Towards ecologically meaningful foundation models

Ross J. Gardiner[1†], Henry Cerbone[2†], Hammed A. Akande[3], Carly H. Batist[4], Amber Cowans[5], Stella M. Felsinger[2], Premdeep S. Gill[6, 7], Taniya Kapoor[8], Eliot T. Miller[9], Rodrigo Oyanedel[10, 11], Steven Reece[12], Yan Ying Tan[13], Jonas Antony[14], Carlos Rodriguez-Pardo[15, 16, 17], Amy Hinsley[2], Micah Bowles[18, 19], Sebastian A. Heilpern[20], Rachel H. Parkinson[2, 21*]

[1]Centre for Ecology and Conservation, University of Exeter, Penryn Campus, United Kingdom.
[2]Department of Biology, University of Oxford, United Kingdom.
[3]Department of Biology, Concordia University, Canada.
[4]Conservation International, Moore Center for Science & Solutions, USA.
[5]Centre for Research into Environmental and Ecological Modelling, School of Mathematics and Statistics, University of St Andrews, United Kingdom.
[6]British Antarctic Survey, United Kingdom.
[7]Scott Polar Research Institute, Department of Geography, University of Cambridge, United Kingdom.
[8]Artificial Intelligence Group, Wageningen University & Research, Netherlands.
[9]American Bird Conservancy, USA.
[10]MAR FUTURA, Chile.
[11]Instituto Milenio en Socio-Ecología Costera (SECOS), Chile.
[12]School of Geography and the Environment, Oxford University Centre for the Environment, University of Oxford, United Kingdom.
[13]Michigan Institute for Data and AI in Society (MIDAS), University of Michigan, USA.
[14]Dian Fossey Gorilla Fund, Rwanda.
[15]Politecnico di Milano, Italy.
[16]Euro-Mediterranean Center on Climate Change (CMCC), Italy.

[17]RFF-CMCC European Institute on Economics and the Environment (EIEE), Italy.
[18]Ellison Institute of Technology, United Kingdom.
[19]Department of Physics, University of Oxford, United Kingdom.
[20]Department of Earth System Science, Stanford University, USA.
[21]School of Biological and Behavioural Sciences, Queen Mary University of London, United Kingdom.

*Corresponding author(s). E-mail(s): rachel.parkinson@biology.ox.ac.uk; r.parkinson@qmul.ac.uk;
†These authors contributed equally to this work.

Abstract

Ecology aims to explain and predict how organisms interact with each other and their environments across space and time. Yet both ecological data and theory are fragmented, leading to models that generalise poorly beyond specific systems or scales. Empirical evidence spans diverse modalities, resolutions and contexts, while theory is distributed across partially overlapping frameworks that are rarely integrated within a single predictive model. We argue that ecological foundation models (ecoFMs), trained on large, multimodal ecological datasets, offer a route toward unifying data and theory within a common framework. Beyond their ecological value, ecoFMs present a challenging and consequential testbed for machine learning, demanding advances in multimodal representation learning, theory-guided modelling, and uncertainty-aware inference. By learning shared representations of organisms, environments and interactions, ecoFMs could improve generalisation, link pattern to process, and enable synthesis across ecological sub-disciplines. We outline a roadmap for developing ecoFMs, including requirements for data infrastructure, model architectures, evaluation strategies and governance, and assess where current machine learning and ecological approaches fall short. If developed responsibly and collaboratively with ecological practitioners and other actors, ecoFMs could enable new modes of analysis and strengthen ecological forecasting, while simultaneously driving advances in machine learning for multimodal integration, theory-guided learning, and generalisation in complex, data-limited systems.

Keywords: ecology, foundation models, multimodal data, species interactions, ecological forecasting

1 Introduction

Ecology seeks to explain how organisms interact with each other and their environments [1]. Uncovering these relationships is difficult because ecological patterns are often context-dependent and arise from the interaction of multiple processes. As a result, ecology has struggled to disentangle the incredible complexity in nature and extract general principles that hold across scales, from genes to ecosystems, and from specific locations to global patterns [2, 3]. This challenge takes two forms. Firstly, the amount of multiscale data that has been generated in long-term projects is immense and has outpaced researchers' ability to fully interpret within a specific context, and critically what can be inferred and generalised about other systems. Efforts such as the Wytham Great Tit Project [4], the Isle Royale Wolf-Moose Project [5], the US Long Term Ecological Research Program [6], and the Park Grass Experiment [7] together account for a combined 250 years of ecological data, spanning molecular samples and genomic sequencing through to population and demographic records. Secondly, the problem of effect size is crucial in ecology [8]. Despite a large amount of global data, data fragmentation often hampers scientists' ability to realise the full potential of the ecological data produced to date. Looking forward, these challenges are intensifying. Ecological data are growing rapidly in both volume and modality, driven by the integration of long-standing field methods and emerging automated and sensing technologies [9]. Addressing the mismatch between data generation and integrative understanding is essential for realising the full value of ecological knowledge. This realisation is necessary to understand Earth system dynamics and for anticipating the impacts of human activity in the Anthropocene.

In the past decade, science has undergone a major shift, with many fields now leveraging artificial intelligence (AI) and machine learning (ML) models to tackle diverse research problems. Foundation models, coined by [10], are large, unsupervised or self-supervised models pre-trained using substantial data and computation, and then adapted to specific tasks, requiring substantially fewer resources than training task-specific models from scratch. In the last three years, the most well-known foundation models, large language models (LLMs), models trained to ingest and output text, have had a widespread impact across domains [11], driven by algorithmic advances, substantial investment in computational resources, and construction of massive training datasets. Although LLMs have been applied in ecology [12, 13], they fall short of both being fundamental ecological tools and of being useful representations for ecological study [13, 14]. What fine-tuned LLMs lack is the ability to interpret and relate ecological phenomena in their native representation, the originally collected format, instead relying on secondary interpretation, i.e., published literature. Outside of language, the range of models released in the last decade in computer vision from CLIP [15] to SAM3 [16] to DINO [17] have ushered in an age of foundation models in vision. The success of large, pre-trained models across input modalities has catalysed the proliferation of foundation models in disparate areas ranging across language [18, 19], computer vision [16, 20, 21], earth science [22], fluid dynamics [23], genomics [24], proteins [25], astrophysics [26], and psychology [27] with calls for the development of more [28]. Despite major advances in these fields, no complete ecological foundation model yet exists. With accelerating climate change, ecological research must scale to meet increasingly complex and urgent challenges. These are new challenges not foreseen by past approaches, making methodological advancement necessary.

In this paper, we argue that an ecological foundation model that incorporates a multimodal representation backbone, graph-based ecological structures, and uncertainty quantification would be transformative for both ecological and computer sciences. We present the first blueprint for a complete ecoFM that, if implemented, would transform the field's ability to reveal connections within the Earth system and synthesise the extensive, albeit fractured, corpus of ecological data. In turn, meeting ecology's unique challenges would also advance foundation models and AI for science-driven challenges [29] by introducing domain-specific demands not found in standard benchmarks, motivating methods that can accommodate ecology's non-standard data, and improving generalisation and interpretability in settings characterised by mechanistic structure, long time horizons and limited supervision. Drawing on an interdisciplinary team, we frame our proposal to be actionable for both computer scientists and ecologists. We begin by surveying related efforts to create foundation models in other fields and highlight points of distinction relevant to ecology. We then outline a vision for what an ecoFM should be, from concept to capabilities. We describe its core requirements and, leveraging extensive discussions across ecology and AI, identify the ethical and societal risks that must be centred in its design. The proposed model would further both ecology and computer science through addressing shared challenges in prediction, generalisation and scaleability.

2 From Ecological Models to ML Models

In ecology, even simple systems can exhibit unexpectedly complex behaviour, as early non-linear population models demonstrated through oscillations, instabilities, and, later, chaotic dynamics [30]. These properties make ecology especially dependent on advances in quantitative modelling, both for inference and for understanding the mechanisms that underlie ecological patterns and dynamics. So, the development of ecological theory has long progressed in tandem with advances in statistical and mathematical methods [31, 32]. This trajectory continues today in more complex models of ecological interactions, including graph-theoretic models of trophic structure, and community stability [33–35].

Ecological modelling spans levels of biological organisation. At the genetic scale, evolutionary processes such as gene selection and mutation were commonly represented using time-stepped Markov chains [36]. More recently, deep learning has also been applied widely in genetics to identify population structure and probe natural selection [37]. Within communities, compositional change and trophic dynamics are examined using methods such as Lotka-Volterra and models based on graph theory [34, 35, 38]. At the macro-ecological scale, a long tradition of linear and linear-additive modelling has defined prediction of patterns over space [39], including species distribution models (SDMs) [40], and models linking body-size and metabolic rate such as Kleiber's law [41]. In many ways, the culmination of these multi-scale modelling efforts were general ecosystem models [42] which sought to model "all of life on earth," [43] first introduced in 2013 [42]. Currently, only the Madingley model has been introduced in this model class to date [44] although there has been little uptake among ecologists [42]. The lack of adoption of these models is largely due to the intense relational requirement;

they are prescriptive rather than data-driven and thus to model at the Earth level, relations need to be known. Across these domains, abstraction is essential for managing complexity, such as using traits as proxies for species, treating species as aggregates of underlying genetic variation or using genetic changes as a proxy for spatial variation [45]. Despite these different levels of abstraction and scale, all ecological processes are ultimately connected through the Earth system, and the data collected across scales are inherently linked.

Modelling in practice is shaped by a trade-off between complexity and effort [46, 47]. Adding more processes can increase realism, but often with diminishing returns as implementation cost grows. Additionally, the "bias-variance" tradeoff is realised in ecology in questions of how many and what kind of interactions to include [48]. As a result, models are typically designed to be just complex enough to address the scientific question, leading to a wide landscape of fragmented and specialised ecological models tailored to specific scenarios [47]. A similar pattern historically occurred in ML: excessive or hand-crafted features often caused over-fitting and poor generalisation [49, 50], prompting the development of highly specialised models for particular datasets and tasks.

This paradigm shifted with the introduction of deep learning-enabled attention mechanisms, which laid the ground-work for foundation models [51]. Self-attention and cross-attention allow models to allocate representational capacity dynamically, focusing computation on the most informative components of input data rather than relying on predefined task-specific structures. Deep learning architectures are typically used either as representation-learning models which learn flexible latent spaces amenable to downstream linear models or clustering, or as generative models which excel at scalable generation rather than producing structured latent embeddings. These mechanisms also scale incredibly well as more data are modelled, enabling models such as GPT to function as generalised text predictors for the majority of written human language [52].

Beyond language, foundation models have emerged across many domains. A defining feature is the ability to integrate multiple modalities, such as text, images, audio, and time series, within a unified deep learning architecture. This is often achieved through cross-attention, which allows representations from one modality to draw information from another; in other cases, modalities are fused through joint embeddings or by applying self-attention over combined input data [53–56]. These mechanisms allow foundation models to relate heterogeneous signals without requiring a bespoke model for each input combination. As a result, a model or collection of models can learn from many data sources and adapt to a wide range of downstream tasks.

The recent advancement in foundation models presents a monumental opportunity for ecology. A unified modelling framework could support more complex, multi-purpose, scalable, and generalisable ecological models, capable of representing cross-scale, multi-modal data in ways that are currently not available. This was simply not possible before the advent of large and complex enough machine learning paradigms. To understand how such a unified framework might be built, we next examine the growing landscape of existing models and ecology-adjacent approaches.

3 Ecology-Adjacent Deep Learning Landscape

Recent advances in machine learning have produced a growing set of deep learning-based models relevant to understanding ecological phenomena. These systems span three major domains: (i) perception, (ii) environmental feature mapping, and (iii) species distribution modelling. Understanding what these models already achieve and where they fall short clarifies the opportunities and constraints for developing an ecoFM.

3.1 Perception

We define perceptive models as those that retrieve relevant biological information from raw ecological data sources. Early work in this area largely centred on fine-grained visual classification benchmarks aimed at identifying visually similar taxa in images [57, 58]. This area has subsequently expanded with the emergence of foundational models, such as BioCLIP2, that learn joint image-text embeddings [59], and general-purpose vision backbones such as DINO, which, when trained towards general objectives to understand images, transfer remarkably well to species identification when fine-tuned or used as feature extractors [60, 61].

Scaling has also extended beyond purely visual inputs. Models such as TaxaBind integrate multiple information streams, including acoustic recordings, satellite imagery, and environmental covariates, to learn a unified embedding of a given observation with each of these data types [62]. In parallel, advances in bioacoustic processing enable robust automated recognition of species calls from noisy field deployments [63, 64].

The current primary utility of these perceptive models in ecology is as data-curation tools: they filter camera-trap data [65–67], classify citizen-science images [60], and detect acoustic events [68], and are highly flexible through zero-shot adaptation, which allows for the identification of unseen species [69]. It is clear that this area will continue to improve as more advancements are made. These feature extraction techniques can form representational layers that can be repurposed for ecoFMs.

3.2 Environmental Feature Mapping

Earth-observation (EO) foundation models constitute the second major group. These systems learn general-purpose representations of environmental states from large volumes of satellite imagery. Models such as TESSERA [70] and AlphaEarth [22] provide dense, globally consistent embeddings, which give insight into ecological processes and patterns such as land use change, vegetation structure, and ecosystem disturbance. Such representations offer an efficient route to integrate remotely sensed information into ecological analyses. By leveraging already existing models, the need to develop specialised EO models is bypassed. However, EO data have inherent limitations: spatial resolution remains coarse relative to fine-scale habitat variation, microhabitats are often unresolved, and performance is not alone sufficient for aquatic environments and beneath dense canopies [71, 72]. EO models therefore provide valuable inputs for ecoFMs but, on their own, do not represent ecological processes.

3.3 Species Distribution Modelling

Species distribution modelling (SDM) has traditionally been grounded in statistical approaches, including hierarchical Bayesian models, occupancy–detection frameworks, point process models, and joint species distribution models (JSDMs) [73]. These methods explicitly model uncertainty, spatial structure, and species co-occurrence, and continue to form the backbone of ecological inference.

More recently, deep learning methods have been introduced into this landscape [74], largely as a response to the growth of global-scale biodiversity datasets and high-resolution environmental feature mapping data. Neural networks trained on occurrence data, environmental rasters, or remote-sensing imagery can achieve predictive performance comparable to, and in some cases exceeding, classical SDMs such as MaxEnt, particularly at large spatial extents. However, these gains are context dependent and often diminish for rare or range-restricted species, under spatially structured validation, or when models are required to extrapolate beyond the training domain [75].

Progress in deep SDMs is expected to continue as deep learning methods mature, particularly through improved architectures, uncertainty estimation, and the integration of further data sources. Multimodality, in particular, offers a promising avenue for incorporating complementary information from remote sensing, citizen science, and other ecological data streams [76–78]. Nevertheless, even recent extensions largely treat prediction tasks in isolation, focusing on mapping species–environment relationships rather than representing the underlying ecology as a series of interconnected interactions and systems.

This limitation motivates the development of ecological foundation models. Rather than modelling species–environment relationships independently, ecoFMs aim to learn shared representations that integrate perception, environmental context, and biological interactions within a unified framework. By enabling transfer across tasks, scales, and taxa, ecoFMs extend beyond the scope of SDMs toward a more general and integrative representation of ecological systems.

4 Vision

Our vision for an ecoFM is a unified intelligent system for ecology that can integrate global data across genes, organisms, communities, environments, space, and time within a single coherent modelling framework. Such a model would learn representations that capture how species interact with each other and with their environments, support tasks that range from species identification to system forecasting and mechanistic inference, and enable synthesis across the currently fragmented landscape of ecological data, models, and theory. In effect, an ecoFM would serve as a general-purpose tool that represents the state and dynamics of ecology on Earth. Below, we outline the datasets needed for training, our proposed architecture and implementation details, and fully define the capabilities of an ecoFM.

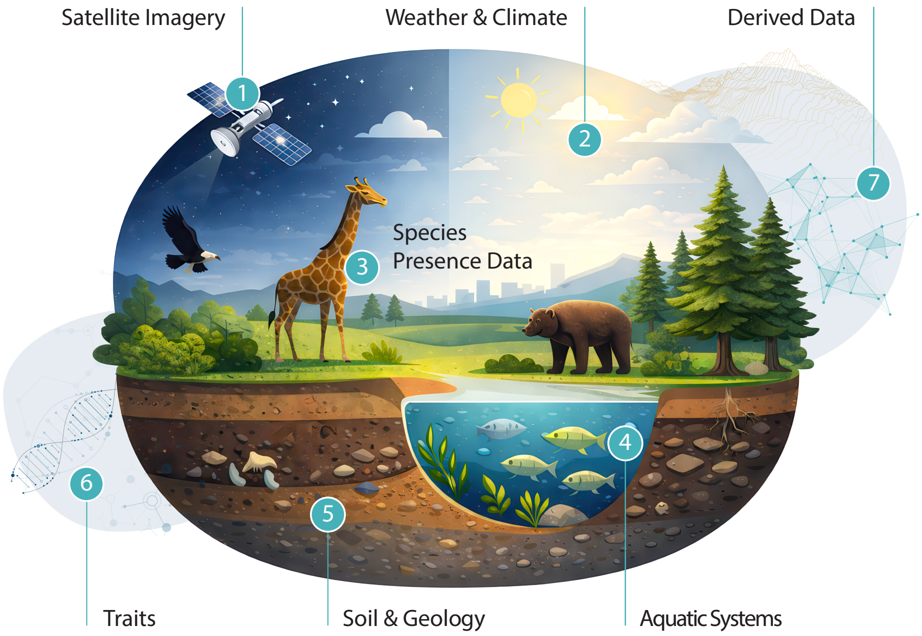

Fig. 1 Overview of ecological data spanning multiple spatial scales, sources, and data types. Numbers correspond to identified large-scale ecological datasets listed in Table A1.

4.1 Data

Large-scale, curated datasets are the core of any foundation model [52, 79]. Ecological data span many biological, spatial and temporal scales and occur in heterogeneous forms. In machine learning, a modality refers to a distinct type of data with its own structure, statistical properties and representation. In ecology, these modalities include images of species or their habitats, acoustic recordings, genetic sequences, trait measurements, environmental time series, satellite imagery, species occurrence records and more (Table A1). Each modality captures different aspects of ecological systems, requires different processing steps and has no standardised format that links it to other modalities.

Constructing a unified dataset for an ecoFM is challenging because these sources differ fundamentally in structure, resolution, sampling design, data quality and coverage. For example, data from acoustic recording devices can provide dense streams of observations in a given location, but can often be noisy and highly biased toward detectable or conspicuous species [80]. In contrast, species occurrence records and trait databases tend to be geographically and/or taxonomically uneven [60, 81], reflecting its own distribution of sampling effort. Environmental variables add further heterogeneity, varying widely in spatial and temporal granularity. Canonicalisation across these disparate data types therefore presents a significant challenge to be overcome for ecoFM development. In addition, the technical challenge of data cleaning as a whole should not be understated. While there are many large datasets in ecology, the amount of uncleaned and unordered data is several orders of magnitude higher still.

To clarify these opportunities and challenges, we identified representative datasets across key ecological modalities and scales (Table A1), visually illustrated by Figure 1. These include remote sensing products that capture environmental structure and dynamics, organismal observations from imagery and acoustics, genetic and trait datasets that characterise biological variation, taxonomic and phylogenetic knowledge graphs, and spatiotemporal environmental records used to contextualise ecological processes. This landscape motivates the need for unified tokenisation and multimodal representation learning methods capable of bridging these heterogeneous sources [26, 82]. In the following section, we detail how such methods integrate within the proposed ecoFM architecture.

4.2 Architecture

It is critical to define a realistic model architecture which can capture the relational structure of ecological systems and produce relevant, interpretable predictions. Our framework proceeds through several key stages: modality-specific encoders to embed diverse data types; multimodal masked modelling to learn shared structure across modalities; downstream task adaptation and interpretability tools to assess and probe the learned representation; uncertainty modelling to quantify confidence; and graph-based representations to express ecological interactions. These components are summarised in Figure 2.

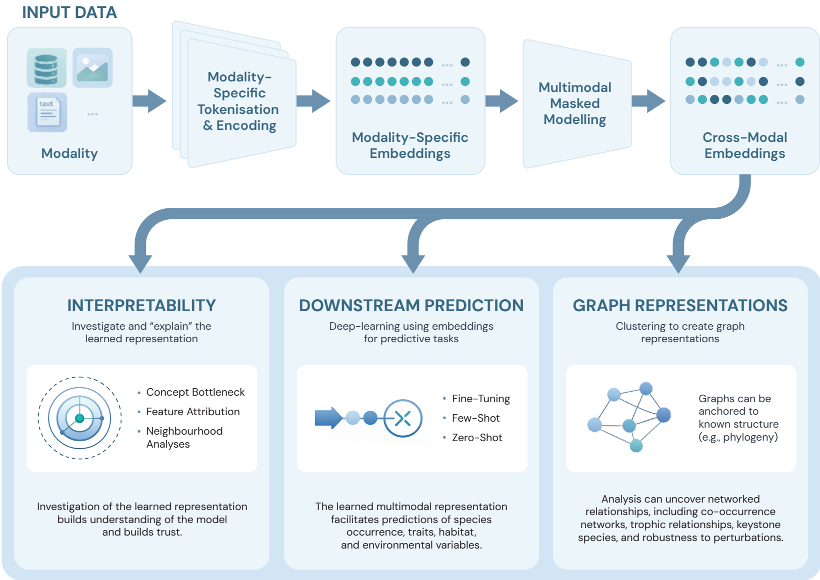

Fig. 2 Workflow illustrating how an ecological foundation model (ecoFM) integrates multimodal ecological data. Raw data from multiple modalities are tokenised and encoded into modality-specific representations, which are jointly encoded using a multimodal masked-modelling objective to produce a shared embedding space. This representation can then be adapted for downstream tasks, including predictive modelling (e.g., species distributions, occupancy, habitat inference). Interpretability methods (e.g., concept bottlenecks, global sensitivity analysis, or attribution analyses) can be applied to interrogate the learned representation. Graph-based tasks may additionally construct ecological networks from the shared embedding using clustering approaches, enabling network-based analyses (e.g., interaction networks, trophic structure, competitive hierarchies)

4.2.1 Modality Encoding

As previously discussed, ecological data span many modalities but are often sparse, imbalanced and incomplete. A core requirement of an ecoFM is therefore a modelling framework that can transform heterogeneous datatypes representing a common ontology into a shared representation space. In such a space, observations describing the same organism or event, e.g., a bird's call, its DNA barcode and its photograph, should map to nearby points. This is very far from the reality of biological data storage currently. Achieving this first requires modality-specific transformers or tokensisation schemes that convert raw data into a common embedding format, enabling feature extraction across modalities. These encoders can build on extensive prior advances in representation learning, particularly by leveraging the growing suite of pretrained, modality-specific models.

The ecoFM can exploit geospatial foundation models such as AlphaEarth [22], SatCLIP [83], GeoCLIP [84], or CLAY [85], which provide rich embeddings of land cover, environmental gradients, and temporal dynamics. This inclusion mirrors the approach taken by many fields of ecology, i.e., connecting shifts in population, movement, and ecosystems to changes in Earth's abiotic features [86, 87]. Incorporating these embeddings can follow multiple strategies: (1) contrastive learning during pre-training to ensure earth observation and biological embeddings share a semantic structure, (2) through conditioning variables that provide environmental context to biological predictions or (3) spatial tokens that represent environmental states at specific locations and times. These embeddings could lower the data requirements from ecological disciplines, reduce overall computational costs by leveraging models trained beyond the resources of a typical ecoFM project, and improve performance in data-poor regions by transferring environmental knowledge captured from globally sampled imagery.

4.2.2 Multimodal Masked Modelling

With these embeddings established, the next step is to learn how information is shared across modalities. A natural approach to learning this is multimodal masked modelling [88, 89], in which the model is trained to reconstruct missing embeddings or tokens from those that remain. This self-supervised objective is particularly well suited to ecological data, which are sparse, unevenly sampled, and rarely contain all modalities for a given observation. By predicting withheld information across modalities, the model is forced to learn the underlying structure that links them, capturing cross-modal dependencies rather than relying on correlations within any single data type. We therefore propose multimodal masked modelling as a central architectural paradigm for ecoFM development, offering a principled way to learn unified ecological representations from heterogeneous and incomplete data.

4.2.3 Downstream Prediction and Interpretability

Once a shared embedding space has been learned, its utility can be assessed through downstream task adaptation. These involve adding lightweight prediction heads or fine-tuning parts of the model to map embeddings onto particular outputs. Examples of this would potentially include species identification in heterogeneous datasets or prediction of traits or habitats. Performance on downstream tasks is the standard way to evaluate foundation models [14, 15], providing an initial measure of how well the multimodal embedding fits predictive tasks.

Complementing downstream prediction, interpretability methods allow users to analyse the embedding space in greater detail and understand how it is structured with respect to both the training data and the underlying model. Neighbourhood analyses [90] can reveal whether the learned representations recover expected ecological organisation, for example clustering by phylogeny, taxonomy or functional traits. Attribution methods, such as Shapley value analysis [91], can quantify the contribution of individual modalities to the resulting embeddings or predictions, indicating how the model integrates disparate data sources. More structured approaches, such as concept bottleneck models [92], allow users to probe whether specific ecological concepts or traits form identifiable axes within the embedding space and how these concepts mediate downstream predictions. We envision interpretability as a critical component of model evaluation and development of the overall training method.

4.2.4 Uncertainty Modelling

An ecoFM must provide not only predictions but also measures of uncertainty. Accounting for uncertainty is critical, as failure to represent model uncertainty can distort predictions, inflate confidence in spurious patterns, and produce ecological inferences that are systematically biased [93]. This goes beyond providing a simple confidence score, which reflects the model's stated certainty about an output but does not account for missing information, modality conflicts, or unfamiliar inputs. Instead, uncertainty should be embedded more fundamentally in the model's internal representations. Recent work such as ProM3E [94] demonstrates how this can be achieved through probabilistic embeddings, in which each latent feature is represented as a distribution with a mean and a variance. This approach allows the model to express uncertainty about the information it has learned, not just about its final prediction: variance increases when inputs are incomplete or ambiguous, and decreases when modalities are informative and consistent. Our ecoFM formulation adopts this framework so that uncertainty propagates through the model and informs downstream tasks, leading to ecological predictions and decisions that reflect data limitations and model confidence.

4.2.5 Graph Representations

Finally, recognising that ecological systems are naturally structured as networks, it is important for compatibility with the wealth of graph-based ecological methods, such as food webs, interaction networks, connectivity networks, and more, that the learned representation can be converted into a graph. This can be achieved through downstream clustering methods, for example by grouping embeddings into nodes that represent functional groups, taxa or locations, and defining edges based on latent similarity or inferred relationships between these groups. Interpretable meaning can be introduced by anchoring the resulting graph to known interaction data, such as the global dataset of species interactions [95], allowing predicted links to be compared with or constrained by established ecological knowledge. Once constructed, such graph representations provide a powerful substrate for analysing system-level properties, including the propagation of perturbations across ecological networks and the modelling of mechanistic behaviours and dynamics.

4.3 Capabilities

Building on our architecture, an ecoFM would support a broad range of ecological tasks, spanning from basic perception of environmental observations to a mechanistic understanding of underlying dynamics and potential unification of ecological theory. We lay these out as formal capabilities which a successful ecoFM should have:

- *Perception.* The model must ingest diverse ecological observations and extract meaningful features. This includes tasks such as identifying species or events from acoustic recordings, camera-trap images, community matrices, genetic barcodes, or other primary data streams.

- *Contextual understanding.* Ecological data are embedded in space, time, environment, and sampling method. The capacity to infer, reconstruct, or integrate this context would reflect a deeper level of understanding than raw perception.
- *Predictive capability.* The model should forecast species distributions, population trends, habitat suitability, biogeochemical cycling or other ecological patterns, mechanisms and dynamics across spatiotemporal scales. This is inherently statistical: learning correlative relationships between inputs and outputs, and mirroring existing species distribution models, but at far greater scale, flexibility, and multimodality.
- *Mechanistic understanding.* The highest capability level involves moving beyond correlations to represent ecological systems as networks of interacting organisms and environments, capturing processes such as trophic dynamics, competition, and stability. Predictions at this level are coupled with interpretable explanations of the ecological mechanisms that generate them.
- *Usability.* While embedding-based outputs such as maps and retrieval tools are increasingly accessible, generative and forecasting foundation models still require substantial technical expertise and computational resources to use. For ecoFMs to reach their full potential, models should be designed not just for scientific accuracy but for usability, allowing for adoption in both the scientific and policy space.
- *Unification.* Similar to other scientific foundation models [24], an ecoFM offers the potential to encode unifying ecological theories in its structure. This encoding of general theory would represent a scientific advancement in its own right.

5 Considerations

Foundation models represent a promising technological frontier for ecology and biodiversity, with the potential to transform how we monitor, understand, and predict ecological dynamics. Their impact, however, will not be determined by technical capability alone, but by how scientific, policy, and practitioner communities approach foresight, ethical design, and governance [96]. To realise the promise of ecoFMs, evaluation must go beyond marginal gains in accuracy on benchmark datasets to systematic auditing of failure modes and the measures taken to mitigate associated harms. Concretely, models should be assessed against principles of ethical AI: benevolence, non-maleficence, autonomy, justice, and explainability [97]. Below, we discuss the potential dangers associated with epistemic, technological, socio-ethical and environmental risks.

5.1 Epistemic Risks

A central epistemic risk of ecoFMs is that, trained on today's incomplete and biased datasets, they may amplify existing blind spots in biodiversity knowledge and entrench distorted understandings of the natural world. Ecological data are heavily skewed geographically and taxonomically: most observations come from the Global North, while biodiversity-rich regions remain undersampled, and certain vertebrate groups (e.g., birds) dominate records relative to insects, fungi, and microbial communities. The problem of mitigating epistemic risk is a common critique of ecology more generally, with effects such as spatial bias [81], artefacts of racist policies [98], and the after-effects of political instabilities [99], creating underrepresentation of taxa [81]. Bias is also introduced through annotation, such as misidentification of rare species, subjective assessments of habitat quality, and inconsistent labelling practices [100]. Additionally, large bodies of ecological knowledge, especially those held by Indigenous and local communities, remain inaccessible to current datasets and models. These biases are therefore socio-technical in origin, reflecting historical asymmetries, infrastructure limitations, and unequal access to scientific resources.

These problems can be partially addressed through targeted investment in under-represented regions and taxa, expanded citizen science and community-based monitoring, and the formal integration of diverse knowledge systems, including Traditional Ecological Knowledge. Integrating these into our proposed model represents a legitimate technical problem in response to an essential social responsibility. While imperfect data is one source of epistemic risk, another issue is understanding the causal mechanism between imperfect data and model outputs. This brings worries about epistemic risk into alignment with interpretability studies as a whole, i.e., understanding the causality of model decisions with respect to training data, which makes biases easier to identify. The integration of these typically unconsidered data sources will only be possible through partnership with knowledge holders in relevant areas and attention to ethical risks (c.f. Section 5.3).

5.2 Technical and Methodological Risks

A methodological risk arises from the limitations of contemporary deep learning itself. Neural networks exhibit well-documented spectral bias, favouring smooth, low-frequency functions [101] thereby underfitting sharp ecological gradients and the rare events that often govern ecological dynamics. Many architectures also implicitly impose Euclidean structure on their latent spaces [102], even when ecological relationships are fundamentally non-Euclidean, i.e. hierarchies of taxonomy or relationships. These inductive characteristics may yield distorted internal representations of ecological systems, undermining model development and introducing errors into downstream scientific analyses. Furthermore, most widely used deep learning models produce point estimates rather than probability distributions [103–105], leading to outputs that are overconfident and insensitive to data scarcity. As a result, they may fit spurious patterns [106, 107], generalise poorly beyond the distribution of training data [108], and offer limited insight into the reliability of their predictions [109], which is especially consequential when ecoFMs are used to support ecological inference and decision-making.

These vulnerabilities are well known within the deep learning community, and each can be mitigated through deliberate architectural and training choices that incorporate ecological prior knowledge. Models with broader spectral capacity, such as SIREN architectures or networks employing random Fourier features, can better capture high-frequency ecological signals and abrupt transitions [84, 110]. Geometry-aware representations, including projections into hyperbolic space, offer improved modelling of hierarchical relationships; however, optimisation for hyperbolic geometry remains computationally inefficient and can complicate convergence [111]. Process-informed neural networks, which embed known ecological dynamics directly into deep learning architectures, provide another route for constraining model behaviour while improving interpretability [112]. Data augmentation and rigorous cross-validation can help limit overfitting and improve performance on out-of-distribution samples. Finally, advances in Bayesian deep learning and ensemble modelling offer principled approaches for quantifying uncertainty and distinguishing data scarcity from underlying ecological variability. Collectively, these strategies are likely to play a central role in building ecoFMs, and real progress will depend on ecologists and computer scientists working in close collaboration.

5.3 Socio-ethical Risks

Without robust ethical safeguards, ecoFMs risk reinforcing existing global power asymmetries. A central concern is data colonialism: the extraction and use of data or knowledge from Indigenous communities and the Global South without appropriate recognition, benefit-sharing, or local governance. Such practices risk commodifying Indigenous knowledge and detaching it from its cultural and ecological context. In addition, the complexity of the AI lifecycle creates ambiguity around accountability: when model outputs lead to harmful downstream decisions, responsibility may be unclear among developers, deployers, and end-users.

Addressing these risks requires participatory and co-designed approaches to AI development, ensuring that affected communities are involved throughout the process and retain authority over how their data and knowledge are used [113]. Several established frameworks offer guidance for structuring governance, such as the Nagoya Protocol, which provides a precedent for recognising sovereignty over genetic and ecological data and ensuring that benefits flow back to providers [114]. The CARE Principles for Indigenous Data Governance and the broader Indigenous Data Sovereignty movement articulate clear expectations regarding collective consent, authority to control, and culturally appropriate data stewardship [115]. Public-sector AI guidelines such as the OECD AI Principles [116] distinguish responsibilities between data owners, model developers, and decision-makers. Translating these principles into ecoFMs will require transparent mechanisms for consent, benefit-sharing and accountability. Organisational structures which ensure that local and Indigenous knowledge holders retain meaningful decision-making power are also an essential part of development.

5.4 Environmental Risks

There is a profound paradox in leveraging energy-intensive technology to address environmental problems. Training state-of-the-art foundation models demands substantial computational resources, and the ensuing energy demand carries a large carbon footprint [117]. Data centres also require substantial water for cooling, and facilities located in water-stressed regions can exacerbate local water scarcity. Beyond contributing to climate change, AI data centres can affect local biodiversity by competing for water and reducing the environmental flows needed to sustain river biodiversity [118]. These well-known impacts of AI on the environment undermine public confidence in the utility of AI for advancing scientific discovery in ecology and its application to natural resource management and conservation problems.

These environmental costs must be central considerations in the design, deployment, and governance of ecoFMs. However, the relationship between AI and environment-focused research remains unclear[119]. Preliminary life-cycle analyses reveal serious carbon, water, and land usage associated with AI models [120]. At the same time, the foundation model paradigm offers a potential efficiency advantage [121]. Although pre-training is resource-intensive, it in principle occurs once; the resulting model can be fine-tuned for diverse downstream tasks using far less data, time, and energy than training separate models for each application. This "pre-train once, fine-tune many" approach represents a shift toward greater computational and environmental efficiency. However, the large upfront energy cost introduces environmental risk: if pre-training needs to be repeated, this cost is incurred again. As understanding of the environmental impacts of AI improves, the development of ecological foundation models should remain responsive to emerging evidence and act accordingly [119, 121].

6 Conclusion

Ecology is entering a phase in which its central scientific challenge—explaining how species interactions and environmental processes give rise to patterns of biodiversity, including their abundance, distribution and ecosystem consequences—aligns with a major inflection point in artificial intelligence. The rise of foundation models offers a path toward unifying disparate ecological data streams and representing ecological systems at the level where they operate: through interactions that scale from genes to ecosystems. Yet realising this potential will require far more than repurposing existing models. It demands multimodal data infrastructure, architectures that encode ecological structure, evaluation metrics grounded in ecological theory, and governance frameworks that ensure equity, transparency, and trust.

We have outlined the conceptual framework and technical requirements for ecoFMs, and provided a roadmap for how they could advance ecological prediction, interaction-centred analysis, and the synthesis of ecological data and theory. Importantly, ecoFMs should be built to complement, not replace, ecologists' existing tools, enabling synthesis across scales rather than obscuring the mechanisms that underpin ecological processes.

If developed responsibly and collaboratively, ecoFMs could mark a step change in ecology's capacity to learn from the accelerating torrent of ecological data and to reveal the structure of the natural world with greater clarity. At the same time, they could position ecology as a driver of new computational advances in learning from multimodal, theory-rich, and data-limited scientific systems. The challenge now is to build the scientific, computational, and ethical foundations required for these models to serve ecology's core goals: understanding, sustaining, and predicting the dynamics of a rapidly changing planet.

Supplementary information. Appendix A: Current Modelling Landscape Table A1: Data modalities, use cases, overlaps, example models, and datasets relevant for ecoFM development.

Acknowledgements. We thank Anil Madhavapeddy and Matteo Fumagalli for comments on an earlier version of this manuscript. We thank Kelly Cassella (MadeBold Design) for graphic design support, including visual refinement of author-generated figures.

Declarations

- Funding: Funding for this project was provided by a Community Initiatives grant from Schmidt Sciences to RHP and AH. RHP, AH, and YYT were supported by Eric and Wendy Schmidt AI in Science Postdoctoral Fellowships, a program of Schmidt Sciences. AC is funded under the NERC SUPER DTP (NE/S007342/1). HC is supported by a Rhodes Scholarship. SR was supported by a grant from the Leverhulme Trust (grant number RC-2021–076) as a member of the Leverhulme Centre for Nature Recovery. HAA was supported by the Canada Research Chair in Spatial Ecology and Biodiversity held by Pedro Peres-Neto, and funding from the NSERC-CREATE program on Leadership in Environmental and Digital Innovation for Sustainability (LEADS) at Concordia University, Montreal, Canada. SMF was supported by Hope Trust funding held by Geraldine Wright. CHB was supported by the Patrick J. McGovern Foundation. CR-P acknowledges support from the European Union European Research Council (ERC) Grant project No 101044703 (EUNICE). PSG is supported by a NERC DTP studentship. RJG is supported by a doctoral training grant awarded as part of the UKRI AI Centre for Doctoral Training in Environmental Intelligence (UKRI grant number EP/S022074/1).
- Competing interests: The authors declare no competing interests.
- Ethics approval and consent to participate: Not applicable.
- Consent for publication: All authors provide consent for publication.
- Data availability: Not applicable.
- Materials availability: Not applicable.
- Code availability: Not applicable.
- Author contributions: RHP, RJG, HC, HAA, CHB, AC, SMF, PSG, ETM, RO, SR, YYT, CR-P, MB, and SH conceptualised the manuscript. RHP, RJG, and SMF designed the figures. RHP and AH coordinated the project. All authors contributed to writing and editing.

Appendix A Data table

Table A1 Data modalities, use cases, and example datasets relevant for ecoFM development.

	Domain	Data type	Example datasets
1	Satellite imagery	Land cover Topography	NASA Harmonised Landsat & Sentinel 2 (HLS)
2	Weather & Climate	Historic & future climate	Copernicus Climate Data Store Temperature 12k
3	Species presence data	Presence-only data	GBIF iNaturalist Natural history museum collections
		Surveys with effort information	BioTIME eBird
		Passively sampled data	FishSounds xeno-canto ecosound-web.de GBIF
4	Aquatic systems	Water quality/pollution	NASA HLS National hydrological surveys
		Bathymetry	IHO Data Centre for Digital Bathymetry
5	Soil & geology	Soil type & chemistry Seismic & volcanic activity	National geological surveys
6	Traits	DNA sequences	GenBank BOLD GoaT/EBG
		Phenotypic traits	Encyclopedia of Life TRY Phenoscape AVONET FishBase
7	Derived data	Taxonomy/Phylogeny	Catalogue of Life Open Tree of Life
		Range maps	IUCN range maps AVONET FishBase
		Foodwebs & interactions	GloBI

References

[1] Taylor, W.P.: What is Ecology and What Good is It? Ecology **17**(3), 333–346 (1936) https://doi.org/10.2307/1931838 1931838

[2] Lawton, J.H.: Are There General Laws in Ecology? Oikos **84**(2), 177 (1999) https://doi.org/10.2307/3546712 3546712

[3] Turchin, P.: Does population ecology have general laws? Oikos **94**(1), 17–26 (2001) https://doi.org/10.1034/j.1600-0706.2001.11310.x

[4] Perrins, C.M.: Birds. In: Savill, P., Perrins, C., Kirby, K., Fisher, N. (eds.) Wytham Woods: Oxford's Ecological Laboratory, pp. 145–172. Oxford University Press, Oxford (2011). https://doi.org/10.1093/acprof:osobl/9780199605187.003.0009

[5] Nelson, M.P., Vucetich, J.A., Peterson, R.O., Vucetich, L.M.: The Isle Royale Wolf–Moose Project (1958-present) and the Wonder of Long-Term Ecological Research. Endeavour **35**(1), 31–39 (2011) https://doi.org/10.1016/j.endeavour.2010.09.002

[6] Hobbie, J.E., Carpenter, S.R., Grimm, N.B., Gosz, J.R., Seastedt, T.R.: The US Long Term Ecological Research Program. BioScience **53**(1), 21–32 (2003) https://doi.org/10.1641/0006-3568(2003)053[0021:TULTER]2.0.CO;2

[7] Silvertown, J., Poulton, P., Johnston, E., Edwards, G., Heard, M., Biss, P.M.: The Park Grass Experiment 1856–2006: Its contribution to ecology. Journal of Ecology **94**(4), 801–814 (2006) https://doi.org/10.1111/j.1365-2745.2006.01145.x

[8] Gould, E., Fraser, H.S., Parker, T.H., Nakagawa, S., Griffith, S.C., Vesk, P.A., Fidler, F., Hamilton, D.G., Abbey-Lee, R.N., Abbott, J.K., Aguirre, L.A., Alcaraz, C., Aloni, I., Altschul, D., Arekar, K., Atkins, J.W., Atkinson, J., Baker, C.M., Barrett, M., Bell, K., Bello, S.K., Beltrán, I., Berauer, B.J., Bertram, M.G., Billman, P.D., Blake, C.K., Blake, S., Bliard, L., Bonisoli-Alquati, A., Bonnet, T., Bordes, C.N.M., Bose, A.P.H., Botterill-James, T., Boyd, M.A., Boyle, S.A., Bradfer-Lawrence, T., Bradham, J., Brand, J.A., Brengdahl, M.I., Bulla, M., Bussière, L., Camerlenghi, E., Campbell, S.E., Campos, L.L.F., Caravaggi, A., Cardoso, P., Carroll, C.J.W., Catanach, T.A., Chen, X., Chik, H.Y.J., Choy, E.S., Christie, A.P., Chuang, A., Chunco, A.J., Clark, B.L., Contina, A., Covernton, G.A., Cox, M.P., Cressman, K.A., Crotti, M., Crouch, C.D., D'Amelio, P.B., de Sousa, A.A., Döbert, T.F., Dobler, R., Dobson, A.J., Doherty, T.S., Drobniak, S.M., Duffy, A.G., Duncan, A.B., Dunn, R.P., Dunning, J., Dutta, T., Eberhart-Hertel, L., Elmore, J.A., Elsherif, M.M., English, H.M., Ensminger, D.C., Ernst, U.R., Ferguson, S.M., Fernandez-Juricic, E., Ferreira-Arruda, T., Fieberg, J., Finch, E.A., Fiorenza, E.A., Fisher, D.N., Fontaine, A., Forstmeier, W., Fourcade, Y., Frank, G.S., Freund, C.A., Fuentes-Lillo, E.,

Gandy, S.L., Gannon, D.G., García-Cervigón, A.I., Garretson, A.C., Ge, X., Geary, W.L., Géron, C., Gilles, M., Girndt, A., Gliksman, D., Goldspiel, H.B., Gomes, D.G.E., Good, M.K., Goslee, S.C., Gosnell, J.S., Grames, E.M., Gratton, P., Grebe, N.M., Greenler, S.M., Griffioen, M., Griffith, D.M., Griffith, F.J., Grossman, J.J., Güncan, A., Haesen, S., Hagan, J.G., Hager, H.A., Harris, J.P., Harrison, N.D., Hasnain, S.S., Havird, J.C., Heaton, A.J., Herrera-Chaustre, M.L., Howard, T.J., Hsu, B.-Y., Iannarilli, F., Iranzo, E.C., Iverson, E.N.K., Jimoh, S.O., Johnson, D.H., Johnsson, M., Jorna, J., Jucker, T., Jung, M., Kačergytė, I., Kaltz, O., Ke, A., Kelly, C.D., Keogan, K., Keppeler, F.W., Killion, A.K., Kim, D., Kochan, D.P., Korsten, P., Kothari, S., Kuppler, J., Kusch, J.M., Lagisz, M., Lalla, K.M., Larkin, D.J., Larson, C.L., Lauck, K.S., Lauterbur, M.E., Law, A., Léandri-Breton, D.-J., Lembrechts, J.J., L'Herpiniere, K., Lievens, E.J.P., de Lima, D.O., Lindsay, S., Luquet, M., MacLeod, R., Macphie, K.H., Magellan, K., Mair, M.M., Malm, L.E., Mammola, S., Mandeville, C.P., Manhart, M., Manrique-Garzon, L.M., Mäntylä, E., Marchand, P., Marshall, B.M., Martin, C.A., Martin, D.A., Martin, J.M., Martinig, A.R., McCallum, E.S., McCauley, M., McNew, S.M., Meiners, S.J., Merkling, T., Michelangeli, M., Moiron, M., Moreira, B., Mortensen, J., Mos, B., Muraina, T.O., Murphy, P.W., Nelli, L., Niemelä, P., Nightingale, J., Nilsonne, G., Nolazco, S., Nooten, S.S., Novotny, J.L., Olin, A.B., Organ, C.L., Ostevik, K.L., Palacio, F.X., Paquet, M., Parker, D.J., Pascall, D.J., Pasquarella, V.J., Paterson, J.H., Payo-Payo, A., Pedersen, K.M., Perez, G., Perry, K.I., Pottier, P., Proulx, M.J., Proulx, R., Pruett, J.L., Ramananjato, V., Randimbiarison, F.T., Razafindratsima, O.H., Rennison, D.J., Riva, F., Riyahi, S., Roast, M.J., Rocha, F.P., Roche, D.G., Román-Palacios, C., Rosenberg, M.S., Ross, J., Rowland, F.E., Rugemalila, D., Russell, A.L., Ruuskanen, S., Saccone, P., Sadeh, A., Salazar, S.M., Sales, K., Salmón, P., Sánchez-Tójar, A., Santos, L.P., Santostefano, F., Schilling, H.T., Schmidt, M., Schmoll, T., Schneider, A.C., Schrock, A.E., Schroeder, J., Schtickzelle, N., Schultz, N.L., Scott, D.A., Scroggie, M.P., Shapiro, J.T., Sharma, N., Shearer, C.L., Simón, D., Sitvarin, M.I., Skupien, F.L., Slinn, H.L., Smith, G.P., Smith, J.A., Sollmann, R., Whitney, K.S., Still, S.M., Stuber, E.F., Sutton, G.F., Swallow, B., Taff, C.C., Takola, E., Tanentzap, A.J., Tarjuelo, R., Telford, R.J., Thawley, C.J., Thierry, H., Thomson, J., Tidau, S., Tompkins, E.M., Tortorelli, C.M., Trlica, A., Turnell, B.R., Urban, L., Van de Vondel, S., van der Wal, J.E.M., Van Eeckhoven, J., van Oordt, F., Vanderwel, K.M., Vanderwel, M.C., Vanderwolf, K.J., Vélez, J., Vergara-Florez, D.C., Verrelli, B.C., Vieira, M.V., Villamil, N., Vitali, V., Vollering, J., Walker, J., Walker, X.J., Walter, J.A., Waryszak, P., Weaver, R.J., Wedegärtner, R.E.M., Weller, D.L., Whelan, S.: Same data, different analysts: Variation in effect sizes due to analytical decisions in ecology and evolutionary biology. BMC Biology **23**(1), 35 (2025) https://doi.org/10.1186/s12915-024-02101-x

[9] Besson, M., Alison, J., Bjerge, K., Gorochowski, T.E., Høye, T.T., Jucker, T., Mann, H.M.R., Clements, C.F.: Towards the fully automated monitoring of ecological communities. Ecology Letters **25**(12), 2753–2775 (2022) https://doi.org/10.1111/ele.14123

[10] Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J.Q., Demszky, D., Donahue, C., Doumbouya, M., Durmus, E., Ermon, S., Etchemendy, J., Ethayarajh, K., Fei-Fei, L., Finn, C., Gale, T., Gillespie, L., Goel, K., Goodman, N., Grossman, S., Guha, N., Hashimoto, T., Henderson, P., Hewitt, J., Ho, D.E., Hong, J., Hsu, K., Huang, J., Icard, T., Jain, S., Jurafsky, D., Kalluri, P., Karamcheti, S., Keeling, G., Khani, F., Khattab, O., Koh, P.W., Krass, M., Krishna, R., Kuditipudi, R., Kumar, A., Ladhak, F., Lee, M., Lee, T., Leskovec, J., Levent, I., Li, X.L., Li, X., Ma, T., Malik, A., Manning, C.D., Mirchandani, S., Mitchell, E., Munyikwa, Z., Nair, S., Narayan, A., Narayanan, D., Newman, B., Nie, A., Niebles, J.C., Nilforoshan, H., Nyarko, J., Ogut, G., Orr, L., Papadimitriou, I., Park, J.S., Piech, C., Portelance, E., Potts, C., Raghunathan, A., Reich, R., Ren, H., Rong, F., Roohani, Y., Ruiz, C., Ryan, J., Ré, C., Sadigh, D., Sagawa, S., Santhanam, K., Shih, A., Srinivasan, K., Tamkin, A., Taori, R., Thomas, A.W., Tramèr, F., Wang, R.E., Wang, W., Wu, B., Wu, J., Wu, Y., Xie, S.M., Yasunaga, M., You, J., Zaharia, M., Zhang, M., Zhang, T., Zhang, X., Zhang, Y., Zheng, L., Zhou, K., Liang, P.: On the Opportunities and Risks of Foundation Models. arXiv (2022). https://doi.org/10.48550/arXiv.2108.07258

[11] HAI, S.: AI Index. https://hai.stanford.edu/ai-index. Accessed: 2025-11-17 (2025)

[12] Moorthy, S.M.K., Qi, M., Rosen, A., Malhi, Y., Salguero-Gomez, R.: Harnessing Large Language Models for Ecological Literature Reviews: A Practical Pipeline (2025)

[13] Dorm, F., Millard, J., Purves, D., Harfoot, M., Aodha, O.M.: Large Language Models Possess Some Ecological Knowledge, but How Much? bioRxiv (2025). https://doi.org/10.1101/2025.02.10.637097

[14] Vendrow, E., Pantazis, O., Shepard, A., Brostow, G., Jones, K.E., Aodha, O.M., Beery, S., Horn, G.V.: INQUIRE: A Natural World Text-to-Image Retrieval Benchmark. arXiv (2024). https://doi.org/10.48550/arXiv.2411.02537

[15] Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., Sutskever, I.: Learning Transferable Visual Models From Natural Language Supervision. arXiv (2021). https://doi.org/10.48550/arXiv.2103.00020

[16] Carion, N., Gustafson, L., Hu, Y.-T., Debnath, S., Hu, R., Suris, D., Ryali, C., Alwala, K.V., Khedr, H., Huang, A., Lei, J., Ma, T., Guo, B., Kalla, A., Marks, M., Greer, J., Wang, M., Sun, P., Rädle, R., Afouras, T., Mavroudi, E., Xu, K., Wu, T.-H., Zhou, Y., Momeni, L., Hazra, R., Ding, S., Vaze, S., Porcher, F., Li, F., Li, S., Kamath, A., Cheng, H.K., Dollár, P., Ravi, N., Saenko, K., Zhang, P., Feichtenhofer, C.: SAM 3: Segment Anything with Concepts. arXiv (2025).

https://doi.org/10.48550/arXiv.2511.16719

[17] Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., Joulin, A.: Emerging Properties in Self-Supervised Vision Transformers. arXiv (2021). https://doi.org/10.48550/arXiv.2104.14294

[18] Guo, D., Yang, D., Zhang, H., Song, J., Wang, P., Zhu, Q., Xu, R., Zhang, R., Ma, S., Bi, X., Zhang, X., Yu, X., Wu, Y., Wu, Z.F., Gou, Z., Shao, Z., Li, Z., Gao, Z., Liu, A., Xue, B., Wang, B., Wu, B., Feng, B., Lu, C., Zhao, C., Deng, C., Ruan, C., Dai, D., Chen, D., Ji, D., Li, E., Lin, F., Dai, F., Luo, F., Hao, G., Chen, G., Li, G., Zhang, H., Xu, H., Ding, H., Gao, H., Qu, H., Li, H., Guo, J., Li, J., Chen, J., Yuan, J., Tu, J., Qiu, J., Li, J., Cai, J.L., Ni, J., Liang, J., Chen, J., Dong, K., Hu, K., You, K., Gao, K., Guan, K., Huang, K., Yu, K., Wang, L., Zhang, L., Zhao, L., Wang, L., Zhang, L., Xu, L., Xia, L., Zhang, M., Zhang, M., Tang, M., Zhou, M., Li, M., Wang, M., Li, M., Tian, N., Huang, P., Zhang, P., Wang, Q., Chen, Q., Du, Q., Ge, R., Zhang, R., Pan, R., Wang, R., Chen, R.J., Jin, R.L., Chen, R., Lu, S., Zhou, S., Chen, S., Ye, S., Wang, S., Yu, S., Zhou, S., Pan, S., Li, S.S., Zhou, S., Wu, S., Yun, T., Pei, T., Sun, T., Wang, T., Zeng, W., Liu, W., Liang, W., Gao, W., Yu, W., Zhang, W., Xiao, W.L., An, W., Liu, X., Wang, X., Chen, X., Nie, X., Cheng, X., Liu, X., Xie, X., Liu, X., Yang, X., Li, X., Su, X., Lin, X., Li, X.Q., Jin, X., Shen, X., Chen, X., Sun, X., Wang, X., Song, X., Zhou, X., Wang, X., Shan, X., Li, Y.K., Wang, Y.Q., Wei, Y.X., Zhang, Y., Xu, Y., Li, Y., Zhao, Y., Sun, Y., Wang, Y., Yu, Y., Zhang, Y., Shi, Y., Xiong, Y., He, Y., Piao, Y., Wang, Y., Tan, Y., Ma, Y., Liu, Y., Guo, Y., Ou, Y., Wang, Y., Gong, Y., Zou, Y., He, Y., Xiong, Y., Luo, Y., You, Y., Liu, Y., Zhou, Y., Zhu, Y.X., Huang, Y., Li, Y., Zheng, Y., Zhu, Y., Ma, Y., Tang, Y., Zha, Y., Yan, Y., Ren, Z.Z., Ren, Z., Sha, Z., Fu, Z., Xu, Z., Xie, Z., Zhang, Z., Hao, Z., Ma, Z., Yan, Z., Wu, Z., Gu, Z., Zhu, Z., Liu, Z., Li, Z., Xie, Z., Song, Z., Pan, Z., Huang, Z., Xu, Z., Zhang, Z., Zhang, Z.: DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning. Nature **645**(8081), 633–638 (2025) https://doi.org/10.1038/s41586-025-09422-z

[19] Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D.M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D.: Language Models Are Few-Shot Learners. arXiv (2020). https://doi.org/10.48550/arXiv.2005.14165

[20] Wiedemer, T., Li, Y., Vicol, P., Gu, S.S., Matarese, N., Swersky, K., Kim, B., Jaini, P., Geirhos, R.: Video Models Are Zero-Shot Learners and Reasoners. arXiv (2025). https://doi.org/10.48550/arXiv.2509.20328

[21] Balestriero, R., LeCun, Y.: LeJEPA: Provable and Scalable Self-Supervised Learning Without the Heuristics. arXiv (2025). https://doi.org/10.48550/arXiv.

2511.08544

[22] Brown, C.F., Kazmierski, M.R., Pasquarella, V.J., Rucklidge, W.J., Samsikova, M., Zhang, C., Shelhamer, E., Lahera, E., Wiles, O., Ilyushchenko, S., Gorelick, N., Zhang, L.L., Alj, S., Schechter, E., Askay, S., Guinan, O., Moore, R., Boukouvalas, A., Kohli, P.: AlphaEarth Foundations: An Embedding Field Model for Accurate and Efficient Global Mapping from Sparse Label Data. arXiv (2025). https://doi.org/10.48550/arXiv.2507.22291

[23] McCabe, M., Mukhopadhyay, P., Marwah, T., Blancard, B.R.-S., Rozet, F., Diaconu, C., Meyer, L., Wong, K.W.K., Sotoudeh, H., Bietti, A., Espejo, I., Fear, R., Golkar, S., Hehir, T., Hirashima, K., Krawezik, G., Lanusse, F., Morel, R., Ohana, R., Parker, L., Pettee, M., Shen, J., Cho, K., Cranmer, M., Ho, S.: Walrus: A Cross-Domain Foundation Model for Continuum Dynamics. arXiv (2025). https://doi.org/10.48550/arXiv.2511.15684

[24] Brixi, G., Durrant, M.G., Ku, J., Poli, M., Brockman, G., Chang, D., Gonzalez, G.A., King, S.H., Li, D.B., Merchant, A.T., Naghipourfar, M., Nguyen, E., Ricci-Tam, C., Romero, D.W., Sun, G., Taghibakshi, A., Vorontsov, A., Yang, B., Deng, M., Gorton, L., Nguyen, N., Wang, N.K., Adams, E., Baccus, S.A., Dillmann, S., Ermon, S., Guo, D., Ilango, R., Janik, K., Lu, A.X., Mehta, R., Mofrad, M.R.K., Ng, M.Y., Pannu, J., Ré, C., Schmok, J.C., John, J.S., Sullivan, J., Zhu, K., Zynda, G., Balsam, D., Collison, P., Costa, A.B., Hernandez-Boussard, T., Ho, E., Liu, M.-Y., McGrath, T., Powell, K., Burke, D.P., Goodarzi, H., Hsu, P.D., Hie, B.L.: Genome Modeling and Design across All Domains of Life with Evo 2. bioRxiv (2025). https://doi.org/10.1101/2025.02.18.638918

[25] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D.: Highly accurate protein structure prediction with AlphaFold. Nature **596**(7873), 583–589 (2021) https://doi.org/10.1038/s41586-021-03819-2

[26] Parker, L., Lanusse, F., Shen, J., Liu, O., Hehir, T., Sarra, L., Meyer, L., Bowles, M., Wagner-Carena, S., Qu, H., Golkar, S., Bietti, A., Bourfoune, H., Casserau, N., Cornette, P., Hirashima, K., Krawezik, G., Ohana, R., Lourie, N., McCabe, M., Morel, R., Mukhopadhyay, P., Pettee, M., Blancard, B.R.-S., Cho, K., Cranmer, M., Ho, S.: AION-1: Omnimodal Foundation Model for Astronomical Sciences. arXiv (2025). https://doi.org/10.48550/arXiv.2510.17960

[27] Binz, M., Akata, E., Bethge, M., Brändle, F., Callaway, F., Coda-Forno, J., Dayan, P., Demircan, C., Eckstein, M.K., Éltető, N., Griffiths, T.L., Haridi, S., Jagadish, A.K., Ji-An, L., Kipnis, A., Kumar, S., Ludwig, T., Mathony, M.,

Mattar, M., Modirshanechi, A., Nath, S.S., Peterson, J.C., Rmus, M., Russek, E.M., Saanum, T., Schubert, J.A., Schulze Buschoff, L.M., Singhi, N., Sui, X., Thalmann, M., Theis, F.J., Truong, V., Udandarao, V., Voudouris, K., Wilson, R., Witte, K., Wu, S., Wulff, D.U., Xiong, H., Schulz, E.: A foundation model to predict and capture human cognition. Nature **644**(8078), 1002–1009 (2025) https://doi.org/10.1038/s41586-025-09215-4

[28] Khan, S.A., Martínez-de-Morentin, X., Alsabbagh, A.R., Maillo, A., Lagani, V., Gomez-Cabrero, D., Lehmann, R., Tegner, J.: Multimodal foundation transformer models for multiscale genomics. Nature Methods, 1–13 (2025) https://doi.org/10.1038/s41592-025-02918-6

[29] Han, B.A., Varshney, K.R., LaDeau, S., Subramaniam, A., Weathers, K.C., Zwart, J.: A synergistic future for AI and ecology. Proceedings of the National Academy of Sciences **120**(38), 2220283120 (2023) https://doi.org/10.1073/pnas.2220283120

[30] May, R.M.: Simple mathematical models with very complicated dynamics. Nature **261**, 459–467 (1976) https://doi.org/10.1038/261459a0

[31] Clark, J.S.: Why environmental scientists are becoming bayesians. Ecology Letters **8**(1), 2–14 (2005) https://doi.org/10.1111/j.1461-0248.2004.00702.x

[32] Gilbert, N.A., Amaral, B.R., Smith, O.M., Williams, P.J., Ceyzyk, S., Ayebare, S., Davis, K.L., Leuenberger, W., Doser, J.W., Zipkin, E.F.: A century of statistical ecology. Ecology **105**(6), 4283 (2024) https://doi.org/10.1002/ecy.4283

[33] May, R.M., Levin, S.A., Sugihara, G.: Complex systems: Ecology for bankers. Nature **451**(7181), 893–895 (2008) https://doi.org/10.1038/451893a

[34] Delmas, E., Besson, M., Brice, M.-H., Burkle, L.A., Dalla Riva, G.V., Fortin, M.-J., Gravel, D., Guimarães Jr., P.R., Hembry, D.H., Newman, E.A., Olesen, J.M., Pires, M.M., Yeakel, J.D., Poisot, T.: Analysing ecological networks of species interactions. Biological Reviews **94**(1), 16–36 (2019) https://doi.org/10.1111/brv.12433

[35] Kalirad, A., Sommer, R.J.: Ecological graph theory: Simulating competition and coexistence on graphs. Methods in Ecology and Evolution **16**(11), 2667–2680 (2025) https://doi.org/10.1111/2041-210X.70121

[36] Wright, S.: EVOLUTION IN MENDELIAN POPULATIONS. Genetics **16**(2), 97–159 (1931) https://doi.org/10.1093/genetics/16.2.97

[37] Huang, X., Rymbekova, A., Dolgova, O., Lao, O., Kuhlwilm, M.: Harnessing deep learning for population genetic inference. Nature Reviews Genetics **25**(1), 61–78 (2024) https://doi.org/10.1038/s41576-023-00636-3

[38] Bonnaffé, W., Sheldon, B.C., Coulson, T.: Neural ordinary differential equations for ecological and evolutionary time-series analysis. Methods in Ecology and Evolution **12**(7), 1301–1315 (2021) https://doi.org/10.1111/2041-210X.13606

[39] Purvis, A., Newbold, T., De Palma, A., Contu, S., Hill, S.L., Sanchez-Ortiz, K., Phillips, H.R.P., Hudson, L.N., Lysenko, I., Börger, L., Scharlemann, J.P.W.: Modelling and projecting the response of local terrestrial biodiversity worldwide to land use and related pressures: The predicts project. In: Bohan, D.A., Dumbrell, A.J., Woodward, G., Jackson, M. (eds.) Advances in Ecological Research vol. 58, pp. 201–241. Academic Press, ??? (2018). https://doi.org/10.1016/bs.aecr.2017.12.003

[40] Elith, J., Leathwick, J.R.: Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annual Review of Ecology, Evolution, and Systematics **40**(Volume 40, 2009), 677–697 (2009) https://doi.org/10.1146/annurev.ecolsys.110308.120159

[41] Kleiber, M.: Body size and metabolism. Hilgardia **6**(11), 315–353 (1932) https://doi.org/10.3733/hilg.v06n11p315

[42] Töpper, J.P., Hoeks, S., Hartig, F., Kolstad, A., Tucker, M.A., Bartlett, J., Bærum, K.M., Chipperfield, J., Harfoot, M., Huijbregts, M.A.J., Svenning, J.-C.: Advancing general ecosystem models (GEMs): Towards a mechanistic understanding of the biosphere in the light of the Anthropocene. Ecological Solutions and Evidence **6**(2), 70053 (2025) https://doi.org/10.1002/2688-8319.70053

[43] Purves, D., Scharlemann, J.P.W., Harfoot, M., Newbold, T., Tittensor, D.P., Hutton, J., Emmott, S.: Time to model all life on Earth. Nature **493**(7432), 295–297 (2013) https://doi.org/10.1038/493295a

[44] Harfoot, M.B.J., Newbold, T., Tittensor, D.P., Emmott, S., Hutton, J., Lyutsarev, V., Smith, M.J., Scharlemann, J.P.W., Purves, D.W.: Emergent Global Patterns of Ecosystem Structure and Function from a Mechanistic General Ecosystem Model. PLOS Biology **12**(4), 1001841 (2014) https://doi.org/10.1371/journal.pbio.1001841

[45] Bradburd, G.S., Ralph, P.L.: Spatial Population Genetics: It's About Time. Annual Review of Ecology, Evolution, and Systematics **50**(1), 427–449 (2019) https://doi.org/10.1146/annurev-ecolsys-110316-022659

[46] Malmborg, C.A., Willson, A.M., Bradley, L.M., Beatty, M.A., Klinges, D.H., Koren, G., Lewis, A.S.L., Oshinubi, K., Woelmer, W.M.: Defining model complexity: An ecological perspective. Meteorological Applications **31**(3), 2202 (2024) https://doi.org/10.1002/met.2202

[47] Geary, W.L., Bode, M., Doherty, T.S., Fulton, E.A., Nimmo, D.G., Tulloch, A.I.T., Tulloch, V.J.D., Ritchie, E.G.: A guide to ecosystem models and their

environmental applications. Nature Ecology & Evolution **4**(11), 1459–1471 (2020) https://doi.org/10.1038/s41559-020-01298-8

[48] Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer Series in Statistics. Springer, New York, NY (2009). https://doi.org/10.1007/978-0-387-84858-7

[49] Lu, Y.: Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics **6**(1), 1–29 (2019) https://doi.org/10.1080/23270012.2019.1570365

[50] Gatla, R.K., Gatla, A., Sridhar, P., Kumar, D.G., Rao, D.S.N.M.: Advancements in Generative AI: Exploring Fundamentals and Evolution. In: 2024 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC), pp. 1–5 (2024). https://doi.org/10.1109/ICECCC61767.2024.10594003

[51] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention Is All You Need. arXiv (2023). https://doi.org/10.48550/arXiv.1706.03762

[52] Kaplan, J., McCandlish, S., Henighan, T., Brown, T.B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., Amodei, D.: Scaling Laws for Neural Language Models. arXiv (2020). https://doi.org/10.48550/arXiv.2001.08361

[53] Zong, Y., Aodha, O.M., Hospedales, T.M.: Self-Supervised Multimodal Learning: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **47**(7), 5299–5318 (2025) https://doi.org/10.1109/TPAMI.2024.3429301

[54] Huang, Q., Jansen, A., Lee, J., Ganti, R., Li, J.Y., Ellis, D.P.W.: MuLan: A Joint Embedding of Music Audio and Natural Language. arXiv (2022). https://doi.org/10.48550/arXiv.2208.12415

[55] Assran, M., Duval, Q., Misra, I., Bojanowski, P., Vincent, P., Rabbat, M., LeCun, Y., Ballas, N.: Self-Supervised Learning From Images With a Joint-Embedding Predictive Architecture. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 15619–15629 (2023)

[56] Alayrac, J.-B., Recasens, A., Schneider, R., Arandjelović, R., Ramapuram, J., De Fauw, J., Smaira, L., Dieleman, S., Zisserman, A.: Self-Supervised MultiModal Versatile Networks. In: Advances in Neural Information Processing Systems, vol. 33, pp. 25–37. Curran Associates, Inc., ??? (2020)

[57] Berg, T., Liu, J., Lee, S.W., Alexander, M.L., Jacobs, D.W., Belhumeur, P.N.: Birdsnap: Large-Scale Fine-Grained Visual Categorization of Birds. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2019–2026. IEEE, Columbus, OH, USA (2014). https://doi.org/10.1109/CVPR.2014.259

[58] Wei, X.-S., Song, Y.-Z., Aodha, O.M., Wu, J., Peng, Y., Tang, J., Yang, J., Belongie, S.: Fine-Grained Image Analysis with Deep Learning: A Survey. arXiv (2021). https://doi.org/10.48550/arXiv.2111.06119

[59] Gu, J., Stevens, S., Campolongo, E.G., Thompson, M.J., Zhang, N., Wu, J., Kopanev, A., Mai, Z., White, A.E., Balhoff, J., Dahdul, W., Rubenstein, D., Lapp, H., Berger-Wolf, T., Chao, W.-L., Su, Y.: BioCLIP 2: Emergent Properties from Scaling Hierarchical Contrastive Learning. arXiv (2025). https://doi.org/10.48550/arXiv.2505.23883

[60] Stevens, S., Wu, J., Thompson, M.J., Campolongo, E.G., Song, C.H., Carlyn, D.E., Dong, L., Dahdul, W.M., Stewart, C., Berger-Wolf, T., Chao, W.-L., Su, Y.: BioCLIP: A Vision Foundation Model for the Tree of Life. arXiv (2024). https://doi.org/10.48550/arXiv.2311.18803

[61] Amini-Naieni, N., Han, T., Zisserman, A.: CountGD: Multi-Modal Open-World Counting. arXiv (2025). https://doi.org/10.48550/arXiv.2407.04619

[62] Sastry, S., Khanal, S., Dhakal, A., Ahmad, A., Jacobs, N.: TaxaBind: A Unified Embedding Space for Ecological Applications. arXiv (2024). https://doi.org/10.48550/arXiv.2411.00683

[63] Merriënboer, B., Dumoulin, V., Hamer, J., Harrell, L., Burns, A., Denton, T.: Perch 2.0: The Bittern Lesson for Bioacoustics (2025). https://arxiv.org/abs/2508.04665

[64] Robinson, D., Miron, M., Hagiwara, M., Weck, B., Keen, S., Alizadeh, M., Narula, G., Geist, M., Pietquin, O.: NatureLM-audio: an Audio-Language Foundation Model for Bioacoustics (2025). https://arxiv.org/abs/2411.07186

[65] Gadot, T., Istrate, S., Kim, H., Morris, D., Beery, S., Birch, T., Ahumada, J.: To crop or not to crop: Comparing whole-image and cropped classification on a large dataset of camera trap images. IET Computer Vision **18**(8), 1193–1208 (2024) https://doi.org/10.1049/cvi2.12318

[66] Gardiner, R.J., Rowlands, S., Simmons, B.I.: Towards scalable insect monitoring: Ultra-lightweight CNNs as on-device triggers for insect camera traps. Methods in Ecology and Evolution **n/a**(n/a) https://doi.org/10.1111/2041-210X.70098

[67] Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., Clune, J.: Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proceedings of the National Academy of Sciences **115**(25), 5716–5725 (2018) https://doi.org/10.1073/pnas.1719367115

[68] Kahl, S., Wood, C.M., Eibl, M., Klinck, H.: BirdNET: A deep learning solution for avian diversity monitoring. Ecological Informatics **61**, 101236 (2021) https:

//doi.org/10.1016/j.ecoinf.2021.101236

[69] Dussert, G., Miele, V., Van Reeth, C., Delestrade, A., Dray, S., Chamaillé-Jammes, S.: Zero-shot animal behaviour classification with vision-language foundation models. Methods in Ecology and Evolution **16**(7), 1460–1472 (2025) https://doi.org/10.1111/2041-210X.70059

[70] Feng, Z., Atzberger, C., Jaffer, S., Knezevic, J., Sormunen, S., Young, R., Lisaius, M.C., Immitzer, M., Jackson, T., Ball, J., Coomes, D.A., Madhavapeddy, A., Blake, A., Keshav, S.: TESSERA: Precomputed FAIR Global Pixel Embeddings for Earth Representation and Analysis. arXiv (2025). https://doi.org/10.48550/arXiv.2506.20380

[71] Singh, J., Ahirwal, S.K., Ramteke, K., Kantharajan, G., Sarma, K.: Remote Sensing Techniques for Monitoring Aquatic Ecosystems. In: Ganie, P.A., Posti, R., Pandey, P.K. (eds.) Information Technology in Fisheries and Aquaculture, pp. 71–107. Springer, Singapore (2025). https://doi.org/10.1007/978-981-97-8553-7_5

[72] Mutanga, O., Masenyama, A., Sibanda, M.: Spectral saturation in the remote sensing of high-density vegetation traits: A systematic review of progress, challenges, and prospects. ISPRS Journal of Photogrammetry and Remote Sensing **198**, 297–309 (2023) https://doi.org/10.1016/j.isprsjprs.2023.03.010

[73] Beery, S., Cole, E., Parker, J., Perona, P., Winner, K.: Species distribution modeling for machine learning practitioners: A review. In: Proceedings of the 4th ACM SIGCAS Conference on Computing and Sustainable Societies. COMPASS '21, pp. 329–348. Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3460112.3471966 . https://doi.org/10.1145/3460112.3471966

[74] Hu, Y., Si-Moussi, S., Thuiller, W.: Introduction to deep learning methods for multi-species predictions. Methods in Ecology and Evolution **16**(1), 228–246 (2025) https://doi.org/10.1111/2041-210X.14466

[75] Kellenberger, B., Winner, K., Jetz, W.: The Performance and Potential of Deep Learning for Predicting Species Distributions. bioRxiv (2024). https://doi.org/10.1101/2024.08.09.607358

[76] Haucke, T., Harrell, L., Shen, Y., Klein, L., Rolnick, D., Gillespie, L., Beery, S.: Deep Multi-modal Species Occupancy Modeling

[77] Trantas, A., Mensio, M., Stasinos, S., Gribincea, S., Khan, T., Podareanu, D., Veen, A.: BioAnalyst: A Foundation Model for Biodiversity. arXiv (2025). https://doi.org/10.48550/arXiv.2507.09080

[78] Davis, C.L., Bai, Y., Chen, D., Robinson, O., Ruiz-Gutierrez, V.,

Gomes, C.P., Fink, D.: Deep learning with citizen science data enables estimation of species diversity and composition at continental extents. Ecology **104**(12), 4175 (2023) https://doi.org/10.1002/ecy.4175 https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1002/ecy.4175

[79] Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., Casas, D.d.L., Hendricks, L.A., Welbl, J., Clark, A., Hennigan, T., Noland, E., Millican, K., Driessche, G., Damoc, B., Guy, A., Osindero, S., Simonyan, K., Elsen, E., Rae, J.W., Vinyals, O., Sifre, L.: Training Compute-Optimal Large Language Models. arXiv (2022). https://doi.org/10.48550/arXiv.2203.15556

[80] Kitzes, J., Chronister, L., Czarnecki, C., Fiss, C., Freeland-Haynes, L., Goodman, B.D., Lapp, S., Lyon, R.P., Nossan, H., Rhinehart, T.A., Ruiz Guzman, S., Syunkova, A., Viotti, L.: Integrating ai models into ecological research workflows: The case of terrestrial bioacoustics. Methods in Ecology and Evolution **n/a**(n/a) https://doi.org/10.1111/2041-210X.70133

[81] Chapman, M., Goldstein, B.R., Schell, C.J., Brashares, J.S., Carter, N.H., Ellis-Soto, D., Faxon, H.O., Goldstein, J.E., Halpern, B.S., Longdon, J., Norman, K.E.A., O'Rourke, D., Scoville, C., Xu, L., Boettiger, C.: Biodiversity monitoring for a just planetary future. Science **383**(6678), 34–36 (2024) https://doi.org/10.1126/science.adh8874

[82] Oord, A., Vinyals, O., Kavukcuoglu, K.: Neural Discrete Representation Learning. arXiv (2018). https://doi.org/10.48550/arXiv.1711.00937

[83] Klemmer, K., Rolf, E., Robinson, C., Mackey, L., Rußwurm, M.: SatCLIP: Global, General-Purpose Location Embeddings with Satellite Imagery. arXiv (2024). https://doi.org/10.48550/arXiv.2311.17179

[84] Cepeda, V.V., Nayak, G.K., Shah, M.: GeoCLIP: Clip-Inspired Alignment between Locations and Images for Effective Worldwide Geo-localization. arXiv (2023). https://doi.org/10.48550/arXiv.2309.16020

[85] Clay Foundation: Clay — AI for Earth. Website. Accessed 2025-11-28. https://madewithclay.org/

[86] Connecting geology to ecology. Nature Geoscience **17**(3), 173–173 (2024) https://doi.org/10.1038/s41561-024-01411-7

[87] At the interface between hydrology and ecology. Nature Water **2**(3), 207–207 (2024) https://doi.org/10.1038/s44221-024-00225-6

[88] Geng, X., Liu, H., Lee, L., Schuurmans, D., Levine, S., Abbeel, P.: Multimodal masked autoencoders learn transferable representations. arXiv preprint arXiv:2205.14204 (2022)

[89] Mizrahi, D., Bachmann, R., Kar, O., Yeo, T., Gao, M., Dehghan, A., Zamir, A.: 4m: Massively multimodal masked modeling. Advances in Neural Information Processing Systems **36**, 58363–58408 (2023)

[90] Boggust, A., Carter, B., Satyanarayan, A.: Embedding comparator: Visualizing differences in global structure and local neighborhoods via small multiples. In: Proceedings of the 27th International Conference on Intelligent User Interfaces, pp. 746–766 (2022)

[91] Lundberg, S., Lee, S.-I.: A Unified Approach to Interpreting Model Predictions (2017). https://arxiv.org/abs/1705.07874

[92] Koh, P.W., Nguyen, T., Tang, Y.S., Mussmann, S., Pierson, E., Kim, B., Liang, P.: Concept bottleneck models. In: International Conference on Machine Learning, pp. 5338–5348 (2020). PMLR

[93] Cowans, A., Lambin, X., Hare, D., Sutherland, C.: Improving the integration of artificial intelligence into existing ecological inference workflows. Methods in Ecology and Evolution **n/a**(n/a) https://doi.org/10.1111/2041-210X.14485

[94] Sastry, S., Khanal, S., Dhakal, A., Lin, J., Cher, D., Jarosz, P., Jacobs, N.: ProM3E: Probabilistic Masked MultiModal Embedding Model for Ecology. arXiv (2025). https://doi.org/10.48550/arXiv.2511.02946

[95] Poelen, J.H., Simons, J.D., Mungall, C.J.: Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets. Ecological Informatics **24**, 148–159 (2014) https://doi.org/10.1016/j.ecoinf.2014.08.005

[96] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., Galstyan, A.: A Survey on Bias and Fairness in Machine Learning. ACM Comput. Surv. **54**(6), 115–111535 (2021) https://doi.org/10.1145/3457607

[97] Lo Piano, S.: Ethical principles in machine learning and artificial intelligence: Cases from the field and possible ways forward. Humanities and Social Sciences Communications **7**(1), 9 (2020) https://doi.org/10.1057/s41599-020-0501-9

[98] Ellis-Soto, D., Chapman, M., Locke, D.H.: Historical redlining is associated with increasing geographical disparities in bird biodiversity sampling in the United States. Nature Human Behaviour **7**(11), 1869–1877 (2023) https://doi.org/10.1038/s41562-023-01688-5

[99] Zizka, A., Rydén, O., Edler, D., Klein, J., Perrigo, A., Silvestro, D., Jagers, S.C., Lindberg, S.I., Antonelli, A.: Bio-Dem, a tool to explore the relationship between biodiversity data availability and socio-political conditions in time and space. Journal of Biogeography **48**(11), 2715–2726 (2021) https://doi.org/10.1111/jbi.14256

[100] Gaston, K.J., May, R.M.: Taxonomy of taxonomists. Nature **356**(6367), 281–282 (1992) https://doi.org/10.1038/356281a0

[101] Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F.A., Bengio, Y., Courville, A.: On the Spectral Bias of Neural Networks. arXiv (2019). https://doi.org/10.48550/arXiv.1806.08734

[102] Arvanitidis, G., Hansen, L.K., Hauberg, S.: Latent Space Oddity: On the Curvature of Deep Generative Models. arXiv (2021). https://doi.org/10.48550/arXiv.1710.11379

[103] Gawlikowski, J., Tassi, C.R.N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R., Zhu, X.X.: A survey of uncertainty in deep neural networks. Artificial Intelligence Review **56**(1), 1513–1589 (2023) https://doi.org/10.1007/s10462-023-10562-9

[104] Lai, Y., Shi, Y., Han, Y., Shao, Y., Qi, M., Li, B.: Exploring Uncertainty in Deep Learning for Construction of Prediction Intervals. arXiv (2021). https://doi.org/10.48550/arXiv.2104.12953

[105] Sbailò, L., Ghiringhelli, L.M.: Uncertainty Quantification in Deep Neural Networks through Statistical Inference on Latent Space. arXiv (2023). https://doi.org/10.48550/arXiv.2305.10840

[106] Sagawa, S., Raghunathan, A., Koh, P.W., Liang, P.: An Investigation of Why Overparameterization Exacerbates Spurious Correlations

[107] Zhang, T., Zhao, C., Chen, G., Jiang, Y., Chen, F.: Feature Contamination: Neural Networks Learn Uncorrelated Features and Fail to Generalize. arXiv (2025). https://doi.org/10.48550/arXiv.2406.03345

[108] Azulay, A., Weiss, Y.: Why do deep convolutional networks generalize so poorly to small image transformations?

[109] Thulasidasan, S., Chennupati, G., Bilmes, J., Bhattacharya, T., Michalak, S.: On Mixup Training: Improved Calibration and Predictive Uncertainty for Deep Neural Networks. arXiv (2020). https://doi.org/10.48550/arXiv.1905.11001

[110] Rahimi, A., Recht, B.: Random features for large-scale kernel machines. In: Proceedings of the 21st International Conference on Neural Information Processing Systems. NIPS'07, pp. 1177–1184. Curran Associates Inc., Red Hook, NY, USA (2007)

[111] Peng, W., Varanka, T., Mostafa, A., Shi, H., Zhao, G.: Hyperbolic deep neural networks: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(12), 10023–10044 (2022) https://doi.org/10.1109/TPAMI.2021.3136921

[112] Wesselkamp, M., Moser, N., Kalweit, M., Boedecker, J., Dormann, C.F.: Process-Informed Neural Networks: A Hybrid Modelling Approach to Improve Predictive Performance and Inference of Neural Networks in Ecology and Beyond. Ecology Letters **27**(11), 70012 (2024) https://doi.org/10.1111/ele.70012

[113] Delgado, F., Yang, S., Madaio, M., Yang, Q.: The Participatory Turn in AI Design: Theoretical Foundations and the Current State of Practice. In: Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization. EAAMO '23, pp. 1–23. Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3617694.3623261

[114] Nations, U.: Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from Their Utilization to the Convention on Biological Diversity. International treaty. Adopted 29 October 2010, entered into force 12 October 2014 (2010). https://www.cbd.int/abs/text

[115] Carroll, S.R., Garba, I., Figueroa-Rodríguez, O.L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J.D., Anderson, J., Hudson, M.: The CARE Principles for Indigenous Data Governance. Data Science Journal **19**, 43 (2020) https://doi.org/10.5334/dsj-2020-043

[116] Organisation for Economic Co-operation and Development (OECD): OECD AI Principles: Recommendation on Artificial Intelligence. OECD Council Recommendation, adopted 22 May 2019. Updated in May 2024; see 2024 revision at OECD.AI website (2019). https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[117] Siddik, M.A.B., Shehabi, A., Marston, L.: The environmental footprint of data centers in the United States. Environmental Research Letters **16**(6), 064017 (2021) https://doi.org/10.1088/1748-9326/abfba1

[118] Lei, N., Lu, J., Shehabi, A., Masanet, E.: The water use of data center workloads: A review and assessment of key determinants. Resources, Conservation and Recycling **219**, 108310 (2025) https://doi.org/10.1016/j.resconrec.2025.108310

[119] Programme, U.N.E.: Artificial Intelligence (AI) End-to-End: The Environmental Impact of the Full AI Lifecycle Needs to Be Comprehensively Assessed - Issue Note, (2024). https://doi.org/10.59117/20.500.11822/46288

[120] Measuring the environmental impacts of artificial intelligence compute and applications: The ai footprint. OECD Digital Economy Papers 341, OECD (November 2022). https://doi.org/10.1787/7babf571-en

[121] Wu, C.-J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., Chang, G., Behram, F.A., Huang, J., Bai, C., Gschwind, M., Gupta, A., Ott,

M., Melnikov, A., Candido, S., Brooks, D., Chauhan, G., Lee, B., Lee, H.-H.S., Akyildiz, B., Balandat, M., Spisak, J., Jain, R., Rabbat, M., Hazelwood, K.: Sustainable AI: Environmental Implications, Challenges and Opportunities. arXiv (2022). https://doi.org/10.48550/arXiv.2111.00364