**, 339–350 (2018).

40. B. Han *et al.*, Data associated with: Beyond AI for X: A synergistic future for AI and ecology. Cary Institute. Dataset. <https://doi.org/10.25390/caryinstitute.22312177.v1>. Deposited 22 March 2023.

41. A. Zador *et al.*, Toward next-generation artificial intelligence: Catalyzing the NeuroAI revolution. *arXiv* [Preprint] (2022). <https://doi.org/10.48550/arXiv.2210.08340> (Accessed 30 June 2023).

42. F. Jabr, *How Brainless Slime Molds Redefine Intelligence* (Nature Publishing Group UK, 2012), <https://doi.org/10.1038/nature.2012.11811> (Accessed 30 June 2023).

43. A. Lalejini, E. Dolson, A. E. Vostinar, L. Zaman, "Selection schemes from evolutionary computing show promise for directed evolution of microbes" in *Proceedings of the Genetic and Evolutionary Computation Conference Companion, GECCO '22* (Association for Computing Machinery, 2022), pp. 723–726.

44. B. T. Martin, S. B. Munch, A. M. Hein, Reverse-engineering ecological theory from data. *Proc. Biol. Sci.* **285**, 20180422 (2018).

45. Y. Chen, M. T. Angulo, Y.-Y. Liu, Revealing complex ecological dynamics via symbolic regression. *Bioessays* **41**, e1900069 (2019).

46. D. Friedman, A. B. Dieng, The Vendi Score: A diversity evaluation metric for machine learning. *arXiv* [Preprint] (2022). <https://doi.org/10.48550/arXiv.2210.02410> (Accessed 30 June 2023).

47. T. Gibbs, S. A. Levin, J. M. Levine, Coexistence in diverse communities with higher-order interactions. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2205063119 (2022).

48. L. H. Gunderson, Ecological resilience—In theory and application. *Annu. Rev. Ecol. Syst.* **31**, 425–439 (2000).

49. Y. Ma, D. Tsao, H.-Y. Shum, On the principles of parsimony and self-consistency for the emergence of intelligence. *arXiv* [Preprint] (2022). <https://doi.org/10.48550/arXiv.2207.04630> (Accessed 30 June 2023).

50. P. W. Koh *et al.*, "WILDS: A benchmark of in-the-wild distribution shifts" in *Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research*, M. Meila, T. Zhang, Eds. (PMLR, July 18–24, 2021), pp. 5637–5664.

51. A. Appling, S. Oliver, J. Read, J. Sadler, J. Zwart, Machine learning for understanding inland water quantity, quality, and ecology. *Earth arXiv* [Preprint] (2022). <https://doi.org/10.31223/x5964s> (Accessed 6 November 2022).

52. A. Karpatne *et al.*, Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Trans. Knowl. Data Eng.* **29**, 2318–2331 (2017).

53. J. S. Read *et al.*, Process-guided deep learning predictions of lake water temperature. *Water Resour. Res.* **55**, 9173–9190 (2019).

54. J. Willard, X. Jia, S. Xu, M. Steinbach, V. Kumar, Integrating scientific knowledge with machine learning for engineering and environmental systems. *ACM Comput. Surv.* **55**, 1–37 (2022). <https://doi.org/10.1145/3514228>.

55. A. Daw *et al.*, "Physics-Guided Architecture (PGA) of neural networks for quantifying uncertainty in lake temperature modeling" in *Proceedings of the 2020 SIAM International Conference on Data Mining (SDM)* (Society for Industrial and Applied Mathematics, 2020), pp. 532–540.

56. S. Shirota, A. E. Gelfand, Approximate Bayesian computation and model assessment for repulsive spatial point processes. *J. Comput. Graph. Stat.* **26**, 646–657 (2017).

57. P. J. Green, K. Latuszynski, M. Pereyra, C. P. Robert, Bayesian computation: A summary of the current state, and samples backwards and forwards. *Stat. Comput.* **25**, 835–862 (2015).

58. T. Beucler, M. Pritchard, P. Gentile, S. Rasp, "Towards physically-consistent, data-driven models of convection" in *IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium* (Waikoloa, HI, 2020), pp. 3987–3990.

59. S. Hochreiter, J. Schmidhuber, Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).

60. P.-J. Hoedt *et al.*, "MC-LSTM: Mass-conserving LSTM" in *Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research*, M. Meila, T. Zhang, Eds. (PMLR, July 18–24, 2021), pp. 4275–4286.

61. X. Jia *et al.*, "Physics-guided recurrent graph model for predicting flow and temperature in river networks" in *Proceedings of the 2021 SIAM International Conference on Data Mining (SDM)* (Society for Industrial and Applied Mathematics, 2021), pp. 612–620.

62. X. Jia *et al.*, Physics-guided machine learning for scientific discovery: An application in simulating lake temperature profiles. *ACM/IMS Trans. Data Sci.* **2**, 1–26 (2021).

63. T. Bao *et al.*, "Partial differential equation driven dynamic graph networks for predicting stream water temperature" in *2021 IEEE International Conference on Data Mining (ICDM)* (2021), pp. 11–20.

64. M. B. Joseph, Neural hierarchical models of ecological populations. *Ecol. Lett.* **23**, 734–747 (2020).

65. E. E. Holmes, Are diffusion models too simple? A comparison with telegraph models of invasion. *Am. Nat.* **142**, 779–795 (1993).

66. V. Dey, How diffusion models unlock new possibilities for generative creativity (2022). <https://venturebeat.com/ai/how-diffusion-models-unlock-new-possibilities-for-generative-creativity/> (Accessed 6 November 2022).

67. J. P. Schmidt *et al.*, Spatiotemporal fluctuations and triggers of ebola virus spillover. *Emerging Infect. Dis.* **23**, 415 (2017).

68. P. A. Marquet *et al.*, On theory in ecology. *Bioscience* **64**, 701–710 (2014).

69. J. Klinger, J. C. Mateos-Garcia, K. Stathopoulos, A narrowing of AI research? *SSRN* (2020). <https://doi.org/10.2139/ssrn.3698698>. Accessed 9 March 2023.

70. N. Ahmed, M. Wahed, N. C. Thompson, The growing influence of industry in AI research. *Science* **379**, 884–886 (2023).

71. L. Cheng, K. R. Varshney, H. Liu, Socially responsible AI algorithms: Issues, purposes, and challenges. *JAIR* **71**, 1137–1181 (2021).

72. I. Baldini, D. Wei, K. Natesan Ramamurthy, M. Singh, M. Yurochkin, "Your fairness may vary: Pretrained language model fairness in toxic text classification" in *Findings of the Association for Computational Linguistics: ACL 2022* (Association for Computational Linguistics, 2022), pp. 2245–2262.

73. K. R. Varshney, *Trustworthy Machine Learning* (Independently Published, 2022).

74. S. Maitra, "Artificial intelligence and indigenous perspectives: Protecting and empowering intelligent human beings" in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, AIES '20* (Association for Computing Machinery, 2020), pp. 320–326.

75. A. Birhane, Algorithmic injustice: A relational ethics approach. *Patterns (NY)* **2**, 100205 (2021).

76. S. Mohamed, M.-T. Png, W. Isaac, Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philos. Technol.* **33**, 659–684 (2020).

77. Y. Tao, K. R. Varshney, Insiders and outsiders in research on machine learning and society. *arXiv* [Preprint] (2021). <https://doi.org/10.48550/arXiv.2102.02279> (Accessed 30 June 2023).

78. C. H. Trios, J. Auerbach, M. Katti, Decoloniality and anti-oppressive practices for a more ethical ecology. *Nat. Ecol. Evol.* **5**, 1205–1212 (2021).

79. Z. Molnár, D. Babai, Inviting ecologists to delve deeper into traditional ecological knowledge. *Trends Ecol. Evol.* **36**, 679–690 (2021).

80. J. M. Lindsay *et al.*, Characteristics of ringed seal (natchiq) *Pusa hispida* denning habitat in Kotzebue Sound during a year of limited sea ice and snow. *Mar. Ecol. Prog. Ser.* **705**, 1–20 (2023), 10.3354/meps14252.

81. C. J. Schell *et al.*, The ecological and evolutionary consequences of systemic racism in urban environments. *Science* **369**, eaay4497 (2020).

82. V. A. Masterson, S. L. Mahajan, M. Tengö, Photovoice for mobilizing insights on human well-being in complex social-ecological systems: Case studies from Kenya and South Africa. *Ecol. Soc.* **23**, 13 (2018).

83. M. J. Bickford, "Traditional knowledge in marine spatial planning: Addressing epistemological and practical challenges in the Arctic." (2017), (Accessed 9 March 2023).

84. S. R. Carroll *et al.*, The CARE principles for indigenous data governance. *Data Sci. J.* **19**, 43 (2020).

85. J. Ross, B. Belgodere, V. Chenthamarakshan, Large-scale chemical language representations capture molecular structure and properties. *Nat. Mach.* **4**, 1256–1264 (2022).

86. D. Y. Ding, S. Li, B. Narasimhan, R. Tibshirani, Cooperative learning for multiview analysis. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2202113119 (2022).

87. S. D. Eigenbrode *et al.*, Employing philosophical dialogue in collaborative science. *Bioscience* **57**, 55–64 (2007).