Paradigms and Principles for Integrating Nature Technologies: Barriers and Opportunities in Real-World Biodiversity Monitoring

Seamus Lombardo [1&16] (seamus.lombardo@planet.com), Amy Rosenthal [1&15], Christopher Anderson [1], Ken Bagstad [2], Carly Batist [3], Christopher Beirne [4], Paul Bunje [5], Marconi Campos-Cerqueira [6], Jeremy Cusack [7], Elsa Carla De Grandi [8], Rachel King [9], Trond Larsen [3], David Luther [10], Jennifer McGowan [11], Adrien Pajot [12], Xavier Raick [13], Vijay Ramesh [13], Chelsea Smith [8], Ali Swanson [3], Tom Walker [14]

Affiliations
1. Planet Labs PBC; San Francisco, CA, US
2. Independent researcher, Charlotte, VT, US
3. Conservation International; Arlington, VA, US
4. Asociación Conservación Osa; Puntarenas, CR
5. Conservation X Labs; Washington, DC, US
6. WildMon, Lewes, DE, US
7. Okala, London, EC2R 8AY, UK
8. Fauna & Flora International, Cambridge, CB2 3QZ, UK
9. National Center for Ecological Analysis and Synthesis, University of California - Santa Barbara, Santa Barbara, CA US
10. Biology Department, George Mason University VA, USA
11. Centre for Biodiversity and Conservation Science, University of Queensland, Queensland, Australia
12. WILDLABS at Fauna & Flora International, Cambridge, CB2 3QZ, UK
13. Center for Avian Population Studies, Cornell Lab of Ornithology, NY USA
14. Cecil Earth, Surry Hills NSW 2010 Australia
15. Field Museum of Natural History; Chicago, IL, USA
16. Gund Institute for Environment, University of Vermont, Burlington, VT, USA

Abstract:
Employing a combination of nature technologies including satellite data, eDNA, camera traps, passive acoustic sensors, and GPS monitors can augment traditional biodiversity measurements to help respond to increasing demand for understanding conservation status and outcomes. While examples exist in the literature that integrate these sensors for biodiversity monitoring, there remains a significant gap between individual cases of early exploration and the development and examination of replicable, scalable, and standardized approaches for real-world decision making. Notably, there is no clear guidance or consensus on how nature technologies are best integrated in practice. To help address this gap, we aim to provide an actionable framework for how nature technologies are being integrated, and present how this enhances understanding at the multiple scales required for growing nature-positive programs. We propose a set of paradigms (survey design, tip-and-cue, validation, consilience, interpolation and extrapolation, and data fusion) and prioritize a set of principles for nature data integration (metadata standards, use rights and licensing, scientific rigor, and utility). We then analyze a diverse set of seven real-world case studies across biomes, use cases, and user types to understand what paradigms are leveraged in practice, how prioritized principles have been implemented, and identify barriers and opportunities for increased technology integration in conservation. Distilling data integration into paradigms helped map the relationships between workflows, such as how low-barrier-to-entry methods (like Consilience) can lead to deeper integration (Extrapolation or Data Fusion) and that there are different purposes and tools needed, even when sensor use is similar. The paradigms also helped elucidate key challenges (such as mismatches in scale and lack of data, reference libraries, and methodological guidance) and opportunities (increasing data resolution, systematizing guidance, and producing new repeatable tools). For prioritized principles,

metadata standards and vehicles for sharing permissions and use-rights could increase impact and reuse. We conclude that a collective, concerted effort is needed to develop a roadmap for guidance, tools, and standards to operationalize these paradigms and principles, providing a foundation for realizing the potential of integrated biodiversity monitoring to help guide future research and conservation action.

Keywords: nature technology; biodiversity; satellite data; eDNA; camera traps; passive acoustic sensors; GPS tags; data fusion

1. Introduction

Sound environmental management and policy decisions require significant, regularly updated information about ecosystem health, ecological processes, and population dynamics. Without robust data, decisions can be made using flawed heuristics, failing to address complex situations on the ground and risk being ineffective and costly. This lack of data-based decision-making has long been a challenge in the conservation community, spurring initiatives like Conservation Evidence (*Conservation Evidence*, 2025) and quantitative conservation status assessments for the IUCN Red List of Threatened Species (*IUCN*, 2025). The global biodiversity framework (CBD, 2022), associated national commitments, corporate reporting efforts (McKenzie *et al.*, 2025; Miller *et al.*, 2025; TNFD, 2025; *NPI*, 2025), and growing biodiversity credit initiatives have only resulted in a higher demand for effective, scalable, and holistic ecosystem monitoring. Overcoming the persistent reality of sparse, fragmented, infrequently updated biodiversity data remains a priority for global conservation efforts.

Employing a combination of nature technologies such as satellite data, eDNA, AI-enabled camera traps, passive acoustic sensors and GPS monitors can augment traditional biodiversity measurements to help respond to increasing demand for understanding ecological status and conservation outcomes (Bell and Malerba, 2025). However, significant technical and operational challenges remain to deploying these technologies and synthesizing the results (Marvin *et al.*, 2016; Anderson, 2025). While previous literature has surveyed integrated biodiversity monitoring (Henry *et al.*, 2008), there remains a significant gap in understanding what works in practice for extracting timely insights at scale, to inform decision making. Notably, there is not yet broad consensus on how nature technologies are best integrated and made interoperable in practice, nor how or whether they can be automated (King and Halpern, 2025). Our contention is that for many conservation solutions – from nature reporting to biodiversity credits – there is a need for replicable, scalable, interpretable, and accessible integrated data from nature technologies, delivered in time for key decisions.

In this paper, we aim to provide an actionable framework for how nature technologies can be integrated for biodiversity data collection in practice, and how and where this enhances our understanding of the state of nature and the outcomes of conservation interventions across multiple scales. First, we identify ways in which nature technologies can be integrated to produce meaningful and actionable insights for practitioners, proposing a set of paradigms. Next, we consider whether and how existing principles for data collection and management apply to the newfound challenges and needs of nature data integration, and prioritize several principles based on expert input. We then present a diverse set of real-world case studies across biomes, use cases, and user types to better understand what paradigms are actively leveraged in practice, how priority principles have been implemented, and whether implementation has delivered enhanced understanding. Drawing on these real examples, we identify current barriers, challenges, and opportunities for harnessing these approaches for effective conservation and decision-making.

2. Paradigms for Nature Technology Integration

Multimodal data integration has benefitted a variety of fields, from medicine (Cai *et al.*, 2019; Xu *et al.*, 2024) to the Internet-of-Things (Al-Fuqaha *et al.*, 2015). In comparison, nature technology data integration is a nascent but growing area of focus for biodiversity monitoring and ecosystem assessment (Besson *et al.*, 2022; Hartig *et al.*, 2024; Sutherland *et al.*, 2025), featuring a wide variety of technologies and limited documentation and research on integration techniques, methodologies, and benefits (Lahoz-Monfort and Magrath, 2021). Literature on nature tech integration is largely restricted to examples of combining passive acoustic monitoring and camera traps (Buxton *et al.*, 2018; Rich *et al.*, 2019; Growcott *et al.*, 2025), with sparser examples integrating other nature technology combinations (Rappaport *et al.*, 2022; Dixon, Baker and Ellis, 2023; Scarpelli *et al.*, 2023). Based on relevant literature, a review of reported projects using the NICFI Satellite Data Program, and expert input from conference sessions with nature technologists (*NTC*, 2025), we identified six archetypes or **paradigms** for integrating multiple nature technologies (Figure 1):

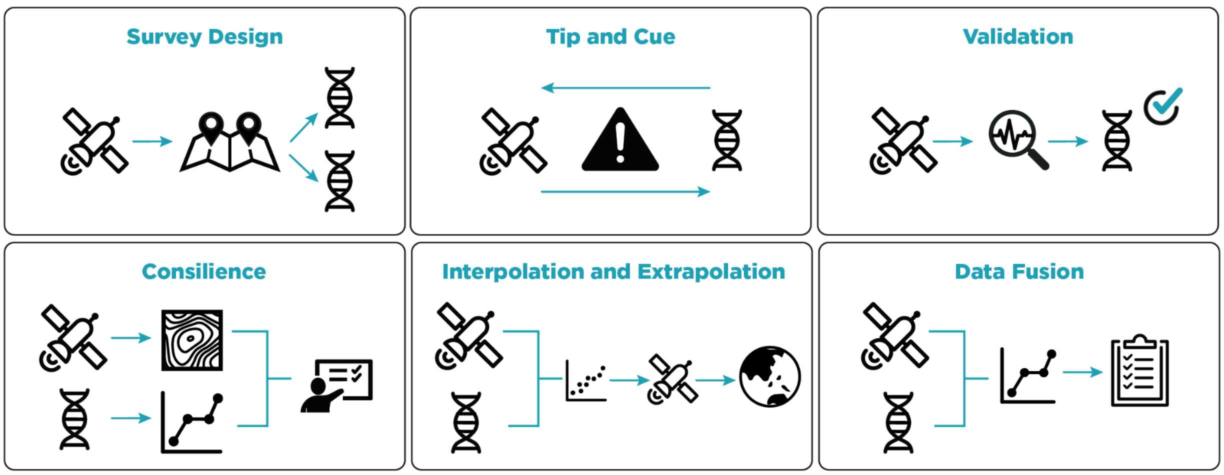

Figure 1. Schematic of paradigms for nature technology integration

Survey Design. In Survey Design, researchers draw on direct or derived data from one sensor to develop their sampling strategy for another type of sensor (Dalton *et al.*, 2024). The purpose is to identify where, when, and sometimes which nature technology to sample to detect a target species presence or behavior. The first sensor helps to delineate and limit the spatial and/or temporal use of the other sensor(s). For example, the site sampling strategy for a camera trap survey could be determined based on satellite imagery-derived habitat mapping.

Tip and Cue. Tip and Cue emerges from the field of remote sensing, where one sensor indicates a possible event of interest, and another satellite, typically with higher spatial resolution, confirms the event (Goana *et al.*, 2024). For conservation, Tip and Cue entails data from one nature technology identifying or alerting users to significant ecological change, pointing to the need for a different technology or sensor to further characterize or confirm the change or its implications for species or ecosystem conditions. The purpose is to provide a warning or indicator of a change in time to react to it, often by taking a management or enforcement action. For example, a satellite monitoring large areas of forest detects evidence of large-tree mortality and using the “tip” from the satellite data, field teams then “cue” an evaluation of already-deployed *in-situ* passive acoustic sensors to detect whether this forest disturbance has affected presence or behavior of bird species and re-assess their conservation status or take management actions.

Validation. Validation occurs when data from one type of nature technology is used to verify the findings of another. The first technology senses a proxy or delivers a low-confidence finding, while the second senses more directly the object of interest for a confirmatory identification or result. The purpose is to build confidence in a finding, such as a species detection or ecological change. For example, satellite data indicate the presence of wildebeest herbivory and eDNA or camera traps confirm wildebeest presence. Alternatively, an eDNA sample could indicate the presence of an endangered species in a site, which is verified with camera trap images or acoustic-derived sound identification.

Consilience. Consilience is when data from multiple nature technologies are analyzed independently. Then the outputs are manually compared to derive new conceptual insights, but the data are never directly integrated to produce wholly new metrics. As an example, bioacoustic data are used to produce findings on bird populations and eDNA is used to derive metrics on aquatic fauna independently; the concurrent increasing trends in both species groups are used together to indicate overall ecosystem health improvement.

Extrapolation and Interpolation. With Extrapolation and Interpolation, a feature in an extensive nature technology dataset correlates with a spatially and/or temporally narrower dataset, and the extensive dataset is leveraged to expand findings from the narrower dataset (Anderson, 2018). This correlation typically uses computational modeling to either fill gaps to create spatially or temporally continuous datasets within the area or timeframe where sparse data have been collected (i.e., interpolate) or expand the spatial or temporal extent to new areas or timeframes (i.e., extrapolate). The purpose of this paradigm is to expand results beyond the range of a narrow, but high-confidence nature technology by correlating with another technology of broader extent. For example, establishing a relationship between satellite-derived environmental variables and site-derived species occurrence to estimate species distribution in areas without *in-situ* sampling.

Data Fusion. Data Fusion uses data from two or more nature technologies, delivered concurrently and synthesized into a single metric or dataset supporting decision-making, and/or used as an input to further analysis. Data Fusion can produce an index or composite metric for decision-making or comparison across space and time, which could not be derived from a single technology (Isaac *et al.*, 2020; Czúcz *et al.*, 2021; Trantas *et al.*, 2025; Yu *et al.*, 2025). A key differentiating factor here is that a new data layer is developed and used as a unit, requiring data from multiple nature technologies to produce and update it. One example of this is producing an ecosystem condition index incorporating both ecosystem structure metrics from satellite data (e.g., canopy height) and compositional intactness derived from *in-situ* sensors. This index is then incorporated into climate and development scenarios to assess how ecosystem condition might change in the future and applied in a company's sustainability reporting.

Although these paradigms can overlap or be interdependent, they are often applied for different conservation purposes and help to indicate the guidelines and principles with special relevance for nature technology data integration.

3. Principles for an integrated nature data future

Three sets of principles are commonly used by the academic and nature technology sectors: FAIR (Wilkinson *et al.*, 2016), CARE (Carroll *et al.*, 2020), and TRUST (Lin *et al.*, 2020). These complementary principles apply across the process of data collection, management, and reuse. As noted by (O'Brien *et al.*, 2024), "FAIR is focused mainly on data clarification and reusability plus simplified retrieval, while TRUST stresses repository operational issues like transparency and sustainability – tenets designed to promote global data openness and awareness. CARE extends awareness to promote equitable participation and outcomes from said data access, use, reuse, or attribution, and emphasizes that the level of openness and accessibility be aligned with Indigenous rights."

While FAIR, CARE, and TRUST address critical aspects of modern scientific data management and governance, they do not include all aspects relevant for integrators of diverse, heterogeneous scientific data, such as the quality, reproducibility, and fit-for-purpose design of data, models, software, and scientific workflows (Royaux *et al.*, 2025; Trochim and Roy, 2025; Wilkinson *et al.*, 2025). To address this, we combined expert opinion with a review of FAIR, CARE, and TRUST principles to identify gaps and areas of special relevance for nature data integration (see Supplementary Material 1).

We found two existing data principles hold special relevance for nature tech integration: metadata standards for data interoperability and interpretation, and use rights and licensing for data access and reuse during integration (see Table S1.2). Solicited experts suggested multiple metadata standards are important for data transfer and machine-readability, particularly as integrated data are reused, become part of derived datasets, or incorporated into integrated global databases. For example, standard annotation using Darwin Core helps maintain consistency and interoperability across species observation data in the Global Biodiversity Information Facility (*GBIF*, 2025), and common standards are emerging for many individual nature tech paradigms. An understanding of licensing terms is also critical when creating derivative datasets, connecting tools, or reusing integrated data and products; for instance, the use of "copyleft" products in workflows requires all derivative products to inherit the same license (de Laat, 2005).

Data integration principles that were consistently underemphasized in existing frameworks include scientific rigor and pragmatism, a core element of utility (Table S1.1; Figure S1.1). Experts indicated that scientific rigor in data integration includes explicitly documenting nature tech data collection methods and study design, clearly communicating appropriate and inappropriate (re)uses of findings based on those methods, and estimating uncertainty in findings. According to experts solicited, utility ensures that data integration is well-suited to policy and management decisions by ensuring access to data and findings for decision-makers, tailoring data integrations to respond to decision needs, and providing guidance and support for responsibly applying data integration to decisions under uncertainty.

The existence of explicit principles does not immediately translate into effective real-world implementation. For example, for FAIR, the Interoperability and Reusability principles have been particularly challenging to achieve, relative to Findability and Accessibility (Jones, Slaughter and Habermann, 2019), in large part due to the slow adoption and expansion of community-endorsed metadata standards(Bagstad *et al.*, 2025). We reviewed how these four broad principles – scientific rigor, utility, metadata, and reuse – have been implemented in the field, in seven real-world case studies, alongside our proposed paradigms (Table S1.2).

4. Case studies

We assessed seven real-world case studies to understand the state of nature data integration for biodiversity monitoring and conservation. These examples represent diverse regions, users, biomes, use cases, and technologies. We compared these experiences with our proposed paradigms and principles (Figure 2).

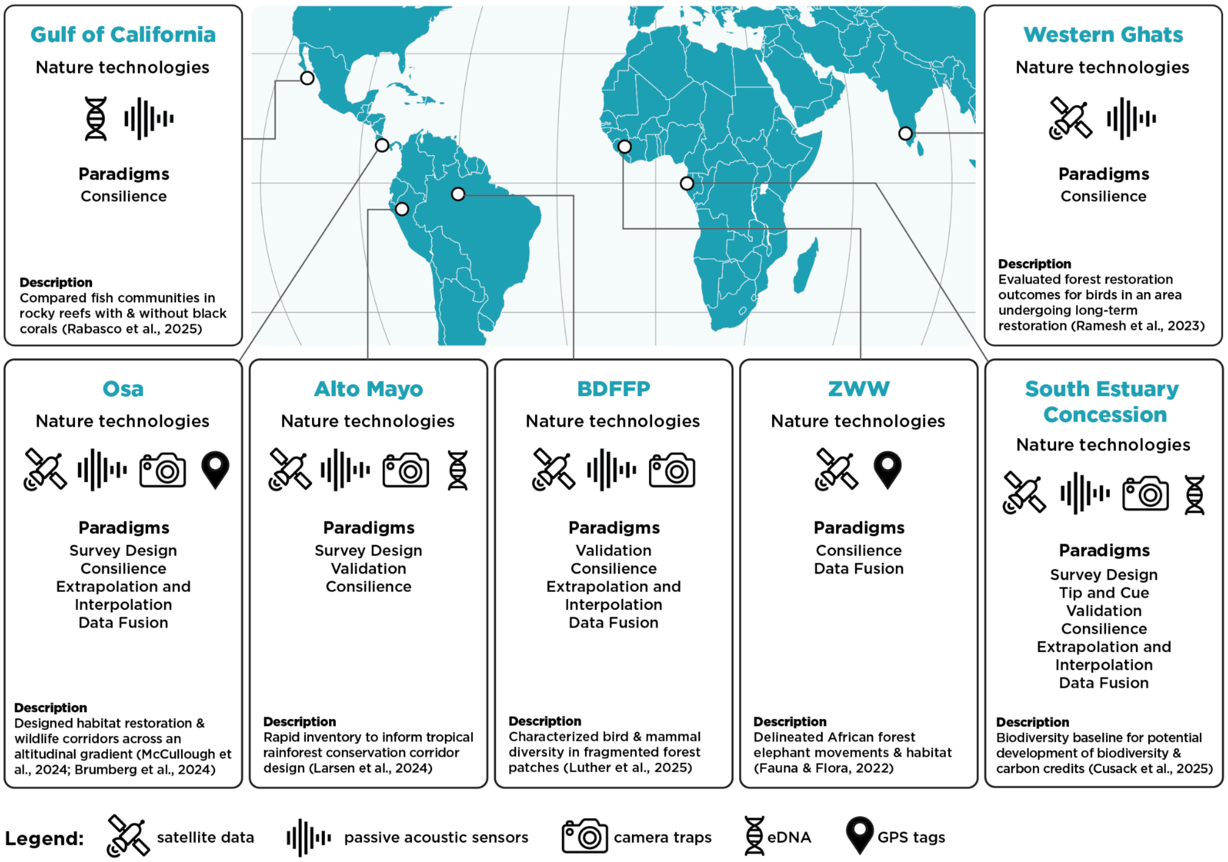

Figure 2: Nature technologies and paradigms for data integration mapped to real-world case studies. Supplement 2 provides detailed case study descriptions, and Supplement 3 offers a comprehensive mapping of case study activities to paradigms.

5. Paradigms and Principles in Action: Case Studies in Nature Technology Integration

5.1 Paradigms in case studies of nature technology integration

We propose that by understanding the paradigmatic approaches to integrating nature technology and data in real-world case studies, rather than theoretical or one-off explorations, we can better identify barriers and opportunities to replicate, scale, and automate timely, robust applications for conservation. Analyzing these paradigms through the lens of real-world case studies provides an indication of either the maturity or novelty of integration approaches, and where there are needs for new tools and guidance to scale, repeat, and build confidence for practical application in business and policy. Table 1 summarizes the barriers and opportunities identified in the real-world case studies through the lens of these paradigms.

Table 1: Barriers and opportunities for nature technology integration by paradigm, as identified by case studies

Barrier	Opportunity	Paradigm(s)
Mismatches in timing, scale and resolution between datasets produced from different technologies	Higher spatial, temporal, and spectral resolution satellite remote sensing datasets available at broad scale, for integration with in-situ technologies	• Consilience • Interpolation and Extrapolation • Data Fusion • Survey Design
Lack of widespread, co-occurring or densely sampled nature technology datasets	Increasing spread and duration of in-situ nature technologies as they continue to decrease in cost; developing and testing more AI models and analytic platforms for big-data analysis of growing datasets	• Interpolation and extrapolation • Tip and Cue • Data Fusion
Lack of methodological guidance and tools for operational data integration and conservation action	Systematizing guidance on workflows and models for conservation questions and goals based on documented case studies and peer review	• Interpolation and extrapolation • Tip and Cue • Data Fusion
Absence of reference libraries for data comparison and emerging AI models associated with lower confidence and higher uncertainty, making it hard to incorporate in operating procedures	Building comprehensive reference libraries, guidance and tools to convey and manage uncertainty, and applying multiple nature technologies for confirmatory results	• Validation • Tip and cue
Manual or complex workflows inhibit the use of integrated nature technologies for timely intervention	Automating workflows that integrate signals from multiple nature technologies to inform decision making and action (e.g., change alerts)	• Data Fusion • Tip and cue

Our seven case studies applied different paradigms to achieve their purpose, and many adopted multiple paradigms (Figure 2).Consilience, the most-used paradigm applied by all seven case studies, most often consisted of the exploration of multiple nature technologies to identify relationships among them (e.g., BDFFP, Gulf of California, Osa); in many cases this meant assessing satellite-based spatial datasets alongside point-based data from *in-situ* sensors. In three cases (ZWW, Osa, Western Ghats), consilience led to delineation of high-quality or active habitat by comparing land cover or ecosystem structure, species movement, and/or species observations. In two others (Alto Mayo, South Estuary Concession), it contributed to the characterization of species diversity and community composition, identifying high conservation value sites. This paradigm presents the lowest-barrier approach, using side-by-side comparison of datasets without systematic or complex integration methods. However, even for this paradigm, barriers were identified for real-world application, including insufficient data, mismatches in timing, scale, or resolution of datasets, and inadequate guidance on summarizing data for intercomparison. The case studies indicate that Consilience will be a primary way biodiversity insights are derived, particularly as in-situ nature technologies proliferate, more studies are published, and satellite remote sensing datasets increase resolution (spatially, temporally, and spectrally).

Consilience also appears to serve as an entry point for more sophisticated data integration. In several cases, Consilience preceded other data integration paradigms, defining relationships that were then be operationalized with Interpolation and Extrapolation or Data Fusion; for example in BDFFP, researchers explored remotely sensed forest structure metrics alongside site-derived mean species occupancy before conducting quantitative interpolation from those established relationships across unsampled areas.

In addition to BDFFP, two case studies (South Estuary Concession, Osa) applied Interpolation and Extrapolation. South Estuary Concession interpolated a spatially continuous biodiversity layer using data from multiple sensors *in situ*. In Osa, nature technology-derived habitat connectivity models are being extrapolated both *spatially* to novel landscapes and *temporally* to predict wildlife habitat preferences in future climate scenarios. In these cases, the purpose of Interpolation and Extrapolation was to fill in data gaps and create an extensive spatial layer beyond the placement of site technologies. A barrier to adopting this paradigm was identified by multiple case studies: a lack of useful tools and clear methodological guidance. There is an opportunity to a systematize guidance on effective workflows and models based on documented case studies and published studies, such as the commonly identified need of extrapolating species diversity and distribution based on ecosystem characteristics, such as structure or floral community composition. Several case studies expressed a strong interest in pursuing this paradigm, but noted that in-situ data were not sampled densely or widely enough to facilitate high-confidence interpolation or extrapolation, pointing to an opportunity for the growth of this paradigm as these technologies decrease in cost and complexity.

Three cases (South Estuary Concession, Osa, ZWW) used Data Fusion, built on Consilience and Extrapolation by joining together nature tech data in a new product that was then used in follow-on analyses or decision-making frameworks, absent the original individual datasets. After extrapolating habitat preferences in Osa, researchers created a new connectivity surface layer to analyze "pinch points" where animal movement is being constrained (McCullough *et al.*, 2024). In ZWW, satellite and GPS tag data were fused to produce a resource-use probability map that was leveraged in a new analysis to understand population-specific habitat preferences for forest elephants. Data fusion seemed to be employed when nature technologies provided complementary data that contributed to an overall indicator; for example, the creation of a multispecies occupancy model that leverages the ability of acoustics to detect vocalizing birds throughout the canopy and the ability of camera traps to detect non-vocalizing mammals on the forest floor. New AI approaches present opportunities for integrating multiple nature technologies via the Data Fusion paradigm (Sastry *et al.*, 2024; Trantas *et al.*, 2025), though, similarly to Consilience, input datasets with coarse spatial and temporal resolution remain a barrier. As with Interpolation and Extrapolation, denser sampling of validated in-situ data inputs and more comprehensive reference libraries will increase the confidence of layers produced by Data Fusion.

With growing use of technology and AI models, Validation was identified by case studies as being increasingly important to confirm the accuracy of species status and trends. Validation was used in three case studies. eDNA findings were verified with passive acoustic monitoring detections (Alto Mayo) and camera trap identifications (South Estuary Concession). The purpose was to confirm a lower-confidence indication in the eDNA data with higher-confidence, location and time-stamped camera trap or acoustic data. BDFFP, in contrast, researchers used this paradigm to confirm a satellite-based proxy for species richness (i.e., structural diversity) with more direct observations from camera traps and acoustics.

Distinct in terms of workflows are the Survey Design and Tip-and-cue paradigms. Three case studies used Survey Design (Atlo Mayo, Osa, South Estuary Concession), primarily by developing a map using satellite data that informed sampling design with other technologies *in situ*, such as passive acoustic sensors, camera traps, and eDNA. in some cases, satellite data were only one input among many to inform site surveying; in one GPS tags were used and in another expert input supplemented satellite data. The purpose tended to be to ensure robust survey design with representative sampling. Opportunities identified by case studies included that higher temporal and spatial resolution maps of landcover, ecosystem condition, or habitat suitability would improve the application of nature technology to Survey Design.

By contrast, only one case study employed Tip and Cue. In the South Estuary Concession, eDNA data identified threatened species, providing a "tip" that "cued" deployment of a new set of camera traps and acoustic sensors targeting species for further management recommendations. Other case studies indicated barriers to applying this paradigm, including lack of access to the desired technologies and an absence of easily applied methodologies or automated workflows. The example from South Estuary Concession described in the case was manual, and the "tip" was detected in data from a nature technology not explicitly established for that purpose; the cued response was an opportunistic response activity. An opportunity for additional research and development includes use of automated "tip" alerts set up specifically to search for certain phenomena, which then mechanistically "cue" additional monitoring or enforcement actions, or the use of multiple sensor types to monitor savanna elephants and rhinos (Dorfling *et al.*, no date).

5.2 Principles in case studies of nature technology integration

Case studies varied in their application of four principles: metadata standards, interoperability, and machine-readability; use rights and licensing for access and reuse of data integrations; scientific rigor in design, documentation, and code; and utility of data integration for real-world decision-making. We found strong operationalization in many areas. Almost all case studies included core metadata fields and archive data and metadata in machine-readable formats. Most cases had clear licensing and use-rights communicated with datasets. All followed established methods for collecting and integrating nature technologies and ensured guidance from scientists and tech specialists. Some cases, like Western Ghats, BDFFP, and Osa, exhibited very detailed documentation on methods and algorithms, enabling responsible reuse, repeatability, and application to conservation decisions. Most case studies were designed for uptake in management or policy decisions and demonstrated active use of findings; two had planned future use, and only one had no clear trajectory for uptake in decision-making.

In some case studies, implementation of data integration principles was not uniform. For metadata, only two of seven case studies (BDFFP and South Estuary Concession) were fully annotated using Darwin Core to enable seamless integration into GBIF and other conservation datasets and databases. Two more were partially aligned (Osa and Western Ghats), and the remainder did not use Darwin Core, though all could be annotated later. Access for reuse was not standardized: three out of seven case studies had internal, proprietary data archiving; in two cases, this was related to data collection for clients and inclusion of sensitive species or locations. Four case studies either archived in public-facing repositories with open licenses or intended to. Sometimes code was shared for AI models and statistics (e.g., via Zenodo for the Gulf of California and Western Ghats, see Supplement 2), but not by most. Although most case studies thoughtfully addressed metadata and scientific rigor, it was not clear that they all effectively communicated constraints on interpreting and reusing resulting data and findings (e.g., limitations on calculating species richness, presence, or movement due to sensor placement, battery life, or technical specifications).

Scientific rigor was clearly considered in the design and use of nature technologies. We did not find evidence for concern that the introduction of nature technologies would reduce scientific rigor. Data access for reuse remains murky despite increasing adoption of licensing and archiving standards. Reuse patterns among the case studies ran the gamut from closed, private storage to re-sharing to public platforms. Importantly, however, two cases that had proprietary internal archiving allow access to data upon request, with documented secondary uses for conservation status assessment.

Looking forward, we support the continued evolution and more widespread use of metadata and archiving standards for newer data formats, like video, derivative datasets, and algorithms, cloud-optimized file formats, and multimodal repositories. It is also valuable to clearly communicate the methodology and purpose of data collection and implications for appropriate reuse (Trochim and Roy, 2025). Opportunities remain for setting community-wide agreements on ethically sharing key data while growing the sector through commercial efforts, reducing barriers to access and capacity in the Global South, and engaging with local and Indigenous knowledge holders.

6. Conclusion and Future Directions

The need for integrated biodiversity monitoring arises from a simple but persistent challenge: no single nature technology can capture the complexity of ecosystems or biodiversity. When used in multi-sensor combinations and with traditional field observations, they enable richer understanding, assessment of trends, and predictive capacity – beyond what any single source offers, and more aligned with the needs of local management, global policy, and corporate reporting. Both the literature and our case studies indicate that integrated data can deliver robust information decision-makers and conservationists seek: locally precise, yet scalable over space and time.

Realizing the potential of these data will depend not only on technological advances, but also on the collective commitment of the broader community to refine and build upon paradigms and principles for integration that informs decisions and permits thoughtful reuse (Ramage, 2021). Distilling data integration into paradigms helped us map the relationships between workflows, such as how low-barrier-to-entry methods (like Consilience) can lead to deeper integration (Extrapolation or Data Fusion) and that there are different purposes and tools needed, even when sensor use is similar. The paradigms also helped elucidate key challenges (such as mismatches in scale and lack of data, reference libraries, and methodological guidance) and opportunities (increasing data resolution, systematizing guidance, and producing new repeatable tools). This kind of integration requires more than interoperability of tools; it depends on alignment across institutions and sectors, convergence on common standards, open or shared infrastructure, and mutual trust.

Dialogue aimed at consensus guidelines, shared digital infrastructure and interoperable semantics is an important step for this evolving sector, which includes stakeholders in civil society, the private sector, philanthropy and government, with different incentives, needs, and preferences. If we can bring these elements together—linking diverse datasets, harmonizing methods, and fostering collaboration—the payoff will be profound. Integrated monitoring could enable high-confidence global assessments, actionable early warning systems, instructive forecasting and management tools, and robust, repeatable metrics to underpin emerging nature finance mechanisms. The principles and paradigms outlined here are intended to provide a foundation for realizing integrated biodiversity monitoring as a framework for all stakeholders to help guide research and action.

Acknowledgements
We are grateful to many thought partners who joined our live discussion sessions at the Nature Tech Collective Unconference and virtually, as well as the teams that contributed to data collection and analysis in each case study.

For the Alto Mayo case study, funding was provided by BHP Foundation. For the BDFFP case study, funding was obtained via NSF DEB 2311046. For the South Estuary Concession case study, data collection, processing and analysis components were funded by Okala and Arise IIP.

For the Gulf of California case study, fieldwork was supported by several fellowships from CONAHCYT (2022-000002-01NACF), the FOCUS Research Unit of the University of Liège (R.CFRA.4142-J-F-C), and UABCS (INV-IN072). We thank all the other researchers involved in the original study: Florian Rabasco, Adrian Munguia-Vega, Eric Parmentier, Luis Hernández, Leonardo Huato-Soberanis, Héctor Reyes-Bonilla, Carlos Sánchez-Ortiz, and Damien Olivier. XR is currently funded by a BAEF-WBI fellowship, and part of his research project is supported by a Rotary International grant.

Osa Conservation's landscape connectivity work is supported by the Gordon and Betty Moore Foundation, Biome Conservation, the Bobolink Foundation, the Krystyna and Dan Houser Foundation, the SENSE Foundation Brussels, Troper Wojcicki Philanthropies, KHR McNeely Family Fund, and the Mazar Family Charitable Foundation Trust.

For the Fauna & Flora case study, the elephant collaring was funded by the African Forest Elephant Foundation (AFEF) and individual donors (311-IND). We would like to acknowledge the contribution of Fauna & Flora's Guinea, Africa and Conservation Technology teams, veterinarians, photographers, Centre Forestier de N'Zerekore rangers and the Forestry Development Agency (FDA) of Liberia who contributed to this work. Special thanks to the Ministere de L'Environment, Des Eaux et Forets – Centre Forestier de N'Zerekore that granted authorization to collar forest elephants in Guinea.

For the Western Ghat's case study, this work was possible through a Rufford Foundation Small Grant (28916-1) awarded to Vijay Ramesh. VR received extensive support and expertise of Dr. Divya Mudappa and Dr. T R Shankar Raman at the Nature Conservation Foundation (NCF). Divya and Shankar Raman's efforts to actively restore degraded forest fragments in the Valparai plateau over the last three decades was instrumental in laying the foundation for this study.

References

Al-Fuqaha, A. *et al.* (2015) "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE Communications Surveys & Tutorials*, 17(4), pp. 2347–2376. Available at: https://doi.org/10.1109/COMST.2015.2444095.

Anderson, C. (2018) "Biodiversity monitoring, earth observations and the ecology of scale." Available at: https://doi.org/10.1111/ele.13106.

Anderson, C.B. (2025) "Designing Multi-Modal Ecosystem Monitoring Technologies: A Network of Networks Approach." Available at: https://ecoevorxiv.org/repository/view/9029/ (Accessed: October 30, 2025).

Bagstad, K.J. *et al.* (2025) "Interoperability for ecosystem service assessments: Why, how, who, and for whom?," *Ecosystem Services*, 72, p. 101705. Available at: https://doi.org/10.1016/j.ecoser.2025.101705.

Bell, K. and Malerba, M.E. (2025) "Biodiversity monitoring for biocredits: a case study comparing acoustic, eDNA, and traditional methods," *Biodiversity and Conservation*, 34(7), pp. 2513–2528. Available at: https://doi.org/10.1007/s10531-025-03083-0.

Besson, M. *et al.* (2022) “Towards the fully automated monitoring of ecological communities,” *Ecology Letters*, 25(12), pp. 2753–2775. Available at: https://doi.org/10.1111/ele.14123.

Buxton, R.T. *et al.* (2018) “Pairing camera traps and acoustic recorders to monitor the ecological impact of human disturbance,” *Global Ecology and Conservation*, 16, p. e00493. Available at: https://doi.org/10.1016/j.gecco.2018.e00493.

Cai, Q. *et al.* (2019) “A Survey on Multimodal Data-Driven Smart Healthcare Systems: Approaches and Applications,” *IEEE Access*, 7, pp. 133583–133599. Available at: https://doi.org/10.1109/ACCESS.2019.2941419.

Carroll, S.R. *et al.* (2020) “The CARE Principles for Indigenous Data Governance,” *Data Science Journal*, 19(1). Available at: https://doi.org/10.5334/dsj-2020-043.

CBD (2022) “15/4. Kunming-Montreal Global Biodiversity Framework.”

Conservation Evidence - Site (no date) *Conservation Evidence*. Available at: https://www.conservationevidence.com/ (Accessed: October 28, 2025).

Czúcz, B. *et al.* (2021) “Selection criteria for ecosystem condition indicators,” *Ecological Indicators*, 133, p. 108376. Available at: https://doi.org/10.1016/j.ecolind.2021.108376.

Dalton, D. *et al.* (2024) *A framework for monitoring biodiversity in protected areas and other effective area-based conservation measures: concepts, methods and technologies*. 1st ed. IUCN, International Union for Conservation of Nature. Available at: https://doi.org/10.2305/HRAP7908.

Dixon, A.P., Baker, M.E. and Ellis, E.C. (2023) “Passive monitoring of avian habitat on working lands,” *Ecological Applications*, 33(5), p. e2860. Available at: https://doi.org/10.1002/eap.2860.

Dorfling, J. *et al.* (no date) “Satellite, Aerial, and Ground Sensor Fusion Experiment for Management of Elephants and Rhinos and Poaching Prevention,” in *AIAA SCITECH 2022 Forum*. American Institute of Aeronautics and Astronautics. Available at: https://doi.org/10.2514/6.2022-1270.

GBIF (2025). Available at: https://www.gbif.org/ (Accessed: October 30, 2025).

Goana, M. *et al.* (2024) “Tool to Generate Deforestation and Illegal Mining Alerts with Remote Sensing,” *Environmental Sciences Proceedings* [Preprint]. Available at: https://www.mdpi.com/2673-4931/28/1/27 (Accessed: October 30, 2025).

Growcott, J. *et al.* (2025) “The secret acoustic world of leopards: A paired camera trap and bioacoustics survey facilitates individual identification of leopards via their roars,” *Remote Sensing in Ecology and Conservation*, 11(4), pp. 361–374. Available at: https://doi.org/10.1002/rse2.429.

Hartig, F. *et al.* (2024) “Novel community data in ecology-properties and prospects,” *Trends in Ecology & Evolution*, 39(3), pp. 280–293. Available at: https://doi.org/10.1016/j.tree.2023.09.017.

Henry, P.-Y. *et al.* (2008) “Integrating ongoing biodiversity monitoring: potential benefits and methods,” *Biodiversity and Conservation*, 17(14), pp. 3357–3382. Available at: https://doi.org/10.1007/s10531-008-9417-1.

Isaac, N.J.B. *et al.* (2020) “Data Integration for Large-Scale Models of Species Distributions,” *Trends in Ecology & Evolution*, 35(1), pp. 56–67. Available at: https://doi.org/10.1016/j.tree.2019.08.006.

IUCN (2025) *IUCN Red List of Threatened Species*. Available at: https://www.iucnredlist.org/en (Accessed: October 30, 2025).

Jones, M.B., Slaughter, P. and Habermann, T. (2019) “Quantifying FAIR: automated metadata improvement and guidance in the DataONE repository network,” 14 September. Available at: https://doi.org/10.5281/zenodo.3408466.

King, R.A. and Halpern, B.S. (2025) “Implementation of automated biodiversity monitoring lags behind its potential,” *Environmental Research Letters*, 20(6), p. 064022. Available at: https://doi.org/10.1088/1748-9326/add02d.

de Laat, P.B. (2005) “Copyright or copyleft?: An analysis of property regimes for software development,” *Research Policy*, 34(10), pp. 1511–1532. Available at: https://doi.org/10.1016/j.respol.2005.07.003.

Lahoz-Monfort, J.J. and Magrath, M.J.L. (2021) “A Comprehensive Overview of Technologies for Species and Habitat Monitoring and Conservation,” *BioScience*, 71(10), pp. 1038–1062. Available at: https://doi.org/10.1093/biosci/biab073.

Lin, D. *et al.* (2020) “The TRUST Principles for digital repositories,” *Scientific Data*, 7(1), p. 144. Available at: https://doi.org/10.1038/s41597-020-0486-7.

Marvin, D.C. *et al.* (2016) “Integrating technologies for scalable ecology and conservation,” *Global Ecology and Conservation*, 7, pp. 262–275. Available at: https://doi.org/10.1016/j.gecco.2016.07.002.

McCullough, I.M. *et al.* (2024) “Integrating high-resolution remote sensing and empirical wildlife detection data for climate-resilient corridors across tropical elevational gradients,” *Biological Conservation*, 298, p. 110763. Available at: https://doi.org/10.1016/j.biocon.2024.110763.

McKenzie, E.J. *et al.* (2025) “Science and technical priorities for private sector action to address biodiversity loss,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, 380(1917), p. 20230208. Available at: https://doi.org/10.1098/rstb.2023.0208.

Miller, B.L. *et al.* (2025) “Corporate Biodiversity Reporting Can Be Scaled With AI and Earth Observation—But Will Miss the Point Without Guidance From Conservation Scientists,” *Conservation Letters*, 18(5), p. e13153. Available at: https://doi.org/10.1111/conl.13153.

NPI (2025) *https://www.naturepositive.org/*. Available at: https://www.naturepositive.org/ (Accessed: October 30, 2025).

NTC (2025) *Nature Tech Collective*. Available at: https://www.naturetechcollective.org/nature-tech-unconference (Accessed: October 30, 2025).

O’Brien, M. *et al.* (2024) “Earth Science Data Repositories: Implementing the CARE Principles,” *Data Science Journal*, 23, p. 37. Available at: https://doi.org/10.5334/dsj-2024-037.

Ramage, S. (2021) “Why people are essential in data interoperability.” Available at: https://www.data4sdgs.org/blog/why-people-are-essential-data-interoperability (Accessed: October 30, 2025).

Rappaport, D.I. *et al.* (2022) “Animal soundscapes reveal key markers of Amazon forest degradation from fire and logging,” *Proceedings of the National Academy of Sciences of the United States of America*, 119(18), p. e2102878119. Available at: https://doi.org/10.1073/pnas.2102878119.

Rich, L.N. *et al.* (2019) “Acoustic and camera surveys inform models of current and future vertebrate distributions in a changing desert ecosystem,” *Diversity and Distributions*, 25(9), pp. 1441–1456. Available at: https://doi.org/10.1111/ddi.12952.

Royaux, C. *et al.* (2025) “Guidance framework to apply best practices in ecological data analysis: lessons learned from building Galaxy-Ecology,” *GigaScience*, 14, p. giae122. Available at: https://doi.org/10.1093/gigascience/giae122.

Sastry, S. *et al.* (2024) “TaxaBind: A Unified Embedding Space for Ecological Applications.” arXiv. Available at: https://doi.org/10.48550/arXiv.2411.00683.

Scarpelli, M.D.A. *et al.* (2023) “Spatial dynamics of soundscapes and biodiversity in a semi-arid landscape,” *Landscape Ecology*, 38(2), pp. 463–478. Available at: https://doi.org/10.1007/s10980-022-01568-9.

Sutherland, W. *et al.* (2025) “Nine changes needed to deliver a radical transformation in biodiversity measurement.” Available at: https://ecoevorxiv.org/repository/view/10320/ (Accessed: October 30, 2025).

TNFD, E. (2025) *TNFD*. Available at: https://tnfd.global/publication/recommendations-of-the-taskforce-on-nature-related-financial-disclosures/ (Accessed: October 30, 2025).

Trantas, A. *et al.* (2025) “BioAnalyst: A Foundation Model for Biodiversity.” arXiv. Available at: https://doi.org/10.48550/arXiv.2507.09080.

Trochim, E. and Roy, S. (2025) “Candid Core Framework.” Zenodo. Available at: https://doi.org/10.5281/zenodo.15265916.

Wilkinson, M.D. *et al.* (2016) “The FAIR Guiding Principles for scientific data management and stewardship,” *Scientific Data*, 3(1), p. 160018. Available at: https://doi.org/10.1038/sdata.2016.18.

Wilkinson, S.R. *et al.* (2025) “Applying the FAIR Principles to computational workflows,” *Scientific Data*, 12(1), p. 328. Available at: https://doi.org/10.1038/s41597-025-04451-9.

Xu, X. *et al.* (2024) “A Comprehensive Review on Synergy of Multi-Modal Data and AI Technologies in Medical Diagnosis,” *Bioengineering*, 11(3), p. 219. Available at: https://doi.org/10.3390/bioengineering11030219.

Yu, R. *et al.* (2025) “Foundation Models for Environmental Science: A Survey of Emerging Frontiers.” arXiv. Available at: https://doi.org/10.48550/arXiv.2504.04280.

Supplement 1: Principles

S1.1 Expert-derived principles and comparisons from the nature tech sector.

Our approach to review and develop principles for data integration was twofold. We first convened experts in the nature technology sector to define principles that are currently or should be used in practice (henceforth “data integration principles”; Table S1.1). We then cross-referenced these data integration principles against FAIR, CARE, and TRUST to identify existing principles needing renewed attention for data integration and gaps identified by the expert-derived principles.

We convened experts at the London Nature Tech Collective Unconference in March 2025 to identify a set of operational principles currently used for data integration by the nature technology sector. Experts represented a range of actors, including academics, nature conservation practitioners, nature technology companies, and actors in the corporate and finance sectors. Participants were asked to propose a set of principles during two round table sessions, which were modified and refined during two online follow-up sessions with additional experts not able to make the live sessions. This process yielded a set of 19 expert-derived principles (Table S1) covering five broad themes: ethical (do no harm), ethical (functional/economic), scientific rigour, technical maturity, and utility (i.e., fit for purpose design and use).

Table S1.1 List and description of data integration principles arrived at by the nature tech expert working group. Highlighted principles are gaps in existing principles deserving more attention for integrated data streams.

Expert-derived data integration principles	**Description**
Ethical (do no harm)	
Safeguarding of nature	Data creation and integration does not harm nature
Data sovereignty (ownership, governance, access rights)	Data creators and integrators retain ownership, governance and access over their data
Open versus commercial data	Data use pays heed to the intent and licensing of all data sources
Protection of IP	Intellectual property used to create and integrate data is respected and honoured
Ethical (functional/economic)	
Benefit sharing (i.e., value attribution)	Agreed value, monetary or otherwise, is passed back to data creators and integrators
System flexibility (around alternative knowledge systems)	Data collection and integration is able to incorporate non-western paradigms of science and data
Transparency (of methods, data)	Methods of data collection and integration are made available and intelligible to all

Auditability (i.e., chain of custody)	Data integration retains a formal chain of custody of data from end user to data source
Scientific rigour	
Objectivity (of methods)	Data collection and integration employ established scientific methods/approaches
Repeatability (of methods)	Methods for data collection and integration are documented in a way that is repeatable by anybody
Accuracy, precision (of data)	Accuracy assessments of data collection and integration are performed and reported on
Uncertainty (i.e., everything is an estimate)	Uncertainty is properly accounted for in data integration and appropriate uncertainty metrics are provided
Technical maturity	
Interoperability (of data types)	Data are delivered in formats that maximise compatibility between types (e.g., images, DNA samples, sound recordings)
Interoperability (of data systems)	Data are stored in systems that maximise compatibility between software and tools
Scalability	Data systems are built to support sufficient scale (e.g., analysis on the entire globe)
Immutability / Permanence of data	Data are stored and audited in perpetuity
Utility (fit for purpose design and use)	
Pragmatism (i.e., operate despite uncertainty)	Suboptimal methods for data collection or integration should still be undertaken if they can yield meaningful insight
Leveling of data sophistication to decision requirements	Data are collected and integrated for a clear, pre-defined, purpose, application, or decision
Accessibility and application (of data)	Data are available to all

We cross-referenced these expert-derived principles against FAIR, CARE, and TRUST principles to identify where they do and do not capture the unique challenges for leveraging integrated data streams into the future. We found existing principles shared the most overlap with "accessibility" in the Utility category, with 19 principles representing the numerous facets of accessibility needed for data integration (Fig. S1.1). Expert-derived principles in the Technical maturity, Ethical (do no harm), and Ethical (functional/economic) categories were also well-covered by existing principles, although the underrepresentation of the "safeguarding of nature" and "interoperability of data types" principles indicates a need for greater attention to these expert-derived principles when integrating data streams. Finally, the four expert-derived principles in the Scientific Rigor category and the "pragmatism" principle (under "Utility") lie outside the scope of the FAIR, CARE, and TRUST principles, requiring their added consideration during nature tech data integration applications (*Biodiversity Data Quality*, no date). These findings justified our targeted questionnaire for the case-study principles review (Table S1.2).

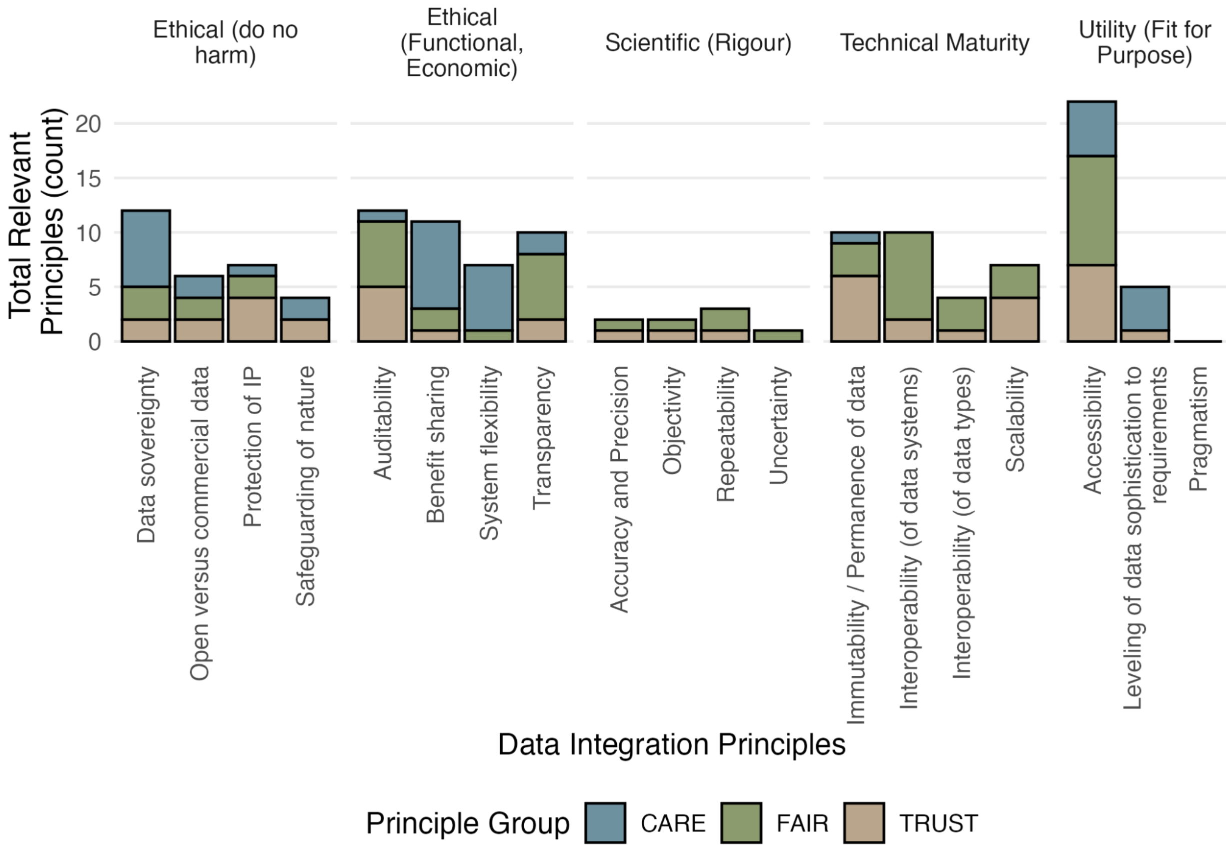

Figure S1.1 Summary of the crosswalk of expert-derived data integration principles with existing data principles. The count shows the total number of principles from CARE, FAIR, and/or TRUST aligned with the data integration principles identified by nature tech experts.

S1.2 Examining the representation of key data integration principles in case studies

The existence of relevant principles for data integration does not mean that their implementation is without challenges. Even within the FAIR Principles, advances in Findability and Accessibility remain uneven, hindered through non–machine-readable outputs, incomplete metadata, and reliance on proprietary standards. Interoperability and Reusability present more profound difficulties: diffusion of community-endorsed metadata standards across disciplines and, critically for nature technology, across modalities, is slow, and licensing restrictions complicate integration of derived datasets (Bagstad *et al.*, 2025). In addition to development and use of data standards for individual nature tech modalities (Bubnicki *et al.*, 2024; Darras *et al.*, 2025; Takahashi *et al.*, 2025), it is equally important that vocabularies and other semantic resources be developed to support interoperability across diverse data types and biodiversity-related scientific disciplines (Bagstad et al., 2025). Modality-specific data standards must account for various real-world complexities - for example, in aquatic bioacoustic studies, where it may be substantially more difficult to attribute a specific sound to a particular species than for terrestrial bioacoustics (Darras *et al.*, 2025). The CARE Principles are frequently cited yet inconsistently implemented, and mechanisms to secure Indigenous authority, benefit sharing, and cultural protocols must be carefully considered and implemented (Carroll *et al.*, 2020; Jennings *et al.*, 2025) ; including use of emerging standards ("Local Contexts – Grounding Indigenous Rights," no date). The TRUST principles highlight the importance of long-term data stewardship, but many biodiversity datasets remain tied to short-term projects without long-term sustainable custodianship (Lin *et al.*, 2020; O'Brien *et al.*, 2024).

Additionally, we note that as data collection and technology rapidly evolve, challenges arise in the uptake and application of principles. For example, the FAIR Principles were originally designed for data (Wilkinson *et al.*, 2016), but have since expanded to include research software (Lamprecht *et al.*, 2020; Barker *et al.*, 2022) and more complex computational workflows (Wilkinson *et al.*, 2025), reflecting critical differences between the characteristics and development and dissemination mechanisms of data and software. These FAIR principles for software and computational workflows, as well as best practices for coding (e.g., Royaux et al., 2025), are valuable to consider when developing open code for nature tech integration workflows. We anticipate similar adoption and evolution challenges with integration principles moving into the future, which is why setting a foundational understanding of current applications and gaps is important to this effort.

The case studies presented in this paper provide an opportunity to assess the implementation of existing principles in nature tech data integration and the presence of some of the aforementioned challenges. We identified four focal areas that warranted special consideration for integration of nature tech data streams during our crosswalk of principles: metadata, use rights, scientific rigor, and utility (Figure S1.1). These focal areas were selected because experts identified that data integration would pose challenges not fully considered by existing principles (metadata, use rights) or because existing principles overlooked key facets of data integration in these categories (scientific rigor, utility). To evaluate how well these principles are being used in practice, case study authors responded to a set of prompts for each category (Table S1.2). The responses were then discussed by experts on the basis of how well each adhered to the criteria for data integration principles. The synthesized results of this discussion form the basis of section 5.2 in the main text.

Finally, we cannot discuss case-study results without acknowledging the structural inequities compounding these issues, especially for nature tech data. In much of the Global South, limited access to hardware, connectivity, and analytical infrastructure prevents organizations from engaging in data integration, interoperability, and reuse at all, entrenching asymmetries in who generates, governs, and benefits from biodiversity data (Ortega, 2024; Doll, 2025). Few workflows explicitly assess whether their resolution, coverage, and uncertainty align with decision-making needs at site, landscape, or national scales. Sensitive contexts introduce further complexity: locations of endangered species, Indigenous knowledge, and enforcement data require governance aligned with the principle of "as open as possible, as closed as necessary." Addressing these constraints demands not only technical innovation but also institutional commitments to equity, transparency, and long-term sustainability.

Table S1.2: **Nature Tech Data Integration Principles: How are they operationalized in case studies?** Case studies documented how they addressed metadata, use rights, scientific rigor, and utility of their data integration, responding to a set of standard questions on case study implementation. Existing principles with special relevance for data integration are listed in parentheses for Metadata and Use Rights.

Principle Group	Specific Principles	Questions
Metadata	FAIR: I1, R1.3; CARE: A1; TRUST: TT1	• Are your nature tech data/metadata machine readable? • Are they aligned with Darwin Core fields or other metadata standards (and if so, which)? • Do they include methods, equipment and sensor type?
Use Rights	FAIR: F1, R1.1, R1.2; CARE: A3, E3; TRUST: R3, S3	• Is your nature tech data being archived, and if so, where? • Will it be made accessible for re-use? • Do your data and associated metadata have a clear data usage license?
Scientific Rigor	None	• How did you ensure scientific rigor in the design and use of nature tech in your case study? • How could your nature tech data streams be re-used with scientific rigor? And how should they *not* be used? • What tools or methods did you use to ensure scientific rigor for your data integration / paradigm in your case study? • How replicable is your nature tech data integration, and why?
Utility	None	• How is your nature tech data integration being used for specific conservation decisions or management actions right now? If not, why not?

S1.3 References

Bagstad, K.J. *et al.* (2025) "Interoperability for ecosystem service assessments: Why, how, who, and for whom?," *Ecosystem Services*, 72, p. 101705. Available at: https://doi.org/10.1016/j.ecoser.2025.101705.

Barker, M. *et al.* (2022) "Introducing the FAIR Principles for research software," *Scientific Data*, 9(1), p. 622. Available at: https://doi.org/10.1038/s41597-022-01710-x.

Biodiversity Data Quality (no date). Available at: https://www.tdwg.org/community/bdq/ (Accessed: October 30, 2025).

Bubnicki, J.W. *et al.* (2024) "Camtrap DP: an open standard for the FAIR exchange and archiving of camera trap data," *Remote Sensing in Ecology and Conservation*, 10(3), pp. 283–295. Available at: https://doi.org/10.1002/rse2.374.

Carroll, S.R. *et al.* (2020) "The CARE Principles for Indigenous Data Governance," *Data Science Journal*, 19(1). Available at: https://doi.org/10.5334/dsj-2020-043.

Darras, K.F.A. *et al.* (2025) "Worldwide Soundscapes: A Synthesis of Passive Acoustic Monitoring Across Realms," *Global Ecology and Biogeography*, 34(5), p. e70021. Available at: https://doi.org/10.1111/geb.70021.

Doll, P.C. (2025) "Addressing disparities in global biodiversity data," *Nature Reviews Biodiversity*, 1(8), pp. 495–495. Available at: https://doi.org/10.1038/s44358-025-00072-7.

Jennings, L. *et al.* (2025) "Governance of Indigenous data in open earth systems science," *Nature Communications*, 16(1), p. 572. Available at: https://doi.org/10.1038/s41467-024-53480-2.

Lamprecht, A.-L. *et al.* (2020) "Towards FAIR principles for research software," *Data Science*, 3(1), pp. 37–59. Available at: https://doi.org/10.3233/DS-190026.

Lin, D. *et al.* (2020) "The TRUST Principles for digital repositories," *Scientific Data*, 7(1), p. 144. Available at: https://doi.org/10.1038/s41597-020-0486-7.

"Local Contexts – Grounding Indigenous Rights" (no date). Available at: https://localcontexts.org/ (Accessed: October 30, 2025).

O'Brien, M. *et al.* (2024) "Earth Science Data Repositories: Implementing the CARE Principles," *Data Science Journal*, 23, p. 37. Available at: https://doi.org/10.5334/dsj-2024-037.

Ortega, R. (2024) "Genetic databases don't fairly capture species richness of Global South, new data show." Available at: https://www.science.org/content/article/countries-global-south-have-more-biodiversity-countries-north-databases-used-study (Accessed: October 31, 2025).

Takahashi, M. *et al.* (2025) "A Metadata Checklist and Data Formatting Guidelines to Make eDNA FAIR (Findable, Accessible, Interoperable, and Reusable)," *Environmental DNA*, 7(3), p. e70100. Available at: https://doi.org/10.1002/edn3.70100.

Wilkinson, M.D. *et al.* (2016) "The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data*, 3(1), p. 160018. Available at: https://doi.org/10.1038/sdata.2016.18.

Wilkinson, S.R. *et al.* (2025) "Applying the FAIR Principles to computational workflows," *Scientific Data*, 12(1), p. 328. Available at: https://doi.org/10.1038/s41597-025-04451-9.

Supplement 2: Case Study Background

The following supplement provides additional background of each of the case studies featured in this work. Case studies all responded to the following prompts:

<u>*Case study background/context*</u>:

- *Please summarize the following aspects of your case study:*
  - *Location*
  - *Project goal*
  - *Nature technologies (including remote sensing)*

<u>*Case study methods and results:*</u>

- *Please summarize at a high-level:*
  - *Procedures for data integration (data collection, data processing, modeling methodology)*
  - *Initial results*

Responses on location background, project goals, data sources, methods, and results are provided for each case study in this supplement.

For more detail on the mapping of each case study to the paradigms for nature technology integration, please see Supplement 3.

S2.1 - Alto Mayo

Location Background:
The Alto Mayo landscape is situated in the upper basin of the Mayo River, spanning approximately 780,700 hectares across the provinces of Moyobamba and Rioja in the San Martín region of northern Peru. This region lies at the interface of the Andes and Amazon, encompassing a diverse elevational gradient (570–2,230 meters above sea level) and a mosaic of ecosystems including white sand forests, montane swamps, and agricultural lands. The area is home to Indigenous Awajún communities and is recognized for its exceptional biodiversity, including numerous endemic and threatened species. Despite its ecological importance, the region faces intense deforestation pressures due to agricultural expansion, logging, and infrastructure development.

Project Goal:
The Alto Mayo RAP (Rapid Assessment Program), led by Conservation International and partners, aimed to address critical biodiversity data gaps in this under-surveyed yet highly threatened landscape. The primary objective was to generate spatially explicit biodiversity and ecosystem health data to inform the design of a conservation corridor linking the Alto Mayo Protected Forest with the Cordillera Escalera Regional Conservation Area. This corridor would enhance ecological connectivity, support sustainable land use planning, and safeguard ecosystem services vital to local communities. The project also sought to engage Indigenous communities in participatory research and to integrate multiple nature tech data streams—including in-situ surveys, remote sensing, and environmental DNA (eDNA)—to produce

actionable conservation insights at scale. As part of Conservation International's global RAP initiative, the Alto Mayo project reflects the program's emphasis on integrating diverse technologies and data streams—such as in-situ sensing, remote sensing, and manual observations—to rapidly generate comprehensive, decision-ready insights.

Data Sources:

- Nature Technologies:
  - Camera traps
  - Acoustic sensors
  - Environmental DNA (eDNA)
  - Satellite data
    - Landsat
    - PlanetScope
    - Digital elevation models
- Other data sources:
  - In-situ biological surveysStream Visual Assessment Protocol (SVAP)
  - Community knowledge and participatory mapping

Methods and Results:
Field surveys were conducted across eight ecologically distinct zones using standardized protocols tailored to each taxonomic group, including plants, vertebrates and select invertebrate groups. In parallel, 60 eDNA samples were collected from freshwater bodies and processed by NatureMetrics using metabarcoding to detect vertebrate species. To prioritize survey site selection, existing biodiversity data were extrapolated using MaxEnt (Maximum Entropy Modeling) to generate habitat suitability models for ecosystem types and species distributions. Data from in-situ surveys, eDNA, camera traps, acoustic sensors and remote sensing were integrated through spatial overlays and cross-validation. The RAP yielded a wealth of biodiversity insights, many of which represent new records or discoveries for the Alto Mayo landscape and Peru more broadly.

Data availability:
Final report, containing open data
https://www.conservation.org/docs/default-source/publication-pdfs/rap-73-alto-mayo-peru.pdf?sfvrsn=e66add1e_3

S2.2 - BDFFP

Location Background:
This case study takes place at the Biological Dynamics of Forest Fragments Project (BDFFP) in central Amazonia, ~80 km north of Manaus, Brazil. The BDFFP was initiated by Thomas Lovejoy in 1978. Originally called the Minimum Critical Size of Ecosystems Project, Dr. Lovejoy coordinated the experimental isolation of replicated 1, 10, and 100 hectare forest fragments, in addition to continuous forest (control) plots at BDFFP, to evaluate the effects of fragment size on rainforest biodiversity and ecological processes as well as assess the efficacy of Island Biogeography in mainland systems. Over time the project turned its focus from edge effects and fragment size, to landscape dynamics and forest regeneration and found that in both fragments of different sizes and mature continuous forest the

communities have changed over time, albeit more drastically in the fragments. Fragment isolation took place in the 1980s, and reisolated by clearing a 100m strip of vegetation around each fragment most recently in 2013 and 2014. In the 1990s the forest began regenerating secondary forest and the site now has 90% or greater forest cover. The site is composed of a collection of historic 1ha, 10ha, and 100ha forest fragments and adjoining secondary forest, ranging from 35 to 45 years old, of different ages and structures, all surrounded by continuous undisturbed *terra firme* forest (Rutt et al., 2017; Stouffer, 2020).

Project Goal:
The goal of the current case study is to investigate the habitat - heterogeneity hypothesis as a means of predicting biodiversity in tropical forests. More specifically we investigate the relationship of forest three-dimensional structure with vertebrate animal diversity, with the prediction that greater vertical forest structural complexity creates more niche spaces for a variety of species and results in a greater diversity of species.

Data Sources:

- Nature Technologies:
    - Camera traps
    - Passive acoustics sensors
    - Satellite Data:
        - The Global Ecosystem Dynamics Investigation (GEDI)

Methods and Results:
During our 4-month sampling period, we detected 201 bird and 35 mammal species, with audio devices identifying 196 species and cameras detecting 61 species. In total, our dataset represents 80% of the forest bird species and 69% of the 51 non-flying mammal species recorded at the BDFFP over the past 40 years. GEDI metrics reliably distinguished forest structure differences between primary, secondary and forest fragments. Mulit-species occupancy models with forest structural metrics as predictors revealed that GEDI metric Foliage Height Diversity (FHD) (a proxy for vertical structural heterogeneity) had the strongest association with animal diversity, with a two-fold increase in multi-species occupancy at sites with greater FHD, which provides support for the habitat heterogeneity hypothesis relationship to biodiversity in tropical forests. However, when species were modeled individually each species had a different forest metric that was a significant predictor of site occupancy, thus predictions of based on species diversity and individual species habitat preferences will be quite different depending on the metrics that are used and a nuanced approach will be needed depending on community or species specific research goals.

Data availability:
The metadata associated with acoustic and camera trap detections (in Darwin Core format), as well as the preprint of the corresponding scientific paper, can be accessed here
https://dashboard.wildmon.ai/project/bdffp

S2.3 - Gabon Estuary

Location Background:
The South Estuary Concession is a 270,000-hectare logging concession situated on the south side of the Gabon Estuary. Its rich variety of habitats, including secondary tropical forests, grassland savannahs, wetlands, and mangroves support a unique assemblage of species, from endemic killy fish to critically endangered forest elephants. The Concession acts as an important corridor between the Wonga Wongue National Park in the south and Pongara National Park in the north.

Project Goal:
Due to the low commercial viability of logging activities, the entity responsible for the management of South Estuary sought to value its natural capital, with a view to promoting sustainable revenue and development from the sale of biodiversity and carbon credits. Okala was hired to carry out an extensive biodiversity and carbon baseline, combining on-the-ground sensors (camera traps, bioacoustics sensors, and eDNA) and remote sensing (land-cover, biomass, and carbon mapping).

Data Streams:

- Nature Technologies:
    - Camera traps
    - Passive acoustic sensors
    - eDNA
    - Satellite data
        - Landcover
        - Biomass

Methods and Results:
Data integration was facilitated by deploying ground level technologies (i.e. camera traps, bioacoustics, and aquatic eDNA) in close proximity to one another, ensuring a given microhabitat was sampled in multiple different ways, simultaneously. Systematic deployment across the Concession was guided by high resolution land-cover maps obtained from remote sensing, thus ensuring sample site representativeness. Although samples and media obtained from the field were processed through separate AI or bioinformatics pipelines, results were compared to complement species lists and confirm occurrence points.All of these overlaps and complementarities helped paint a fuller picture of the species present at each sampling point. In terms of modelling methodology, presence-absence points were used within a stacked species distribution modelling approach to generate multi-taxa predictive maps, with covariates derived from remote sensing layers, including landcover, hydrology, and biomass.

S2.4 - Gulf of California

Location Background:
This study takes place in the Bay of La Paz, Gulf of California (Baja California Sur, Mexico). While scleractinian corals are found at very shallow depths, rocky reefs with and without black corals are found at deeper depths. Black coral ecosystems are found in isolated patches and mainly consist of *Antipathes galapagensis*.

Project Goal:
This study compared reef fish assemblages between the two by using a combination of three techniques: eDNA, passive acoustic monitoring, and underwater visual census.The aim of the study was to investigate the relationship between reef fish assemblages and black coral habitat.

Data Sources:

- Nature Technologies:
    - Passive acoustic sensors
    - eDNA
- Other data sources:

  - Underwater visual census

Methods and Results:
Passive acoustics was used for nocturnal soniferous species, eDNA was used for cryptobenthic species, and underwater visual census for conspicuous fish species. Fish assemblages are similar in sites with and without black corals, except for snappers, which were more abundant in black coral reef sites. A higher richness of small benthic fish was also found in black coral sites. Passive acoustics analyses showed site differences but no clear relationship with black corals. eDNA failed to support the diversity of cryptobenthic fish but did support the role of black corals for snappers.These results underscore the ecological significance of black coral ecosystems for reef fish and highligh the need of more research on this underexplored ecosystem.

Data availability:
- The raw genetic sequences can be found in GenBank: https://www.ncbi.nlm.nih.gov/sra/PRJNA1260186
- All the files and codes related to the original study can be found at: https://doi.org/10.5281/zenodo.16540403

S2.5 - Osa

Location Background:
The southern pacific region of Costa Rica has been identified as one of the high priority landscapes which have the potential to provide climate resilience (connectivity between low and high elevation protected areas) for wildlife populations in central America.

Project Goal:
The goal of Osa Conservation's 'climate lifeboat' project is to map and validate the environmental conditions, habitat integrity and wildlife connectivity within this region in order to identify high priority locations for habitat restoration and connectivity improvement interventions (such as arboreal bridges and overpasses).

Data Sources:
- Nature technologies:
  - Camera traps
  - Passive acoustic sensors
  - Solar-powered GPS location trackers
  - Satellite data
    - Land cover classification: Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A and Sentinel-1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected, log scaling
    - Planet's Land Surface Temperature product

Methods and Results:

Sentinel derived habitat layer was used to extract 'pure' habitat pixels 100% of a single land cover type matched to the resolution of the LST product. We then used a stratified random sampling approach to select pixels to compare the interaction between landcover, elevation and LST. The next steps are to 1) characterise the surrounding context of the 'pure' pixels, to determine how that drives variation in LST ; 2) Link the thermal map of the focal region to existing predictions of climate connectivity and wildlife movement behaviour/presence on the landscape. For example, to the movement behaviour of GPS tagged king vultures in the region.

Data availability:

- Data and code from "Integrating high-resolution remote sensing and empirical wildlife detection data for climate-resilient corridors across tropical elevational gradients": https://zenodo.org/records/11122373. Data and code for "Mapping climate adaptation corridors for biodiversity—A regional-scale case study in Central America": https://zenodo.org/records/11150568. Landcover maps from "Increasing Forest Cover and Connectivity Both Inside and Outside of Protected Areas in Southwestern Costa Rica": https://figshare.com/articles/dataset/Osa_Peninsula_LULC_Maps_1987_1998_2019/19337912?file=34343183

S2.6 - Western Ghats

Location Background:
In the Valparai plateau (nestled in the Anamalai hills of the Western Ghats biodiversity hotspot), the Nature Conservation Foundation (NCF), a non-governmental wildlife conservation and research organization, established a rainforest restoration programme in 2002. Despite supporting a high diversity of wildlife, rainforest fragments in the Valparai plateau became degraded due to intensive land-use change, weed invasion, and selective tree felling for fuelwood. Such large-scale changes also resulted in the fragmentation of contiguous rainforest tracts, and the region currently has >60 rainforest fragments (ranging from 1 ha to 300 ha in size). Over the last three decades, many of these rainforest fragments have been ecologically restored by NCF in collaboration with three plantation companies. Monitoring efforts that have assessed the effectiveness of ecological restoration had often relied on the response of a single taxonomic group, owing to the logistical and financial hurdles associated with sampling multiple taxa simultaneously.

Project Goals:
Ramesh et al. (2023) used a combination of passive acoustic monitoring and systematic measurements of habitat structure to examine the impacts of ecological restoration on all vocalizing fauna. In the near future, this study could be replicated by fusing high-resolution satellite imagery alongside acoustic monitoring to better understand the impacts of ecological restoration on vocalizing fauna.

Data Sources:

- Nature Technologies:
    - Passive acoustic monitoring
    - Satellite data (future)
        - Forest structure

- Other data sources:
    - Tree measurements

Methods and Results:
At each of the three 'treatment' types (naturally regenerating, actively restored and benchmark forests), the authors deployed passive acoustic recorders (AudioMoths) to collect continuous recording of audio for two weeks in summer (March 2020 - May 2020) and winter (November 2020 - January 2021). Across sites corresponding to each treatment type, vegetation measurements were extracted by sampling 20 × 20 m quadrats located at the centre of each site. Measurements included tree height, basal area, canopy cover, tree species richness and tree density and were sampled as detailed in Osuri et al. (2019). We propose the replication of this study by relying on either derived or raw estimates of vegetation metrics from high-resolution satellite imagery. All vocalizing bird species were identified manually across a subset of the audio data (detailed in Ramesh et al. 2023). Using Raven Pro, each audio recording was visually inspected and heard to note the presence/absence of each species. Community ecological analyses were carried out by estimating first-order jackknife estimates using species richness data to estimate the total rainforest and open-country bird species across treatment types.

Data Availability:
Code and analyses necessary to reproduce this study: https://github.com/vjjan91/acoustics-Restoration

S2.7 - ZWW

Location Background:
The Ziama-Wonegizi-Wologizi transboundary landscape spans the border of Guinea and Liberia. The landscape forms part of a proposed protected area mosaic, ranging from Pic de Fon in Guinea southwesterly through to the Gola landscape on the border of Liberia and Sierra Leone. The locations of the Massif du Ziama Biosphere reserve, and the Wonegizi and Wologizi proposed protected areas (PPAs) form an integral portion of the 'ecological stepping stones' of protected areas that aim to protect the Guinean forest ecosystem, and the rich fauna and flora that reside in this biodiversity hotspot.

Project Goal:
The project concerned two male elephants who regularly migrate from the South-East of Guinea through Liberia to the Ivory Coast. Along this route, the elephants have been found to raid crops, damaging farmers' livelihoods and worsening the relationship between the local people and conservationists. By collaring the elephants, the project would gain more information about the locations of the elephants. This in turn, allows the community to better protect crops and work with members of the affected communities to reduce incidences of human-wildlife conflict. Movement data could be used to better understand the use of bees on elephant behaviour and whether it is an effective deterrent. The project could also result in a greater understanding of elephant movement. More information here.

Data Sources:
- Nature Technologies:
    - GPS tags

    - Satellite data:
        - MODIS Landcover Type Yearly Global product
        - Elevation
- Other data sources:
    - Location of beehives (GPS points)

Methods and Results:
Two elephants, Little John and Big John, were collared in June 2022 (press release). The GPS collars have collected GPS locations of the two elephants for the first five months at a transmission frequency of every 12 hours. Beehives were fully installed and operational from March 2023. Home ranges were estimated for both elephants using the adehabitatHR package in R with kernel density estimation.

To provide an initial assessment of which landcover the elephants are using, the GPS locations of both elephants were compared against the MODIS Landcover Type dataset. Resource selection functions (RSFs) were used to assess which habitat characteristics are important to a specific population by assessing the difference between use of habitat and the available habitat.

Data Availability:
Code necessary to reproduce this study is adapted from Prof Elie Gurarie (University of Maryland), Working with Movement and Spatial Data in R:
https://terpconnect.umd.edu/~egurarie/research/NWT/Step06_RSF_PartI.html

Supplement 3: Case Study to Paradigm Mapping

The following supplement provides a more detailed description of what work in each case study led it to be mapped to a given paradigm.

		Paradigm					
		Survey Design	Tip and Cue	Validation	Consilience	Extrapolation and Interpolation	Data Fusion
Case Study	Alto Mayo	X		X	X		
	BDFFP			X	X	X	X
	Gabon Estuary	X	X	X	X	X	X
	Gulf of California				X		
	Osa	X			X	X	X
	Western Ghats				X		
	ZWW				X		X

Table S3: Mapping of case studies to paradigms.

S3.1 - Survey Design

In Alto Mayo, a satellite-derived ecosystem map was used to guide the in-situ sampling of bioacoustics, camera traps, and eDNA. For example, the map was used to select eDNA/in-situ sampling sites to ensure comprehensive coverage across the different habitat zones (montane forest, alluvial forest, etc.). However, the ecosystem map was not the only criteria driving survey design, which utilized other layers including existing biodiversity data to identify important places and data gaps, as well as modeling of various ecosystem services. Additional layers were developed using existing data to model biodiversity priorities (using a species benefit index based on threatened species ranges and generalized dissimilarity models) and ecosystem services (water provisioning, sediment retention and carbon stocks), and these layers were combined and weighted to identify high priority survey areas, also taking biodiversity data gaps into account.

In Osa, a habitat connectivity model for wildlife that was derived from satellite, camera trap, and GPS tag nature technologies was used to identify areas for additional in-situ sampling.

In the Gabon Estuary, satellite-derived landcover maps guided the in-situ sampling of camera traps, acoustics, and eDNA.

S3.2 - Tip and Cue

Tip and cue was only used in one case study - the Gabon Estuary. In this case study, eDNA provided a "tip" by identifying threatened and endangered species. This tip "cued" the deployment of a new subset of camera traps and acoustics sensors that were focused on those species.

S3.3 - Validation

In Alto Mayo, species detected via eDNA were cross-checked against acoustic and camera trap detections, as well as other in-situ observations using traditional methods.

In the Gabon Estuary, Camera Trap data was used to confirm low-confidence eDNA detections.

In BDFFP, correlation between forest structural diversity (measured via space-based LiDAR) and fauna species occupancy (derived via camera trap and acoustics) helped further validate the habitat heterogeneity hypothesis relationship to biodiversity in tropical forests.

S3.4 - Consilience

In Alto Mayo, localized in situ data from camera traps, acoustics, eDNA,were spatially overlaid with remote sensing to more fully understand species richness, composition, diversity indices.

In the Gabon Estuary, biodiversity layers derived from in situ sensors and eDNA were spatially overlaid on top of biomass and carbon estimates derived from remote sensing.

In ZWW, satellite-derived land cover was compared to home range derived from GPS-derived elephant location to determine elephant habitat preference.

In Osa, the relationship between satellite-derived thermal and landcover data and wildlife metrics derived from camera traps, acoustics, and GPS tags derived fauna were used to better understand species habitat

In BDFFP, multiple metrics derived from space-based LiDAR (e.g. canopy height, above ground carbon) were explored alongside mean species occupancy derived from camera traps and acoustics to better understand the relationship between remotely sensed forest structure metrics and biodiversity.

In Western Ghats, project plans include comparing satellite-derived forest structure and acoustic-derived bird species metrics to assess how restoration affects bird diversity.

In the Gulf of California, eDNA, acoustic data, and visual scuba-diving transects streams were analyzed separately, to study different facets of the overall unifying investigation into reef fish assemblages (nocturnal soniferous species for acoustics, transects for conspicuous, and cryptobenthic species for eDNA).

S3.5 - Extrapolation and Interpolation

In the Gabon Estuary, using only in-situ sensors, a spatially continuous biodiversity layer was interpolated from camera trap, acoustic, and eDNA-derived presence/absence points.

In BDFFP, the consilience of remote sensed forest structure metrics and camera trap and acoustic derived species metrics led to the interpolation of species occurrence in unsampled areas. The team in BDFFP plans to leverage the relationship between space-based LiDAR-derived predictors and acoustic/camera trap species occurrence to extrapolate occurrence in new spatial areas.

In Osa, habitat connectivity models derived from satellite data, camera traps, and GPS tags are being extrapolated to novel landscapes. The Osa case study has also employed temporal extrapolation - extrapolating the relationship between satellite-derived land surface temperature and GPS-derived species movement model wildlife habitat preferences in future climate scenarios.

S3.6 - Data Fusion

In the Gabon Estuary, camera trap, acoustic, and eDNA-derived species estimates were fused to create a new unified and predictive biodiversity layer.

In the Osa case study, satellite derived landcover and thermal data is being fused with camera trap/GPS tag-derived habitat preference data to create a new connectivity surface layer. This connectivity layer is now being used to analyze "pinch points" where animal movement is being constrained, a clear example of a new metric being produced from the fusion of two nature technology data streams.

In ZWW, satellite-derived landcover and GPS collar data were fused via a resource selection function to produce a resource use probability map. The novel habitat preference map was then used to understand the elephants' relationship to human settlements, elevation, distance to beehives and distance to rivers in a way that was not possible using only one data stream.