Trends in
Ecology & Evolution

CellPress
OPEN ACCESS

Feature Review

The potential for AI to revolutionize conservation: a horizon scan

Sam A. Reynolds [1,*], Sara Beery [2], Neil Burgess [3,4], Mark Burgman [5], Stuart H.M. Butchart [6,7], Steven J. Cooke [8], David Coomes [9], Finn Danielsen [10], Enrico Di Minin [11,12], América Paz Durán [13], Francis Gassert [14], Amy Hinsley [15], Sadiq Jaffer [16], Julia P.G. Jones [17], Binbin V. Li [18,19], Oisin Mac Aodha [20], Anil Madhavapeddy [21], Stephanie A.L. O'Donnell [22], William M. Oxbury [23], Lloyd Peck [24], Nathalie Pettorelli [25], Jon Paul Rodríguez [26,27,28], Emily Shuckburgh [29], Bernardo Strassburg [30,31], Hiromi Yamashita [32], Zhongqi Miao [33], and William J. Sutherland [1]

Artificial Intelligence (AI) is an emerging tool that could be leveraged to identify the effective conservation solutions demanded by the urgent biodiversity crisis. We present the results of our horizon scan of AI applications likely to significantly benefit biological conservation. An international panel of conservation scientists and AI experts identified 21 key ideas. These included species recognition to uncover 'dark diversity', multimodal models to improve biodiversity loss predictions, monitoring wildlife trade, and addressing human–wildlife conflict. We consider the potential negative impacts of AI adoption, such as AI colonialism and loss of essential conservation skills, and suggest how the conservation field might adapt to harness the benefits of AI while mitigating its risks.

Highlights

Following a horizon scan methodology with a panel of 27 conservation scientists and artificial intelligence (AI) experts, 21 ideas likely to significantly impact on the success of global biodiversity conservation were identified from a long list of 104.

Our 21 issues include novel interpretation of image and audio data, digital twins for ecosystems, improving species distribution models, and AI-powered conservation advisors.

We believe that adoption of AI in conservation will lead to beneficial outcomes for conservation effectiveness and improve our understanding of the natural world. However, it is not a panacea, and will not wholly replace established conservation techniques, education, and on-the-ground research. Moreover, careful and creative measures must be taken to ensure equitable access, development, and deployment.

The intersection between conservation and AI

Biological conservation science is concerned with understanding the causes and consequences of biodiversity loss, and developing and testing solutions to halt and reverse that loss. Given the urgency of the biodiversity crisis [1], more effective solutions are rapidly required. Biodiversity is complex and multifaceted [2], meaning that understanding status and trends often requires processing large quantities of data, which are time-consuming to collect and analyze. In addition, conservation operates in a coupled human and natural system in which there are complex feedbacks between social and ecological components [3] that can be difficult to fully understand and model accurately [4], obstructing decision-making.

One hope of lifting these barriers is to look to methods of **artificial intelligence** (**AI**, see Glossary) which have been transformative in many domains such as healthcare [5] and international security [6]. The key to this success has been the use of **machine learning (ML)** which, given sufficient data, can demonstrate high predictive performance on tasks where good mathematical **models** are lacking [7]. Even in scientific areas that do rely on well-understood mathematical models, such as weather forecasting, ML can often increase modeling performance at reduced computational cost [8].

Decades of systematic and opportunistic data collection by ecologists and conservation scientists (whether using citizen/community science, expert, or remote sensing approaches) have provided a wealth of data, much of which is underutilized [9]. Moreover, the increasing availability of novel sensors (e.g., for remote sensing, biotelemetry, and biologging) and hardware (including smart phones) for conservation applications has led to the rapid accumulation of new data,

[1]Conservation Science Group, Department of Zoology, Cambridge University, Cambridge CB2 3QZ, UK
[2]Faculty of Artificial Intelligence and Decision Making, Massachusetts Institute of Technology, Cambridge, MA, USA
[3]United Nations Environment Programme World Conservation Monitoring Centre (UNEP-WCMC), Cambridge, CB3 0DL, UK
[4]Center for Macroecology, Evolution, and Climate, University of Copenhagen, DK-2100, Copenhagen, Denmark
[5]School of Life Sciences, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA

and analytical tools to keep pace with these advances are required. Combining these data with the power of AI represents a potentially revolutionizing force to increase the effectiveness of conservation approaches, and thus accelerate efforts to the levels necessary to meet the targets in the recently agreed Kunming–Montreal Global Biodiversity Framework [10].

Public interest in AI has exploded in recent years, and platforms such as ChatGPT and Midjourney have captured our imaginations through their generation of human-like text and image content. However, AI is not a novel topic in scientific research and has been adopted by researchers in some form for decades [11]. Nonetheless, the rapid investment in AI within academic and industrial research has irrevocably altered the scientific landscape. The functions of AI in scientific understanding can be outlined as follows [12].

(i) Providing information through advanced simulation and data representation, that cannot be obtained through experimentation, to reveal properties of a physical system that are otherwise difficult or even impossible to probe. In conservation, AI systems are making a rapidly growing contribution to providing better information and, in some cases, more effective actions such as 'Skylight.global' – which identifies illegal and unreported fishing in near real time from satellite data[i].
(ii) Providing information that expands the scope of human imagination or creativity, such as identifying surprises in data, models, and the literature. Recent projects allow researchers to automatically gain ecological and animal behavioral insights from audio[ii] and image data[iii]. AI technology is also helping to model species distributions from partial observations [13].
(iii) Providing insights to human experts by translating observations into new knowledge. Platforms such as CAPTAIN [14] take the first steps to allow policymakers to identify optimal areas for protection to preserve species and habitats.

These advances not only improve current practices in conservation (e.g., by improving methods for monitoring trends in species, habitats, and threats) but also open completely novel areas for R&D. However, new technology also brings challenges. There is growing awareness that the deployment of wildlife monitoring technologies might have unintended consequences and could potentially undermine conservation efforts [15], and wider use of drones, audio recording, camera traps, and electronic tagging and tracking tools brings risks of misuse by bad actors. There are also justifiable concerns about the energy-use requirements of AI [16] and the way in which use of AI could aggravate existing inequalities and biases [17]. Furthermore, current **large language models (LLMs)** can 'hallucinate' – generating inaccurate or nonsensical outputs which are presented as fact.

The aim of this paper is to identify the key areas where AI can help to improve the effectiveness of conservation, as revealed through a horizon scan methodology in consultation with both conservation and AI experts, as well as to reflect on the challenges remaining to ensure that this powerful emerging technology is a force for good in improving the effectiveness of conservation.

Recent developments in AI and machine learning

The potential for AI methods to process and analyze large datasets, identify subtle patterns, and generate novel insights offers promising opportunities for conservation and ecology [18]. An overview of AI and ML is available in Box 1. Much of the media excitement in recent years has been the promise of LLMs and generative models. These have been shown to be generalizable even with no additional training data for a given task (zero shot learning) [19] and can also reduce the cost of training new models across a wide range of domains through fine-tuning or **transfer learning**. LLMs in particular have seen rapid adoption through chat interfaces such as ChatGPT[iv] and

[6]BirdLife International, Cambridge CB2 3QZ, UK
[7]Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK
[8]Department of Biology and Institute of Environmental and Interdisciplinary Science, Carleton University, Ottawa, ON K1S 5B6, Canada
[9]Conservation Research Institute and Department of Plant Sciences, Cambridge CB2 3QZ, UK
[10]Nordic Foundation for Development and Ecology (NORDECO), Copenhagen DK-1159, Denmark
[11]Department of Geosciences and Geography, and Helsinki Institute of Sustainability Science (HELSUS), FI-00014, University of Helsinki, Finland
[12]School of Life Sciences, University of KwaZulu-Natal, Durban 4041, South Africa
[13]Instituto de Ecología y Biodiversidad (IEB-Chile), Santiago, Chile
[14]Vizzuality, 28010 Madrid, Spain
[15]Department of Biology, University of Oxford, Oxford, OX1 3SZ, UK
[16]Department of Computer Science and Technology, University of Cambridge, Cambridge CB3 0FD, UK
[17]School of Environmental and Natural Sciences, Bangor University, Bangor, UK
[18]Environmental Research Centre, Duke Kunshan University, Kunshan, Jiangsu 215316, China
[19]Nicholas School of the Environment, Duke University, Durham, NC 27708, USA
[20]School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, UK
[21]Department of Computer Science and Technology, University of Cambridge, Cambridge CB3 0FD, UK
[22]Fauna & Flora International, Cambridge, CB2 3QZ, UK
[23]School of Mathematical Sciences, Lancaster University, Lancaster LA1 4YQ, UK
[24]British Antarctic Survey, Natural Environment Research Council (NERC), Cambridge CB3 0ET, UK
[25]Institute of Zoology, Zoological Society of London, London NW1 4RY, UK
[26]International Union for Conservation of Nature (IUCN) Species Survival Commission, Calle La Joya, Edificio Unidad Técnica del Este, Chacao, Caracas 1060, Venezuela
[27]Provita, Calle La Joya, Edificio Unidad Técnica del Este, Chacao, Caracas 1060, Venezuela
[28]Centro de Ecología, Instituto Venezolano de Investigaciones Científicas (IVIC), Miranda 1204, Venezuela

Box 1. An overview of AI and machine learning

Although AI covers the broad field of intelligent software systems, today the term is most often used to refer to systems that implement ML algorithms. An ML algorithm is one that incorporates a data structure (called a model) obtained as a compression of the 'training' data, usually with the aim of approximating probabilities of interest.

Neural networks are the best-known example of such models, and the field of deep learning has shown outstanding success in applying large-scale neural networks to many previously intractable tasks in image and audio processing, natural language processing (NLP), and other areas.

In its simplest form, the training of ML algorithms is supervised (i.e., **supervised learning**). This means that the training data comprise (usually human-labeled) pairs such as image pixels and the type of object in the image or audio spectrogram (e.g., the bird species). Example applications include predicting taxonomic identity from a photo or an audio recording of an animal, or predicting land-cover class from remote sensing tiles.

The process of generating trustworthy labeled data for supervised learning is, and remains, labor-intensive. However, there are powerful variants of the supervised approach: the 'training signal' on which model training depends need not always come from human-generated labels. It may come from structures observed within the data itself (e.g., next-term prediction in sequential data such as text) or between datasets (e.g., machine translation models learn from comparison of many parallel texts). This is commonly referred to as self-supervised learning. It may also come from transfer learning, where a model trained on one task (e.g., recognizing road traffic objects) is fine-tuned with limited data to solve another task (e.g., alerting to the presence of a pedestrian).

Reinforcement learning solves decision problems (such as in game-playing) and takes its training signal from direct exploratory interaction with its environment. It learns a decision policy as a ML model, often a neural network, and can converge to superhuman skill levels (e.g., in chess, Go, and some video games).

Sequential models, in particular, have reached a high degree of sophistication in LLMs (such as ChatGPT) which use a combination of large-scale neural networks and transformer architectures, and can be further fine-tuned with reinforcement learning.

Finally, the idea of transfer learning can potentially be taken a long way with the proposal of foundation models (FMs) [21]. The idea here is to invest one-off (for an application domain of interest) in very large-scale and ideally self-supervised, models. The resulting FM can then be retuned at relatively low cost for specific predictive tasks in the domain. However, it is important to note that the current success of foundation models is derived from their access to large high-quality datasets for training, which are often manually annotated in an expensive and time-consuming process and are therefore not based purely on unsupervised learning.

[29]Department of Computer Science and Technology, University of Cambridge, Cambridge CB3 0FD, UK
[30]Re.Green, Rio de Janeiro 22470-060, Brazil
[31]Rio Conservation and Sustainability Science Centre, Department of Geography and the Environment, Pontifícia Universidade Católica, Rio de Janeiro 22451-900, Brazil
[32]Ritsumeikan Asia Pacific University, Jyumonji-baru, Beppu City, Oita, 874-8577, Japan
[33]AI for Good Lab, Microsoft, Redmond, WA, USA

*Correspondence:
sar87@cam.ac.uk (S.A. Reynolds).

assistive **copilots** such as Github Copilot[v]. Further, their reasoning capabilities are starting to be used to drive AI agents [20] that can sense their environment, make decisions, and take action.

These **foundation models** [21] take unlabeled input and exploit structure from within the data to learn. A common and general approach is to hide part of each input example and train the model to predict the missing parts. This process is part of the training of the model, not a model evaluation method such as cross-validation in conventional statistics. The next-word prediction task used to train LLMs such as ChatGPT is one example of **self-supervised learning**.

Freed of the need for labels, models can be trained with large amounts of data and through their self-supervised task can learn generic representations. For example, examination of large visual models trained through self-supervision reveals specialized structures for detecting different classes of objects or animals [22]. Once trained, these models can be fine-tuned to a specific task. Consider a visual model that can detect various tree species and leaf shapes. It will require fewer new labeled examples of a specific endangered tree species of interest to be able to classify new examples compared to training a new visual model from scratch [23].

Current successful deployments of LLMs often take the form of copilots, where AI systems improve user productivity by assisting decision-makers, surfacing information, suggesting changes,

Box 2. Conducting the horizon scan

In July 2024 a panel of 27 experts accepted an invitation to participate in the horizon scan, comprising 18 conservation scientists and nine computer scientists and AI specialists. Some experts sit across both areas of expertise. These experts are the authors of the present paper.

Participants were required to submit 2–8 ideas detailing how AI could be leveraged in conservation across a range of applications, including image and audio recognition, environmental DNA (eDNA) analysis, modeling, and data interpretation and integration. Participants canvassed their networks and colleagues to broaden the perspectives captured in the horizon scan.

Figure 1 provides an overview of the process used to identify and score ideas. For this horizon scan, 104 ideas were submitted by participants for consideration (Table S1). Participants then confidentially and independently reviewed each idea submitted, and ranked them by assigning a score of between 1 (least significant) and 1000 (most significant) to each idea. The large range of scores was simply to make it easy for scorers to rank items, and the scores themselves were not used. Participants assigned a single score to each idea by subjectively combining two criteria: (i) its potential to have an important impact on the field of conservation science and/or practice, and (ii) its likelihood of coming to fruition. Notes were added by participants to provide further information as to whether ideas should be retained for the second round. To counter the effect of voter fatigue [91], participants were randomly assigned to one of three lists of ideas presented in different orders. Ideas were clustered into broad topics (audio, images, reviewing literature, modeling, data, generating text, negative consequences, citizen science, eDNA, remote sensing, robotics, society, and other) by the lead author. Participant scores were then converted into ranks (1–104) where the highest score denotes the highest rank. The median rank across all participants for each idea was calculated, and the 30 top-ranked ideas were brought forward for discussion in the workshop. All rank calculations were conducted in R statistical software [92], and the code is available via GitHub[xi]. Two ideas in the top 30 were deemed by the organizing committee to be similar to two other ideas suggested, and were therefore presented together in the shortlist. This led to 28 distinct ideas being shortlisted.

A meeting was held with 13 participants attending in person and 12 attending online. Two participants were unable to attend the workshop but participated in the initial long-list scoring and preparation of the manuscript.

To account for participants attending from different time zones, the meeting was held in three sessions. Session 1 with 25 participants in attendance discussed audio, reviewed the literature, and generated text, whereas session 2 with 25 participants discussed images and remote sensing, including a focused discussion on the potential negative consequences and pitfalls of AI. The final session with 24 participants focused on modeling and data followed by a discussion of how the field of conservation may need to adapt to take full advantage of AI.

Each idea was discussed for 10 minutes. After each discussion, participants scored each idea on a scale of 1–1000 (low to high). High scores denote the most significant ideas, low scores the least significant. The scores of each participant were converted into ranks at the end of the workshop (the highest-scoring idea being assigned to rank 1), and the 20 ideas with the highest median ranks (therefore the most significant ideas) were identified.

Not all online participants could attend the full workshop. To incorporate their partially scored lists with the lists of participants who scored all ideas, we computed a 'division factor'. The division factor is the ratio of total number of ideas to the number of ideas scored, plus one (Equation I):

$$DF = \frac{TI}{SI + 1} \quad [I]$$

where DF is the division factor, TI is the total number of ideas, and SI is the number of ideas scored by the expert. The partially scored lists were then ranked (e.g., if only 10 ideas were scored, they were assigned ranks 1 through 10, where rank 1 is the most significant idea). To calculate the appropriate final rank for integration with the fully ranked score sheets, we used the following equation (Equation II):

$$FR = PR \times DF \quad [II]$$

where FR is the final rank, PR is the rank given when all partially scored ideas are ranked, and DF is the division factor.

This creates a buffer in the ranking scale that allows for the possibility that ideas unranked by a participant could potentially be ranked higher than their partially ranked ideas if they were to be evaluated. This adjustment expands the range of the partial rankings to align with a full ranking scale, leaving space for potentially unranked items between them. Ideas that were not scored by a participant did not have a rank imputed and were left blank. Therefore, participants only contributed to the ranking of items they were able to score.

Two ideas were related to identifying the expansion frontiers of human disturbance, and it was proposed that these ideas should be amalgamated into a single idea. In addition, there was a tied rank at 20, and it was decided that both ideas should be included, therefore 21 ideas are included in the final list.

Glossary

Artificial intelligence (AI): the broad field of creating systems that can execute tasks traditionally associated with human cognitive abilities.
Copilot: an AI system that improves user productivity by assisting them, surfacing information, suggesting changes, and receiving corrections.
Digital twin: a model that is coupled to and learns from data generated by a physical system.
Federated learning: a field of ML in which participants can keep their data private while collaboratively training a model with other participants.
Foundation model: also termed 'foundational model', a type of ML model that is usually pretrained via self-supervision on extremely large datasets to perform a range of tasks. Such models include all data types, including text, images, audio, and data from satellite sensors.
Human in the loop: a system in which human experts provide evaluation and feedback during the training process to improve accuracy.
Large language models (LLMs): foundation models that are pretrained specifically on textual material.
Machine learning (ML): a subfield of AI in which a computer is algorithmically trained to perform a task without following explicit instructions.
Model: an abstraction of a system, process, or phenomenon that is designed to capture its essential features and relationships. Models can take various forms such as mathematical equations, statistical distributions, algorithms, or neural networks, and are used to make predictions, explain patterns, or simulate behaviors based on input data or parameters.
Multimodal models: AI models that are able to work with inputs of different modalities such as images, text, video, or audio.
Neural network: a computational deep-learning model, inspired by the network of neurons in the human brain, that is designed to recognize patterns and make decisions based on data.
Physics-inspired AI: an AI approach that incorporates principles from physics to improve model performance, efficiency, and interpretability by leveraging known physical laws and constraints.

CellPress
OPEN ACCESS

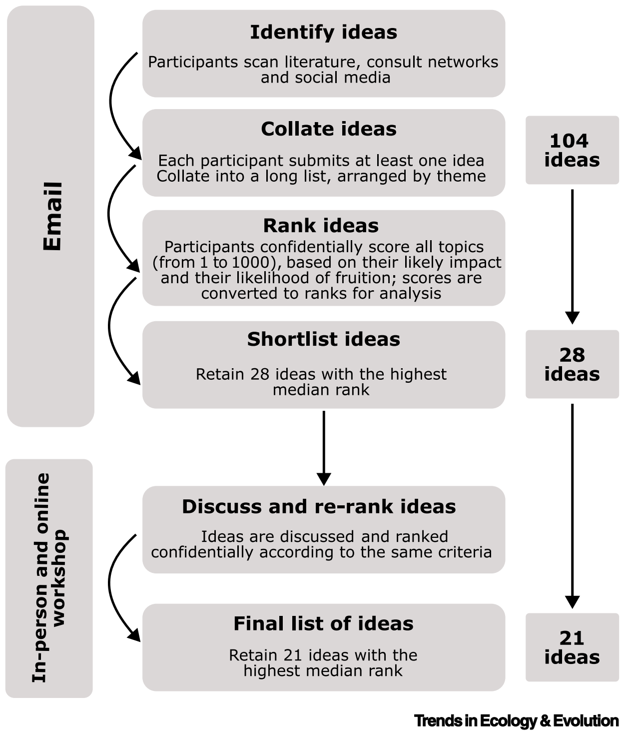


Figure 1. The process used and the resulting filtering of ideas to produce a ranked list of how AI could revolutionize conservation effectiveness. Ideas were generated and ranked by a panel of 27 experts (18 conservation scientists, and nine computer scientists and AI specialists).

Reinforcement learning: a type of ML in which a system makes decisions in a potentially changing environment and may receive only intermittent signals as to the effectiveness of its decisions in reaching its intended goal.
Retrieval augmented generation (RAG): a technique where an AI model retrieves relevant information from a large dataset to enhance the generation of more accurate and contextually appropriate responses.
Self-supervised learning: learning that takes place without explicit labels by exploiting some property or structure of the input data.
Species distribution models (SDMs): models that use environmental variables and species occurrence data to predict the distribution of a species across geographic space and time.
Supervised learning: a method of learning in which a model is trained with both the data inputs and the corresponding desired outputs/labels.
Transfer learning: a model developed for one task is reused as the starting point for a model on a second related task. Leveraging knowledge already gained from solving one problem to a new but similar problem.

and receiving corrections. This approach leverages the strengths of AI by synthesizing large amounts of information while mitigating risks associated with potential errors or biases. However, challenges remain, including the need to address biases in training data and the importance of human oversight in key applications.

Scanning the horizon for future applications of AI to improve conservation

To identify the areas where AI has considerable potential to revolutionize conservation, we applied a modified Delphi technique to select and rank suggestions from experts [24,25] (Box 2). This technique maintains the transparency, repeatability, and inclusivity of the process [26]. Participants in the horizon scan were selected based on consultation of the professional networks of the organizing committee and internet searches, and attempted to produce a mix of subject area expertise and geographical representations. The organizing committee produced an initial document summarizing recent advances in AI, the needs of conservation, and the main areas where AI is relevant to conservation. This provided a grounding in AI for conservation experts who may have had little exposure to AI advances, and gave AI experts an understanding of the interests of conservationists to facilitate discussion at the workshop. Authors were asked to suggest 2–8 potential developments, each with a short explanation. In some cases the authors gathered ideas from within their organizations, thus further expanding the sample of experts consulted and the geographical spread of idea generation.

We recognize that there may be limitations to ideas generated in the horizon scan process and that a different group of experts may identify a different set of ideas. Our methodology of inviting participants from a range of subject backgrounds and global regions, and asking them to canvass their network of colleagues and collaborators, aims to identify as broad a set of issues as possible and limit bias towards a particular discipline or study area. We note also that Sutherland *et al.* [27] reported no significant correlation between the areas of research expertise of the participants and the top issues selected in a horizon scan conducted in 2009. Therefore, horizon scans do not necessarily represent ideas that reflect the expertise of participants.

Potential applications of AI in conservation

In the following we present the 21 ideas where AI has considerable potential to revolutionize biological conservation (Figure 2). The ideas are presented in thematic groups rather than in rank order. The full list of 104 ideas is included in Table S1 in the supplemental information online. Ideas in the full list vary in their detail, relevance, and completeness, but may provide a useful insight for additional applications of AI within conservation. They include ideas relating to the use of robotics and citizen science, and understanding the connections between people and nature.

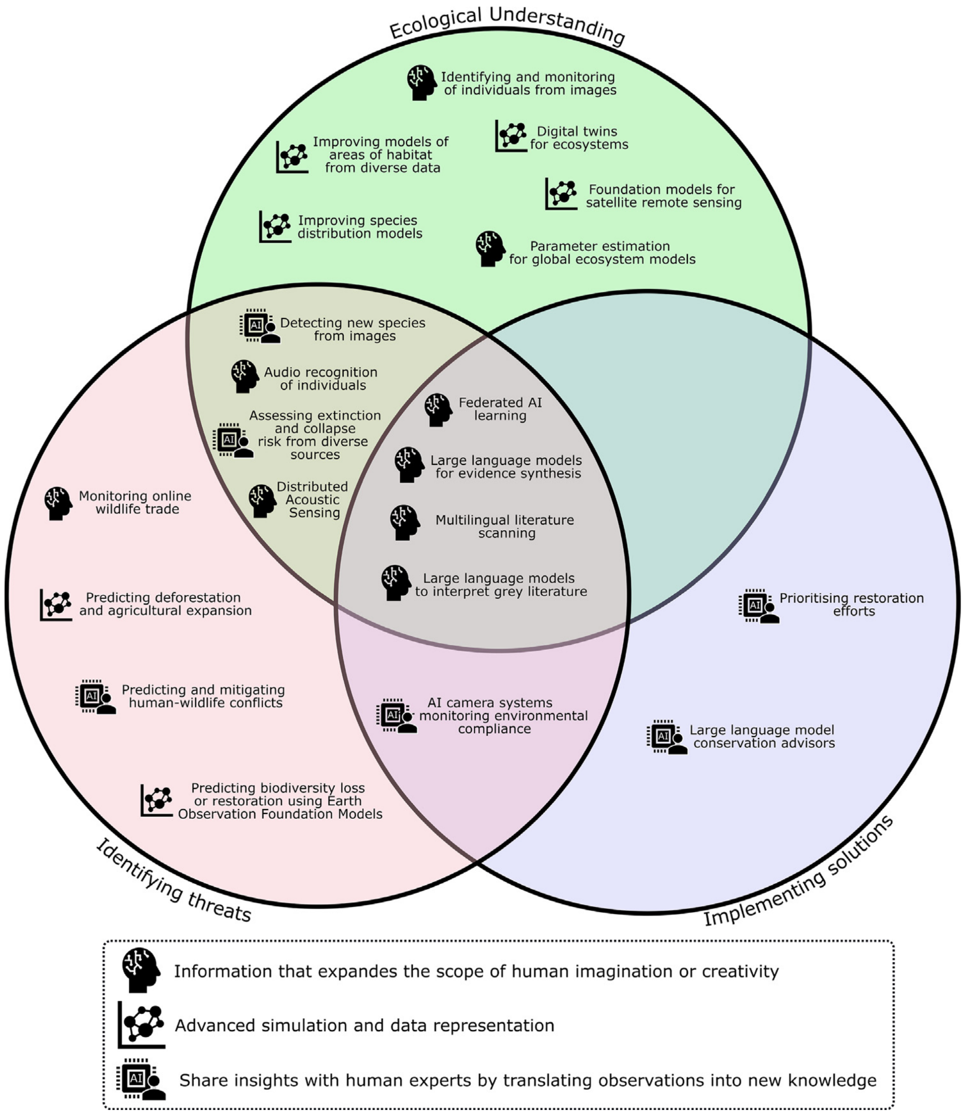
Interpretation of images collected by ground-based sensors

Automated species recognition from images collected by devices such as camera traps and mobile phones is well advanced and is available through applications such as iNaturalist [28], Pl@ntnet [29], ObsIdentify[vi], Google Lens[vii], and Merlin[viii]. However, improvements in AI will enable substantial acceleration of its implementation and application, including real-time identification systems that can send alerts, for example, when specific species, large numbers of individuals, or threats are detected [30]. Scaling up image acquisition could be achieved through citizen science, community-based monitoring, harvesting images from social media, or repurposing existing datasets (e.g., Google Street View). This supports applications, such as mapping species distributions and range extensions, monitoring the establishment or spread of invasive alien species, and detecting and identifying illegal imports of traded species at customs [31].

Identifying and monitoring of individuals from images. Scaling up automated recognition of individual animals from images could enable more widespread and accurate assessment of population sizes, for example by using mark–recapture [32], leading to more accurate assessment of status and trends, as well as other opportunities such as quantifying home ranges and identifying movement patterns.

Detecting new species from images. AI workflows have recently been developed for the detection and confirmation of unknown species identity from images, also known as novel category discovery [33]. This technique could accelerate the documentation of 'dark diversity', especially in combination with DNA analysis. These approaches could be used both on images captured from the field and from digitization of museum collections [34].

AI camera systems for monitoring environmental compliance. AI systems for monitoring compliance could be as broad as measuring biodiversity gains promised by developers, analyzing water quality measurements from treatment plants, and detecting illegal deforestation from satellite imagery. An obvious candidate area would be monitoring compliance with measures to mitigate commercial fishery bycatch. Bycatch of seabird species are resulting in mortality rates that are driving some species towards extinction [35]. Regional Fisheries Management Organizations already require implementation of measures to counter this [36]. However, detecting compliance relies on boat-based observers, which is resource-intensive and dangerous. AI-enabled onboard camera systems can facilitate monitoring seabird presence, interactions with fishing gear (e.g., cable strikes), bycatch (i.e., ensnared birds), and the use of mitigation measures, thus enabling safer, cheaper, and more reliable monitoring of compliance.



Trends in Ecology & Evolution

Figure 2. The 21 artificial intelligence (AI) applications identified in the horizon scan placed into a framework adapted from Krenn *et al.* [12]. Each circle represents a different area of study within conservation science – ecological understanding, identifying threats, and implementing solutions – whereas each icon represents a different way in which AI can contribute to scientific understanding.

Audio

AI is already being used extensively to identify species from audio recordings, principally for birds and bats. For example, using ML, BirdNET can currently identify ~3000 of the most common bird species worldwide [37]. However, there is considerable potential for scaling up to cover the remaining 75% of bird taxa, as well as other vocalizing animals including invertebrates. In combination with distribution of autonomous recording units or other devices running identification algorithms locally, this could transform our ability to monitor biodiversity [38]. AI is already being used to assess soundscapes to generate insights into the state of ecosystems such as coral reefs [39] and tropical forests [40].

Audio recognition of individuals. In some cases it is already possible to distinguish individuals within a species from their sounds [41]. There is likely to be considerable opportunity to extend this given that (i) voice recognition in humans is well understood (using 'voiceprint' vectors) [42], and (ii) many species are observed to recognize individuals from their vocalizations [41]. Potential applications include estimating population sizes, tracking individual movements, and quantifying dispersal.

Distributed acoustic sensing. Novel AI-enabled technologies for detecting sounds in the marine environment could transform monitoring of marine fauna. For example, distributed acoustic sensing (DAS) can detect sounds in real time using the dense network of fiber-optic telecommunication cables that cover both deep ocean and coastal areas worldwide. It could be used for audio recognition of marine fauna such as cetaceans, seals, and fish, as well as of human activities (shipping and deep-sea mining) that may impact on them [43,44]. This would enable monitoring even in the most remote areas or at oceanic features that are difficult to access, such as seamounts. Similarly, scaling up the application of hydrophones with built-in low-powered AI audio classifiers could revolutionize our understanding of the distribution and abundance of marine animals [45,46].

Satellite and airborne remote sensing

ML is commonly used to interpolate data from satellite sensors to classify ecosystems and track degradation and land-use change [47]. There is also a rapidly advancing application of this technology to identify species from space [48,49]. Increasingly, deep learning approaches can detect patterns not easily picked up by conventional approaches [50]. The surface of the planet can be obscured by clouds, and imagery quality is affected by illumination angle and atmospheric effects, requiring the inclusion of complex preprocessing steps: AI is more effective than conventional approaches at making these corrections [51].

Foundation models for satellite remote sensing. Pretrained foundation models (Box 1) can facilitate the monitoring of areas with restricted availability of training and validation data [52]. However, the fine tuning of these models for specialized tasks can be highly challenging [53]. AI can facilitate the fusing of multiple data modalities, such as optical and radar imagery, as well as the use of time-series analyses for improved land cover classification, threat detection [54], and monitoring [55,56].

Predicting biodiversity loss or restoration using Earth observation foundation models. Global biodiversity monitoring uses a series of metrics, but there are often challenges in keeping the agreed metrics up to date and as close to real time as possible. AI could facilitate the automated production of Earth observation-based datasets used in policy and business applications, such as the Human Footprint Index and the Biodiversity Intactness Index which assess the degree to which terrestrial ecosystems are affected by land-use change and intensification. However, applying this will require some care to avoid error propagation and its associated impact on biodiversity science [57].

Combining datasets for new insights

Primary observational datasets are the 'fuel' for ML algorithms, but AI also creates opportunities to improve higher-level datasets by combining diverse types of input data. For example, training models on photo catalogs together with remote sensing images has facilitated the development and validation of land-cover change products such as Dynamic World [58]. Gains are likely when bringing multiple datasets together. For example, the CAPTAIN project [14] uses **reinforcement learning** to train models to identify areas for conservation prioritization by integrating species distribution data, anthropogenic disturbance data, human population densities, phylogenetic diversity, land-use data, and climate change projections, while another project combines similar datasets to estimate extinction risks for 89% of known tree species [59].

Federated AI learning. Several biodiversity data systems contain 'data islands' that cannot be directly shared for the purposes of training AI models (e.g., critically endangered species locations, or illegal hunting and trade data). Using expertise in managing sensitive data from other fields, such as healthcare [60], it may be possible through **federated learning** to train a distributed model for biodiversity characteristics which is an accurate representation of the underlying raw sensitive data and delivers real-time insights into the status and trends in biodiversity, pressures, responses, and benefits. If possible, this would allow the unification of numerous distributed databases into models that facilitate decision-making without revealing sensitive raw data. However, much work will be necessary to ensure that these models can be made robust to attacks that leak the original sensitive training data [61].

Monitor online wildlife trade. Computer vision (i.e., automated processing and understanding of images), natural language processing (i.e., automatic processing of textual content), and **multi-modal models** can be used to understand where, when, how, why, and what species and wildlife products are being traded on which online platforms [62]. For example, a recent study of global trade in chameleons used multiple lines of evidence to understand trade patterns and the impacts of trade bans [63], but automation and AI-assisted insights could have dramatically speeded up the process. Research using these methods must prioritize adherence to the highest standards of data privacy and protection [64].

Improving models of areas of habitat from diverse data. Our understanding of the areas of habitat (AoH) of different species currently relies on extremely patchy sampling, estimated range maps, and incomplete knowledge of species ecology and habitat preferences. AI could improve AoH maps by integrating occurrence data on specimen locations, habitat and cohabitation preferences, environmental and ecological data, remote sensing data, and spatial datasets on human impacts. These approaches are already being tested [65] and could move from relatively simple correlative AoH models (e.g., masking estimated species range maps with elevation and climate) to more sophisticated covariates that describe species habitat selection across diverse ecosystems, and deep learning could make more comprehensive predictions of species ranges and perhaps also populations [66,67].

Predicting and mitigating human–wildlife conflicts. Human–wildlife conflict, where people themselves, or their property, crops, or livestock, are harmed by wildlife is a serious problem for some human communities and, when it results in retaliatory killing, is an important threat to some species [68]. AI-powered cameras are already being used in various countries to warn communities when particular mammal species are nearby, for example elephants in Africa or tigers in Asia [69]. Similar technology can also track humans within landscapes set aside for animals[ix]. However, there are important ethical and privacy concerns with such technology which potentially may exacerbate conflict between conservation and local communities [15]. AI-enabled systems combining short- and long-term data on factors correlated with human–wildlife conflicts will provide more accurate predictions and opportunities to take action earlier.

Modeling and causal inference

Using datasets, usually in combination, there is the potential to apply AI to build system models that can be used to test hypotheses or to build credible counterfactuals for the evaluation of conservation policy and practice. Uses could include simulating the outcomes of policy decisions and predicting projections based on a range of data sources. Large-scale modeling of biodiversity outcomes from different policy interventions, using general ecosystem models, are increasingly available [70,71]. However, they are time-consuming and resource-intensive, and are typically used to inform international policy formulation – for example the Global Biodiversity Framework [72] and the EU New Green Deal[x]. AI may offer opportunities to accelerate these large-scale modeling efforts and observe emerging patterns that are important for policy decision-making.

Digital twins for ecosystems. **Digital twins** are used in engineering and Earth sciences [73], but they can also be applied to the problem of predicting outcomes of conservation interventions. Digital twins for ecosystem modeling and prediction could be developed, and potentially paired across multiple domains (e.g., physics, fish demography, and seabird ecology), to understand and model the likely outcomes of different interventions [74,75]. Moreover, where we seek to apply AI to optimize intervention policies (as in CAPTAIN), simulation from a digital twin can provide an additional training signal.

Prioritizing restoration efforts. Restoring degraded habitats is essential to avert extinctions, recover populations, and minimize climate change. AI could help us to direct such efforts more efficiently by identifying the most important locations for restoration given habitat loss and degradation to date, the distribution of species and ecosystems, projected climate change, and shifts in energy production, food production, and human population distributions. Artificial **neural networks** have been previously explored for prioritizing areas for wetland restoration (utilizing multiple inputs including elevation), as well as for generating soil texture and permeability maps, identifying protected areas for waterbirds, and estimating the likelihood of dust storms and urbanization [76].

Improving species distribution models. **Species distribution models (SDMs)** are valuable for conservation decision making such as assessing land-use change impacts. Using species distribution records and remotely sensed data, SDMs produce maps of the (relative) probability of occurrences. ML models are well suited to integrate heterogeneous and multi-fidelity datasets efficiently, and can handle missing data, noise, and non-linear relationships; they can automatically learn relevant features from raw spatial data and uncover complex patterns and interactions. However, current limitations include the scarcity of labeled data (especially for plants and non-vertebrate animals), spatial biases in data, temporal mismatches between field and remotely sensed data, data accuracy, over-reliance on abiotic conditions, limited consideration of biotic interactions, and overfitting. New AI approaches [e.g., **physics-inspired AI** and explainable AI (xAI)] and rapidly increasing citizen science datasets may help to address these problems.

Predicting deforestation and agricultural expansion. Deep learning can be used to analyze satellite imagery and other geospatial data to predict where land-use change is most likely by learning spatiotemporal patterns from the remote sensing data [77] instead of using the mechanistic modeling approaches that are currently used. Such models could be used to predict where deforestation would have occurred under business-as-usual scenarios to allow better estimates of the effectiveness of interventions aimed at slowing land conversion [such as zero deforestation commitments or reducing emissions from deforestation in developing countries (REDD+) projects]. Similar techniques can also be applied to training predictive models from time-series maps of agricultural expansion which has seen rapid acceleration in the past two decades [78].

Parameter estimation for global ecosystem models. Global ecosystem models (GEMs) simulate the dynamics of life on Earth, including the interactions between plants, animals, and the environment. GEMs are valuable for modeling the impacts of climate change on nature. However, predictions of dynamic GEMs do not align well with field measurements (e.g., of fluxes of greenhouse gas); they fail to make reliable predictions of large-scale fires and droughts, and they rarely include plant–animal interactions. Many processes in dynamic GEMs are represented by semi-empirical semi-theoretical equations. AI optimization techniques could automate the calibration of the parameters in these models across multiple scales. Physics-inspired AI techniques, such as reduced-order modeling, could potentially find low-dimensional representations that capture the essential dynamics while being much faster to compute [75].

Reviewing the literature

There is considerable potential for using AI to improve the efficiency of extracting, screening, and collating literature for conservation [79], as is already underway to some extent in medicine [80]. For example, the creation of systematic reviews could be automated such that they can be carried out in days rather than months or years, enabling more timely guidance drawing on the full range of evidence.

LLMs for evidence synthesis. Evidence syntheses for biodiversity conservation are key for effective decision-making, but are challenged by increasingly time-consuming tasks, a broad evidence base, and persistent underfunding. Moreover, most evidence syntheses produced in the conservation space are static and fail to incorporate new evidence as it is generated. AI has the potential to be harnessed to identify relevant (new) evidence and integrate it into the existing evidence base and synthesize key messages rapidly (i.e., living evidence syntheses [81]). Doing so will ensure decision-makers have access to the best available evidence to guide them.

Multilingual literature scanning. LLMs enable multilingual literature searches for non-English evidence [82], which is key given that non-English languages typically dominate in areas of most conservation concern [83], although there is a growing quantity of non-English evidence [84].

LLMs to interpret gray literature. LLMs can be used for identifying gray literature and text matching to cluster relevant evidence. They can accelerate evidence synthesis and make it more timely, equitable, and inclusive in terms of the evidence base and the perspectives considered.

Assessing extinction and collapse risk from diverse sources. Assessments of extinction risk for species and collapse risk for ecosystems currently rely on manual compilation of information on the population sizes of species and their distributions and trends, as well as trends in areas and biotic/abiotic factors for ecosystems, to produce parameter estimates that are applied to IUCN Red List criteria. AI could accelerate and expand the process by scanning the scientific literature and other online materials to locate relevant information in published and unpublished sources in all languages, including real-time land-cover change. This would substantially accelerate and improve the process of assessing extinction and collapse risk while significantly reducing costs. More ambitiously, extinction and collapse risk could be estimated directly in the future by combining relevant literature-derived parameter estimates with spatially explicit predictive models informed by remote sensing.

Generating text

The success of LLMs in creating novel text, images, and videos is well recognized. There has long been a distinction between the use of generative AI for error-tolerant commercial applications such as marketing and social media, as well as for safety-critical or politically sensitive applications in specialist use cases. However, their recent success has led to their application being debated in medicine [85]. Their ability to generate personalized outputs and recommendations based on a synthesis of available evidence has led to suggestions that they could play a role as policy advisors [86]. This could include synthesis of multiple data sources for a specific location to provide bespoke management recommendations, or application to complex numerical datasets to provide non-specialist text summaries for decision-makers.

LLM conservation advisors

LLMs could synthesize the evidence of impact for different conservation management options [79], and combine this with data on species, land-use, or socioeconomic factors for a specific context to produce easily understandable and evidence-based information for practitioners and policymakers. The inclusion of **human in the loop**, whereby human experts are involved in providing feedback and evaluating advisor outputs during model training, would be essential to limit biases and hallucinations [86], as would fine-tuning and **retrieval augmented generation (RAG)** based on curated, robust evidence and data.

Negative consequences of AI

AI will undoubtedly lead to substantial changes in conservation in the coming years. Although many will be positive, there is also the potential for unintended negative outcomes. These need to be understood and mitigated.

There is a danger that AI could have a polarizing effect on conservation and conservation funding. If AI-supported conservation becomes the 'gold standard', because people are impressed by the novelty, claims of large impacts, or the promise to revolutionise conservation, then we may risk seeing a shift in funding and leadership in conservation. Support for conventional experimentation and on-the-ground practices, which already struggle to attract resources, could be redirected towards financially wealthy institutions which are able to undertake AI work. This could be especially true in conservation research where grounded field-based and participatory studies, which have a role in advancing understanding and local ownership, may become ever more difficult to fund. This could undermine efforts to improve the diversity of voices, knowledge, and approaches in conservation. Hence it is important that funders recognize the importance of supporting a spectrum of conservation research and practice that embraces both conventional and AI approaches.

There is a fear that we could see a loss of essential skills in conservation if people in the field pivot towards implementing AI over conventional techniques. Retaining species, ecosystem, and community experts will be integral to creating reliable AI technology. Data is the fuel of AI, and data collected by conservation experts will be essential for producing better models, and this crucial data-collection work must be appropriately recognized. Moreover, information itself, however it is obtained, does not lead to better conservation, and it is important that any recommendations are designed to work in the real world and are not detached from the social and ecological reality on the ground.

AI colonialism is a central concern – data potentially extracted from the Global South might be forwarded predominantly to data centers in the Global North for training AI models, followed by AI-driven mandates being issued to the Global South on how land and resources should be managed. This would undermine the efforts of the conservation community to address the colonial legacy of contemporary conservation and recognize the importance of indigenous rights and voices. Furthermore, there is a risk that AI contributes to a militarization of conservation – where computer systems, developed far from the area concerned, identify infractions and trigger enforcement without understanding the local context. Given that local perceived legitimacy is essential to promote compliance with conservation rules [87], this could create or exacerbate conflict. To address digital inequalities and injustices, and to produce less biased, fairer, and more robust information for conservation actions, there is a need to integrate epistemic feedback

CellPress
OPEN ACCESS

loops into black box models. This can be achieved by leveraging human-in-the-loop designs as well as political agencies and democratic decision-making [88].

Combining AI models can create inherent biases and propagate errors, leading to 'AI pollution' if the outputs of biased models are used to train new models. This may lead to increasingly poor representation of understudied species or ecosystems, potentially pushing people away from considering these understudied areas if we are over-reliant on AI for decision-making, and this needs to be explicitly considered in any AI-based approaches.

Much of the promise of AI lies in bigger and better models. However, the bigger the models become, the more expensive they are to run in terms of computation, bandwidth requirements, power, and expertise. There is already significant concern about the environmental implications of the power consumption and cooling requirements of AI infrastructures, and these are likely to increase. The cost of using these models may push AI beyond the current resources (financial and human) of conservation. However, in conservation there is less call for large generalist AI models such as Chat GPT, and the focus is instead on smaller and more specialized models for specific use cases. It would be helpful for the conservation community to adopt a code of practice to address the sustainability considerations associated with AI research.

Finally, although AI models to further the effectiveness of conservation are built with good intentions, it must be remembered that they could also be used by bad actors. For example, image and audio tools used by conservationists to track and locate endangered or protected species can equally be exploited by poachers or be coopted by government regimes to monitor the movement of people. Similarly, remote sensing data could be contaminated or poisoned by malicious private companies to exaggerate restoration efforts so as to attract greater revenue from carbon or biodiversity credit schemes. It is important that significant steps are taken to protect data and to ensure that tools are used only for their intended purpose.

How might conservation be organized to take advantage of AI and reduce problems?

To take advantage of the potential ability of AI to identify new patterns, generate accurate predictions, run more accurate simulations drawing on diverse data sources, and help decision makers to better assess the efficacy of potential interventions, society and the field of conservation will need to adapt. Reproducibility of results, one of the pillars of the scientific method, becomes challenging when there is insufficient documentation, limited access to the underlying codes and data, and a lack of understanding of how AI tools reach conclusions, which makes it difficult to scrutinize, verify, and replicate experiments [11]. Improved literacy concerning AI will help to counteract its misinterpretation and recognize its limitations. Improved AI literacy will also be essential to ensure sufficient peer review of papers using AI.

To counteract the influx of outputs from generative AI models, it will be important for society to identify a way to segregate human-produced content from that produced by AI. One such example could be digital kitemarks, where individuals, organizations, and institutions are assigned digital signatures that can be used to authenticate the human-generated origin a document or dataset.

Interdisciplinary collaboration between AI specialists, domain experts in the natural and social sciences, and indigenous and local knowledge will be central to building accurate AI models. There is a danger of decoupling between the creators and users of AI tools. Hallucinations in outputs are more likely to be detected by subject experts. This also extends to ensuring that there is close consultation between the developers of AI tools and the experts who manage underlying datasets to ensure that data limitations are understood and the data are not used inappropriately because inaccurate use of data can lead to misleading advice for decision-makers [89]. Moreover, researchers should not incorporate advanced AI techniques at the cost of appropriate conventional methodologies. It should be remembered that AI is only one of the tools in the toolbox.

There are additional barriers to broad uptake of new approaches which will need to be considered. There is a large inequality of data availability between the Global South and the Global North. This needs to be acknowledged and understood to avoid embedding biases in AI models that are intended to predict outputs on a global scale. This inequality extends also to ecosystems and species where there is a disparity in data availability – for example, data on birds in the Amazonian rainforest are more numerous than for marine bryozoans in Antarctica.

Access to the infrastructure and resources that are necessary to train and run these models is currently limited and is disproportionately held by elite academic institutions, governments, and technology companies in the Global North. Training foundation AI models is expensive and requires a large amount of computing power, as well as IT expertise, high internet bandwidth, and reliable power supplies. Methods of access to these resources, as well as overcoming language barriers, need to be thoughtfully considered to widen participation in AI research by conservation practitioners, researchers, and governance institutions including civil society associations, non-governmental organizations (NGOs), and government conservation managers based in the Global South.

The role of expert data labelers will be crucial for maintaining data quality, preventing biased outputs, and addressing data fabrication. Many species and ecosystem experts are based in the Global South, and it is imperative that their skills and labor are equitably used and recognized. It also highlights the value of conventional conservation expertise in the production of these models. Moreover, concerted efforts will be needed to substantially strengthen the connection between the outputs of AI models and the ecological and social realities of implementing and sustaining on-the-ground conservation actions [90].

Although AI systems have potential to make monitoring of infractions of conservation rules cheaper and therefore more widespread, they cannot by themselves overcome social and governance challenges which strongly influence compliance. For example, regulators can put AI-enabled cameras on board ships to monitor bycatch, but the utility of the technology will depend on compliance and follow-through.

Concluding remarks

Following a horizon scan methodology, we identified 21 areas where AI can help to revolutionize conservation. These represent both a cross-section of research areas within conservation science and practice, as well as methods through which AI contributes to scientific understanding.

AI stands to rapidly improve our ability to understand distributions of species; locate rare or so far unknown species; identify, model, and monitor threats; identify priority areas for conservation; model the effectiveness of planned conservation actions; monitor adherence to environmental legislation; and assimilate and interact with scientific evidence. However, we need to ensure that AI-supported conversation does not replace valuable established conservation techniques, education, and on-the-ground research.

The intersection within conservation between scientists, practitioners, governments, local communities, and indigenous peoples is nuanced and complex. It is important that AI technologies

Outstanding questions

How do researchers and practitioners ensure that funders recognize the importance of supporting a spectrum of research embracing both traditional and AI approaches? There is a danger that AI could have a polarizing effect on conservation funding, and directs funds away from effective traditional methods of conservation.

How can the field of conservation redress the unbalanced access to computing resources required for AI research between the Global North and Global South? AI colonialism should be a central concern of AI technology in conservation. With data potentially being sent from the Global South to data centers in the Global North where they are used to train AI models, we must be careful not to issue AI-driven mandates to the Global South on how land and resources should be managed.

What protections can be put in place to prevent exploitation by bad actors? These technologies could be coopted by bad actors to track human populations, locate endangered and/or valuable species, or identify and pollute data sources for financial gain. How this can be prevented requires careful consideration.

Should the conservation community create and adopt a code of practice to address sustainability, equity, equality, and data privacy and security in AI research? To ensure that these issues are considered in the pursuit of producing AI technologies for conservation, the creation of a centralized code of practice may lead to more robust and considered applications.

are developed and deployed with understanding of local contexts. Moreover, the fact that many of the intact ecosystems of our planet reside in the Global South poses many significant challenges to the equitable implementation of AI technologies because at present these are predominantly trained and developed in the Global North.

AI will be an invaluable tool to support conservation, but it is not a panacea. Proactive and creative efforts to embrace AI, while also ensuring that proper protections and attention to equitable and just conservation practices are in place, will be necessary for AI to reach its transformative potential (see Outstanding questions).

Acknowledgments

This work was undertaken as part of the AI@cam project, which we thank for funding. N.B. was funded by the UK Research and Innovation Global Challenges Research Fund (UKRI GCRF) through the Trade, Development, and the Environment Hub project (project ES/S008160/1). F.D. was supported by the EU projects FRAMEwork, European Citizen Science (ECS), BESTLIFE2030, and more4nature (grants 862731, 101058509, 101096887, 101133983), UArctic, and Danida (Maarifa). Support for A.H. was provided by Schmidt Sciences LLC. N.P. was supported by Research England.

Declaration of interests

The authors declare no competing interests.

Resources

[i] https://www.skylight.global/

[ii] https://arbimon.org/

[iii] https://sentinel.conservationxlabs.com/

[iv] https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

[v] https://github.blog/news-insights/product-news/github-copilot-is-generally-available-to-all-developers

[vi] https://observation.org/

[vii] https://lens.google/

[viii] https://merlin.allaboutbirds.org

[ix] https://news.mongabay.com/2018/06/this-tiny-camera-aims-to-catch-poachers-before-they-kill/

[x] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

[ix] https://zenodo.org/records/13170385

Supplemental information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tree.2024.11.013.

References

1. Ripple, W.J. *et al.* (2017) World scientists' warning to humanity: a second notice. *BioScience* 67, 1026–1028
2. Pollock, L.J. *et al.* (2020) Protecting biodiversity (in all its complexity): new models and methods. *Trends Ecol. Evol.* 35, 1119–1128
3. Ferraro, P.J. *et al.* (2019) Causal inference in coupled human and natural systems. *Proc. Natl. Acad. Sci. U. S. A.* 116, 5311–5318
4. Nuno, A. *et al.* (2014) Managing social-ecological systems under uncertainty: implementation in the real world. *Ecol. Soc.* 19, 52
5. Jiang, F. *et al.* (2017) Artificial intelligence in healthcare: past, present and future. *Stroke Vasc. Neurol.* 2, 230–243
6. Gill, A.S. (2019) Artificial intelligence and international security: the long view. *Ethics Int. Aff.* 33, 169–179
7. Jordan, M.I. and Mitchell, T.M. (2015) Machine learning: trends, perspectives, and prospects. *Science* 349, 255–260
8. Xu, L. *et al.* (2023) Enhanced persistent orientation of asymmetric-top molecules induced by cross-polarized terahertz pulses. *Phys. Rev. Res.* 7, 013201
9. Binley, A.D. *et al.* (2024) Making the most of existing data in conservation research. *Perspect. Ecol. Conserv.* 22, 122–128
10. Convention on Biological Diversity (2022) *Kunming–Montreal Global Biodiversity Framework*, United Nations Environment Programme
11. Science in the Age of AI Working Group (2024) *Science in the Age of AI: How Artificial Intelligence Is Changing the Nature and Method of Scientific Research*, The Royal Society
12. Krenn, M. *et al.* (2022) On scientific understanding with artificial intelligence. *Nat. Rev. Phys.* 4, 761–769
13. Cole, E. *et al.* (2023) Spatial implicit neural representations for global-scale species mapping. *ArXiv*, Published online June 5, 2023. https://doi.org/10.48550/arXiv.2306.02564
14. Silvestro, D. *et al.* (2022) Improving biodiversity protection through artificial intelligence. *Nat. Sustain.* 5, 415–424
15. Simlai, T. and Sandbrook, C. (2021) Digital surveillance technologies in conservation and their social implications. In *Conservation Technology* (Wich, S.A. and Piel, A.K., eds), pp. 239–254, Oxford University Press
16. Spelda, P. and Stritecky, V. (2020) The future of human-artificial intelligence nexus and its environmental costs. *Futures* 117, 102531

17. García-Roselló, E. *et al.* (2023) The biased distribution of existing information on biodiversity hinders its use in conservation, and we need an integrative approach to act urgently. *Biol. Conserv.* 283, 110118
18. Christin, S. *et al.* (2019) Applications for deep learning in ecology. *Methods Ecol. Evol.* 10, 1632–1644
19. Xian, Y. *et al.* (2017) Zero-shot learning – the good, the bad and the ugly. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3077–3086, IEEE
20. Xi, Z. *et al.* (2023) The rise and potential of large language model based agents: a survey. *ArXiv*, Published online September 14, 2023. https://doi.org/10.48550/arXiv.2309.07864
21. Bommasani, R. *et al.* (2021) On the opportunities and risks of foundation models. *ArXiv*, Published online August 16, 2021. https://doi.org/10.48550/arXiv.2108.07258
22. Oquab, M. *et al.* (2023) DINOv2: learning robust visual features without supervision. *ArXiv*, Published online April 14, 2023. https://doi.org/10.48550/arXiv.2304.07193
23. Chen, T. *et al.* (2020) Big self-supervised models are strong semi-supervised learners. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, pp. 22243–22255, NIPS
24. Sutherland, W.J. *et al.* (2014) A horizon scan of global conservation issues for 2014. *Trends Ecol. Evol.* 29, 15–22
25. Sutherland, W.J. *et al.* (2024) A horizon scan of global biological conservation issues for 2024. *Trends Ecol. Evol.* 39, 89–100
26. Mukherjee, N. *et al.* (2015) The Delphi technique in ecology and biological conservation: applications and guidelines. *Methods Ecol. Evol.* 6, 1097–1109
27. Sutherland, W.J. *et al.* (2019) Ten years on: a review of the first global conservation horizon scan. *Trends Ecol. Evol.* 34, 139–153
28. Van Horn, G. *et al.* (2018) The iNaturalist Species Classification and Detection Dataset. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8769–8778, IEEE
29. Goëau, H. *et al.* (2013) Pl@ntNet mobile app. In *Proceedings of the 21st ACM International Conference on Multimedia*, pp. 423–424
30. Chalmers, C. *et al.* (2019) Conservation AI: live stream analysis for the detection of endangered species using convolutional neural networks and drone technology. *ArXiv*, Published online October 16, 2019. https://doi.org/10.48550/arXiv.1910.07360
31. Pirotta, V. *et al.* (2022) Detecting illegal wildlife trafficking via real time tomography 3D X-ray imaging and automated algorithms. *Front. Conserv. Sci.* 3, 757950
32. de Lorm, T.A. *et al.* (2023) Optimizing the automated recognition of individual animals to support population monitoring. *Ecol. Evol.* 13, e10260
33. Zhao, B. *et al.* (2024) Labeled data selection for category discovery. *ArXiv*, Published online June 7, 2024. http://dx.doi.org/10.48550/arXiv.2406.04898
34. Lu, Y. *et al.* (2024) Accelerating the discovery of biodiversity by detecting 'new' species based on machine learning method. *Res. Sq.*, Published online January 9, 2024. https://doi.org/10.21203/rs.3.rs-3832815/v1
35. Dias, M.P. *et al.* (2019) Threats to seabirds: a global assessment. *Biol. Conserv.* 237, 525–537
36. Da Rocha, N. *et al.* (2021) Reduction in seabird mortality in Namibian fisheries following the introduction of bycatch regulation. *Biol. Conserv.* 253, 108915
37. Kahl, S. *et al.* (2021) BirdNET: A deep learning solution for avian diversity monitoring. *Ecol. Inform.* 61, 101236
38. Manzano-Rubio, R. *et al.* (2022) Low-cost open-source recorders and ready-to-use machine learning approaches provide effective monitoring of threatened species. *Ecol. Inform.* 72, 101910
39. Williams, B. *et al.* (2024) Unlocking the soundscape of coral reefs with artificial intelligence: pretrained networks and unsupervised learning win outbio. *BioRxiv*, Published online May 8, 2024. https://doi.org/10.1101/024.02.02.578582
40. Müller, J. *et al.* (2023) Soundscapes and deep learning enable tracking biodiversity recovery in tropical forests. *Nat. Commun.* 14, 6191
41. Linhart, P. *et al.* (2022) The potential for acoustic individual identification in mammals. *Mamm. Biol.* 102, 667–683
42. Snyder, D. *et al.* (2018) X-Vectors: robust DNN embeddings for speaker recognition. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5329–5333, IEEE
43. Bouffaut, L. *et al.* (2022) Eavesdropping at the speed of light: distributed acoustic sensing of baleen whales in the arctic. *Front. Mar. Sci.* 9, 901348
44. Wilcock, W.S.D. *et al.* (2023) Distributed acoustic sensing recordings of low-frequency whale calls and ship noise offshore Central Oregon. *JASA Express Lett.* 3, 026002
45. Parijs, S.M.V. *et al.* (2009) Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales. *Mar. Ecol. Prog. Ser.* 395, 21–36
46. Sanguineti, M. *et al.* (2021) Real-time continuous acoustic monitoring of marine mammals in the Mediterranean Sea. *J. Mar. Sci. Eng.* 9, 1389
47. Helber, P. *et al.* (2019) EuroSAT: a novel dataset and deep learning benchmark for land use and land cover classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 12, 2217–2226
48. Duporge, I. *et al.* (2021) Using very-high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes. *Remote Sens. Ecol. Conserv.* 7, 369–381
49. Cubaynes, H.C. and Fretwell, P.T. (2022) Whales from space dataset, an annotated satellite image dataset of whales for training machine learning models. *Sci. Data* 9, 245
50. Pettorelli, N. *et al.* (2024) Deep learning and satellite remote sensing for biodiversity monitoring and conservation. *Remote Sens. Ecol. Conserv.*, Published online June 17, 2024. https://doi.org/10.1002/rse2.415
51. Lisaius, M.C. *et al.* (2024) Using Barlow twins to create representations from cloud-corrupted remote sensing time series. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 17, 13162–13168
52. Jakubik, J. *et al.* (2023) Foundation models for generalist geospatial artificial intelligence. *ArXiv*, Published online October 28, 2023. https://doi.org/10.48550/arXiv.2310.18660
53. Diao, S. *et al.* (2024) LMFlow: an extensible toolkit for finetuning and inference of large foundation models. *ArXiv*, Published online May 5, 2024. https://doi.org/10.48550/arXiv.2306.12420
54. Ortega Adarme, M. *et al.* (2022) Improving deforestation detection on tropical rainforests using Sentinel-1 data and convolutional neural networks. *Remote Sens.* 14, 3290
55. Ye, W. *et al.* (2022) Pine pest detection using remote sensing satellite images combined with a multi-scale attention-UNet model. *Ecol. Inform.* 72, 101906
56. López, S. *et al.* (2020) Landscape change in Southern Ecuador: an indicator-based and multi-temporal evaluation of land use and land cover in a mixed-use protected area. *Ecol. Indic.* 115, 106357
57. Mittelstadt, B. *et al.* (2023) To protect science, we must use LLMs as zero-shot translators. *Nat. Hum. Behav.* 7, 1830–1832
58. Brown, C.F. *et al.* (2022) Dynamic World, Near real-time global 10 m land use land cover mapping. *Sci. Data* 9, 251
59. Silva, S.V. *et al.* (2022) Global estimation and mapping of the conservation status of tree species using artificial intelligence. *Front. Plant Sci.* 13, 839792
60. Rieke, N. *et al.* (2020) The future of digital health with federated learning. *NPJ Digit. Med.* 3, 119
61. Nasr, M. *et al.* (2023) Scalable extraction of training data from (production) language models. *ArXiv*, Published online November 28, 2023. https://doi.org/10.48550/arXiv.2311.17035
62. Kulkarni, R. and Di Minin, E. (2023) Towards automatic detection of wildlife trade using machine vision models. *Biol. Conserv.* 279, 109924
63. Isaac, M.C. *et al.* (2024) Status and trends in the international wildlife trade in Chameleons with a focus on Tanzania. *PLoS ONE* 19, e0300371
64. Di Minin, E. *et al.* (2021) How to address data privacy concerns when using social media data in conservation science. *Conserv. Biol.* 35, 437–446
65. O'Malley, W.C. *et al.* (2024) Machine learning allows for large-scale habitat prediction of a wide-ranging carnivore across diverse ecoregions. *Landsc. Ecol.* 39, 106
66. Estopinan, J. *et al.* (2024) Modelling species distributions with deep learning to predict plant extinction risk and assess climate

change impacts. *ArXiv*, Published online January 10, 2024. https://doi.org/10.48550/arXiv.2401.05470

67. Zhang, H.-T. *et al.* (2024) A novel hybrid model for species distribution prediction using neural networks and Grey Wolf Optimizer algorithm. *Sci. Rep.* 14, 11505
68. Gross, E. *et al.* (2021) *A Future for All: The Need for Human–Wildlife Coexistence*, World Wildlife Fund https://doi.org/10.1038/s41598-024-62285-8
69. Dertien, J.S. *et al.* (2023) Mitigating human–wildlife conflict and monitoring endangered tigers using a real-time camera-based alert system. *BioScience* 73, 748–757
70. Leclère, D. *et al.* (2020) Bending the curve of terrestrial biodiversity needs an integrated strategy. *Nature* 585, 551–556
71. Newbold, T. *et al.* (2020) Non-linear changes in modelled terrestrial ecosystems subjected to perturbations. *Sci. Rep.* 10, 14051
72. Díaz, S.M. *et al.* (2019) *The Global Assessment Report on Biodiversity and Ecosystem Services: Summary for Policy Makers*, Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
73. Tzachor, A. *et al.* (2023) Digital twins: a stepping stone to achieve ocean sustainability? *NPJ Ocean Sustain.* 2, 16
74. Keith, D.A. *et al.* (2022) A function-based typology for Earth's ecosystems. *Nature* 610, 513–518
75. Kochkov, D. *et al.* (2024) Neural general circulation models for weather and climate. *Nature* 632, 1060–1066
76. Maleki, S. *et al.* (2020) Wetland restoration prioritization using artificial neural networks. *Wetlands* 40, 179–192
77. Ball, J.G.C. *et al.* (2022) Using deep convolutional neural networks to forecast spatial patterns of Amazonian deforestation. *Methods Ecol. Evol.* 13, 2622–2634
78. Potapov, P. *et al.* (2022) Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nat. Food* 3, 19–28
79. Berger-Tal, O. *et al.* (2024) Leveraging AI to improve evidence synthesis in conservation. *Trends Ecol. Evol.* 39, 548–557
80. Blaizot, A. *et al.* (2022) Using artificial intelligence methods for systematic review in health sciences: a systematic review. *Res. Synth. Methods* 13, 353–362
81. Elliott, J.H. *et al.* (2014) Living systematic reviews: an emerging opportunity to narrow the evidence–practice gap. *PLoS Med.* 11, e1001603
82. Zenni, R.D. *et al.* (2023) Multi-lingual literature searches are needed to unveil global knowledge. *J. Appl. Ecol.* 60, 380–383
83. Christie, A.P. *et al.* (2021) The challenge of biased evidence in conservation. *Conserv. Biol.* 35, 249–262
84. Chowdhury, S. *et al.* (2022) Growth of non-English-language literature on biodiversity conservation. *Conserv. Biol. J. Soc. Conserv. Biol.* 36, e13883
85. Thirunavukarasu, A.J. *et al.* (2023) Large language models in medicine. *Nat. Med.* 29, 1930–1940
86. Tyler, C. *et al.* (2023) AI tools as science policy advisers? The potential and the pitfalls. *Nature* 622, 27–30
87. Duffy, R. *et al.* (2019) Why we must question the militarisation of conservation. *Biol. Conserv.* 232, 66–73
88. Debnath, R. *et al.* (2023) Harnessing human and machine intelligence for planetary-level climate action. *NPJ Clim. Action* 2, 20
89. Challender, D.W.S. *et al.* (2022) Mischaracterizing wildlife trade and its impacts may mislead policy processes. *Conserv. Lett.* 15, e12832
90. Danielsen, F. *et al.* (2021) The concept, practice, application, and results of locally based monitoring of the environment. *BioScience* 71, 484–502
91. Danziger, S. *et al.* (2011) Extraneous factors in judicial decisions. *Proc. Natl. Acad. Sci.* 108, 6889–6892
92. R Core Team (2018) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing