Reconciling top-down and bottom-up conservation priorities with reinforcement learning from human feedback

Daniele Silvestro[1,2,3,*], Stefano Goria[3,4], E-Ping Rau[2], Thomas Sterner[5,*], Alexandre Antonelli[2,3,5,6,7,*]

[1]Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland; [2]Gothenburg Global Biodiversity Centre, Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden; [3]Captain Technologies Ltd, London, UK; [4]Thymia Ltd, London, UK; [5]Department of Economics, University of Gothenburg, Gothenburg, Sweden; [6]Royal Botanic Gardens, Kew, Richmond, UK; [7]Antonelli Foundation for Biodiversity Research and Conservation, Nova Friburgo, Brazil; [8]Department of Biology, University of Oxford, Oxford, UK.

*Corresponding authors.

Email: daniele.silvestro@bsse.ethz.ch, stefano@thymia.ai, epingchris@gmail.com, thomas.sterner@gu.se, a.antonelli@kew.org

To meet the Kunming-Montreal Global Biodiversity Framework, spatial planning must upscale area-based conservation, while ensuring equitable governance and local community integration. Historically, a deep divide has persisted between data-driven, top-down systematic conservation planning, often performed at large spatial scales, and participatory, bottom-up local initiatives. Here, we propose a novel framework that bridges this gap by harnessing the power of human and artificial intelligence. By deploying a technique called Reinforcement Learning with Human Feedback in our software CAPTAIN (Conservation Area Prioritization through Artificial INtelligence), our approach can process vast biophysical and socioeconomic data to create fine-tuned conservation strategies that are both data-driven and sensitive to local realities. The methodology outlined here provides a promising avenue for scalable, effective, and equitable biodiversity conservation.

Global conservation targets

Facing the accelerating global crisis of biodiversity loss, we must significantly scale up effective and inclusive conservation. This has often translated to pledges for targets of area-based protection. Notably, Target 3 of the Kunming-Montreal Global Biodiversity Framework, also known as the “30 by 30” target, sets the ambitious target of protecting 30% of the world’s terrestrial and inland water areas, and of marine and coastal areas, by 2030 – and has been committed by nearly 200 nations. Target 3 also explicitly expands area-based protection beyond traditional protected areas to include "other effective area-based conservation measures", mandating that these systems be ecologically representative and equitably governed, while recognizing and respecting the rights of Indigenous Peoples and Local Communities (IPLCs).

However, implementing these high-level area-based targets presents several shortcomings. These include a general lack of impact evaluation and distinction between different categories of protected areas (Martínez and Melo 2025), discrepancies between reported and genuinely implemented actions (Pedersen *et al* 2026), and failure to operationalize metrics of representativeness, connectivity, management effectiveness, durability, and equitable governance (Geldmann 2026). Because area-based conservation must be implemented rapidly in accordance with the Global Biodiversity Framework’s timelines, a recurring debate centers around the relative merits and pitfalls of top-down versus bottom-up approaches. To avoid them becoming mutually exclusive, tools are urgently required which enable operationalizing the biodiversity targets and help bridge large-scale prioritization and local realities.

Top-down planning

Historically, the designation of protected areas and other conservation measures has been performed by government bodies with limited or no local engagement and buy-in (Gissibl *et al* 2012). They also often proceeded without a prior assessment of how individual interventions across a region would best complement, rather than duplicate each other. This is where systematic conservation planning came in – a more formalized, data-informed approach for locating and designing conservation areas that explicitly attempts to maximize complementarity (Margules and Pressey 2000). Under this umbrella, many methods and metrics have been developed to analyze large data sets of various biodiversity indices and create priority maps for conservation (e.g. Ball et al 2009, Hanson et al 2025). Such analyses have resulted in influential proposals at the global, regional, and national levels (e.g., Strassburg et al 2020, Jung et al 2021).

Unfortunately, conservation is also steeped in colonial practices that prioritized, among other things, non-biological factors such as elite sport hunting and recreation for expats (Garland 2008).

Some early protected areas were established by forcibly displacing local communities (Kabra 2018). These practices that assert a racialized control over land and nature, stemming from a view of human presence as a threat to nature and of hard dichotomy between biodiversity and livelihood outcomes have often resulted in "conservation refugees" while asserting western, racialized control over land and nature, often leading to strong local resentment to biodiversity conservation. Such practices have not only been deeply unjust, but also counterproductive, since it can be difficult to achieve sustainable conservation outcomes without the long-term support of local stakeholders. Modern conservation, in contrast, seeks to thoroughly embrace the opinions and participation of IPLCs (Layden et al., 2025), for the identification and implementation of protected areas as well as other effective area-based conservation measures (Dudley et al. 2018).

While systematic conservation planning can generate priority maps in cost-effective ways, they often lack direct contact with local stakeholders, which risks leading to tension and conflicts when high-level conservation priorities might clash with local needs and preferences. This disconnect has led to fierce criticism from some conservationists, proposing that “conservation needs to break free from global priority mapping” since “global maps embody a technocratic view of environmental decision-making” erasing “local context and difference” (Wyborn and Evans 2021).

Bottom-up initiatives

Several conservation organizations, such as the World Wide Fund for Nature (WWF), are instead committed to amplifying “locally-led conservation”, a concept that is gaining traction in alignment with the increased recognition of the roles and contributions of IPLCs as custodians of biodiversity and as integral partners in its conservation (including both protection and restoration) and sustainable use (Sze *et al* 2021). The upfront participation of IPLCs can contribute critical context-specific knowledge of the territory being considered for conservation, ranging from cultural values of specific sites to a better understanding of nature’s contributions to people at a local scale. Bottom-up conservation strategies co-created by IPLCs can also lead to more effective and equitable use of resources allocated for conservation (supporting for instance education and knowledge dissemination), long-term ownership by the communities involved and additional benefits for biodiversity (Reyes-García *et al* 2022).

However, bottom-up approaches also face challenges that limit how much they can contribute to global or national conservation targets: (1) without strong institutional support, bottom-up efforts may lack resources to successfully counter economic incentives for practices that harm nature, or to stay resilient to large-scale environmental risks and climate change; (2) without coordination and planning, isolated initiatives may fail to take into consideration macro-scale goals such as connectivity, prioritization of globally endangered species, complementarity and representativeness, and risk giving focus to more populated and accessible areas instead of where the most abundant and threatened biodiversity can be found; (3) locally-led initiatives may not be effective where social cohesion and traditional governance structure within the local communities have been fragilized (e.g., due to war, displacement or assimilation, or severe impoverishment). These are not inherent flaws with bottom-up approaches, but rather limitations due to the global scale and immense complexity of the direct and indirect drivers underpinning biodiversity loss.

Bridging divergent perspectives

Rather than being an irreconcilable divide, our proposal is that these divergent perspectives should be integrated: there is indispensable value in both large-scale priority planning (because the challenge of biodiversity loss is global) and in local knowledge and perspectives (because effective conservation needs to benefit people and nature). We need a fair and effective integration of data from modern science (such as remotely sensed information, functional and genetic diversity) and local, traditional and indigenous knowledge (Araújo 2025).

In recent years, artificial intelligence (AI) has become a promising technology that can revolutionize conservation science (Reynolds *et al* 2025, Ullah *et al* 2025), providing new tools for biodiversity monitoring and data analysis, and powerful ways to process high-dimensional and high resolution data (Pollock *et al* 2025, Dorm *et al* 2026). In this context, the software Conservation Area Prioritization Through Artificial Intelligence (CAPTAIN) (Silvestro *et al* 2022, 2025), uses spatially explicit ecosystem simulations and Reinforcement Learning to identify conservation priorities through time and in space, while considering both biodiversity targets and implementation costs. The selection of conservation areas is performed by an agent, i.e. a virtual policy maker, who observes the state of the environment (the spatial distribution of species, their extinction risk, land use, costs) and how it evolves through time and then implements protection strategies, which will for instance prevent habitat degradation or be selected for restoration. The agent is rewarded based on one or multiple criteria describing its objectives, e.g., reducing extinction risk and minimizing implementation and opportunity costs. While the agent can process a highly dimensional environment, with potentially thousands of species and their spatial variation through time (Bütikofer et al. 2026) to optimize top-down spatial planning, it currently still lacks an effective way to integrate local realities, including the knowledge, needs and preferences of local stakeholders.

To fill in this gap, we introduce the architecture of a new framework based on Reinforcement Learning from Human Feedback (RLHF) to integrate available biological and socioenvironmental data at large scales with input from local stakeholders. RLHF has become an integral component in developing and refining AI systems, enabling models to better reflect human preferences and expectations (Amershi *et al* 2014), gaining increasing traction to improve AI models in different domains such as robot locomotion and alignment of large language models to human values (Christiano *et al* 2017). We consider RLHF particularly promising to devise conservation policies that incorporate large-scale multidimensional data and internationally agreed biodiversity targets, while also learning from the knowledge and input of the local communities involved, through an iterative and interactive process (Fig. 1).

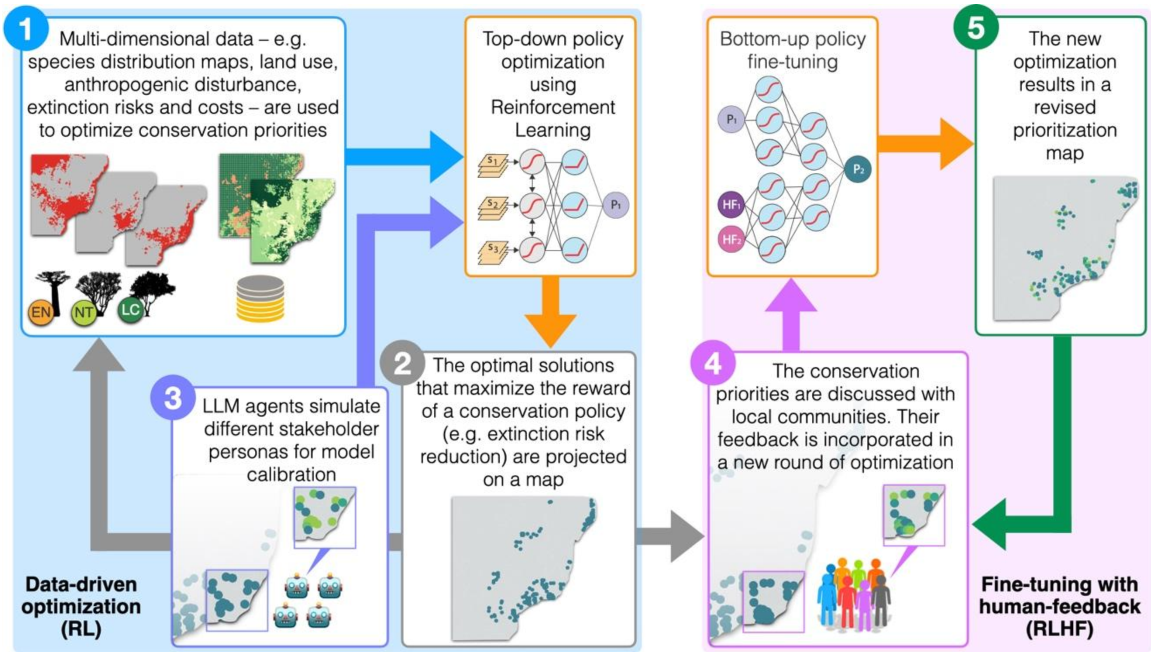

Figure 1 | Identifying conservation priorities through Reinforcement Learning from Human Feedback. Recent advances in AI could help break down the divide between top-down and bottom-up approaches to spatial conservation prioritization. Within the CAPTAIN framework for spatial planning, we use RL to find priority areas for conservation based on high dimensional data that maximize a standard reward, e.g. minimizing extinction risks and implementation costs (steps 1-2). The behavior of the model can be tested through LLM agents to ensure it is architecturally able to meaningfully handle additional, potentially contrasting feedback (step 3). Proposed solutions are evaluated by local stakeholders providing input from a new round of optimization based on RLHF (steps 4-5). These steps can be iterated until a satisfactory solution is found.

Prioritization with RLHF

In contrast to standard Reinforcement Learning where the agent behavior is optimized to maximize a static reward function *R(s, a)* dependent solely on biophysical and cost data from the environment, our proposed framework incorporates human feedback through a dynamic, aggregate reward structure (Bıyık *et al* 2022):

$$R_{\text{total}} = R_{\text{standard}} + \alpha \cdot R_{\text{human}}(s, a, H)$$

where *s* represents the spatial ecosystem state, *a* is the conservation action allocation (e.g., protection or restoration), *H* represents the aggregated human feedback stream, and $\alpha \geq 0$ is a scaling hyperparameter that determines the relative weight of local human preferences against top-down biological targets. We transform qualitative community insights into quantitative signals (*H*) using distinct methods depending on the feedback format:

1. Pairwise preferences: using Bradley-Terry preference models, stakeholders are presented with alternative spatial plannings (a_1 vs a_2) generated by the baseline agent. Their preferences are compiled to calculate a probability distribution of acceptable spatial trade-offs.
2. Spatial annotations: when communities identify point-based or polygon-based feedback (e.g., traditional hunting grounds or sacred sites), spatial kernel density estimation or fuzzy classification algorithms are applied to smooth and generalize these points into continuous spatial penalty or reward surfaces (Ding et al. 2023).
3. Textual and verbal insights: qualitative feedback from community surveys is converted via semantic embeddings (using domain-specific language models) into a preference vector to reflect preference patterns.

The scaling factor α can be systematically varied during model calibration to evaluate trade-offs between biological conservation gains and local stakeholder utility (Silvestro *et al* 2025).

Aggregating heterogeneous feedback

A core challenge in socially engaged conservation is that local communities are rarely homogeneous; they frequently comprise distinct groups with competing, and sometimes irreconcilable, interests and priorities (McShane et al. 2011). If human feedback is integrated naively, preference-learning pipelines risk "elite capture", where the viewpoints of the most vocal, accessible, or digitally literate stakeholders dominate the learned reward signal (Wilder and

Walpole 2008). An intuitive remedy would be to formulate R_{human} as weighted function rather than a simple majority vote:

$$R_{\text{human}}(s, a, H) = \sum_{i=1}^{M} w_i \cdot R_i(s, a, H_i)$$

where w_i represents an equity weight assigned to stakeholder group i, and M is the total number of distinct local groups identified via prior stakeholder mapping. To protect marginalized demographics and uphold indigenous sovereignty as mandated by the GBF, vulnerable or historically displaced groups the function could be used to assign higher structural weights (w_i).

However, this solution may be inadequate for resolving value conflicts. This is because a weighted mean collapses divergent preferences, i.e. a multi-modal distribution, onto a single numerical score. This represents a utilitarian definition of the reward aiming to maximize the average satisfaction while potentially leading to a compromise solution that satisfies none of the stakeholders. Aggregating divergent human feedback into a single objective is a problem of social choice rather than of engineering (Conitzer et al., 2024) and single-reward learning is unable to represent diverse populations (Siththaranjan et al., 2024).

To preserve the true plurality of interests that motivates bottom-up engagement, we instead formulate the aggregation of divergent feedback as a problem of social choice. Rather than optimizing for an average reward, we adopt an inequality-averse, MaxMin RLHF principle (Chakraborty et al., 2024). Under this formulation, the optimization framework evaluates a candidate spatial conservation policy (π) against the group-specific rewards (R_i) computed independently based on each stakeholder group's feedback. The final conservation policy ($\pi^\star$) is optimized to maximize the welfare of whichever group is least served by the spatial allocation:

$$\pi^\star \in \arg\max_{\pi} \min_{i \in \{1,\ldots,M\}} R_i(\pi)$$

Following the MaxMin RLHF construction, this architecture prevents the algorithm from sacrificing the livelihoods of a minority group to achieve a high aggregate score for a dominant group. By focusing optimization on lifting the satisfaction floor of the least-served stakeholders, the agent is therefore forced to identify mutually tolerable conservation planning. While deeply divergent opinions still imply that no single policy can satisfy all groups simultaneously, this optimization can return a set of admissible plans, which can be used as explicit options for deliberation through a transparent participatory process. This approach should therefore preserve, rather than erase, the plurality of interests that motivates bottom-up engagement (Chakraborty et al., 2024).

Expected benefits

Our proposed model combines the key advantages of top-down and bottom-up approaches. Firstly, it allows conservation planners to maintain a broad geographical context while optimizing conservation priorities. Secondly, it allows the integration of granular information about local preferences, cultural values, and implementation feasibility. Thirdly, it promotes the active participation of relevant stakeholders in the decision-making process – increasing trust and engagement towards conservation. Taken together, these advantages should increase the probability of achieving long-lasting societal and biodiversity improvements.

We envision an integrated workflow within the CAPTAIN platform, initiated by a top-down optimization that synthesizes multi-dimensional variables—including species distributions, anthropogenic threats, and climate sensitivity—to propose initial conservation priorities (Fig. 1, steps 1–2). These steps are based on standard RL optimization and allow the agent to learn about the features of the environment and to identify the areas that are, together, most likely to yield the best outcomes. The spatial extent of the initial top-down optimization can be potentially larger than the focal area where human feedback is collected and potentially even global, if that helps capture biodiversity aspects (connectivity, complementarity) that might be missed at smaller scales. This proposal then enters a collaborative "*human feedback*" phase (step 4) with IPLCs and other relevant stakeholders. The specific implementation of this phase depends on the local social context, spatial extent of the area, and available resources. It can entail engagement with local landowners, citizen juries or assemblies, online surveys, and futuring exercises (Jansen et al. 2023; Vollan et al. 2024; Wyborn et al. 2021). Unlike the biodiversity data used in steps 1-2, this feedback provides granular, qualitative, and spatially sparse insights, identifying unforeseen implementation constraints or alternative conservation solutions. These local values constitute the basis for a dynamic reward signal with human feedback and are then integrated into a subsequent round of AI-driven optimization (step 5) to refine the plan. This iterative cycle between model calibration and stakeholder engagement (steps 4-5) can be repeated until a satisfactory, inclusive conservation scenario is achieved.

Simulating human feedback for system calibration

Social and ecological systems share a bewildering local complexity. Before real-world deployment, we propose validating the mathematical stability of our RLHF framework using Large Language Model (LLM) agents to simulate stakeholder feedback (Fig. 1, step 3). This approach, rooted in human-AI interaction research, uses "silicon sampling" to stress-test model behavior across diverse hypothetical scenarios before unnecessarily exposing human communities to the tool (Park *et al* 2023, Dubois *et al* 2023). Specifically, LLM agents conditioned on distinct stakeholder personas (e.g., local officials, community representatives) allow for the systematic optimization of software hyperparameters. This allows researchers to determine: (i) the most computationally actionable feedback format (e.g., ranked alternatives vs. spatial annotations), (ii) the minimum feedback volume required to shift data-driven priorities, and (iii) how the algorithm behaves when fed intentionally contradictory or high-variance rewards.

Regardless of how carefully LLM agents are prompted or fine-tuned on regional policy documents and contexts, simulated feedback cannot capture the nuanced authority, lived experiences, and contextual complexity of IPLCs. This "sim2real gap" (Tobin *et al* 2017) is exacerbated by severe under-representation or erasure of indigenous knowledge systems within the training corpora of commercial LLMs (Perera *et al* 2025, Rosser and Spennemann 2025). Consequently, this simulation phase must be viewed strictly as an engineering calibration step for software pipeline verification, to protect people from excessive requests until the model is deemed ready for real application, and never as a substitute for genuine stakeholder engagement. By pre-testing the feedback ingestion mechanism in a controlled sandbox, we ensure that when authentic human feedback is subsequently collected (Step 3), the system is computationally optimized to process it smoothly. This minimizes stakeholder fatigue and ensures that the final policy remains rooted in the authenticity and primacy of community voices.

From theory to practice

Bridging top-down and bottom-up priorities is fundamentally a challenge of institutional design and political economy (Ando and Langpap 2018, Muchapondwa and Ntuli 2024). From a resource economics perspective, spatial conservation planning often treats local regulatory compliance as a given, neglecting the underlying transaction costs and incentive structures required for long-term policy durability (Mirzabaev and Wuepper 2023). This oversight is critical, as local communities frequently face highly asymmetric trade-offs where global or national biodiversity benefits are diffuse, but economic opportunity costs, such as restricted resource extraction or land-use constraints, are highly concentrated locally. Beyond macro-economics, local perceptions of these biodiversity policies reflect sensitive political dynamics regarding whose interests come to the fore, often involving competing and contradictory priorities. For instance, a single area may simultaneously support indigenous hunter-gatherers, informal miners, fishers, and extractive industries. Because resource restrictions impose direct livelihood costs on these diverse groups, community opinions on conservation programs remain highly contingent upon the alternative opportunities and compensatory frameworks offered, including benefit-sharing models such as community-managed tourism revenue sharing programs (Muchapondwa and Ntuli 2024, Snyman et al. 2023).

If local communities are to accept corridors for species migration or expected to provide labor to fight invasive species, or if restrictions are imposed on hunting, they will probably want something in return. This may be, for instance, a share of tourism revenue as in the Campfire project in Zimbabwe or payments for ecosystem services if they protect a watershed for the benefit of downstream users. The CAPTAIN framework can expose the economic trade-offs within the standard reward function, thus helping co-design conservation strategies that are economically viable (Silvestro *et al* 2025). Yet, the quid pro quo is not always monetary, as local populations may be looking for protection, for health or security benefits, or even strengthened property rights. These aspects can be integrated within R_{human} term through the proposed RLHF framework.

Deploying AI for biodiversity conservation requires accounting for its ecological footprint and data ethics. The RLHF models proposed here are tiny by current AI standards and involve relatively small energy requirements. Yet, quantifying and disclosing the carbon footprint associated with conservation planning optimizations is valuable and can be facilitated through future integrations of open-source carbon-auditing tools (e.g., CodeCarbon[1]) within CAPTAIN. Ethical data governance should be maintained, e.g. under the CARE Principles for Indigenous Data Governance (Carroll et al. 2020). The relatively lightweight nature of the proposed model facilitates this task, because if the analyses can be executed locally and offline, sensitive data (e.g., poaching

[1] https://codecarbon.io/

targets) and private community preferences can remain strictly under local custody without needing to be uploaded to public or centralized cloud servers.

Real-life validations of the proposed RLHF approach should include data compilation and analyses, followed by stakeholder engagement through in-person community workshops or surveys[2], where results can be presented and input gathered in the local languages. While the implementation of the new RLHF models proposed here requires technical and scientific expertise, the core framework should be implemented as an open-source well-documented tool deployed through boundary-spanning organizations, e.g., conservation NGOs, academic extension services, and civil society groups, acting as neutral facilitators. These intermediaries could both host the participatory workshops and run the RLHF analyses, supporting conservation planning without creating technological barriers for local communities. This process should be done in synergy with, rather than instead of, well-tested workflows, such as those established by the Important Plant Areas program (Darbyshire *et al* 2017), which has so far resulted in the identification of over 2,000 priority areas for plant conservation in nearly 40 countries. By grafting RLHF optimization onto existing conservation programs, our proposed approach holds the potential to support governments, non-governmental organizations, and policymakers with the spatial allocation of conservation resources in ways that are holistic, effective, and inclusive.

[2] E.g., https://www.biodiversa.eu/2026/04/03/peace/

Acknowledgments

We thank Gabriel K. Ameka, Iain Darbyshire, and Björn Vollan for valuable discussions on the approach proposed here. D.S. acknowledges funding from ETH Zurich and from the PEACE research program funded by Swiss National Science Foundation and the European Biodiversity Partnership (www.biodiversa.eu/2026/04/03/peace; 31BD-0_233170). A.A. and D.S. are members of the BIOPATH research program funded by the Swedish Foundation for Strategic Environmental Research MISTRA (F 2022/1448) and acknowledge additional funding from the Swedish Research Council (2024-04303). A.A. is also funded by the Royal Botanic Gardens, Kew. TS acknowledges support from the Familjen Kamprad Foundation, project 20230142.

Data availability

The CAPTAIN software is available at: www.captain-project.net

References

1. Amershi S, Cakmak M, Knox W B and Kulesza T 2014 Power to the people: The role of humans in interactive machine learning *AI Mag.* 35 105–20
2. Araújo M B 2025 Conflicting rationalities limit the uptake of spatial conservation prioritizations *Nat. Rev. Biodivers.* 1 279–81
3. Ball I R, Possingham H P and Watts M 2009 Marxan and relatives: software for spatial conservation prioritisation Spatial conservation prioritisation: *Quantitative methods and computational tools* 14 185–96
4. Bıyık E, Losey D P, Palan M, Landolfi N C, Shevchuk G and Sadigh D 2022 Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences *Int. J. Rob. Res.* 41 45–67
5. Bütikofer, Luca, Daniele Silvestro, M. Luisa Rubio Teso, Ada S. Molina, Carlos Lara-Romero, Raul García Valdés, Olivier Broennimann et al. "Future-proofing agrobiodiversity: climate and niche-aware conservation planning using reinforcement learning." *bioRxiv* (2026): 2026-05.
6. Chakraborty, S., Qiu, J., Yuan, H., Koppel, A., Huang, F., Manocha, D., Bedi, A.S. and Wang, M., 2024. MaxMin-RLHF: Alignment with diverse human preferences. *arXiv* doi:arXiv:2402.08925.
7. Christiano P, Leike J, Brown T B, Martic M, Legg S and Amodei D 2017 Deep reinforcement learning from human preferences *arXiv*. Doi: http://arxiv.org/abs/1706.03741
8. Conitzer, V., Freedman, R., Heitzig, J., Holliday, W.H., Jacobs, B.M., Lambert, N., Mossé, M., Pacuit, E., Russell, S., Schoelkopf, H. and Tewolde, E., 2024. Social choice should guide ai alignment in dealing with diverse human feedback. *arXiv* doi: arXiv:2404.10271.

9. Darbyshire I, Anderson S, Asatryan A, Byfield A, Cheek M, Clubbe C, Ghrabi Z, Harris T, Heatubun C D, Kalema J, Magassouba S, McCarthy B, Milliken W, de Montmollin B, Lughadha E N, Onana J-M, Saïdou D, Sârbu A, Shrestha K and Radford E A 2017 Important Plant Areas: revised selection criteria for a global approach to plant conservation *Biodivers. Conserv*. 26 1767–800
10. Ding, Zhaowei, Hua Zheng, Jun Wang, Patrick O'Connor, Cong Li, Xiaodong Chen, Ruonan Li, and Zhiyun Ouyang. "Integrating top-down and bottom-up approaches improves practicality and efficiency of large-scale ecological restoration planning: insights from a social–ecological system." *Engineering* 31 (2023): 50-58.
11. Dorm F, Millard J, Purves D, Harfoot M and Mac Aodha O 2026 Large language models possess some ecological knowledge, but how much? *Ecol. Inform*. 95 103699
12. Dubois Y, Li X, Taori R, Zhang T, Gulrajani I, Ba J, Guestrin C, Liang P and Hashimoto T B 2023 AlpacaFarm: A simulation framework for methods that learn from human feedback ed A Oh, T Naumann, A Globerson, K Saenko, M Hardt and S Levine. *arXiv* 30039–69 doi: 2305.14387
13. Dudley, Nigel, Holly Jonas, Fred Nelson, Jeffrey Parrish, Aili Pyhälä, Sue Stolton, and James EM Watson. "The essential role of other effective area-based conservation measures in achieving big bold conservation targets." Global ecology and conservation 15 (2018): e00424.
14. Garland, E., 2008. The elephant in the room: Confronting the colonial character of wildlife conservation in Africa. *African Studies Review*, 51(3), pp.51-74.
15. Geldmann J 2026 The world could reach 30% protection by 2030, and still fail to conserve biodiversity without effective and well-located protected areas *Conserv. Lett.* 19 Online: http://dx.doi.org/10.1111/con4.70030
16. Gissibl, B., Höhler, S. and Kupper, P. eds., 2012. Civilizing nature: national parks in global historical perspective (Vol. 1). *Berghahn Books.*
17. Hanson J O, Schuster R, Strimas-Mackey M, Morrell N, Edwards B P M, Arcese P, Bennett J R and Possingham H P 2025 Systematic conservation prioritization with the prioritizr R package *Conserv. Biol.* 39 e14376
18. Janssen, M.A., Falk, T., Meinzen-Dick, R. and Vollan, B., 2023. Using games for social learning to promote self-governance. Current Opinion in Environmental Sustainability, 62, p.101289.
19. Jung M, Arnell A, de Lamo X, García-Rangel S, Lewis M, Mark J, Merow C, Miles L, Ondo I, Pironon S, Ravilious C, Rivers M, Schepaschenko D, Tallowin O, van Soesbergen A, Govaerts R, Boyle B L, Enquist B J, Feng X, Gallagher R, Maitner B, Meiri S, Mulligan M, Ofer G, Roll U, Hanson J O, Jetz W, Di Marco M, McGowan J, Rinnan D S, Sachs J D, Lesiv

M, Adams V M, Andrew S C, Burger J R, Hannah L, Marquet P A, McCarthy J K, Morueta-Holme N, Newman E A, Park D S, Roehrdanz P R, Svenning J-C, Violle C, Wieringa J J, Wynne G, Fritz S, Strassburg B B N, Obersteiner M, Kapos V, Burgess N, Schmidt-Traub G and Visconti P 2021 Areas of global importance for conserving terrestrial biodiversity, carbon and water *Nat. Ecol. Evol.* 5 1499–509

20. Kabra, A., 2018. Dilemmas of conservation displacement from protected areas. In *Challenging the prevailing paradigm of displacement and resettlement* (pp. 117-140). Routledge.
21. Layden, T. J., David-Chavez, D. M., García, E. G., Gifford, G. L., Lavoie, A., Weingarten, E. R., & Bombaci, S. P. (2025). Confronting colonial history: toward healing, just, and equitable Indigenous conservation futures. *Ecology and Society*, 30(1).
22. Margules C R and Pressey R L 2000 Systematic conservation planning Nature 405 243–53
23. Martínez O A and Melo O 2025 To increase or not to increase? The impact of terrestrial protected areas on tree cover and the risk of species extinction *J. Environ. Manage*. 382 125386
24. McShane, T.O., Hirsch, P.D., Trung, T.C., Songorwa, A.N., Kinzig, A., Monteferri, B., Mutekanga, D., Van Thang, H., Dammert, J.L., Pulgar-Vidal, M. and Welch-Devine, M., 2011. Hard choices: Making trade-offs between biodiversity conservation and human well-being. *Biological conservation,* 144(3), pp.966-972.
25. Muchapondwa, E., & Ntuli, H. (2024). Managing trade-offs between communities' welfare and nature conservation: the case of wildlife management systems in and outside protected areas in Africa. *Annual Review of Resource Economics*, 16(1), 143-166.
26. Park J S, O'Brien J, Cai C J, Morris M R, Liang P and Bernstein M S 2023 Generative agents: Interactive simulacra of human behavior Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology UIST '23: The 36th Annual ACM Symposium on User Interface Software and Technology (New York, NY, USA: ACM) Online: http://dx.doi.org/10.1145/3586183.3606763
27. Pedersen P B M, Petersen A H, Corcoran D, Madsen N, Svenning J-C, Timmermann K, Kaae B C, Kragh T, Olsen B E, Strange N, Normand S and Rahbek C 2026 Identifying conservation gaps: A framework for evaluating national contributions to the 30 × 30 target. *One Earth* 101620
28. Perera M, Vidanaarachchi R, Chandrashekeran S, Kennedy M, Kennedy B and Halgamuge S 2025 Indigenous peoples and artificial intelligence: A systematic review and future directions Big Data Soc. 12 Online: http://dx.doi.org/10.1177/20539517251349170

29. Pollock L J, Kitzes J, Beery S, Gaynor K M, Jarzyna M A, Mac Aodha O, Meyer B, Rolnick D, Taylor G W, Tuia D and Berger-Wolf T 2025 Harnessing artificial intelligence to fill global shortfalls in biodiversity knowledge *Nat. Rev. Biodivers.* 1 166–82
30. Reyes-García V, Fernández-Llamazares Á, Aumeeruddy-Thomas Y, Benyei P, Bussmann R W, Diamond S K, García-Del-Amo D, Guadilla-Sáez S, Hanazaki N, Kosoy N, Lavides M, Luz A C, McElwee P, Meretsky V J, Newberry T, Molnár Z, Ruiz-Mallén I, Salpeteur M, Wyndham F S, Zorondo-Rodriguez F and Brondizio E S 2022 Recognizing Indigenous peoples' and local communities' rights and agency in the post-2020 *Biodiversity Agenda Ambio* 51 84–92
31. Reynolds S A, Beery S, Burgess N, Burgman M, Butchart S H M, Cooke S J, Coomes D, Danielsen F, Di Minin E, Durán A P, Gassert F, Hinsley A, Jaffer S, Jones J P G, Li B V, Mac Aodha O, Madhavapeddy A, O'Donnell S A L, Oxbury W M, Peck L, Pettorelli N, Rodríguez J P, Shuckburgh E, Strassburg B, Yamashita H, Miao Z and Sutherland W J 2025 The potential for AI to revolutionize conservation: a horizon scan *Trends Ecol. Evol.* 40 191–207
32. Rosser E and Spennemann D 2025 Indigenous erasure in the age of generative artificial intelligence Social Science Research Network Online: https://papers.ssrn.com/abstract=5689603
33. Silvestro D, Goria S, Rau E-P, Ferreira de Lima R A, Groom B, Jacobsson P, Sterner T and Antonelli A 2025 Using artificial intelligence to optimize ecological restoration for climate and biodiversity *bioRxiv* 2025.01. 31.635975 doi: 10.1101/2025.01.31.635975
34. Silvestro D, Goria S, Sterner T and Antonelli A 2022 Improving biodiversity protection through artificial intelligence *Nat. Sustain.* 5 415–24
35. Siththaranjan, A., Laidlaw, C. and Hadfield-Menell, D., 2024, May. Distributional preference learning: Understanding and accounting for hidden context in RLHF. In *International Conference on Learning Representations* (Vol. 2024, pp. 27448-27471).
36. Strassburg B B N, Iribarrem A, Beyer H L, Cordeiro C L, Crouzeilles R, Jakovac C C, Braga Junqueira A, Lacerda E, Latawiec A E, Balmford A, Brooks T M, Butchart S H M, Chazdon R L, Erb K-H, Brancalion P, Buchanan G, Cooper D, Díaz S, Donald P F, Kapos V, Leclère D, Miles L, Obersteiner M, Plutzar C, de M Scaramuzza C A, Scarano F R and Visconti P 2020 Global priority areas for ecosystem restoration *Nature* 586 724–9
37. Snyman, Susan (Sue) S., Kathleen Fitzgerald, Anastasiya Bakteeva, Telesphore Ngoga, and Benjamin Mugabukomeye (2023). "Benefit-sharing from protected area tourism: A 15-year review of the Rwanda tourism revenue sharing programme." *Frontiers in Sustainable Tourism* 1:1052052. DOI: 10.3389/frsut.2022.1052052.
38. Sze J S, Carrasco L R, Childs D and Edwards D P 2021 Reduced deforestation and degradation in Indigenous Lands pan-tropically Nat. Sustain. 5 123–30

39. Tobin J, Fong R, Ray A, Schneider J, Zaremba W and Abbeel P 2017 Domain randomization for transferring deep neural networks from simulation to the real world 2017 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 2017 pp 23–30
40. Ullah F, Saqib S and Xiong Y-C 2025 Integrating artificial intelligence in biodiversity conservation: bridging classical and modern approaches *Biodivers. Conserv.* 34 45–65
41. Vollan, Björn, Max Burger, and Marco Nilgen. "Applying citizens' juries to the Global South." (2024).
42. Wilder, L. and Walpole, M. (2008) Measuring social impacts in conservation: experience of using the Most Significant Change method, *Oryx*, 42(4), pp. 529–538. doi:10.1017/S0030605307000671.
43. Wyborn C and Evans M C 2021 Conservation needs to break free from global priority mapping *Nat. Ecol. Evol.* 5 1322–4
44. Wyborn, Carina, Elena Louder, Mike Harfoot, and Samantha Hill. Engaging with the science and politics of biodiversity futures: a literature review. *Environmental Conservation* 48, 1 (2021): 8-15.
45. Carroll, S.R., Garba, I., Figueroa-Rodríguez, O.L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R. and Sara, R., 2020. The CARE principles for indigenous data governance. Data science journal, 19(1).