ENVIRONMENTAL RESEARCH
LETTERS

PERSPECTIVE • OPEN ACCESS

The conservation metadata gap: why AI classification is a symptom, not a solution

To cite this article: Chris McCarthy *et al* 2026 *Environ. Res. Lett.* **21** 031001

IOP Publishing Environ. Res. Lett. 21 (2026) 031001 https://doi.org/10.1088/1748-9326/ae3335

ENVIRONMENTAL RESEARCH
LETTERS

Check for updates

OPEN ACCESS

RECEIVED
26 September 2025

REVISED
10 December 2025

ACCEPTED FOR PUBLICATION
5 January 2026

PUBLISHED
30 January 2026

PERSPECTIVE

The conservation metadata gap: why AI classification is a symptom, not a solution

Chris McCarthy[1,2,*], **Troy Sternberg**[3] and **Cassandra Brooks**[4,5]

[1] Zanvyl Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, MD 21218, United States of America
[2] American Association for the Advancement of Science, Science and Technology Policy Fellowship, Washington, DC 20005, United States of America
[3] CEI Centre for International Studies ISCTE—University Institute Lisbon, Avenida das Forças Armadas, 1649 Lisbon, Portugal
[4] Department of Environmental Studies, University of Colorado Boulder, Boulder, CO 80303, United States of America
[5] Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, CO 80303, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: cmccar27@jh.edu and cassandra.brooks@colorado.edu

Keywords: conservation metadata, evidence synthesis, policy frameworks, scientific publishing, artificial intelligence

Conservation science needs structured metadata captured at submission, not reconstructed afterward by artificial intelligence (AI). Each year, thousands of studies are published that could inform decisions under the United Nations Sustainable Development Goals (SDGs), the Kunming–Montreal Global Biodiversity Framework, the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR), and National Biodiversity Strategies and Action Plans (NBSAPs). Authors know their study species, locations, methods, and often their work's policy relevance, yet this information remains buried in article text rather than searchable metadata. While AI classification tools accelerate evidence synthesis compared to manual efforts, they attempt to extract this information post-publication, turning a simple data entry task into a complex natural language processing challenge.

Consider a researcher studying Antarctic krill (*Euphausia superba*) population dynamics in the Southern Ocean. Depending on where their study takes place, their work might help inform CCAMLR's ongoing krill management in Area 48 (Antarctic Peninsula and Scotia Sea), or could be important to the Ross Sea region marine protected area (MPA) research and monitoring in the MPA's krill research zone. Further, almost any work towards marine conservation in the CCAMLR region would contribute to SDG Target 14.2 on marine ecosystem management. None of this precision appears in searchable metadata. Instead, it remains buried in the text sections, perhaps under recommendations in the discussion, or often is missing entirely. While some in the conservation science community may be unfamiliar with specific policy instruments, the discipline has been oriented toward informing management since its founding (Soulé 1985), and many researchers explicitly frame their work around policy relevance.

Major conservation assessments reveal the scale of this problem. The IPBES (2019) Global Assessment required 145 authors processing 15,000 sources over three years (IPBES 2019). Individual systematic reviews require 164–211 days and approximately $141,000 each (Haddaway and Westgate 2019, Michelson and Reuter 2019). Each organization solves the same puzzle separately because no standardized system captures these connections at publication.

The inefficiency might be tolerable if it only wasted time and money, but research integration lags behind emerging threats by years. Capturing policy-aware metadata at submission would eliminate this bottleneck entirely, making evidence discoverable upon publication. This challenge intensifies as new frameworks like the Biodiversity Beyond National Jurisdiction (BBNJ) Agreement add classification dimensions.

1. Limitations of post hoc classification

Publishers, technology companies, and researchers have recognized this problem and deployed various AI powered tools towards finding a solution (Berger-Tal *et al* 2024). Elsevier tags articles with relevant SDGs, Dimensions offers its own SDG classifier, and platforms like OpenAlex use topic modeling to organize literature (Jayabalasingham *et al* 2019). These tools achieve reasonable accuracy for their specific classification schemes (Elsevier 2021). Recent developments explore how large language models can improve evidence retrieval (Iyer *et al* 2025) with ongoing debates about whether AI will ultimately help or hinder literature reviews (Reynolds *et al* 2025).

Our own experience developing SciBERT models for Antarctic conservation research classification demonstrates both the potential and severe limitations of this approach (McCarthy *et al* 2025). While the enhanced model achieved strong performance on conservation objective alignment, this success required extensive expert-annotated training data from weeks of manual classification effort, custom model architecture for multi target classification, and computational resources that many institutions lack.

More fundamentally, these AI approaches fail at the core tasks conservation science requires. Natural language processing struggles particularly with geographic information extraction, as papers routinely mention multiple locations including author institutions, comparative study sites, and regional context, making it difficult for AI to reliably identify where research actually occurred. A paper discussing 'similar patterns were observed in temperate forests of Brazil' leaves AI guessing whether Brazil was a study site, a comparison drawn from literature, or merely contextual information. When conservation implementation requires precise geographic boundaries, this ambiguity becomes unacceptable.

Similar challenges arise with temporal information. AI cannot reliably determine if 'recent decades' means 1990–2010 or 2000–2020, or whether 'historical data' refers to the 1970s or 1870s. Species identification fails when common names vary regionally or papers use outdated taxonomy. Most critically, AI struggles to distinguish actual research findings from literature review content, potentially misclassifying decades old results as new findings.

These AI systems address symptoms while metadata capture would prevent the underlying problem. They attempt to reconstruct metadata after publication that could be captured directly during submission. Each system uses different classification schemes and methods, producing inconsistent results. A paper might be tagged with SDG 14 (Life Below Water) by one system and SDG 13 (Climate Action) by another, depending on algorithmic interpretation. This inconsistency compounds when different organizations rely on different tools, creating multiple incompatible classifications of the same literature. When BBNJ targets emerge, or NBSAPs update every few years, these models require complete retraining on newly annotated datasets, perpetuating the manual classification burden they claim to eliminate.

Unlike standardized metadata, post hoc AI solutions cannot guarantee the precision conservation decision making requires, though they may complement human provided metadata for quality assurance.

2. Costs of manual classification

The true cost of missing metadata extends beyond individual classification efforts. When IPBES, the European Environment Agency, and national agencies each independently classify the same papers, duplication consumes millions that could fund conservation action. Research shows 67% of meta-analyses overlap with other published reviews (Page *et al* 2016), and managing this overlap remains challenging (Lunny *et al* 2021). Recent calls for improving the conservation evidence landscape acknowledge these inefficiencies (O'Connell *et al* 2024), yet proposed solutions focus on better synthesis methods rather than preventing the problem at its source.

The geographic metadata gap creates particularly severe problems. Without precise location data, from coordinates to watershed boundaries to administrative regions, research cannot be effectively applied to local management. A forest restoration study might be relevant to dozens of protected areas, but managers cannot identify this relevance when location information is buried in text.

Expert time represents an even greater loss. Senior conservation scientists spend days categorizing papers, work eliminable with basic structured metadata. Early career researchers dedicate months to literature reviews that still miss relevant studies because keyword searches fail to capture semantic relationships, while practitioners struggle to stay current because discovery tools fail to surface applicable work (Haddaway *et al* 2020).

The inequity deserves attention. Well-funded institutions can afford dedicated classification staff; universities in biodiversity hotspots lack such resources, meaning research from the most biodiverse regions faces the highest barriers to reaching decision-makers. Language barriers compound this challenge, with non-English research often excluded despite its importance (Amano and Berdejo-Espinola 2025, Hannah *et al* 2025).

3. A standardized metadata approach

The solution does not require new technology but rather coordinated use of existing tools. Other scientific fields have solved similar problems through standardized metadata captured at publication. GenBank transformed genomic research by requiring structured data submission (One Thousand Plant Transcriptomes Initiative 2019). ClinicalTrials.gov changed medical research transparency through mandatory registration with standardized fields (Zarin *et al* 2011). The FAIR principles (findable, accessible, interoperable, reusable) have guided data management across disciplines, Dublin Core provides a foundation for metadata interoperability,

Table 1. COMPASS (Conservation Objectives Metadata for Policy Alignment and Systematic Search) v1.0.

Field	What it encodes	Source/ID system	Example
Framework_alignment	Targets/objectives addressed	Hierarchical policy framework identifiers (to be standardized)	CCAMLR, SDG Target 14.2, CBD Target 3, Australia Marine Parks Management Plan 2020, Section 3.1
Geo_scope	Study location(s), multiple entries supported	GADM (terrestrial); Marine Regions (marine); coordinates	GADM:BRA.1.1 (Acre, Brazil), MRGID:1901 (Ross Sea), −77.5°S 166.0°E
Taxon_scope	Species/assemblage studied	GBIF taxonKey/WoRMS AphiaID	GBIF:2441 176 (Euphausia superba)
Temporal_coverage	When data were collected	ISO 8601 date range	1998-01-01/2003-12-31
Method	Research approach	Controlled vocabulary	Field study, laboratory experiment, simulation model, literature review, intervention trial

Darwin Core enables standardized biodiversity data sharing, and infrastructure directives like INSPIRE have demonstrated how policy mandates can drive metadata standardization across jurisdictions (Craglia and Annoni 2006, Wilkinson *et al* 2016). Recent implementations demonstrate the feasibility of standardized approaches, from community-centric reporting formats in Earth sciences (Crystal-Ornelas *et al* 2022) to FAIR-enabled biodiversity monitoring systems (Menger *et al* 2024). While conservation research presents unique challenges with diverse study designs and complex socio ecological systems, these examples demonstrate that standardized metadata can transform discovery.

Conservation science needs equivalent infrastructure. We propose that journals add five essential metadata fields during submission, what we term COMPASS (Conservation Objectives Metadata for Policy Alignment and Systematic Search). Unlike existing dataset-level standards like Darwin Core (for biodiversity occurrence data) or (Ecological Metadata Language, for dataset documentation), COMPASS operates at the article submission level, specifically linking research to policy frameworks (table 1). COMPASS could accommodate both international agreements and national policy instruments through hierarchical framework identifiers, allowing papers to indicate relevance at multiple governance levels. For example, a study could align with both SDG Target 14.2 and Australia's Marine Parks Management Plan, or both CBD Target 3 and Brazil's National Biodiversity Strategy.

These five fields represent the minimum structured information necessary for policy-relevant discovery while remaining feasible for author completion during submission. Each addresses a specific gap: framework_alignment enables direct linkage to policy targets that keyword searches cannot reliably identify; geo_scope provides machine-readable boundaries rather than ambiguous place names; taxon_scope resolves taxonomic synonymy; temporal_coverage distinguishes data currency from publication date; and method supports filtering by research approach. Together, they complement rather than duplicate existing standards: Darwin Core captures specimen-level occurrence data, while COMPASS captures article-level policy alignment.

COMPASS leverages existing infrastructure rather than creating new standards. Geographic identifiers from Marine Regions (the standard gazetteer for marine place names and boundaries) and GADM (Global Administrative Areas database for terrestrial jurisdictions) provide machine-readable boundaries enabling local managers to find research relevant to their specific conservation areas. For studies spanning ecological gradients, migration routes, or discontinuous sites, multiple geo_scope entries accommodate complex spatial patterns. The GBIF (Global Biodiversity Information Facility) taxonomic backbone and World Register of Marine Species (WoRMS) identifiers automatically handle synonyms while supporting both species and ecosystem classifications. Temporal coverage separates study period from publication date using ISO 8601 (the international standard for date notation), critical for understanding data currency.

A 2025 paper analyzing 1990s population data would be clearly distinguished from one with current observations. The system must accommodate evolution; conservation frameworks change, new SDG targets will emerge post-2030, and management plans update regularly. COMPASS would use versioned vocabularies maintained by authoritative bodies such as IUCN and UN, with papers linked to specific framework versions. This hierarchical structure could support frameworks at all governance levels, creating a historical record while enabling forward compatibility.

Existing literature presents a parallel challenge. While COMPASS cannot retroactively fix millions of published papers, AI systems could verify author-provided metadata and assist with historical classification using COMPASS-compliant papers as training data. High-priority collections like IPBES source documents could receive expert annotation, creating reference datasets for broader application. This positions AI as a quality assurance tool rather than primary solution.

Major scholarly databases including Web of Science, Scopus, Google Scholar, and others already index structured metadata when available. These platforms collectively serve millions of researchers across disciplines and have demonstrated capacity to incorporate new metadata standards, as seen with ORCID integration and data availability statements. Technical implementation would leverage existing infrastructure: (Journal Article Tag Suite), the standard XML format for scholarly publishing metadata; Crossref for DOI registration and distribution, including their established funder registry; and DataCite's geolocation fields for spatial data.

4. From concept to practice

Success requires aligning incentives across the conservation research ecosystem. Journals could implement COMPASS gradually, starting with optional fields that demonstrate value before becoming standard. Success depends on intuitive interfaces with dropdown menus and autocomplete that minimize administrative burden. Initial resistance is addressable through demonstrated benefits like increased citation rates and policy applicability. Quality control would occur during peer review, with reviewers verifying framework alignments just as they verify methodological claims.

Research funders hold the key leverage point. Major conservation funders like NSF's biological sciences programs, EU LIFE programme, and private foundations such as Moore, MacArthur, and Packard could require COMPASS compliance in final reports. Once researchers know they must eventually provide this information, they will pressure journals to support it during submission. This approach mirrors successful adoption patterns for data availability statements and ORCID identifiers.

The framework_alignment field directly supports emerging policy reporting requirements. CBD Target 21 calls for ensuring biodiversity data, information, and knowledge are accessible to decision-makers and the public; COMPASS-tagged literature would be immediately discoverable for national reporting. Similarly, SDG indicator 14.c.1 tracks implementation of ocean-related international law; research tagged with relevant CCAMLR measures or BBNJ provisions would surface automatically for progress assessments. This alignment transforms COMPASS from an academic convenience into policy infrastructure.

5. Moving forward

The metadata gap in conservation publishing represents a solvable problem that continues causing massive inefficiency due to lack of coordination. Every year of delay means thousands more conservation relevant papers published without proper metadata, adding to the already overwhelming backlog requiring classification. The conservation community has achieved comparable coordination before: CITES standardized wildlife trade reporting across 184 parties, and the CBD established common biodiversity assessment frameworks globally. Standardizing five metadata fields should be achievable. Conservation journals should partner with society publications for pilots, implementing optional COMPASS fields to demonstrate value. Research funders should add COMPASS requirements to 2026 grant reporting guidelines while supporting vocabulary maintenance. Most importantly, researchers themselves must advocate for these changes at editorial boards and professional societies.

The question is not whether to implement standardized metadata but how quickly we can coordinate action while maintaining scientific rigor. COMPASS offers a pragmatic starting point for essential infrastructure conservation science urgently needs.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

The authors thank Dr Andrew Titmus (Office of Polar Programs, National Science Foundation) for his valuable contributions to the conceptualization and development of this research.

Author contributions

Chris McCarthy 0000-0001-5563-3583
Conceptualization (equal), Formal analysis (equal), Investigation (equal), Methodology (equal), Project administration (equal), Writing – original draft (equal), Writing – review & editing (equal)

Troy Sternberg 0009-0001-9230-8461
Conceptualization (equal), Formal analysis (equal), Investigation (equal), Methodology (equal), Project administration (equal), Writing – original draft (equal), Writing – review & editing (equal)

Cassandra Brooks
Conceptualization (equal), Writing – original draft (equal)

References

Amano T and Berdejo-Espinola V 2025 Language barriers in conservation: consequences and solutions *Trends Ecol. Evol.* **40** 273–85

Berger-Tal O *et al* 2024 Leveraging AI to improve evidence synthesis in conservation *Trends Ecol. Evol.* **39** 548–57

Craglia M and Annoni A 2006 INSPIRE: an innovative approach to the development of spatial data infrastructures in Europe *Conf. Proc.* pp 6–10

Crystal-Ornelas R *et al* 2022 Enabling FAIR data in Earth and environmental science with community-centric (meta)data reporting formats *Sci. Data* **9** 700

Elsevier 2021 SDG research mapping initiative

Haddaway N R, Bethel A, Dicks L V, Koricheva J, Macura B, Petrokofsky G, Pullin A S, Savilaakso S and Stewart G B 2020 Eight problems with literature reviews and how to fix them *Nat. Ecol. Evol.* **4** 1582–9

Haddaway N R and Westgate M J 2019 Predicting the time needed for environmental systematic reviews and systematic maps *Conserv. Biol.* **33** 434–43

Hannah K, Fuller R A, Smith R K, Sutherland W J and Amano T 2025 Language barriers in conservation science citation networks *Conserv. Biol.* **39** e70051

IPBES 2019 Global assessment report on biodiversity and ecosystem services of the intergovernmental science-policy platform on biodiversity and ecosystem services (Zenodo) (https://doi.org/10.5281/ZENODO.3831673)

Iyer R, Christie A P, Madhavapeddy A, Reynolds S, Sutherland W, Jaffer S and Carrasco-Farré C 2025 Careful design of large language model pipelines enables expert-level retrieval of evidence-based information from syntheses and databases *PLoS One* **20** e0323563

Jayabalasingham B, Boverhof R, Agnew K and Klein L 2019 Identifying research supporting the United Nations sustainable development goals *Mendeley Data, v1* (https://doi.org/10.17632/87txkw7khs.1)

Lunny C, Pieper D, Thabet P and Kanji S 2021 Managing overlap of primary study results across systematic reviews: practical considerations for authors of overviews of reviews *BMC Med. Res. Methodol.* **21** 140

McCarthy C, Brooks C M, Sternberg T, Shaney K and Hoshino B 2025 Where are the research gaps? Ai-assisted multi-target classification for research-policy alignment in conservation science (https://doi.org/10.2139/ssrn.5390788)

Menger J, Magagna B, Henle K, Harpke A, Frenzel M, Rick J, Wiltshire K and Grimm-Seyfarth A 2024 FAIR-EuMon: a FAIR-enabling resource for biodiversity monitoring schemes *Biodivers. Data J.* **12** e125132

Michelson M and Reuter K 2019 The significant cost of systematic reviews and meta-analyses: a call for greater involvement of machine learning to assess the promise of clinical trials *Contemp. Clin. Trials Commun.* **16** 100443

O'Connell M J *et al* 2024 A vision for the future conservation evidence landscape *Ecol. Solut. Evidence* **5** e12397

One Thousand Plant Transcriptomes Initiative 2019 One thousand plant transcriptomes and the phylogenomics of green plants *Nature* **574** 679–85

Page M J *et al* 2016 Epidemiology and reporting characteristics of systematic reviews of biomedical research: a cross-sectional study *PLoS Med.* **13** e1002028

Reynolds S A, Christie A P, Dicks L V, Jaffer S, Madhavapeddy A, Smith R K and Sutherland W J 2025 Will AI speed up literature reviews or derail them entirely? *Nature* **643** 329–31

Soulé M E 1985 What is conservation biology? *BioScience* **35** 727–34

Wilkinson M D *et al* 2016 The FAIR guiding principles for scientific data management and stewardship *Sci. Data* **3** 160018

Zarin D A, Tse T, Williams R J, Califf R M and Ide N C 2011 The clinicaltrials.gov results database—update and key issues *New Engl. J. Med.* **364** 852–60