Challenges and solutions for ecologists adopting AI

Jessica M. Kendall-Bar[1*], Allison R. Payne[2], Max F. Czapanskiy[3], Nathan Fox[4], Jessica Leivesley[5], Michaela Alksne[1], Beverly French[1], Alice Malivert[7], Gerard O'Leary[8], Avik Biswas[9], Isaac Nealey[10], Rachel H. Parkinson[11], Christopher Lowrie[3], Ismael Perez[12], Eliot Miller[13], Ian Costello[3], Timothy Sainburg[14], Dominique Rissolo[15], Lauren Severance[16], Ritika Mukherji[17], Felipe Montealegre-Mora[18], Kevin H. Nguyen[1], Carla Calvo Tusell[19], William Oestreich[20], Michael Sorenson[1,21], Sergey Gurvich[12], Joseph Walker[1], Allison Kellum[1], Tanya Berger-Wolf[22], Ilkay Altintas[21], Stuart Sandin[1], Nicolas Lecomte[24,25]

[1]Scripps Institution of Oceanography, University of California, San Diego. 9500 Gilman Drive, La Jolla, CA 92093
[2]Department of Ecology & Evolutionary Biology, University of California, Santa Cruz. 100 McAllister, Santa Cruz, CA 95060
[3]Department of Ocean Sciences, University of California, Santa Cruz. 1156 High St, Santa Cruz, CA 95064
[4]Michigan Institute for Data Science, University of Michigan. 500 Church Rd., Suite 600, Ann Arbor, MI 48109
[5]Department of Statistics, University of Toronto. 700 University Avenue, 9th Floor. Toronto, ON M5G 1Z5. Canada.
[7]Bioengineering, IX, Imperial College London. Exhibition Road London SW7 2AZ, United Kingdom.
[8]Department of Computer Science, University of Toronto. 40 St George St, Toronto, ON M5S 2E4, Canada
[9]Salk Institute for Biological Studies, University of California, San Diego. 10010 N Torrey Pines Rd La Jolla, CA 92037
[10]Department of Computer Science and Engineering, University of California, San Diego. 3235 Voigt Dr, La Jolla, CA 92093
[11]Department of Biology, University of Oxford. 11a Mansfield Road, Oxford, OX1 3SZ, United Kingdom
[12]San Diego Supercomputer Center, University of California, San Diego. 9836 Hopkins Dr, La Jolla, CA 92093
[13]American Bird Conservancy. P.O. Box 249. The Plains, VA 20198
[14]Department of Neurobiology, Harvard University. 26 Oxford Street, Cambridge, MA 02138
[15]Qualcomm Institute, University of California, San Diego. 3195 Voigt Dr, La Jolla, CA 92093
[16]Department of Bioengineering, University of California San Diego. 9500 Gilman Drive, La Jolla, CA 92093
[17]Department of Physiology, Anatomy and Genetics, University of Oxford. Sherrington Rd, Oxford, UK
[18]Department for Environmental Science, Policy & Management, University of California, Berkeley. Mulford Hall, 130 Hilgard Way, Berkeley, CA 94720
[19]School of Biological Sciences, University of California, San Diego. 9500 Gilman Drive, La Jolla, CA 92093
[20]Monterey Bay Aquarium Research Institute. 7700 Sandholdt Rd, Moss Landing, CA 95039
[21]Halıcıoğlu Data Science Institute, University of California, San Diego. 3234 Matthews Ln, La Jolla, CA 92093
[22]Department of Computer Science and Engineering, Ohio State University. 2015 Neil Avenue, Columbus, OH 43210-1277
[23]Halıcıoğlu Data Science Institute, University of California, San Diego
[24]Department of Biology, University of Moncton. 18 Antonine-Maillet Ave Moncton, NB E1A 3E9 Canada
[25]Centre d'Études Nordiques. Pavillon Abitibi-Price, Room 1202, Quebec City, Quebec, Canada, G1V 0A6

*Corresponding author: jkendallbar@ucsd.edu

Keywords: artificial intelligence; ecology; cyberinfrastructure; visualization; education

Abstract

Artificial Intelligence (AI) can rapidly process large ecological datasets, uncover patterns, and inform conservation decisions. However, its adoption by ecologists is often hindered by steep learning curves, overwhelming model options with varying transparency, and uneven access to data, code, and technical skills. We led a workshop, EcoViz+AI: Visualization and AI for Ecology, that brought together 35 experts to synthesize a guide for ecologists as they navigate, implement, interpret, and contribute to the fast-evolving AI landscape. Using workshop discussions and experiences as a foundation, this review article synthesizes the opportunities and risks for AI in ecology as well as practical challenges and solutions for adopting AI. Four solutions include: (1) educational resources to help researchers assess the opportunity cost associated with AI compared to traditional methods, (2) communities of practice to combat the overwhelming landscape of AI with knowledge, technical skills, collaboration, and inclusivity, (3) effective visualizations to address the transparency deficit of AI for understanding and communicating results including model outputs, performance, and functionality, and (4) computational resources to ease the implementation burden of AI through shared data, modifiable code, and accessible computing. Our workshop compiled resources, including science communication videos for five AI use cases and repositories for ecology-related AI models and communities of practice. Aligning AI initiatives with broader movements towards interdisciplinary open science and computational literacy will promote inclusivity and the ecological relevance of novel tools, advancing basic research and impactful translational ecology.

Impact Statement
AI can rapidly process large datasets, uncover patterns, and inform conservation decisions. However, its adoption by ecologists is often hindered by steep learning curves, overwhelming model options, and uneven access to data, code, and technical skills. When the benefits of AI outweigh the risks, we argue that ecologists are likely dissuaded from using AI due to practical challenges. We propose four solutions: (1) educational resources to assess the opportunity cost associated with AI compared to traditional methods, (2) communities of practice to foster technical skills and interdisciplinary collaboration, (3) effective visualizations to mitigate AI's transparency deficit by communicating results and functionality, and (4) computational resources to ease the implementation burden of AI through open data, extensible code, and accessible computing. Aligning AI initiatives with broader movements towards interdisciplinary open science and computational literacy will increase the ecological relevance of novel tools, advancing basic research and impactful translational ecology.

1. Introduction

Applications of AI in ecology and evolution are now increasing the speed, scale, and resolution of computational analysis (Christin et al. 2019). AI can help predict species responses based on environmental conditions, facilitate the integration of models and theory for complex systems-level understanding, and generate novel hypotheses to pave the way for conservation (Han et al. 2023). While AI is not necessary in every, or even most, situations, it is nonetheless a powerful, adaptable tool available to ecologists for answering pressing research questions. However, the significant practical challenges to AI-adoption, such as a lack of domain-specific tutorials and communities of practice, and the considerable ethical concerns associated with AI (well described in the literature [Chapman et al., 2024; Cooper et al., 2024; Scoville et al., 2021; Tabassi, 2023]) have generated skepticism and trepidation among ecologists and likely slowed AI adoption. To address this issue, we gathered 35 experts in ecology and artificial intelligence for a week-long facilitated workshop (*EcoViz+AI: Visualization and AI for Ecology; ecoviz-ai.github.io* (Kendall-Bar 2024a); see Supplementary Material for workshop details and outputs). We defined AI broadly to include the spectrum of machine learning and deep learning models available to ecologists, although strict definitions of AI may preclude machine learning and even deep learning (Wang 2019, Sheikh et al. 2023).

For this overview, we first provide a summary of the current opportunities and risks for using AI in ecological research. We then present a review of the key challenges to AI adoption for ecologists followed by specific examples and ethical considerations for key solutions to these challenges. We organized these solutions around access to and development of educational resources, communities of practice, effective visualizations, and computational resources. Overall, we hope this review will help ecologists who are considering adopting AI methods navigate the tools available to them and apply them with responsibility and rigor.

Challenges regarding model selection, implementation, interpretation, and communication are not unique to AI, but its use exacerbates challenges posed by traditional methods and computational analysis more broadly. Therefore, we emphasize that AI's increased transparency deficit and implementation burden call for a renewed emphasis on technical training, community, and science communication. Our solutions section details AI-specific initiatives, cultural shifts, and best practices that help make AI in ecology more approachable, relevant, and understandable, while aligning synergistically with broader movements towards open science and technical literacy.

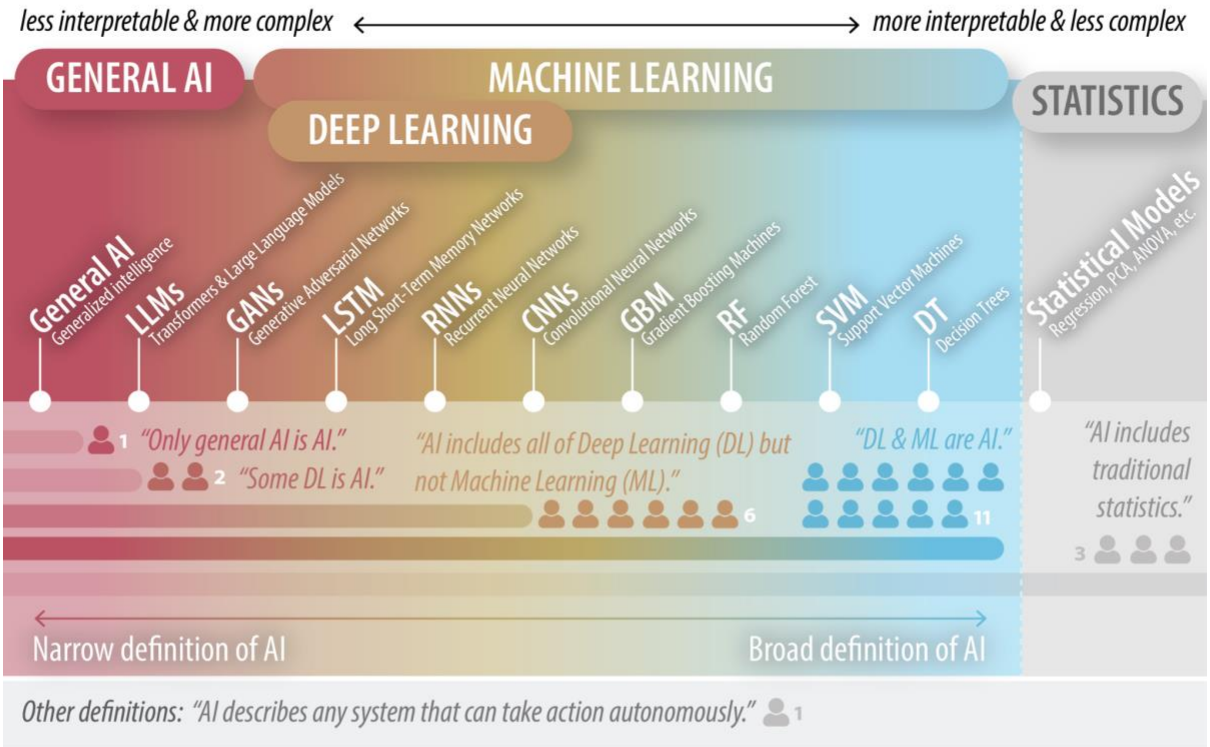

Fig. 1. A diagram describing our workshop participants' definitions of AI, showing a spectrum of models from less interpretable and more complex models on the left (darker red values) to more interpretable and less complex on the right (lighter blue values). Icons next to a definition represent votes from workshop participants. The spectrum is divided into four categories: **General AI** (or strong AI), **Deep Learning** (DL including Large Language Models [LLM], Generative Adversarial Networks [GAN], Long Short-Term Memory Networks [LSTM], Recurrent Neural Networks [RNN], and Convolutional Neural Networks [CNN]), **Machine Learning** (encompassing DL models as well as Gradient Boosting Machines [GBM], Random Forest [RF], Support Vector Machines [SVM], and Decision Trees [DT]), and **Statistics** (including e.g. linear, logistic, and multivariate regressions, Principal Component Analysis [PCA], and Analysis of Variance [ANOVA]).

2. Opportunities and risks of AI applications in ecology

AI, defined broadly as a spectrum of machine learning to deep learning models (Fig. 1), can be applied to ecological research through: (1) ***data processing***, such as labelling, annotating, clustering, or filtering raw data, (2) ***inference***, where processed data can be used to answer ecological or evolutionary questions via testable hypotheses, and (3) ***decision-making***, where the answers to these questions can serve as the basis for policy or management recommendations. For example, sleep studies of wild animals could apply AI to help assign labeled sleep states based on electrophysiological data (AI for data processing), as demonstrated by Allocca et al. (2019) and Vallat & Jajcay (2020). Whether or not AI was used to assist sleep scoring, the resulting labelled data could be fed to an AI-based habitat suitability model to identify the locations within an animal's range that are best suited for sleep (AI for inference; e.g., (Saarenmaa et al. 1988, DŲeroski 2009). Finally, a conservation-focused AI model could incorporate this knowledge with other aspects of animal behavior to recommend the prioritization of habitats for conservation (AI for decision-making) (Silvestro et al. 2022).

We discuss the opportunities and risks of AI applications across these three stages: data processing, inference, and decision-making. Early in the field of ecology, inference primarily relied on strong experimental design and statistical methods, while data processing and decision-making were led by human experts. While ecologists sought to use statistical models to describe biological systems and observational processes (Hilborn and Mangel 1997), relatively small datasets allowed manual labeling and annotation (Besson et al. 2022). Traditionally, conservation and management decisions have also relied on the consensus of domain experts (Kirlin et al. 2013). However, the increasing size and complexity of datasets have augmented the burden of manual processing, making it harder for experts to manually process data or to consider all sources of information relevant to conservation (Besson et al. 2022). These complex, multidimensional datasets also present unique opportunities for AI even in the traditionally statistics-dominated space of inference. For each of these three tasks, we discuss emerging uses of AI, particularly pertaining to machine learning and deep learning, as alternatives for more traditionally implemented statistical or manual methods.

2.1. AI for data processing

Opportunities. With relatively fast and inexpensive models, AI approaches allow ecologists to rapidly process large datasets with algorithms to detect or classify habitat types, species, or behaviors (Schirpke et al. 2023). AI accelerates tasks like labelling sleep states, annotating the edges of coral colonies, and detecting anomalies in time series data, enabling more data to be processed (i.e., increases scalability), but can lower prediction accuracy compared to manual methods (Mosqueira-Rey et al. 2022). When compared to statistical or signal processing techniques for processing data, AI can improve accuracy in addition to speed, introducing the possibility of edge computing, where models can be deployed to perform reliable real-time detection in the field or while sensors are deployed on animals (Wasimuddin et al. 2020, Yu et al. 2024).

The use of computer vision convolutional neural networks (CNNs) has improved scientists' capacity to process large amounts of video, acoustic, and movement data (Christin et al. 2019, Lauer et al. 2022, Yu et al. 2022) with promise for improving conservation and biodiversity monitoring (Galaz García et al. 2023). For example, new developments in near-real-time automated processing of ship-based thermal imagery can alert captains to potential ship strikes in time to avoid endangered whales (Baille and Zitterbart 2022). AI tools including WildBook, HappyWhale, FathomVerse, Seek and iNaturalist, and Merlin engage the public in the detection and classification of individual plants and animals, enabling citizen science and local community engagement (Sullivan et al. 2014, Berger-Wolf et al. 2017, Katija et al. 2022, Manderfield 2022, Cheeseman et al. 2023). The cyberinfrastructure supporting new AI techniques also accelerates and enables traditional signal processing of large datasets. For example, the National Data Platform was built for open and equitable data access and computing for AI, but its benefits extend to researchers using other methods as well (Parashar and Altintas 2023).

Risks. Even commonplace AI methods for processing ecological data can have significant downstream consequences. For instance, AI models that process video to track locomotion and analyze the sublethal effects of pesticides on insect behavior (Parkinson et al. 2022) carry the risk of a behavioral anomaly being overlooked. AI models that automate species detection for camera traps (images) or automated recording units (audio) (Beery et al. 2019, Tuia et al. 2022) carry the risk of overlooking the presence of an endangered species. Issues with low recall – the ability of a model to identify all relevant instances – could lead to the exclusion of that area from conservation efforts. Conversely, a "cry wolf" effect may also undermine these models and erode public trust. Consistent false alerts from automated whale detection systems could result in those systems being ignored (Baille and Zitterbart 2022). This highlights the importance of optimizing precision – the proportion of correctly identified positive cases out of all positive detections – to reduce false alarms. Furthermore, if automated detections are the basis for further inference, the impacts of a false result are amplified. For example, a boosted decision tree that aims to identify a rare behavior could boast high accuracy despite low sensitivity. Models to detect events should therefore consider the tradeoff between sensitivity/recall and specificity/precision. For example, models to detect sleep when rare, such as for seals at sea, can optimize for sensitivity to prevent the oversight of a critical resting habitat (Kendall-Bar et al. 2023).

Promising solutions. Algorithmic bias can be addressed through averaging techniques that correct for class imbalances where certain types of data are overrepresented (Beery et al. 2021). "Human-in-the-loop" methods include human input in the model training process, incorporating user feedback to refine and supervise classifications (Mosqueira-Rey et al. 2022, Wu et al. 2022). These methods can improve the accuracy, precision, and transparency of automated tools and foster trust in the model predictions. For example, TagLab, an AI-driven tool for coral imagery segmentation, uses semi-automatic methods to maintain high accuracy and ease the burden of manual annotation (Pavoni et al. 2021). Similar manual review is enabled in sleep scoring software (Allocca et al. 2019) and camera trap annotation (Miao et al.

2021). The inclusion of human review helps mitigate the consequences of model error and allows for an opportunity to consider the practical, scientific, and ethical implications of the model output.

2.2. AI for inference

Opportunities. AI methods can provide an alternative to traditional statistics for establishing relationships between environmental variables and biological systems. These relationships can be studied at many ecological scales, from individuals (physiology and behavior) and populations (abundance and density) to communities (species interactions) and ecosystems (species and their environment). For instance, species distribution models (SDMs) have been extensively used in ecology to quantify occupancy, density, and distribution changes, limits, and expanses (Elith and Leathwick 2009, Grace 2024). SDMs that associated species' presence and absence with the environment using statistical methods (e.g., logistic regression) originated in the 1980s and gained broader adoption over the following decades (Elith and Leathwick 2009). However, these tools often struggled to capture the complex, non-linear relationships between species' distributions and their environment, leading many ecologists to adopt ML algorithms including boosted regression trees (Elith et al. 2008) and MAXENT (Phillips et al. 2006). Though generally considered a machine learning algorithm, Renner and Warton demonstrated the equivalence of MAXENT to a Poisson regression model (Renner and Warton 2013). More recently, deep learning algorithms (e.g., convolutional neural networks) have been applied to the relatively few occurrence datasets large enough for these data-hungry methods (Benkendorf and Hawkins 2020). These algorithms hold potential for advancing SDM methodologies by learning novel features from multidimensional environmental data (Beery et al., 2021).

Risks. AI for inference in ecosystem monitoring is rapidly becoming a reality (Sethi et al. 2020, Galaz García et al. 2023), raising important concerns about how systematic errors in AI alter the derived ecological insights. As causal inference methods develop alongside AI, the ability to make predictions may outpace the ability to explain the mechanisms behind the predictions (Grace 2024). SDMs require true observations of species and robust environmental sampling, each of which has various sources of error and uncertainty that must be clearly identified, explained, and controlled for (Beale and Lennon 2012, Beery et al. 2021). Black-box deep learning models further exacerbate the challenges associated with aggregating multimodal ecological data such as spatial and temporal autocorrelation, differences in sampling protocols, and other caveats related to environmental feature generation and ground-truthed datasets (Beery et al. 2021). Challenges regarding how to fit models and interpret their results are not unique to AI, but its use can exacerbate issues already present with more traditional statistical methods. Even with traditional statistics, ecologists sometimes select improper methods for estimating parameters of mixed effects models (Bolker et al. 2009). For example, the most used package for fitting mixed effects models in R, lme4 (Bates et al. 2015), does not provide functionality for predicting intervals because of disagreements regarding standard error estimation. These challenges for interpretation may be compounded when using AI methods.

Promising solutions. Improvements in AI interpretability are providing insight into the internal mechanisms of models. For example, explainable AI methods can rank feature importance to provide global explanations for overall model predictions and local explanations for individual model predictions ((Ryo et al. 2021, Alicioglu and Sun 2022, Molnar 2024, Zhang et al. 2025); see Fig. 3 and *Effective Visualizations* section (Ryo et al. 2021, Alicioglu and Sun 2022, Molnar 2024, Zhang et al. 2025). Initiatives to encode domain knowledge explicitly into algorithms (knowledge-guided and model-based AI) aim to improve trust in AI models and address these inherent risks and complexities (Doll et al. 2012, Bishop 2013, Swischuk et al. 2019). Similar recent efforts within ecology leverage the rich biological structure that underlies taxonomy and phylogeny to build upon the generic OpenAI model CLIP to create the biology-specific BioCLIP model (Stevens et al. 2023). In addition to improving the accuracy and context of image-based species identification, knowledge-guided tools, that seek to incorporate domain-specific logic, can lead to the generation of new evolutionary hypotheses. They can reveal missing links in high-dimensional networks in complex systems, such as suggesting intermediate phenotypes based on phylogeny (Han et al. 2023, Stevens et al. 2023).

2.3. AI for decision-making

Opportunities. Given the complexity and scale of ecosystem management, AI is increasingly being used to support decision-making for conservation (Scoville et al. 2021, Lapeyrolerie et al. 2022) and sustainable management of natural resources (Lindkvist et al. 2017, Ebrahimi et al. 2021, Montealegre-Mora et al. 2023). Conservation prioritization has traditionally involved a spectrum of approaches, from nonalgorithmic processes like the California MPA Network Blue Ribbon Task Force, which heavily relied on expert knowledge and stakeholder engagement (Kirlin et al. 2013), to algorithmic tools such as MARXAN, a widely used decision-support software for optimizing reserve design (Ball et al. 2009). Newer AI-driven platforms like CAPTAIN leverage AI and specifically reinforcement learning to model and predict biodiversity outcomes under varying conservation strategies (Silvestro et al. 2022). For instance, reinforcement learning, where the model learns by receiving feedback in the form of rewards or punishments for its actions, has been identified as particularly useful in fishery science when fish stocks display complex patterns in recruitment dynamics (Lapeyrolerie et al. 2022, Chapman et al. 2023, Kühn et al. 2024, Montealegre-Mora et al. 2024). Automated methods, repeatable workflows, and transdisciplinary research are being used to review and aggregate large, multimodal datasets into biodiversity syntheses that assess the status of ecosystems at a global scale (Galaz García et al. 2023, Berger-Tal et al. 2024). These global analyses have the potential to inform international policy, supporting initiatives like 30 by 30 and other worldwide conservation efforts (Scoville et al. 2021).

Risks. When applied to conservation decisions, AI carries serious implications, especially if it prioritizes areas and communities based on biased and uneven data (Chapman et al. 2021, Scoville et al. 2021). In forming the design and objectives of these models, researchers optimize for biodiversity or conservation outcomes that often reinforce predominant colonial conservation paradigms that neglect the needs of local communities and generational knowledge from Indigenous Knowledge Systems (Al-Mansoori and Hamdan 2023, Layden et al. 2024). As mentioned for SDMs, black-box models compound these issues by adding a layer of obscurity. These algorithmic approaches can encode biases that increase the likelihood of colonialist research that neglects the parties most affected by management decisions (Han et al. 2023). These ethical risks also arise in the application of AI data processing outputs to conservation and management decisions. For example, the identification of illegal environmental activities (e.g. poaching, fishing in no-take marine reserves) could eventually serve as the basis for future predictive policing practices (Mporas et al. 2020, Swartz et al. 2021). Though this could be viewed as an opportunity, AI for policing practice is fraught with ethical challenges (Davis et al. 2022), including but not limited to perpetuating biases and the extreme negative consequences of errors in detection. Similar to AI for conservation prioritization, the use of AI for automated surveillance and decision-making poses significant social and ethical risks beyond ecology.

Potential solutions. When AI is involved in decision-making, long-term partnerships with local communities become even more important so that local voices can review what goes into and comes out of models to provide feedback and adjust recommended decisions. Clear data-sharing agreements are needed to ensure informed public engagement that supports environmental justice and improves conservation governance of local communities (Layden et al. 2024). To mitigate compounded ethical issues, models can explicitly include data from social science methods like participatory mapping, surveys, and interviews to promote the inclusion of more diverse knowledge, values, and identities (Bennett et al. 2017). Once algorithmic solutions are suggested by AI, distribution equity assessments using similar social science methods, like focus group discussions or community-based research on potential outcomes, can help select among several near-optimal solutions to prioritize an equitable distribution of cost and benefits (Kockel et al. 2020). Promoting equity in the use of AI for conservation and management will require the active integration of co-design principles, social science, procedural justice, and equity assessments (Benyei et al. 2020, Chapman et al. 2021, Hsu et al. 2022, Oestreich et al. 2024).

3. Practical challenges: barriers to AI adoption in ecology

Alongside the ethical and scientific concerns outlined previously, ecologists encounter practical challenges in understanding when and how to implement AI over more traditional manual or statistical methods (Fig. 2C). While many of these challenges are present in learning any technical skill, we highlight how AI can exacerbate existing burdens. We describe key practical challenges an ecologist faces when applying AI to assist with labelling, clustering, or filtering datasets (AI for data processing), transforming data into knowledge (AI for inference), or translating knowledge into action (AI for decision-making):

3.1. Opportunity cost

Based on the ecological question at hand, what are the costs and benefits of using an AI model as opposed to traditional methods? While it takes time and effort to learn any new skill or method, AI models have a steeper learning curve for ecologists. AI relies on complex computational infrastructure and can require specialized knowledge of algorithms to navigate the rapidly evolving landscape of tools and methods. Additionally, most resources use programming languages that are less familiar to ecologists (e.g., Python instead of R). The scale of investment in learning to use AI can vary based on the researcher's career stage, level of experience, access to collaborators, and the availability of open-access educational materials. Current academic incentive structures (publications, funding, and job opportunities) favor the development and use of novel AI models over the creation of accessible educational resources (i.e., tutorials, blogs, free online courses). The people who are best suited to create educational materials may therefore not have the time or resources to do so.

3.2. Overwhelming landscape

Considering your data and question, which AI model, if any, is relevant and could provide value? A rapidly evolving landscape of AI models can be overwhelming to an ecologist seeking to responsibly analyze their data. They may be tempted to use traditional statistical methods, which may be more familiar and invite less skepticism than AI models, therefore missing opportunities to use AI to advance their research. Using AI requires keeping pace with rapidly advancing technologies, new architectures, and a proliferation of models with varied baseline assumptions and caveats. Researchers must carefully consider the benefits and drawbacks of increasingly complex models in terms of decreased transparency and increased computational load (Pichler and Hartig 2023). This can be especially overwhelming for researchers who cannot draw on the collective knowledge of a community of experts. For cutting-edge technical tools such as AI, students often learn informally from more junior mentors, such as senior grad students and postdocs, who are not always recognized for their efforts (Higino et al. 2023). Academic research incentivizes mentorship at the lab or institutional level, rather than to the broader, interdisciplinary research community.

3.3. Transparency deficit

How can we understand the model's performance? How can we understand and trust how the model came to its answer? AI models, particularly those using deep learning, introduce unique challenges in interpretability and explainability. This makes it difficult to fully understand the reasoning behind their predictions compared to traditional methods, which are often perceived as being more transparent and grounded in well-understood statistical principles. Tools for interpretability in AI are rapidly evolving but can be challenging to navigate without technical expertise. These tools often lack standardization, leaving researchers with limited guidance on how to evaluate and trust AI predictions effectively. Because ecologists are not typically formally trained in AI methods, they may not be familiar with visualization tools that can support model interpretation or the science communication surrounding the data collection and the model's functionality, caveats, and performance.

3.4. Implementation burden

How do we acquire the resources and data management systems needed to run our models? How hard is it to run the model on new data? How should we share the workflows and models we produce? Unlike traditional statistical approaches, AI models often require complex preprocessing steps, significant computational power, and fine-tuning of hyperparameters, which can create barriers for researchers with limited technical expertise, indecipherable code repositories, or lack of access to computational resources. Once a researcher has decided to use a particular model, they must alter the model to fit the structure and scale of their dataset. This may mean reconfiguring data pipelines and workflows, engineering features, iteratively evaluating model performance, and eventually scaling up this analysis to larger datasets. This work can be limited by unavailable or unclear data and code, as well as lack of access to computational resources like cloud computing and data management systems (Allen and Mehler 2019). Once researchers have created useful tools, they also face several obstacles when seeking to share the code, workflows, and models related to reuse concerns, disincentives, and knowledge barriers (Gomes et al. 2022). This perpetuates the cycle and creates challenges for incoming researchers seeking to understand the opportunity cost associated with using AI in ecology.

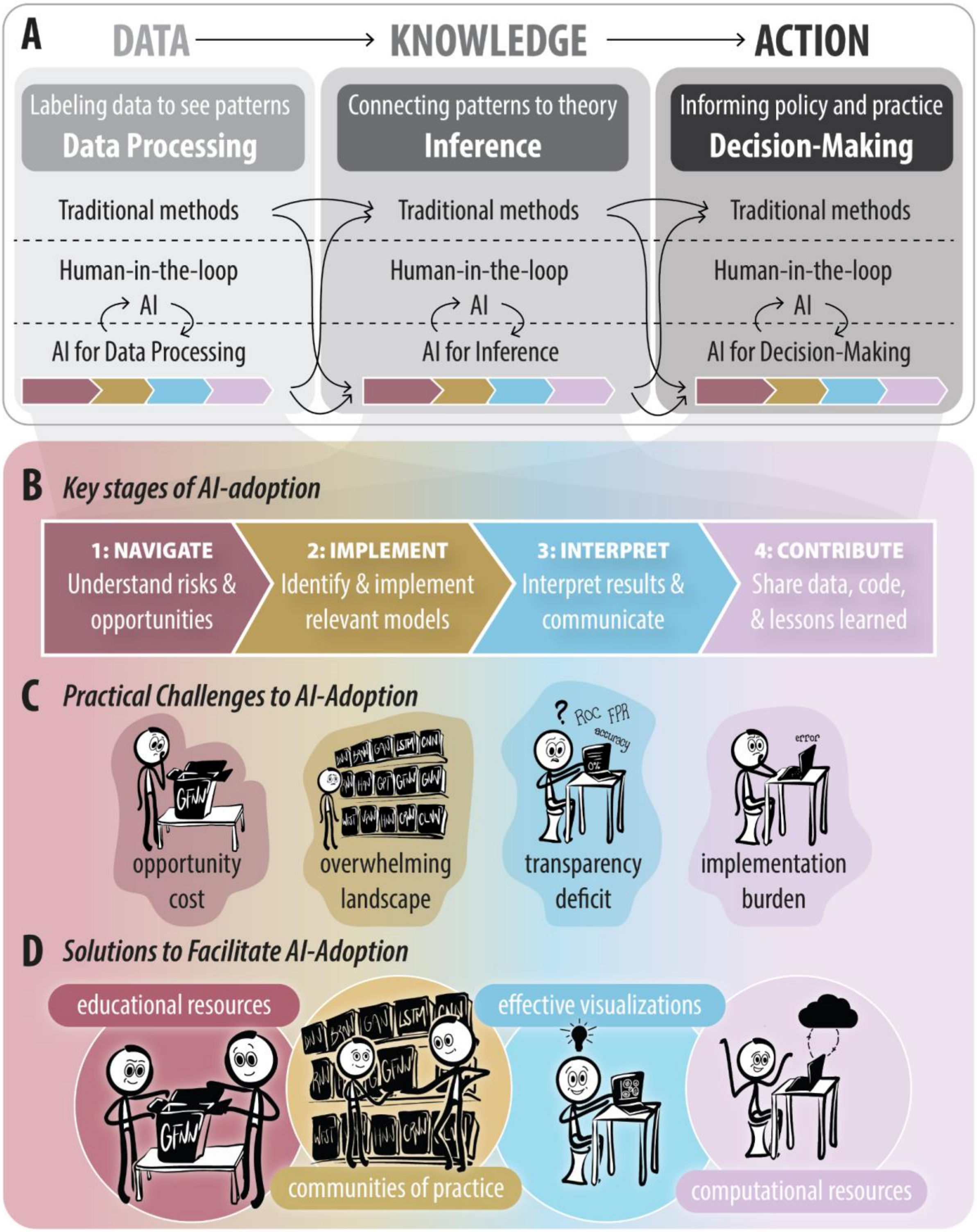

Fig. 2. This synthesis figure outlines a roadmap for AI-adoption in ecology as it is used for data processing, inference, and decision-making. A) Often, human-in-the-loop technologies improve workflows that transform data into knowledge that can inform action by mitigating the scientific and ethical pitfalls of AI. B) The key stages for ecologists adopting AI: (1) Navigate the landscape of risks and opportunities, (2) Implement identified relevant models, (3) Interpret and communicate the results, and (4) Contribute data, code, and lessons learned back to the research community. These stages map (C) key practical challenges onto (D) key solutions: (1) opportunity cost (unclear benefits and costs) of using AI: educational resources, (2) overwhelming landscape of model choices: communities of practice, (3) transparency deficit: effective visualizations, and (4) implementation burden: computational resources.

4. Practical solutions: facilitating AI adoption in ecology

These practical challenges, while considerable, must be overcome if ecologists want to responsibly leverage opportunities offered by AI. We used the wealth of expertise across career stages at our workshop to identify practical solutions that collectively ameliorate key challenges. These solutions align four key challenges—opportunity cost, overwhelming landscape, transparency deficit, and implementation burden—with corresponding interventions: educational resources, communities of practice, effective visualizations, and computational resources. The solutions aim to guide ecologists as they navigate AI-related risks and opportunities, implement relevant models, interpret and communicate their results, and contribute data, code, and lessons learned with their research community.

For each challenge, we identify an overarching solution and provide ecologists new to AI with a starting point for finding resources and initiatives. We pair these examples with broader recommendations for AI practitioners across experience levels to collectively advocate for cultural shifts that will facilitate the responsible adoption of AI in ecology. Ecologists encountering AI for the first time benefit most from accessible educational resources and collaborative communities of practice that demystify AI and provide mentorship. For researchers scaling up their use of AI, visualizations play a critical role in helping researchers understand and communicate data and the outputs, performance, and function of AI models. As researchers advance, they gain valuable expertise that they can share back with their research community by sharing data, code, and lessons learned. For each solution, we invite researchers experienced in AI to lower barriers for future ecologists by advocating for specific cultural shifts and incentives that reward open sharing of data, tools, and expertise.

4.1. Educational resources

When ecologists begin to use AI, they often benefit from educational resources, especially those that are open-access and available online. Educational resources to learn AI range in accessibility, investment, and impact from informal (i.e., blog posts, YouTube videos, tutorials, review papers) to formalized courses, workshops, and fellowships. Informal resources are an excellent entry point for students, allowing them to learn for free at a flexible pace. While informal resources to learn AI are plentiful, high-quality ecology-specific resources are rare. Some examples include extensive reviews of how deep learning is being used in ecology with practical guides for model selection (Borowiec et al. 2022, Pichler and Hartig 2023), conceptual tutorials of deep learning for biologists (Aurisano et al. 2017), or specific coding examples of using AI and machine learning in R (Lefcheck 2015) and Python (Gray 2024). Ecologists may want to explore tools such as OpenSoundscapes (Lapp et al. 2023), which provide extensive documentation and tutorials to walk ecologists through the process of training a Convolutional Neural Network to identify sounds in audio data (Lapp et al. 2024). When employed cautiously, ChatGPT and other generative AI tools can lower the barrier to entry to Python for ecologists or biologists who may lack formal training, especially for simple tasks such as translating syntax from another more familiar language like MATLAB or R (Lubiana et al. 2023). However, while helpful to get started, ChatGPT alone is inadequate to guide the responsible selection and implementation of a model. Informal, exploratory learning is an important first step for ecologists to understand the opportunity cost associated with implementing AI.

Formal educators, courses, and programs can provide ecologists with a nuanced understanding of the field, as well as tailored guidance for bespoke data processing pipelines. Ecologists may consider formal educational resources such as Massively Open Online Courses (also known as MOOCs) on machine learning (Ng 2024), but they are not typically aimed at ecologists and may not be accessible in terms of pricing or prior knowledge. If funding is available, in person courses such as the Oxford Research Software Engineering program can introduce rigorous best practices for Python programming to scientists, which can provide the skills needed to implement AI and produce useful scientific software (Course website: (OxRSE 2024a); Github: (OxRSE 2024b)). Ecology-specific opportunities with in-person engagement, while harder to come by, can be transformational for ecologists new to AI, by allowing learners to engage directly with instructors and adapt models to their own datasets. In particular, the Computer Vision for Ecology (CV4E) workshop led by Beery et al. (Beery et al. 2023, Cole et al. 2023) combines formal instruction with coding support so that students come away with a conceptual understanding of the model to accompany their code and model outputs. Ecologists who are considering significant use of AI in their work may seek fellowships such as those offered by Schmidt Sciences and the

Allen Institute (Allen Institute 2020, Schmidt Sciences 2022). These opportunities can also give ecologists the time and resources to receive formal training and connect with expert mentors in the field of AI. However, as educational opportunities increase in support, they decrease in accessibility in the form of greater costs and fewer spots available.

Cultural shifts towards greater emphasis on programming education are already underway, as ecology departments increasingly hire dedicated teaching faculty, research software engineers, and data science educators (Harlow et al. 2020). Moving forward, academia should reward the creation of accessible educational resources by recognizing these contributions in tenure and promotion decisions. Increased value and opportunities for interdisciplinary co-instruction by ecology and computer science educators can improve the quality and availability of formal and informal educational resources for ecologists. Finally, education and the resulting increase in understanding will provide ecologists with tools to critically evaluate research using AI. These shifts are essential not only for AI adoption but also for advancing computational literacy more broadly, enabling ecologists to confidently integrate AI, when applicable, alongside other scalable computational methods in their research.

4.2. Communities of practice

Educational opportunities provide a launching point, but ecologists who are new to AI may struggle to set their work within a broader, collaborative context. In ecology and other fields, communities of practice have been a valuable tool for scaling and tailoring education and mentorship opportunities. Communities of practice are social structures composed of individuals who share a common domain of interest and collectively enhance their expertise through sustained interactions and knowledge exchange (Wenger 2011). We argue that ecologists new to AI, especially "advanced beginners," may benefit from joining a community of practice where they can interact with domain experts (Stevens et al. 2018). Communities of practice for AI in ecology allow members to share technical knowledge, provide interdisciplinary expertise, and create inclusive environments across expertise levels.

Communities of practice can provide important support for scientists in fields that, like ecology, require intensive technical skill-building (Stevens et al. 2018). For example, organizations like PyOpenSci (pyOpenSci 2024) and rOpenSci (2024) create supportive environments where scientists can learn how to practice programming and open science. Communities of practice are also key in interdisciplinary fields to understand gaps and areas of synergy between fields. In the AI domain, interdisciplinary communities like Climate Change AI (Climate Change AI 2024), the NSF-funded COnvergence REsearch (CORE) Institute at San Diego Supercomputer Center (NSF CICORE 2024), and the NSF- and NSERC-funded AI and Biodiversity Change Center (ABC Global Climate Center, 2024) bridge disciplines between computer science, climate science, and ecology. Such links have improved methods to monitor, analyze, and assess changes in global biodiversity (MacWilliams et al. 2024). Organizations like the Turing Institute and professional societies like NeurIPS also provide structures for interdisciplinary collaboration to establish guiding principles for the ethical use of AI (NeurIPS 2024, Turing Institute 2024). While not specific to AI, the National Center for Ecological Analysis & Synthesis (NCEAS) seeks to intentionally foster the Environmental Data Science community through events like their inaugural Summit in 2023 (NCEAS 2023). Ecologists who would like to use AI can benefit from engaging in communities relevant to their interests and goals that have strong community agreements (Bates et al. 2024), dedicated facilitators (Cravens et al. 2022), and inclusive, engaging events (Woodley and Pratt 2020). Choosing to join but also contribute to intentional, inclusive spaces can help counteract pervasive challenges associated with the impostor syndrome and STEM (Bates et al. 2024).

While vital for ecologists navigating an evolving AI landscape, in-person opportunities with ample funding for travel and accommodation are inherently exclusive and involve difficult ethical decisions regarding who gets invited. This is especially important to consider in cases where participatory decision-making informs conservation through communities of practice focused on translational ecology (Lawson et al. 2017). It is also important to consider the pros and cons of social learning, as has been well studied in the field of behavioral ecology, where there is the potential for stagnation and inertia without active inclusion of diverse perspectives (Laland and Williams 1998, Johnstone et al. 2002, Barrett et al. 2019). To prevent this stagnation, AI in ecology can serve as an opportunity to invite expertise and best practices across disciplines, cross-pollinating across groups, including within a single university. At multiple scales, we hope that communities of practice can be increasingly used to foster science identity and agency as new programmers learn to leverage AI.

4.3. Effective visualizations

As computational analyses scale and AI models become more complex, ecologists gaining familiarity with AI can benefit from effective data visualization to understand patterns in the data, interpret model functionality, communicate model outputs, and foster transparency with stakeholders. The use of AI, especially black-box deep learning methods, can exacerbate the lack of transparency associated with scientific research; this calls for a renewed emphasis on effective visualizations for diverse audiences. While visualization is important to master at all career stages, leveraging its impact for applied AI has the potential to better engage scientists, decision makers, and the public (Kendall-Bar et al. 2024).

The design and intent of these visualizations depend heavily on an ecologist's target audience. We present two primary purposes for the visualization of AI in ecology: exploration and explanation (Fig. 3). Exploratory visualizations for AI include those dedicated to exploring the data and the model to a narrow audience of experts, intimately familiar with the data and questions. These visualizations are used to uncover patterns in the data, identify key features, understand model performance, and diagnose model functionality (Fig. 3A). For example, an ecologist seeking to visualize data prior to fitting an AI-based species distribution model may first examine satellite imagery or maps with color-coded sensor measurements (Fig. 3A1a) to obtain processed features for model inputs (Fig. 3A1c). Visualizations of ground-truth presence/absence data from manual censuses can help visually assess model accuracy (Fig. 3A1b). After fitting and visualizing the model (Fig. 3A1d), AI predictions of habitat suitability can be assessed against this ground truth, e.g. through a receiver operating characteristic (ROC) curve (Fig. 3A2a). Such a curve helps identify a habitat suitability threshold for the model (above which it is considered habitable) that optimizes for tradeoffs in model performance, between a sensitive model (measured via true positive rate) and one with low false positive rate (or high specificity). Model performance for a given suitability threshold can be visualized with a confusion matrix (Fig. 3A2b). Colors for these performance metrics (true/false positives/negatives) can then be arranged across space (Fig. 3A2b: Spatial accuracy) or time, in the case of time series data. Overall model functionality as well as individual model predictions can then be explained through bar plots that rank the relative contributions of each feature (Fig. 3A3a-b; see supplemental text for more details on Explainable AI methods).

Explanatory visualizations offer a curated presentation of data, key results, model outputs, and implications paired with contextual information to effectively guide a broader audience less familiar with the dataset and question (see example in Fig. 3B). Explanatory visualizations build upon standalone versions of plots, line charts, or heatmaps useful for data exploration, often by adding annotations, infographics, scientific illustrations, voiceover narration, or data-driven animations. The perceived complexity of AI models may alienate or foster distrust with local community partners or decision makers, making it more important to visually explain the scientific basis of the model's use and its proposed decisions. Interactive web-based data browsers can increase trust and transparency regarding the use of AI in ecology by allowing direct engagement with the public (HappyWhale: (Cheeseman et al. 2017); FlukeBook: (Blount et al. 2022)) or decision makers through decision support tools designed for dynamic management (Welch et al. 2020). While interaction can be valuable for those closely involved, short videos can incorporate visualizations and narration provide a wide-reaching, standalone overview of a topic (Kendall-Bar et al. 2021, Kendall-Bar 2023).

Shaping a narrative through visualizations involves ethical decisions about what data to highlight, simplify, or omit (Walsh 2015). Researchers can accurately depict results and uncertainty with responsible visualizations that foster trust in science and broaden who has access to information about AI in ecology, supporting the critical role of science communication (Longdon 2023). While not exclusive to AI, visualizations can present valuable opportunities for AI-related science communication and stakeholder engagement with the wide array of inherently visual datasets in ecology such as computer vision for camera traps and aerial imagery or physics-based AI models for weather, flood, or fire simulation (Kendall-Bar et al. 2024). To promote technical literacy of AI among ecologists and collaborators, institutions and funding agencies must more formally incentivize science communication (Swain 2023). Recognizing visual storytelling as a valued contribution—on par with traditional metrics like publications—can incentivize researchers to invest time and effort in creating widely accessible, high-quality visualizations that responsibly and effectively communicate their use of AI.

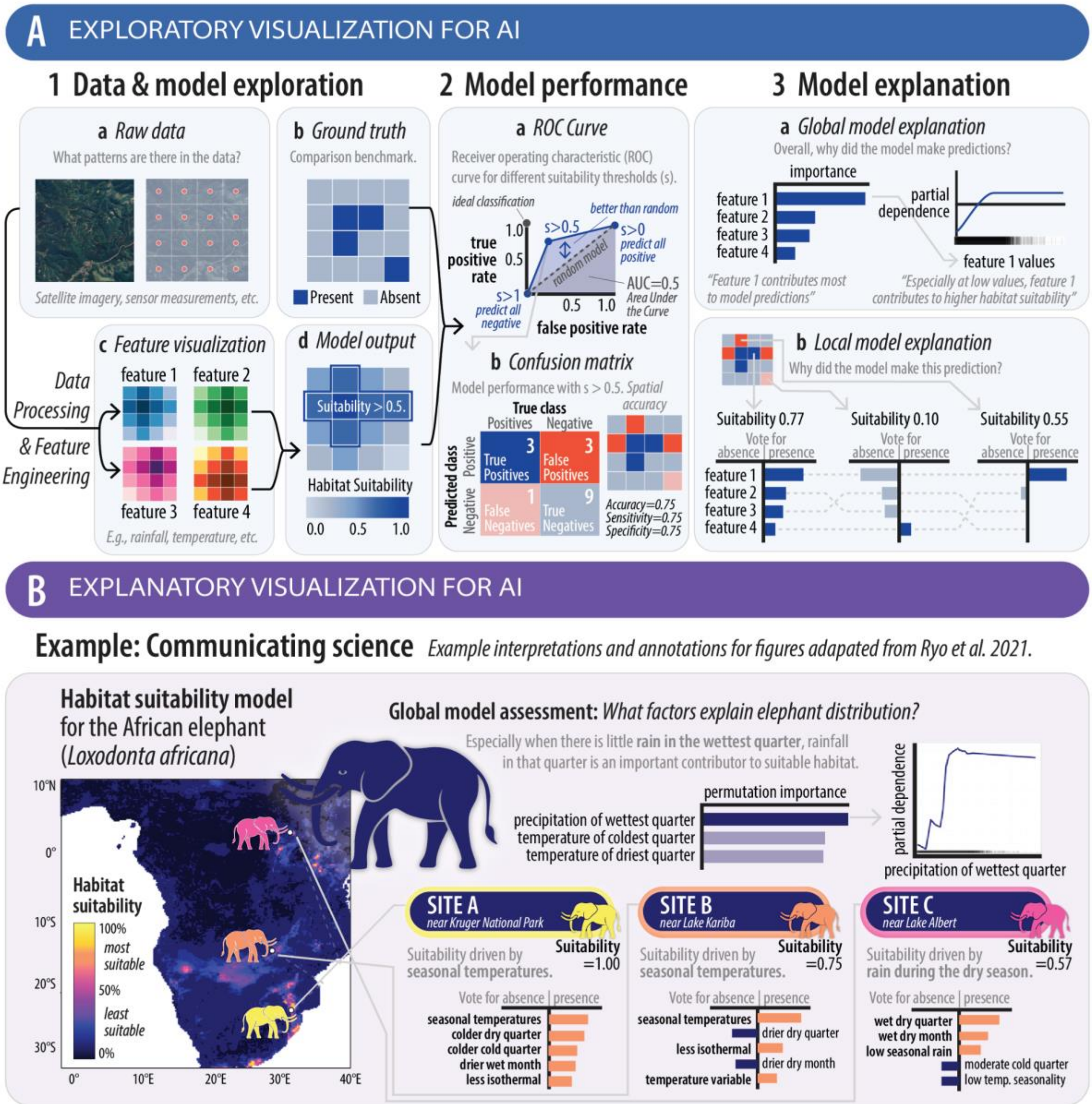

Fig. 3. **(A) Exploratory visualizations** to understand AI models include: **(1) Data and model exploration:** (a) Raw data visualization, including satellite imagery and geospatial representations of species' presence/absence, (b) Feature visualizations (e.g., rainfall, temperature) used as model inputs, (c) Model output geospatial predictions of habitat suitability; **(2) Model performance:** (a) Receiver Operating Characteristic (ROC) curve illustrating the tradeoff between true positive rate and false positive rates at different habitat suitability thresholds (s), (b) Confusion matrix for a suitability threshold of 0.5, showcasing true/false positives and negatives, with accuracy ((TP+TN)/(P+N)), sensitivity (i.e., true positive rate; TP/(TP+FN)), and specificity (TN/(FP+TN)); **(3) Model explanation:** (a) Global explanations highlighting feature importance and partial dependence plots to interpret the contributions of key variables, (b) Local explanations illustrating feature-level contributions for individual predictions (possible using explainable AI methods like LIME or SHAP with bar plots to rank feature importance for specific predictions). **(B) Explanatory visualization** composite infographic with plots and data adapted from Ryo et al. 2021 to provide example graphics, annotations, and interpretations that can guide the viewer to better understand AI model outputs. For additional details and references for LIME and SHAP, see the Supplemental Text.

4.4. Computational resources

As researchers narrow in on the methods essential to their question, their key limitation may shift to their access and expertise with computation. Here, we define computational resources broadly to encompass the hardware and software to train and run AI models, including openly available labelled datasets, transferrable AI models (i.e. usable code repositories), data management systems, and cloud computing resources. Ecologists' use of computational methods, not just AI, is hindered by the lack of formal training on sharing data, curating code repositories, managing datasets, and accessing supercomputers (Stockwell et al. 2000). Open science and its growing support by funding agencies aim to democratize AI and accelerate computational science (Parashar and Altintas 2023, Würthwein 2024). Ecologists can act as partners in these efforts to better connect domain-specific needs and existing initiatives with new tools and best practices from AI and computer science.

Due to the large size of datasets and associated computing requirements, the use of AI is limited without cloud computing. Ecologists who want to use AI should familiarize themselves with broadly accessible cloud computing services such as Nautilus, the National Research Platform, designed to democratize AI internationally (Parashar and Altintas 2023, NRP 2024, Würthwein 2024). Nautilus and other government-funded initiatives like ACCESS (NSF 2024) leverage academic institutions to offer low-cost and scalable computational resources. Industry tools, such as Amazon Web Services or Google, can be more expensive but may offer more technical support. As ecologists leverage supercomputing to scale analyses, the non-negligible environmental impacts of AI (Strubell et al. 2020) can be partially alleviated by adjusting the extent, timing, and location of resource use (Dodge et al. 2022).

After adapting AI models to specific use cases, or even developing new models, ecologists often aim to share models and their training data with others, whether to meet publication requirements or to contribute to their communities of practice. Ecology-specific databases may not be designed to enforce existing best practices for AI datasets, such as benchmarking or datasheets for datasets (paper: (Gebru et al. 2021); Overleaf template: (Garbin 2021)). However, ecologists can learn and adopt AI-specific documentation methods, including dataset datasheets as well as AI model cards (paper: (Mitchell et al. 2019); Markdown template: Garbin, 2020 & 2024). BioTrove is a large, well-documented benchmark dataset based on iNaturalist's Open Dataset (iNaturalist 2024) images, scientific and common names, and taxonomic hierarchies with code and example cards for the data and model (paper: (Yang et al. 2024a); website: (Yang et al. 2024b)). Standardized AI-specific documentation describes caveats associated with models and datasets, specifies appropriate downstream use, and facilitates open sharing via platforms like Hugging Face (Jain 2022). For instance, the Cookiecutter Data Science framework also provides guidance for sharing AI models in Python; and these structures are flexible to accommodate complex data processing pipelines and model workflows (Rybicki 2019). Ecologists who want to use AI can learn more about these best practices for Python as well as the recommendations for sharing ecological analyses made in R via research compendia, e.g. (Marwick et al. 2018). There is a growing need for educational materials and explicit recommendations for systematic AI model sharing for ecological audiences that may have less technical expertise or familiarity with Python- or AI-specific best practices.

For ecologists or computer scientists ready to start sharing their AI tools with others, we have curated a list of practical recommendations for how to best facilitate the adoption of these tools by ecologists with little technical training (Fig. S1). We have also illustrated what type of tool may best serve tool-adopters at different levels of technical proficiency and familiarity with ecological datasets and questions (Fig. S2.). Future work in AI in ecology can incorporate model cards and dataset datasheets into browsable model zoos, similar to the one for microscopy computer vision models with the BioImage Model Zoo (Ouyang et al. 2022). We have curated a starter-pack Model Zoo for AI models in ecology on our website (ecoviz-ai.github.io/modelzoo (Kendall-Bar, 2024a)) which can receive new contributions via Github (Kendall-Bar 2024b).

As more ecologists begin to use AI, the careful and generous sharing of models and datasets will help reduce the need to train models or relabel datasets. As data sharing and data availability statements become more prevalent (Jiao et al. 2024), journals will need to formally incentivize the review of data and code. For instance, the Journal of Open Source Software (JOSS 2024) and Methods in Ecology and Evolution have dedicated editors for reviewing software, code, and data; we are also aware that other journals have prioritized maintaining data editors as key staff members (Muench 2023). Practitioners of AI in ecology should be mindful of the ethical considerations associated with sharing code and data. Open-sourcing datasets or models used in large language models like ChatGPT present serious ethical concerns outside the scope of this manuscript (Liesenfeld et al. 2023, Cooper et al. 2024). We recommend that any ecologist new to AI familiarize themselves with practical guides set forth by the British Ecological Society (2025) and ethical guidelines set forth by NeurIPS and others as they begin to implement and share AI models (NeurIPS 2024). Within the scope of environmental science, data management plans co-designed with Indigenous and local knowledge holders have innovated upon open data frameworks like FAIR and CARE to provide local context labels that indicate provenance, protocols, or permission tied to disseminated materials that could contain culturally sensitive or sacred information (Anderson and Christen 2013, Carroll et al. 2021). Overall, a cultural shift towards incentivizing conscientiously open, modular, and expandable tools moves away from redundant, proprietary, or opaque analyses and contributes to more transparent, robust, and defensible science (Brunsdon and Comber 2021, Czapanskiy and Beltran 2022).

5. Conclusion

AI in ecology is quickly gaining momentum, offering unprecedented opportunities to speed and scale ecological research (Christin et al. 2019). There are several important challenges to leveraging AI for ecology, ranging from a lack of trust in AI approaches to the risk of overeager, undiscerning, and potentially dangerous implementation of existing models. Compared to traditional statistical methods, AI often entails a higher transparency deficit and implementation burden, increasing the risk that training data could encode biases or lead to misleading results. Despite these risks, there are many cases where AI presents significant opportunities and low risk for automating tedious manual tasks or leveraging large datasets (Besson et al. 2022, Galaz García et al. 2023, Han et al. 2023). Here we reviewed the key challenges and solutions facing ecologists seeking to leverage AI in their research. When the benefits of AI outweigh the risks, we argue that ecologists are likely to be dissuaded from using AI due to practical challenges such as: (1) the **opportunity costs** while understanding the risks and opportunities of AI, (2) an **overwhelming landscape** while selecting and implementing a model, (3) a **transparency deficit** when interpreting model performance and function, and (4) the **implementation burden** when attempting to modify models, scale their use, and share tools with others. Addressing and/or alleviating these challenges likely requires a multifaceted approach combining: (1) **educational resources** to create openly available informal and formal learning resources, (2) **communities of practice** to create interdisciplinary and inclusive environments for technical social learning, (3) **effective visualizations** to interpret and communicate the functionality and performance of models, and (4) **computational resources** for adapting models, scaling deployments to large datasets, and sharing data, code, and lessons learned with the research community.

Our initiative, *EcoViz+AI*, has created a website that collates several AI-related resources for ecological researchers (ecoviz-ai.github.io; (Kendall-Bar 2024a)). There, we have curated a list of communities of practice to connect ecological researchers to initiatives in the field of ecology and AI. To reduce the time spent looking for models, we have also curated a list of AI tools into a model zoo. We describe five case studies for AI in ecology with science communication videos (see supplement for more information). We invite others to contribute additional models or communities of practice via Github (github.com/ecoviz-ai/ecoviz-ai.github.io; (Kendall-Bar 2024b)).

Looking to the future, a cultural shift is needed to emphasize and reward efforts to produce open and reproducible science that promotes the responsible reuse of data, code, and models (Gundersen et al. 2018, Czapanskiy and Beltran 2022). This cultural shift is already underway, as ecologists replace perceived barriers to sharing data and code with the recognition that these efforts will ultimately save us time, help us establish explicit data sharing agreements, avoid proprietary formats, and help us contribute to communities of practice (Gomes et al. 2022). While the benefits of open science and technical training are not exclusive to AI, the rapid growth, opacity, and implementation challenges of AI highlight the need to prioritize technical education, community development, and science communication. Efforts to improve approachable and responsible AI adoption are strengthened when deployed synergistically with broad, interdisciplinary initiatives to increase computational and scientific literacy, such as the AAAS Vision and Change Action Plan for undergraduate education (Woodin et al. 2009). AI presents an opportunity to harness new momentum, cyberinfrastructure, and computational techniques to incentivize responsible and generous sharing of resources to educate, train, and empower the next generation of ecologists.

Acknowledgments. We would like to extend our gratitude to Schmidt Sciences for generously providing the funding that made the EcoViz+AI workshop possible. Our sincere thanks also go to Brittany Aguilar and Jenna Metcalf for their invaluable logistical support both leading up to and during the event. Additionally, we are appreciative of the Center for Marine Biodiversity and Conservation at Scripps and the San Diego Supercomputer Center for graciously hosting the workshop.

Funding statement. This research was supported by grants from the Schmidt Sciences AI Collaboration Fund.
Competing interests. None
Data availability statement. Website code can be found at https://github.com/ecoviz-ai/ecoviz-ai/github.io
Ethical standards. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.
Author contributions. Conceptualization: J.K-B., A.R.P., M.F.C., N.F., J.L, M.A., B.F., A.M., G.O., A.B., I.N., R.H.P, C.L., I.P., E.M., I.C., T.S., D.R., L.S., R.M., F.M-M., K.H.N., C.C.T., W.O., M.S., S.G., J.W., A.N.K., T.B-W., I.A., S.S., N.L. Data curation: J.M.K-B., M.F.C., N.F., J.L., M.A., A.M., A.B., R.H.P., C.L. Formal analysis: J.M.K-B., M.F.C., N.F., J.L., M.A., A.M., A.B., R.H.P., C.L., R.M., N.L. Funding acquisition: J.M.K-B., E.M., A.N.K, I.A., N.L. Investigation: J.M.K-B., A.R.P., N.F., J.L., M.A., B.F., R.M., I.A., N.L. Methodology: J.M.K-B., A.R.P., M.F.C., N.F., J.L., M.A., A.M., G.O., A.B., R.H.P., C.L., M.S., I.A., N.L. Project administration: J.M.K-B., A.R.P., A.N.K., I.A., S.S. Resources: J.M.K-B., J.L., M.A., A.N.K., T.B-W., I.A., S.S., N.L. Software: J.M.K-B., M.F.C., G.O., R.H.P., C.L., M.S., T.B-W., I.A., N.L. Supervision: J.M.K-B., A.R.P., M.F.C., J.W., I.A., S.S., N.L. Validation: J.M.K-B, A.B., M.S., N.L. Visualization: J.M.K-B., M.F.C., J.L., M.A., B.F., A.M., A.B., I.N., R.H.P., I.C., K.H.N., M.S., N.L. Writing original draft: J.M.K-B. A.R.P. M.F.C., N.F., J.L., B.F., G.O., I.C., R.M., I.A., S.S., N.L. All authors approved the final submitted draft.

Supplementary material. Four supplemental files are also submitted for review.

References

ABC Global Climate Center. (2024). ABC Global Climate Center. ABC AI and Biodiversity Change Center. Available at: https://www.abcclimate.org/; last accessed 28 December 2024.

Alicioglu G, Sun B. (2022). A survey of visual analytics for Explainable Artificial Intelligence methods. Computers & Graphics 102, 502–520.

Allen C, Mehler DMA. (2019). Open science challenges, benefits and tips in early career and beyond. PLOS Biology 17, e3000246.

Allen Institute. (2020). Young investigators program: Allen Institute for Artificial Intelligence (Ai2). Available at: https://allenai.org/young-investigators; last accessed 24 December 2024.

Allocca G, Ma S, Martelli D, Cerri M, Del Vecchio F, Bastianini S, Zoccoli G, Amici R, Morairty SR, Aulsebrook AE, Blackburn S, Lesku JA, Rattenborg NC, Vyssotski AL, Wams E, Porcheret K, Wulff K, Foster R, Chan JKM, Nicholas CL, Freestone DR, Johnston LA, Gundlach AL. (2019). Validation of 'Somnivore', a Machine Learning Algorithm for Automated Scoring and Analysis of Polysomnography Data. Frontiers in Neuroscience 13.

Al-Mansoori F, Hamdan A. (2023). Integrating Indigenous Knowledge Systems into Environmental Education for Biodiversity Conservation: A Study of Sociocultural Perspectives and Ecological Outcomes. AI, IoT and the Fourth Industrial Revolution Review 13, 61–74.

Anderson J, Christen K. (2013). 'Chuck a Copyright on it': Dilemmas of Digital Return and the Possibilities for Traditional Knowledge Licenses and Labels. Museum Anthropology Review 7, 1–126.

Aurisano J, Naqib A, Reiman D. (2017). CS 502 Computational Biology - Deep learning for biology: a tutorial. Available at: https://sites.google.com/view/cs502project/deep-learning-for-biology-a-tutorial; last accessed 23 December 2024.

Baille LMR, Zitterbart DP. (2022). Effectiveness of surface-based detection methods for vessel strike mitigation of North Atlantic right whales. Endangered Species Research 49, 57–69.

Ball IR, Possingham HP, Watts ME. (2009). Marxan and Relatives: Software for Spatial Conservation Prioritization. Pages 185–195 in Moilanen A, Wilson KA, and Possingham HP, eds. Spatial Conservation Prioritization. Oxford University Press.

Barrett B, Zepeda E, Pollack L, Munson A, Sih A. (2019). Counter-Culture: Does Social Learning Help or Hinder Adaptive Response to Human-Induced Rapid Environmental Change? Frontiers in Ecology and Evolution 7.

Bates AE, Davies MA, Stuart-Smith RD, Lazzari N, Lefcheck JS, Ling SD, Mellin C, Mouillot D, Bernard ATF, Bennett S, Brown CJ, Burrows MT, Butler CL, Cinner J, Clausius E, Cooper A, Costello MJ, Denis-Roy L, Edgar GJ, Fuchs YH, Johnson OJ, Gordó-Vilaseca C, Hautecoeur C, Harper LM, Heather FJ, Jones TR, Markey AC, Oh E, Rose M, Ruiz-Ruiz PA, Sanabria-Fernandez JA, Schuster JM, Schmid JK, Baker SC. (2024). Overcome imposter syndrome: Contribute to working groups and build strong networks. Biological Conservation 293, 110566.

Bates D, Mächler M, Bolker B, Walker S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1–48.

Beale CM, Lennon JJ. (2012). Incorporating uncertainty in predictive species distribution modelling. Philosophical Transactions of the Royal Society B: Biological Sciences 367, 247–258.

Beery S, Cole E, Parker J, Perona P, Winner K. (2021). Species Distribution Modeling for Machine Learning Practitioners: A Review. 23 September 2021, New York, NY, USA.

Beery S, Morris D, Yang S. 2019. Efficient Pipeline for Camera Trap Image Review.

Beery S, Parham J, Stathatos S, Kellenberger B, Cole E, Lutjens B, Sharma T, Kay J. (2023). CV4Ecology Course Details. Available at: https://cv4ecology.caltech.edu/course_content2023.html; last accessed 23 December 2024.

Benkendorf DJ, Hawkins CP. (2020). Effects of sample size and network depth on a deep learning approach to species distribution modeling. Ecological Informatics 60, 101137.

Bennett NJ, Roth R, Klain SC, Chan K, Christie P, Clark DA, Cullman G, Curran D, Durbin TJ, Epstein G, Greenberg A, Nelson MP, Sandlos J, Stedman R, Teel TL, Thomas R, Veríssimo D, Wyborn C. (2017). Conservation social science: Understanding and integrating human dimensions to improve conservation. Biological Conservation 205, 93–108.

Benyei P, Arreola G, Reyes-García V. (2020). Storing and sharing: A review of indigenous and local knowledge conservation initiatives. Ambio 49, 218–230.

Berger-Tal O, Wong BBM, Adams CA, Blumstein DT, Candolin U, Gibson MJ, Greggor AL, Lagisz M, Macura B, Price CJ, Putman BJ, Snijders L, Nakagawa S. (2024). Leveraging AI to improve evidence synthesis in conservation. Trends in Ecology & Evolution.

Berger-Wolf TY, Rubenstein DI, Stewart CV, Holmberg JA, Parham J, Menon S, Crall J, Van Oast J, Kiciman E, Joppa L. (2017). Wildbook: Crowdsourcing, computer vision, and data science for conservation.

Besson M, Alison J, Bjerge K, Gorochowski TE, Høye TT, Jucker T, Mann HMR, Clements CF. (2022). Towards the fully automated monitoring of ecological communities. Ecology Letters 25, 2753–2775.

Bishop CM. (2013). Model-based machine learning. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 371, 20120222.

Blount D, Gero S, Van Oast J, Parham J, Kingen C, Scheiner B, Stere T, Fisher M, Minton G, Khan C, Dulau V, Thompson J, Moskvyak O, Berger-Wolf T, Stewart CV, Holmberg J, Levenson JJ. (2022). Flukebook: an open-source AI platform for cetacean photo identification. Mammalian Biology 102, 1005–1023.

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White J-SS. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology & Evolution 24, 127–135.

Borowiec ML, Dikow RB, Frandsen PB, McKeeken A, Valentini G, White AE. (2022). Deep learning as a tool for ecology and evolution. Methods in Ecology and Evolution 13, 1640–1660.

Brunsdon C, Comber A. (2021). Opening practice: supporting reproducibility and critical spatial data science. Journal of Geographical Systems 23, 477–496.

Carroll SR, Herczog E, Hudson M, Russell K, Stall S. (2021). Operationalizing the CARE and FAIR Principles for Indigenous data futures. Scientific Data 8, 108.

Chapman M, Xu L, Lapeyrolerie M, Boettiger C. (2023). Bridging adaptive management and reinforcement learning for more robust decisions. Philosophical Transactions of the Royal Society B: Biological Sciences 378: 20220195.

Chapman MS, Goldstein BR, Schell CJ, Brashares JS, Carter NH, Ellis-Soto D, Faxon HO, Goldstein JE, Halpern BS, Longdon J, Norman KEA, O'Rourke D, Scoville C, Xu L, Boettiger C. (2024). Biodiversity monitoring for a just planetary future. Science 383, 34–36.

Chapman MS, Oestreich WK, Frawley TH, Boettiger C, Diver S, Santos BS, Scoville C, Armstrong K, Blondin H, Chand K, Haulsee DE, Knight CJ, Crowder LB. (2021). Promoting equity in the use of algorithms for high-seas conservation. One Earth 4, 790–794.

Cheeseman T, Johnson T, Southerland K, Muldavin N. (2017). Happywhale: Globalizing Marine Mammal Photo Identification via a Citizen Science Web Platform. International Whaling Commission SC/67A/PH/02.

Cheeseman T, Southerland K, Acebes JM, Audley K, Barlow J, Bejder L, Birdsall C, Bradford AL, Byington JK, Calambokidis J, Cartwright R, Cedarleaf J, Chavez AJG, Currie JJ, De Weerdt J, Doe N, Doniol-Valcroze T, Dracott K, Filatova O, Finn R, Flynn K, Ford JKB, Frisch-Jordán A, Gabriele CM, Goodwin B, Hayslip C, Hildering J, Hill MC, Jacobsen JK, Jiménez-López ME, Jones M, Kobayashi N, Lyman E, Malleson M, Mamaev E, Martínez Loustalot P, Masterman A, Matkin C, McMillan CJ, Moore JE, Moran JR, Neilson JL, Newell H, Okabe H, Olio M, Pack AA, Palacios DM, Pearson HC, Quintana-Rizzo E, Ramírez Barragán RF, Ransome N, Rosales-Nanduca H, Sharpe F, Shaw T, Stack SH, Staniland I, Straley J, Szabo A, Teerlink S, Titova O, Urban R. J, van Aswegen M, de Morais MV, von Ziegesar O, Witteveen B, Wray J, Yano KM, Zwiefelhofer D, Clapham P. (2023). A collaborative and near-comprehensive North Pacific humpback whale photo-ID dataset. Scientific Reports 13, 10237.

Christin S, Hervet É, Lecomte N. (2019). Applications for deep learning in ecology. Methods in Ecology and Evolution 10, 1632–1644.

Climate Change AI. (2024). Climate Change AI 2024 Innovation Grants. Climate Change AI. Accessed at: https://www.climatechange.ai/calls/innovation_grants_2024; last accessed 26 December 2024.

Cole E, Stathatos S, Lütjens B, Sharma T, Kay J, Parham J, Kellenberger B, Beery S. (2023). Teaching Computer Vision for Ecology.

Cooper N, Clark AT, Lecomte N, Qiao H, Ellison AM. (2024). Harnessing Large Language Models for Coding, Teaching, and Inclusion to Empower Research in Ecology and Evolution.

Cravens AE, Jones MS, Ngai C, Zarestky J, Love HB. (2022). Science facilitation: navigating the intersection of intellectual and interpersonal expertise in scientific collaboration. Humanities and Social Sciences Communications 9, 1–13.

Czapanskiy MF, Beltran RS. (2022). How Reproducibility Will Accelerate Discovery Through Collaboration in Physio-Logging. Frontiers in Physiology 13.

Davis J, Purves D, Gilbert J, Sturm S. (2022). Five ethical challenges facing data-driven policing. AI and Ethics 2, 185–198.

Dodge J, Prewitt T, Tachet des Combes R, Odmark E, Schwartz R, Strubell E, Luccioni AS, Smith NA, DeCario N, Buchanan W. (2022). Measuring the Carbon Intensity of AI in Cloud Instances. 20 June 2022, New York, NY, USA.

Doll BB, Simon DA, Daw ND. (2012). The ubiquity of model-based reinforcement learning. Current Opinion in Neurobiology 22, 1075–1081.

DŲeroski S. (2009). Machine Learning Applications in Habitat Suitability Modeling. Pages 397–411 in Haupt SE, Pasini A, and Marzban C, eds. Artificial Intelligence Methods in the Environmental Sciences. Springer Netherlands.

Ebrahimi SH, Ossewaarde M, Need A. (2021). Smart Fishery: A Systematic Review and Research Agenda for Sustainable Fisheries in the Age of AI. Sustainability 13, 6037.

Elith J, Leathwick JR. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annual Review of Ecology, Evolution, and Systematics 40, 677–697.

Elith J, Leathwick JR, Hastie T. (2008). A working guide to boosted regression trees. Journal of Animal Ecology 77, 802–813.

Galaz García C, Bagstad KJ, Brun J, Chaplin-Kramer R, Dhu T, Murray NJ, Nolan CJ, Ricketts TH, Sosik HM, Sousa D, Willard G, Halpern BS. (2023). The future of ecosystem assessments is automation, collaboration, and artificial intelligence. Environmental Research Letters 18, 011003.

Garbin C. (2021). Datasheet for dataset template. Accessed at: https://www.overleaf.com/latex/templates/datasheet-for-dataset-template/jgqyyzyprxth; last accessed 26 December 2024.

Garbin C. (2024). Template for model cards.

Gebru T, Morgenstern J, Vecchione B, Vaughan JW, Wallach H, III HD, Crawford K. (2021). Datasheets for Datasets.

Gomes DGE, Pottier P, Crystal-Ornelas R, Hudgins EJ, Foroughirad V, Sánchez-Reyes LL, Turba R, Martinez PA, Moreau D, Bertram MG, Smout CA, Gaynor KM. (2022). Why don't we share data and code? Perceived barriers and benefits to public archiving practices. Proceedings of the Royal Society B: Biological Sciences 289, 20221113.

Grace JB. (2024). An integrative paradigm for building causal knowledge. Ecological Monographs 94, e1628.

Gray P. (2024). deep_learning_ecology.

Gundersen OE, Gil Y, Aha DW. (2018). On Reproducible AI: Towards Reproducible Research, Open Science, and Digital Scholarship in AI Publications. AI Magazine 39, 56–68.

Han BA, Varshney KR, LaDeau S, Subramaniam A, Weathers KC, Zwart J. (2023). A synergistic future for AI and ecology. Proceedings of the National Academy of Sciences 120, e2220283120.

Harlow A, Lo SM, Saichaie K, Sato BK. (2020). Characterizing the University of California's tenure-track teaching position from the faculty and administrator perspectives. PLOS ONE 15, e0227633.

Higino G, Barros C, Bledsoe E, Roche DG, Binning SA, Poisot T. (2023). Postdoctoral scientists are mentors, and it is time to recognize their work. PLOS Biology 21, e3002349.

Hilborn R, Mangel M. (1997). The Ecological Detective. Princeton University Press.

Hsu Y-C, Huang T-H 'Kenneth', Verma H, Mauri A, Nourbakhsh I, Bozzon A. (2022). Empowering local communities using artificial intelligence. Patterns 3, 100449.

iNaturalist. (2024). iNaturalist Open Dataset: Observations. iNaturalist. Accessed at: https://www.inaturalist.org/observations; last accessed 29 December 2024.

Jain SM. (2022). Hugging Face. Pages 51–67 in Jain SM, ed. Introduction to Transformers for NLP: With the Hugging Face Library and Models to Solve Problems. Apress.

Jiao C, Li K, Fang Z. (2024). Data sharing practices across knowledge domains: A dynamic examination of data availability statements in PLOS ONE publications. Journal of Information Science 50, 673–689.

Johnstone RA, Dall SRX, Giraldeau L, Valone TJ, Templeton JJ. (2002). Potential disadvantages of using socially acquired information. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 357, 1559–1566.

JOSS. (2024). Journal of Open Source Software. Accessed at: https://joss.theoj.org; last accessed 26 December 2024.

Katija K, Orenstein E, Schlining B, Lundsten L, Barnard K, Sainz G, Boulais O, Cromwell M, Butler E, Woodward B, Bell KLC. (2022). FathomNet: A global image database for enabling artificial intelligence in the ocean. Scientific Reports 12, 15914.

Kendall-Bar J. (2023). Data-Driven Animation for Science Communication. Coursera. Accessed at: https://www.coursera.org/learn/data-driven-animation/; last accessed 4 June 2024.

Kendall-Bar J, Kendall-Bar N, Forbes AG, McDonald G, Ponganis PJ, Williams C, Horning M, Hindle A, Klinck H, Beltran RS, Friedlaender AS, Wiley D, Costa DP, Williams TM. (2021). Visualizing Life in the Deep: A Creative Pipeline for Data-Driven Animations to Facilitate Marine Mammal Research, Outreach, and Conservation. 2021 IEEE VIS Arts Program (VISAP) 1–10.

Kendall-Bar JM. (2024a). EcoViz+AI: Visualization and AI for Ecology – EcoViz+AI. EcoViz+AI Online Hub. Accessed at: https://ecoviz-ai.github.io/; last accessed 5 August 2024.

Kendall-Bar JM. (2024b). Github: Code for EcoViz+AI Website. Code for EcoViz+AI Online Hub. Accessed at: https://github.com/ecoviz-ai/ecoviz-ai.github.io; last accessed 5 August 2024.

Kendall-Bar JM, Nealey I, Costello I, Lowrie C, Nguyen KH, Ponganis PJ, Beck MW, Altintas I. (2024). EcoViz: co-designed environmental data visualizations to communicate ecosystem impacts, inform management, and envision solutions. Paper presented at IEEE VIS 2024. October 2024.

Kendall-Bar JM, Williams TM, Mukherji R, Lozano DA, Pitman JK, Holser RR, Keates T, Beltran RS, Robinson PW, Crocker DE, Adachi T, Lyamin OI, Vyssotski AL, Costa DP. (2023). Brain activity of diving seals reveals short sleep cycles at depth. Science 380, 260–265.

Kirlin J, Caldwell M, Gleason M, Weber M, Ugoretz J, Fox E, Miller-Henson M. (2013). California's Marine Life Protection Act Initiative: Supporting implementation of legislation establishing a statewide network of marine protected areas. Ocean & Coastal Management 74, 3–13.

Kockel A, Ban NC, Costa M, Dearden P. (2020). Addressing distribution equity in spatial conservation prioritization for small-scale fisheries. PLOS ONE 15, e0233339.

Kühn B, Cayetano A, Fincham JI, Moustahfid H, Sokolova M, Trifonova N, Watson JT, Fernandes-Salvador JA, Uusitalo L. (2024). Machine Learning Applications for Fisheries—At Scales from Genomics to Ecosystems. Reviews in Fisheries Science & Aquaculture 1–24.

Laland KN, Williams K. (1998). Social transmission of maladaptive information in the guppy. Behavioral Ecology 9, 493–499.

Lapeyrolerie M, Chapman MS, Norman KEA, Boettiger C. (2022). Deep reinforcement learning for conservation decisions. Methods in Ecology and Evolution 13, 2649–2662.

Lapp S, Rhinehart T, Freeland-Haynes L, Khilnani J, Syunkova A, Kitzes J. (2023). OpenSoundscape: An open-source bioacoustics analysis package for Python. Methods in Ecology and Evolution 14, 2321–2328.

Lapp S, Rhinehart T, Freeland-Haynes L, Khilnani J, Syunkova A, Viotti L, Ruiz Guzman S, Kitzes J. (2024). Train a CNN. OpenSoundscape. Accessed at: https://opensoundscape.org/en/latest/tutorials/train_cnn.html; last accessed 29 December 2024.

Lauer J, Zhou M, Ye S, Menegas W, Schneider S, Nath T, Rahman MM, Di Santo V, Soberanes D, Feng G, Murthy VN, Lauder G, Dulac C, Mathis MW, Mathis A. (2022). Multi-animal pose estimation, identification and tracking with DeepLabCut. Nature Methods 19, 496–504.

Lawson DM, Hall KR, Yung L, Enquist CA. (2017). Building translational ecology communities of practice: insights from the field. Frontiers in Ecology and the Environment 15, 569–577.

Layden TJ, Fernández S, Sandoval-Lemus M, Sonius KJ, David-Chavez D, Bombaci SP. (2024). Shifting Power in Practice: Implementing Relational Research and Evaluation in Conservation Science. Social Sciences 13, 555.

Lefcheck J. (2015). A practical guide to machine learning in ecology. Accessed at: https://jonlefcheck.net/2015/02/06/a-practical-guide-to-machine-learning-in-ecology/; last accessed 23 December 2024.

Liesenfeld A, Lopez A, Dingemanse M. (2023). Opening up ChatGPT: Tracking openness, transparency, and accountability in instruction-tuned text generators. 19 July 2023, New York, NY, USA.

Lindkvist E, Ekeberg Ö, Norberg J. (2017). Strategies for sustainable management of renewable resources during environmental change. Proceedings of the Royal Society B: Biological Sciences 284, 20162762.

Longdon J. (2023). Visualising Forest Sound: Justice-led Ecoacoustic Data Interaction. 19 April 2023, New York, NY, USA.

Lubiana T, Lopes R, Medeiros P, Silva JC, Goncalves ANA, Maracaja-Coutinho V, Nakaya HI. (2023). Ten quick tips for harnessing the power of ChatGPT in computational biology. PLOS Computational Biology 19, e1011319.

MacWilliams R, Kim S, Trueman RJ. (2024). Bugs and bots: how technology is changing the game in biodiversity monitoring. Biodiversity 25, 295–296.

Manderfield M. (2022). Seek, Picture Insect, Google Lens: An Analysis of Popular Insect Identification Apps Using Photos of Realistic Quality. Department of Entomology, Distance Master of Science Projects.

Marwick B, Boettiger C, Mullen L. (2018). Packaging Data Analytical Work Reproducibly Using R (and Friends). The American Statistician 72, 80–88.

Miao Z, Liu Z, Gaynor KM, Palmer MS, Yu SX, Getz WM. (2021). Iterative human and automated identification of wildlife images. Nature Machine Intelligence 3, 885–895.

Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, Spitzer E, Raji ID, Gebru T. (2019). Model Cards for Model Reporting. 29 January 2019.

Molnar C. (2024). 9.6 SHAP (SHapley Additive exPlanations) | Interpretable Machine Learning. Interpretable Machine Learning: A Guide to Making Black Box Models Explainable.

Montealegre-Mora F, Boettiger C, Walters CJ, Cahill CL. 2024. Using machine learning to inform harvest control rule design in complex fishery settings.

Montealegre-Mora F, Lapeyrolerie M, Chapman M, Keller AG, Boettiger C. (2023). Pretty Darn Good Control: When are Approximate Solutions Better than Approximate Models. Bulletin of Mathematical Biology 85, 95.

Mosqueira-Rey E, Hernández-Pereira E, Alonso-Ríos D, Bobes-Bascarán J, Fernández-Leal Á. (2022). Human-in-the-loop machine learning: a state of the art. Artificial Intelligence Review 56, 3005–3054.

Mporas I, Perikos I, Kelefouras V, Paraskevas M. (2020). Illegal Logging Detection Based on Acoustic Surveillance of Forest. Applied Sciences 10, 7379.

Muench A. (2023). The Roles of Data Editors in Astronomy. Science Editor.

NCEAS. (2023). Environmental Data Science Summit. EDS Summit. Accessed at: https://eds-summit.github.io/; last accessed 28 December 2024.

NeurIPS. (2024). NeurIPS Code of Ethics. Accessed at: https://neurips.cc/public/EthicsGuidelines; last accessed 28 December 2024.

Ng A. (2024). Deep Learning. Coursera. Accessed at: https://www.coursera.org/specializations/deep-learning/; last accessed 28 December 2024.

NRP. (2024). Nautilus – National Research Platform. Accessed at: https://nationalresearchplatform.org/nautilus/; last accessed 28 December 2024.

NSF. (2024). ACCESS | National Science Foundation. Access. Accessed at: https://access-ci.org/; last accessed 28 December 2024.

NSF CICORE. (2024). CORE Institute. Core Institute. Accessed at: https://www.core-institute.org; last accessed 28 December 2024.

Oestreich WK, Oliver RY, Chapman MS, Go MC, McKenna MF. (2024). Listening to animal behavior to understand changing ecosystems. Trends in Ecology & Evolution 0.

Ouyang W, Beuttenmueller F, Gómez-de-Mariscal E, Pape C, Burke T, Garcia-López-de-Haro C, Russell C, Moya-Sans L, de-la-Torre-Gutiérrez C, Schmidt D, Kutra D, Novikov M, Weigert M, Schmidt U, Bankhead P, Jacquemet G, Sage D, Henriques R, Muñoz-Barrutia A, Lundberg E, Jug F, Kreshuk A. (2022). BioImage Model Zoo: A Community-Driven Resource for Accessible Deep Learning in BioImage Analysis. 2022.06.07.495102.

OxRSE. (2024a). OxRSE Training Course. Accessed at: https://train.rse.ox.ac.uk/; last accessed 28 December 2024.

OxRSE. (2024b). Course-material: OxRSE Training Github. Accessed at: https://github.com/OxfordRSE; last accessed 28 December 2024.

Parashar M, Altintas I. (2023). Toward Democratizing Access to Science Data: Introducing the National Data Platform. Paper presented at 2023 IEEE 19th International Conference on e-Science (e-Science). October 2023.

Parkinson RH, Fecher C, Gray JR. (2022). Chronic exposure to insecticides impairs honeybee optomotor behaviour. Frontiers in Insect Science 2.

Pavoni G, Corsini M, Ponchio F, Muntoni A, Cignoni P. 2021. TagLab: A human-centric AI system for interactive semantic segmentation.

Phillips SJ, Anderson RP, Schapire RE. (2006). Maximum entropy modeling of species geographic distributions. Ecological Modelling 190, 231–259.

Pichler M, Hartig F. (2023). Machine learning and deep learning—A review for ecologists. Methods in Ecology and Evolution 14, 994–1016.

pyOpenSci. (2024). pyOpenSci. pyOpenSci. Accessed at: https://www.pyopensci.org/; last accessed 28 December 2024.

Raphael Vallat, Nikola Jajcay. (2020). raphaelvallat/yasa, v0.4.0. Accessed at: https://github.com/raphaelvallat/yasa; last accessed 28 December 2024.

Renner IW, Warton DI. (2013). Equivalence of MAXENT and Poisson Point Process Models for Species Distribution Modeling in Ecology. Biometrics 69, 274–281.

Rybicki J. (2019). Best Practices in Structuring Data Science Projects. 2019, Cham.

Ryo M, Angelov B, Mammola S, Kass JM, Benito BM, Hartig F. (2021). Explainable artificial intelligence enhances the ecological interpretability of black-box species distribution models. Ecography 44, 199–205.

Saarenmaa H, Stone ND, Folse LJ, Packard JM, Grant WE, Makela ME, Coulson RN. (1988). An artificial intelligence modelling approach to simulating animal/habitat interactions. Ecological Modelling 44, 125–141.

Schirpke U, Ghermandi A, Sinclair M, Van Berkel D, Fox N, Vargas L, Willemen L. (2023). Emerging technologies for assessing ecosystem services: A synthesis of opportunities and challenges. Ecosystem Services 63, 101558.

Schmidt Sciences. (2022). Schmidt AI in Science Postdoctoral Fellowship. Schmidt Sciences. Accessed at: https://www.schmidtsciences.org/schmidt-ai-in-science-postdocs/; last accessed 24 December 2024.

Scoville C, Chapman M, Amironesei R, Boettiger C. (2021). Algorithmic conservation in a changing climate. Current Opinion in Environmental Sustainability 51, 30–35.

Sethi SS, Jones NS, Fulcher BD, Picinali L, Clink DJ, Klinck H, Orme CDL, Wrege PH, Ewers RM. (2020). Characterizing soundscapes across diverse ecosystems using a universal acoustic feature set. Proceedings of the National Academy of Sciences 117, 17049–17055.

Sheikh H, Prins C, Schrijvers E. (2023). Artificial Intelligence, Definition and Background. Pages 15–41 in Sheikh H, Prins C, and Schrijvers E, eds. Mission AI, The New System Technology. Springer International Publishing.

Silvestro D, Goria S, Sterner T, Antonelli A. (2022). Improving biodiversity protection through artificial intelligence. Nature Sustainability 5, 415–424.

Stevens S, Kuzak M, Martinez C, Moser A, Bleeker P, Galland M. (2018). Building a local community of practice in scientific programming for life scientists. PLOS Biology 16, e2005561.

Stevens S, Wu J, Thompson MJ, Campolongo EG, Song CH, Carlyn DE, Dong L, Dahdul WM, Stewart C, Berger-Wolf T, Chao W-L, Su Y. (2023). BioCLIP: A Vision Foundation Model for the Tree of Life.

Stockwell D, Arzberger P, Fountain T, Helly J. (2000). An Interface between Computing, Ecology and Biodiversity : Environmental Informatics. The Korean Journal of Ecology 23, 101–106.

Strubell E, Ganesh A, McCallum A. (2020). Energy and Policy Considerations for Modern Deep Learning Research. Proceedings of the AAAI Conference on Artificial Intelligence 34, 13693–13696.

Sullivan BL, Aycrigg JL, Barry JH, Bonney RE, Bruns N, Cooper CB, Damoulas T, Dhondt AA, Dietterich T, Farnsworth A, Fink D, Fitzpatrick JW, Fredericks T, Gerbracht J, Gomes C, Hochachka WM, Iliff MJ, Lagoze C, La Sorte FA, Merrifield M, Morris W, Phillips TB, Reynolds M, Rodewald AD, Rosenberg KV, Trautmann NM, Wiggins A, Winkler DW, Wong W-K, Wood CL, Yu J, Kelling S. (2014). The eBird enterprise: An integrated approach to development and application of citizen science. Biological Conservation 169, 31–40.

Swain D. (2023). Climate researchers need support to become scientist-communicators. Nature 624, 9–9.

Swartz W, Cisneros-Montemayor AM, Singh GG, Boutet P, Ota Y. (2021). AIS-based profiling of fishing vessels falls short as a "proof of concept" for identifying forced labor at sea. Proceedings of the National Academy of Sciences 118, e2100341118.

Swischuk R, Mainini L, Peherstorfer B, Willcox K. (2019). Projection-based model reduction: Formulations for physics-based machine learning. Computers & Fluids 179, 704–717.

Tabassi E. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). National Institute of Standards and Technology (U.S.). Report no. NIST AI 100-1.

Tuia D, Kellenberger B, Beery S, Costelloe BR, Zuffi S, Risse B, Mathis A, Mathis MW, van Langevelde F, Burghardt T, Kays R, Klinck H, Wikelski M, Couzin ID, van Horn G, Crofoot MC, Stewart CV, Berger-Wolf T. (2022). Perspectives in machine learning for wildlife conservation. Nature Communications 13, 792.

Turing Institute. (2024). Artificial intelligence (Safe and ethical). The Alan Turing Institute. Accessed at: https://www.turing.ac.uk/research/research-programmes/artificial-intelligence-ai/safe-and-ethical; last accessed 28 December 2024.

Walsh L. (2015). The visual rhetoric of climate change. WIREs Climate Change 6, 361–368.

Wang P. (2019). On Defining Artificial Intelligence. Journal of Artificial General Intelligence 10, 1–37.

Wasimuddin M, Elleithy K, Abuzneid A-S, Faezipour M, Abuzaghleh O. (2020). Stages-Based ECG Signal Analysis From Traditional Signal Processing to Machine Learning Approaches: A Survey. IEEE Access 8, 177782–177803.

Welch H, Brodie S, Jacox MG, Bograd SJ, Hazen EL. (2020). Decision-support tools for dynamic management. Conservation Biology 34, 589–599.

Wenger E. (2011). Communities of practice: A brief introduction.

Woodin T, Smith D, Allen D. (2009). Transforming Undergraduate Biology Education for All Students: An Action Plan for the Twenty-First Century. CBE—Life Sciences Education 8, 271–273.

Woodley L, Pratt K. (2020). The CSCCE Community Participation Model – A framework to describe member engagement and information flow in STEM communities.

Wu X, Xiao L, Sun Y, Zhang J, Ma T, He L. (2022). A survey of human-in-the-loop for machine learning. Future Generation Computer Systems 135, 364–381.

Würthwein F. (2024). National Research Platform: Open Cyberinfrastructure for an Open Society. Accessed at: https://nationalresearchplatform.org/presentations/open-cyberinfrastructure-for-an-open-society-with-frank-wurthwein-2024/; last accessed 26 December 2024.

Yang C-H, Feuer B, Jubery TZ, Deng ZK, Nakkab A, Hasan MZ, Chiranjeevi S, Marshall KO, Baishnab N, Singh AK, Singh A, Sarkar S, Merchant N, Hegde C, Ganapathysubramanian B. (2024a). BioTrove: A Large Curated Image Dataset Enabling AI for Biodiversity. Paper presented at The Thirty-eight Conference on Neural Information Processing Systems Datasets and Benchmarks Track. 13 November 2024.

Yang C-H, Feuer B, Jubery TZ, Deng ZK, Nakkab A, Hasan MZ, Chiranjeevi S, Marshall KO, Baishnab N, Singh AK, Singh A, Sarkar S, Merchant N, Hegde C, Ganapathysubramanian B. (2024b). BioTrove: A Large Curated Image Dataset Enabling AI for Biodiversity. BioTrove. Accessed at: https://baskargroup.github.io/BioTrove/; last accessed 29 December 2024.

Yu H, Amador GJ, Cribellier A, Klaassen M, Knegt HJ de, Naguib M, Nijland R, Nowak L, Prins HHT, Snijders L, Tyson C, Muijres FT. (2024). Edge computing in wildlife behavior and ecology. Trends in Ecology & Evolution 39, 128–130.

Yu H, Klaassen CAJ, Deng J, Leen T, Li G, Klaassen M. (2022). Increasingly detailed insights in animal behaviours using continuous on-board processing of accelerometer data. Movement Ecology 10, 42.

Zhang T, Guo H, Song L, Yuan H, Sui H, Li B. (2025). Evaluating the importance of vertical environmental variables for albacore fishing grounds in tropical Atlantic Ocean using machine learning and Shapley additive explanations (SHAP) approach. Fisheries Oceanography 34, e12701.

rOpenSci (2024). Accessed at: https://ropensci.org/; last accessed 28 December 2024.

Better Science Guides. (2025). British Ecological Society. Accessed at: https://www.britishecologicalsociety.org/content-type/better-science-guides/; last accessed 28 December 2024.