https://doi.org/10.1038/s44185-026-00121-0

A global biodiversity use data infrastructure acknowledging indigenous and local knowledge

Cristiane Julião Pankararu[1,2,47], Irene Teixidor-Toneu[3,47], Guillaume Odonne[4,47], Frederick Asante[5], Salomão O. Bandeira[6], Ángela Marcela Barrera-Bello[7], Francisco Javier Benitez-Capistros[8], Farid Dahdouh-Guebas[9,10,11], Eduardo Dalcin[12], Zoë Honoria Dennehy-Carr[13], Khady Diallo[14], Huran Tonalli Drouet-Cruz[15], Viviane Stern Fonseca-Kruel[16], Sandrine Gallois[17], Setondé Constant Gnansounou[18,19], Amina Juma Hamza[20], Jean Hugé[21,22], Fiona M. Jordan[23], Raivo Kalle[24], Nabiul Islam Khan[25], Inge Kuijper[26], Carolina Levis[27], Adriana S. Lima[28], Giulia Mattalia[3,29], William Milliken[30], Cosmas N. Munga[31], Nemer E. Narchi[32], Magdalene N. Ngeve[33], Samuel Appiah Ofori[14], Shyam S. Phartyal[34], Nivaldo Peroni[35], Samuel Pironon[30,36,37], Jaime Polanía[38], Julia Prakofjewa[39], Marciano Toledo Silva[40,41], Renata Sõukand[39], Michael B. Thomas[42], Tiziana Ulian[43], Yadav Uprety[44], Ina Vandebroek[45], Arimatéa C. Ximenes[46], Sofia Zank[35] & Natalia Hanazaki[35,47] ✉

Many global biodiversity datasets overlook or misrepresent the knowledge of Indigenous Peoples, Local Communities, and Afro-Descendants (IPLCAD). We propose minimum data and metadata standards for a global data infrastructure on biodiversity knowledge and use, co-designed with IPLCAD, including information on language, community attribution and consent, to ensure data traceability and ethical use. This initiative integrates ancestral and academic sciences to advance inclusive biodiversity governance, addressing historical inequities for global sustainability.

The urgent need for respectful, ethical, and inclusive data management is widely acknowledged[1–4]. Here we aim to highlight the challenges of improving visibility of biological species' use datasets, as well as strengthening collaboration with IPLCAD– i.e., Indigenous Peoples, Local Communities and Afro-Descendants, hereafter IPLCAD, see Box 1 for definitions. Indigenous and local knowledge (ILK) systems are expressions of traditional and ancestral sciences (Box 1) recognized as crucial for tackling complex, interconnected global challenges and harnessing biodiversity's potential to drive sustainable development[5,6], https://www.wipo.int/treaties/en/ip/gratk/]. ILK about the uses of biodiversity is rarely restricted to a single cultural group, as sharing knowledge is a fundamental aspect of cultural dynamics[7]. Species can be used and valued by multiple groups in similar or different ways. Ecological, technological, and utilitarian ILK (along with other aspects of knowledge about biodiversity) have been passed down within and across groups for centuries and travel alongside people during migration events, which are pervasive throughout human history[8]. Global datasets of ILK regarding biodiversity hold undeniable value for analyzing trends and impacts of these knowledge systems worldwide, by upscaling local evidence relevant to conservation, ecology, and bioeconomy, as well as understanding the historical ecology of social-ecological systems[9,10].

The importance of global datasets has grown significantly with the rise of big data and powerful processing tools and technologies, such as artificial intelligence, and the increasing push to make publicly held data more accessible[1,11]. However, although these datasets may incorporate information from various historical and contemporary sources[12,13], challenges related to data quality, interoperability, and completeness remain[14]. These data often rely on original sources that did not provide information from Indigenous Peoples and other custodians of ancestral sciences–i.e., IPLCAD —or even neglect the original data's intentionality and context. For example, specific contexts of use and knowledge custodians are frequently not associated with information about the uses of biodiversity[12,15]. Thus, the paradox of scarcity and abundance in IPLCAD's data[16] arises from the limited availability of information that respects and aligns with community rights and interests, contrasted with a vast amount of data on biodiversity uses rooted in ILK systems that are mislabeled, unacknowledged, or held by non-Indigenous entities–such as information on species uses that fails to credit ancestral sciences or ILK, even when derived from them[16].

A full list of affiliations appears at the end of the paper. ✉e-mail: natalia.hanazaki@ufsc.br

Box 1. | Definitions adopted for indigenous peoples, local communities and Afro-descendants and indigenous knowledge

Indigenous Peoples, Local Communities and Afro-Descendants (IPLCAD) include holders, protectors, guardians, and custodians of ancestral sciences and traditional knowledge and wisdom associated with biodiversity. These groups are not homogeneous, and their definition is constantly evolving according to national policies and international concerns. The 16th Conference of the Parties to the Convention on Biological Diversity (CBD) in 2024 explicitly includes Afro-descendants in official CBD documents, formally acknowledging their crucial role in biodiversity conservation and the value of their traditional knowledge. In a broad sense, IPLCAD includes communities, local groups, and nations who self-identify as Indigenous to the territories they occupy (or occupied), and whose organization is based wholly or partially on their own customs, traditions, and laws, with historical continuity[53]. Cultural diversity must be considered for the representation and inclusion of their knowledge systems in decision-making, particularly those shared by multiple groups. Without ignoring the depth of these debates, for simplicity, this article will use IPLCAD to refer to this diverse group.

Indigenous and local knowledge (ILK) systems are cumulative bodies of knowledge, practices, and beliefs, evolving through adaptive processes and passed across generations through cultural transmission within a particular territory[54]. While nuances exist between traditional, local, and indigenous knowledge, they all possess a core, unifying trait: their adaptive nature. They are dynamic, constantly integrating novel elements and engendering innovations[54,55], expressing ancestral or Indigenous sciences[56,57], or ancestral knowledge systems, in contrast to the academic science of Western tradition–which marginalized and neglected ILK throughout history. Unlike utilitarian or instrumentalist views of indigenous environmental knowledge, there is growing recognition that localized, adaptive, and context-dependent sciences persist globally. Indigenous science evolves in response to shifting environmental, political, economic, and social conditions, extending beyond the notion of 'traditional' knowledge[58].

Over a decade ago, the Nagoya Protocol established international requirements for governing access to genetic resources and associated traditional knowledge, ensuring a fair benefit-sharing system from their use. The Protocol sparked a debate regarding the adequate management of databases related to ILK and sciences[17]. Central to these arguments is the requirement for the mandatory and necessary involvement of IPLCAD in decision-making processes concerning access and benefit-sharing of genetic resources and associated knowledge (e.g., articles 7, 12, 16 of the Nagoya Protocol). This includes the need to ensure effective mechanisms in the context of synthetic biology and digital sequencing of information on genetic resources (DSI[18,19]; and the Kunming-Montreal Global Biodiversity Framework [https://www.cbd.int/gbf]]). A significant step forward was recently made by the new World Intellectual Property Organization Treaty [https://www.wipo.int/treaties/en/ip/gratk/], which requires implementing measures to disclose the origin of genetic resources and traditional knowledge in patent applications from signatory countries[20]. Being able to trace IPLCAD knowledge about uses of species is necessary to implement such a requirement. Traceability of IPLCAD knowledge associated with species is essential for ensuring fair and equitable benefit-sharing[17,21]. This would address gaps in the access and benefit-sharing system established by the Nagoya Protocol. The protocol does not specify the need to consider holders and custodians of traditional knowledge in research reporting, however, this information is critical for traceability. Traceability is also a requirement for targets 13 (fairly share the benefits from genetic resources, digital sequence information, and traditional knowledge), 21 (make knowledge readily available to guide biodiversity action), and 22 (guarantee participation, justice, and information access for all regarding biodiversity) of the Kunming-Montreal Global Biodiversity Framework [https://www.cbd.int/gbf]. Co-creating databases and other information systems on genetic resources and associated knowledge with knowledge holders and guardians, in line with national laws, would represent major progress, according to the WIPO.

Within the context of open data and open science, IPLCAD's control and sovereignty over databases is a crucial step toward recognizing their rights and interests in data related to their knowledge, culture, territories, and livelihoods[2]. Still, institutions–which are still influenced by a colonial legacy–have yet to uphold IPLCAD rights as they apply to digital environments[17].

Many local, regional, and national collectives have already decided to develop databases that aggregate species information and IPLCAD´s knowledge. For example, India has broad expertise in documenting traditional knowledge, drawing on efforts to combat biopiracy and initiatives such as the People´s Biodiversity Register[22], a participatory tool for reporting and managing India's biodiversity and traditional knowledge, and maintaining the Traditional Knowledge Digital Library [22, http://nbaindia.org/content/105/30/1/pbr.html]. The "Portal de Conocimientos Tradicionales de los Pueblos Indígenas" is a database on Peruvian ILK on biodiversity, with limited information available to the general public. This initiative addresses the concern of Indigenous Peoples to maintain the confidentiality of their knowledge, aligning with the *sui generis* Peruvian legislation on access and benefit sharing[23]. At the regional level, in Brazil, the Useflora [https://useflora.ufsc.br/] database will serve as a regional synthesis of data concerning the use, management, and domestication of Neotropical plants. A core requirement of this database is the explicit acknowledgement of the knowledge holders and guardians. This database is being built through participatory governance, involving both IPLCAD researchers and non-IPLCAD researchers and aims to become an instrument to support the recognition and protection of ILK associated with socio-biodiversity and the IPLCAD custodians. The EthnoFlora DB-French Guiana project, developed by the LEEISA Lab, seeks to create a computer storage tool—whose structure and code will soon be offered in open access—for ethnobotanical data sourced from the literature, concerning IPLCAD uses and vernacular names of plants in this territory. Other potential examples include existing databases that compile ILK from published literature and other sources (e.g., ref. 14, https:mackiki.uqat.ca/) or the compilation of archival (folklore) texts (e.g., https://herba.folklore.ee). However, in such cases, the objective of making ILK openly accessible raises concerns in light of the Nagoya Protocol on Access and Benefit Sharing, as well as ongoing debates about the right of IPLCAD to exercise agency over the ownership, use, and control of their knowledge[24]. In addition, these and similar databases should be aligned with existing standards and initiatives – such as Indigenous Metadata Bundle[25], Dublin Core for web resources[26], Plinian Core for aspects of biological species information[27], Darwin Core for biodiversity information[28], or other TDWG (Biodiversity Information Standards) standards – and interoperable with GBIF as a global data infrastructure.

How might a global data infrastructure that respects IPLCAD data sovereignty be structured and implemented?

We propose a decentralized global data infrastructure with interoperable guidelines, protocols and metadata standards agreed upon through inclusive IPLCAD governance at global, regional, national, and local scales (see Table 1 for definitions of data infrastructure dimensions, or characteristics).

Table 1 | Definitions of data infrastructure dimensions

Dimension	Meaning	Proposed aspects to be implemented in a global data infrastructure of IPLCAD biodiversity-associated knowledge.
Governance	Data management systems and practices that enforce data process, ownership, control, and access requirements[17].	Robust participatory and collaborative governance framework, primarily–but not exclusively–at the local, regional, and national scale. This dimension supports CARE principles for inclusive development and innovation, improved governance, and recognition of Indigenous rights and interests.
Metadata	The information that describes the content, context, quality, structure and other characteristics of data, making it easier to find, identify, select, and obtain information[49].	Minimally, holders' or guardians' identity, with controlled vocabulary at the national/regional level and language[35], biological species name, vernacular names, and the Free Prior Informed Consent process. Indigenous Metadata Bundles refer to information about governance, provenance, physical space, protocols, data rights and authority to be recorded alongside data on biodiversity use and knowledge[25]. This dimension adheres to CARE principles about recognition of rights and interests, responsibility for positive relationships, and respect for IPLCAD languages and worldviews.
Accessibility	Refers to the ability of users to access, share, and use the database effectively. For example, academic researchers typically have greater access to databases than individuals from IPLCAD without an academic background. Thus, accessibility is also influenced by socio-cultural factors, intellectual property agreements, institutional reputation, and the perceived value of the database[50].	Interface translated to local languages, if desired. Some parts of the database can have restricted access (see Permissions, below). This dimension adheres to CARE principles for recognition of rights and interests, respect for IPLCAD languages and worldviews, ethics for justice and future use, and collective benefits for equitable outcomes.
Permissions	Refers to the access rights granted to specific contents of a database, which can require written requests following conditions for unrestricted or restricted use.	Profile log-in with varying accessibility, TK (traditional knowledge) labels. This dimension adheres to CARE principles for recognition of rights and interests, and all principles related to governance and authority to control.
Accountability	Accountability involves being responsible for one's actions and being able to explain or justify them to others, or the process of being held accountable for one's actions[51].	Governance framework ensuring transparency, traceability, and inclusivity across all scales (from local/regional to global). This dimension relates, specifically, to CARE principles on authority to control and governance.
Interoperability	The ability of data or tools from non-cooperating resources to integrate or work together with minimal effort[30].	Refers to the use of shared metadata standards and data pipelines (sequences for moving data from different sources through interconnected stages, to a given output); and since it must be critically aligned with other databases and data structures, this dimension relates to the majority of CARE principles focused on Authority to Control and Ethics.

This structure refers to the organization and arrangement of data within the database, including how data is stored, accessed, and interconnected, which is essential for efficient retrieval and management[52], and guided by common metadata standards.

Box 2. | Definitions adopted for digital sequencing information, CARE and FAIR

Digital Sequencing Information (DSI) refers to digital biological data associated with or derived from genetic resources, including nucleotide sequences as well as epigenetic, protein, and metabolite data[19,59]. Access to genetic sequences has facilitated research and economic exploitation of biodiversity, and part of this information is based on ILK, but without recognizing IPLCAD´s rights. Multilateral arrangements for using DSI should be enforced to enable fair benefit-sharing among IPLCAD and other stakeholders, thereby strengthening biodiversity conservation and sustainable use[60].

CARE - Collective Benefit, Authority, Responsibility, and Ethics. CARE principles focus on the stewardship of ILK by IPLCAD, advocating for the fair and equitable treatment of these knowledge systems[3,16,17].

FAIR - Findable, Accessible, Interoperable, and Reusable. FAIR principles emphasize the need for data to be discoverable and usable by both humans and machines, ensuring accessibility and long-term utility[3,16,17,26].

To promote the interoperability of existing datasets (and future ones), data aggregators can guide data integration and exchange, ensuring traceability while adhering to local determinations and decisions on data shareability. Legally, it is essential to create a framework that clearly recognizes and protects IPLCAD's knowledge internationally, ensuring that it is not misappropriated or exploited academically or otherwise[29], which is covered by the Nagoya Protocol but still lacks operability. Technically, we need a system that allows for storing and sharing such information in a FAIR - and CARE- compliant manner (see Box 2), ensuring data consistency (e.g., standardization or homogeneity in reporting). Operationalizing both CARE (Collective Benefit, Authority to Control, Responsibility, and Ethics) and FAIR (Findable, Accessible, Interoperable, and Reusable) (see Box 2) principles within existing data infrastructures would ensure that the knowledge of IPLCAD is both respected and utilized effectively, thereby addressing the historical injustices associated with data governance[3]. The CARE principles intend to balance the technical FAIR data management principles with social and ethical considerations. While these principles share common goals and are intended to be complementary, there are notable incongruences in their application, such as the focus on people versus data: CARE prioritizes the ethical and cultural dimensions[16], whereas FAIR emphasizes technical aspects of data management, enhancing machine usability[30].

Operationalizing these principles must be aligned with the complementary principles of the Codes of Ethics of institutions such as the ISE (International Society for Ethnobiology, https://www.ethnobiology.net/code-of-ethics/), the Tkarihwaié:ri (Code of Ethical Conduct to Ensure Respect for the Cultural and Intellectual Heritage of Indigenous and Local Communities Relevant to the Conservation and Sustainable Use of Biological Diversity[31]), and Article 31.1 of the United Nations Declaration on the Rights of Indigenous Peoples[32], which states the right of Indigenous Peoples to maintain, control, protect, and develop their cultural heritage, ILK, and other cultural expressions of ancestral sciences. To accomplish this fair operationalization, IPLCAD must be able to contribute to decision-making on all data access and management aspects, including database architecture

definition and its long-term governance (throughout the data management lifecycle).

Academically and politically, such a data infrastructure can be a valuable step towards decolonization, by supporting IPLCAD´s rights, contributing to data sovereignty if designed in such a way that user groups are both IPLCAD and non-IPLCAD, and including both academics and civil society. Such a data infrastructure could also serve as a foundation for IPLCAD to develop their own research projects, based on their self-identified needs and ideas, with or without the involvement of non-IPLCAD researchers. For example, in the U.S., the National Ecological Observatory Network, a large-scale ecological monitoring initiative, recognizes Indigenous sovereignty in relation to data, Indigenous governance, and promotes more equitable inclusion of Indigenous Peoples in environmental research[17].

Therefore, a global data infrastructure for the knowledge and uses of biodiversity is more likely to be seen as a network of regional, national, and local databases, with each adhering to shared guidelines and common standards. A good model already exists for cross-linguistic and cross-cultural datasets, e.g., the Cross Linguistic Data Format (ref. 33, https://cldf.clld.org/), which allows researchers to link datasets and analyze them together with the same pipelines, while allowing the datasets to stand alone for different purposes (and on different servers, with different access levels). Similarly, a global data infrastructure can have differentiated mechanisms tailored to the local context, such as compliance with national access and benefit sharing legislation and IPLCAD´s rights.

Technically, we suggest a set of data standards including a) minimal taxonomic information (e.g., biological species name); b) vernacular names in the language of the species' user(s) who originally provided the data; and metadata standards comprising the inclusion of: i) the language name in the local language and identified using the alphanumeric Glottocode (ref. 34, https://glottolog.org/); ii) community or IPLCAD identity, at the individual or group level, considering their preferences and supporting self-determination. For this last point, we recommend using the self-declared name of the group in their own language (e.g., autonym or endonym) and non-pejorative exonyms[35]. The use of controlled vocabularies in some of these fields can ensure interoperability between different initiatives. Examples of fields with controlled vocabulary include: biological species names, languages, and the identity of community or IPLCAD (holders) if they wish to disclose this information. When developing the controlled vocabulary for identifying holders, it is essential that it operates at the national and/or regional level. This strategy requires careful consideration of specific national and regional rights legislation related to Access and Benefit Sharing and intellectual property, alongside the cultural and biological similarities shared (if any) across countries in the region.

Whenever possible, it is desirable to include metadata information about Free Prior and Informed Consent regarding the shared data and, when applicable, about biocultural community protocols or other mechanisms for community biodiversity registers. Information about uses (purposes of use), may be included only when appropriate and transparently negotiated with IPLCAD´s knowledge holders and guardians, and their leadership, where applicable.

The definition of these minimum standards is essential for future ethnobiological and biodiversity-related field studies; however, when dealing with secondary or tertiary sources of data (e.g., data already published in scientific papers), we recognize that much data on species uses or occurrence lacks information about the language, community, and original knowledge holders or guardians. These data may have been used as sources for existing databases as mentioned above[13]. When these data are lacking, we suggest including a clear notice that this information was not recorded in the original studies, highlighting gaps that should be corrected in future studies, as suggested by the Local Contexts Labels and Notices system (https://localcontexts.org/labels/traditional-knowledge-labels/). Local, national, and regional databases must establish a protocol for use, reuse, and sharing, along with the necessary permissions. Other information related to this minimal set, such as detailed information about uses (e.g., traits of a given medicinal plant), may be included only when appropriate and transparently negotiated with IPLCAD to define what is relevant. Standards for cultural metadata and provenance of information can be directly embedded into the digital infrastructure of cataloguing, classification, labeling, and content management systems, with metadata fields suitable for such information. One promising tool is the use of Local Contexts Traditional Knowledge (TK) Labelling system, an extra-legal digital mechanism that repositions Indigenous cultural authority and governance over data and collections. TK Labels are digital tags that convey cultural protocols and IPLCAD rights tied to specific data and materials[36]. They support proper attribution, recognition, and control. TK Labels also correct misrepresentations and clarify the cultural significance and origins of collected materials across diverse types of cultural content.

Indigenous governance experts widely recognize the importance of establishing and maintaining mindful and transparent relationships with IPLCAD knowledge holders and guardians throughout the entire process, starting from their involvement in decisions about data structure and database governance[2,3,16,17,21,36–38]. These efforts should bridge the gap between knowledge production and database governance, ensuring that these groups are represented and actively involved whenever possible. By integrating diverse perspectives and promoting co-governance, such partnerships can strengthen the equitable sharing and sustainable use of biodiversity information. It is important to highlight that, since the inclusion of IPLCAD is an essential condition for building such databases, their participation also expands the beneficiary users of these databases beyond non-IPLCAD academics.

A global data infrastructure that follows these suggested minimal technical and data standards, and which assumes the participation and collaboration of IPLCAD in decision making about the architecture of the databases and in all steps of its governance could be scientifically valuable to: a) identify regions where little research has so far been recorded to support the development of future geographical priority areas[39,40]; b) identify regions where research efforts have been done in the past which should be revisited to provide proper attribution of knowledge rights (e.g., data previously collected about species' use with no clear linkage with who were the users, see also[8,15,40]; c) conduct global analyses enhancing our fundamental understanding of nature-culture interactions (e.g., refs. 10,41); and d) identify regions and species important for both biodiversity and cultural diversity conservation and restoration efforts (e.g., ref. 42).

On a global scale, it is essential to develop a strategic action plan that is clear and operational within the framework of the CBD (Convention on Biological Diversity) and accessible across cultures, with provisions for accurate and meaningful translation into multiple languages. This plan must be co-designed through a deeply collaborative and transdisciplinary process, ensuring meaningful participation from diverse stakeholders[43]. While its foundation should be grounded in local realities, aligned with national legislations, and, where applicable, guided by local biocultural protocols[44], it must also maintain a clear focus on the interconnectedness of global challenges, bridging local actions to broader worldwide goals. However, advancing an inclusive approach implies addressing several obstacles, including language barriers, biases that marginalize underrepresented voices, and insufficient institutional support for meaningful inclusivity[44]. These barriers also include access to technology, understanding of legislation, organizational level and support for leadership, and political representation of IPLCAD. Limited interdisciplinary training and awareness among researchers from different fields also emerge as an impending limitation.

Potential benefits and inherent challenges

A global data infrastructure can help achieve social-environmental justice, reduce knowledge inequalities, protect ILK and biodiversity from industrial misappropriation and ensure equitable benefit-sharing and Free, Prior and Informed Consent. Such a tool also supports local reappropriation and rematriation/repatriation efforts to prevent the loss of ILK, notably by helping lawyers and NGOs trace the knowledge back to its cultural origin. Furthermore, it can stimulate the integration of academic and ancestral

sciences and other forms of knowledge in managing natural resources for the future[45].

The debate about a global data infrastructure weaving IPLCAD knowledge of biodiversity also has special considerations for historical data and collections. This multi-layered approach incorporates the protection, reappropriation, and potential rematriation[24,46] of IPLCAD knowledge and data to ensure that it does not disappear or become lost due to top-down systems of knowledge extraction. Historically, ILK has often been extracted, commodified, and used by external entities (governments, corporations, researchers) without the proper consent of IPLCAD communities[16,24]. Reappropriation involves not only regaining access to this knowledge but also the right to manage and govern its use. Databases that allow the recognition of original knowledge holders are a powerful tool to support reappropriation and rematriation, in alignment with the Global Indigenous Data Alliance, which advocates for global awareness about the importance of safeguarding Indigenous data from misuse and ensuring that Indigenous Peoples remain the primary beneficiaries of their own data (ref. 37, GIDA Global Indigenous Data Alliance, https://www.gida-global.org/).

The diverse nature of ILK presents several challenges that need to be further addressed at various levels, including the accurate identification of knowledge custodians and support for community protagonism. A significant challenge in attributing ILK to biodiversity use data lies in accurately identifying its ILK custodians, holders, and guardians. This complexity is inherent due to ILK's communal and intergenerational nature. Even so, another challenge hindering the development of a global ILK data infrastructure on biodiversity uses is the publishing policy in many journals and ethics review committees that contributes to the invisibility of local collaborators and knowledge holders[35], which conflicts with principles of attribution and transparency. Conversely, external recognition of expertise can reshape IPLCAD's community values, within a new system of meaning altered by Western/Euro-Christian moralities, which can have a marked impact on power, knowledge, and wealth (see ref. 47).

An additional challenge is the absence of robust community-level governance over culturally significant species, which can create vulnerability to appropriation and biopiracy. Effective protection against misuse primarily relies on locally grounded biocultural protocols. Where these protocols are lacking, assessing the ethical and equitable adequacy becomes exceptionally challenging. Consequently, the risk of misuse of ILK about biodiversity uses remains a significant concern. A comprehensive approach to improving data quality and ethical practices must therefore encompass both the improvement of ethical data collection efforts and the retrospective incorporation of missing attribution information into existing published works, ensuring a consistent standard across all biodiversity-related data.

To move beyond these challenges, we highlight the need for a dedicated, multi-level working group to ensure the debate's continuity, moving from foundational discussions to concrete implementation and interoperability. Securing the necessary financial, technical, and human resources for long-term engagement is an immediate priority for this global data infrastructure. Crucially, this work must emphasize community data governance mechanisms and include support for existing Indigenous data governance initiatives (e.g., the Indigenous Data Governance Task Group of GBIF).

To genuinely reflect a vision of such a data infrastructure of biodiversity knowledge and use, it is crucial to tackle not only the technical database challenges but also to confront and reshape core aspects of the academic system (e.g., publishing norms, ethical norms). For this, we must recognize the importance of moving beyond the perspectives of researchers, who are mostly academic-science oriented or, using the words of Bispo-dos-Santos[48], "with a Western/Euro-Christian monotheistic background". Rather, this data infrastructure should reflect, from its inception, the participation of transdisciplinary partners with diverse perspectives, including members and researchers from IPLCAD communities, who are guardians and custodians of ancestral and Indigenous sciences. Finally, more important than safeguarding ILK of biodiversity uses in a global data infrastructure—which can help mitigate language and knowledge loss for future generations—is guaranteeing the necessary conditions for these ancestral sciences to be maintained in their original territories.

Data availability

No datasets were generated or analyzed during the current study.

Received: 28 August 2025; Accepted: 13 January 2026;
Published online: 25 February 2026

References

1. Walter, M. et al. Indigenous data sovereignty in the era of big data and open data. *Aust. J. Soc. Issues* **56**, 143–156 (2021).
2. Carroll, S. R., Herczog, E., Hudson, M., Russell, K. & Stall, S. Operationalizing the CARE and FAIR Principles for Indigenous data futures. *Sci. Data* **8**, 108 (2021).
3. Jennings, L. et al. Applying the 'CARE Principles for Indigenous Data Governance' to ecology and biodiversity research. *Nat. Ecol. Evol.* **7**, 1547–1551 (2023).
4. Hart, R. et al. Repositories of biocultural diversity: toward best practices for empowering ethnobotany in digital herbaria. *People Plants Planet* https://doi.org/10.1002/ppp3.70052 (2025).
5. Reyes-García, V. et al. The contributions of Indigenous Peoples and local communities to ecological restoration. *Restor. Ecol.* **27**, 3–8 (2019).
6. Fajardo, P. et al. Aichi Target 18 beyond 2020: mainstreaming Traditional Biodiversity Knowledge in the conservation and sustainable use of marine and coastal ecosystems. *PeerJ* **9**, e9616 (2021).
7. Tareau, M.-A., Bonnefond, A., Palisse, M. & Odonne, G. Phytotherapies in motion: French Guiana as a case study for cross-cultural ethnobotanical hybridization. *J. Ethnobiol. Ethnomed.* **16**, 54 (2020).
8. Fernández-Llamazares, Á et al. Scientists' warning to humanity on threats to indigenous and local knowledge systems. *J. Ethnobiol.* **41**, 144–169 (2021).
9. Fernández-Llamazares, Á et al. The global relevance of locally grounded ethnobiology. *J. Ethnobiol. Ethnomed.* **20**, 53 (2024).
10. Pironon, S. et al. The global distribution of plants used by humans. *Science* **383**, 293–297 (2024).
11. Stepp, J. R. & Thomas, M. B. Managing ethnopharmacological data: Herbaria, relational databases, literature. In *Ethnopharmacology* (eds, Elisabetsky, E. & Etkin, N.) 27–43 (John Wiley & Sons, 2007).
12. Levis, C. et al. Safeguard stewards of biodiversity knowledge. *Science* **385**, 504–506 (2024).
13. Pironon, S. et al. Safeguard stewards of biodiversity knowledge—Response. *Science* **385**, 506–506 (2024).
14. Ningthoujam, S. S., Talukdar, A. D., Potsangbam, K. S. & Choudhury, M. D. Challenges in developing medicinal plant databases for sharing ethnopharmacological knowledge. *J. Ethnopharmacol.* **141**, 9–32 (2012).
15. Zank, S., Hanazaki, N., Gonçalves, M. C., Ferrari, P. A. & de Morais, B. P. Threats and opportunities for sustainable use of medicinal plants in Brazilian Atlantic forest based on the knowledge of indigenous peoples and local communities. In *Medicinal Plants: Biodiversity, Biotechnology and Conservation* 95–113 (Springer Nature Singapore, 2023).
16. Carroll, S. R., Rodriguez-Lonebear, D. & Martinez, A. Indigenous data governance: strategies from United States native nations. *Data Sci. J.* **18**, 1–15 (2019).
17. Jennings, L. et al. Governance of Indigenous data in open earth systems science. *Nat. Commun.* **16**, 572 (2025).
18. Aubry, S. et al. Bringing access and benefit sharing into the digital age. *Plants People Planet* **4**, 5–12 (2022).
19. Halewood, M., Bagley, M., Wyss, M. & Scholz, A. New benefit-sharing principles for digital sequence information. *Science* **382**, 520–522 (2023).

20. Jefferson, D. J. The World Intellectual Property Organization Treaty on genetic resources and traditional knowledge: implications for plant science. *Plants People Planet* **7**, 901–905 (2025).
21. Souza de Lima, A. et al. Proposals of indigenous peoples and local communities from Brazil for multilateral benefit-sharing from digital sequence information. *Nat. Commun.* **15**, 8933 (2024).
22. Gadgil, M. People's biodiversity registers: lessons learnt. *Environ., Dev. Sustain.* **2**, 323–332 (2000).
23. Reyes-Tagle, Y. Indigenous peoples' participation and the creation of databases managed by the intellectual property rights office in Peru. In *Digital Indigenous Cultural* Heritage (eds. Linkola-Aikio, I. A., Keskitalo, P., Ballardini, R. & Sarantou, M.) 207–229 (Springer, 2025).
24. McAlvay, A. C. et al. Ethnobiology phase VI: decolonizing institutions, projects, and scholarship. *J. Ethnobiol.* **41**, 170–191 (2021).
25. Taitingfong, R., Martinez, A., Carroll, S. R., Hudson, M. & Anderson, J. Indigenous Metadata Bundle Communiqué. Collaboratory for Indigenous Data Governance, ENRICH: Equity for Indigenous Research and Innovation Coordinating Hub, and Tikanga in Technology (2023).
26. Baker, T. Libraries, languages of description, and linked data: a Dublin Core perspective. *Libr. Hi Tech.* **30**, 116–133 (2012).
27. Pando, F., Mora-Cross, M. A., Plata, C., Vargas, M. & Martínez-Sagarra, G. Plinian core: a data specification for species pages in the real world. *Biodivers. Inf. Sci. Stand.* **8**, e135064 (2024).
28. Wieczorek, J. et al. Darwin Core: an evolving community-developed biodiversity data standard. *PloS one* **7**, e29715 (2012).
29. Dahdouh-Guebas, F., Ahimbisibwe, J., Van Moll, R. & Koedam, N. Neo-colonial science by the most industrialised upon the least developed countries in peer-reviewed publishing. *Scientometrics* **56**, 329–343 (2003).
30. Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 1–9 (2016).
31. Convention on Biological Diversity. *Tkarihwaié:ri: the Code of Ethical Conduct to Ensure Respect for the Cultural and Intellectual Heritage of Indigenous and Local Communities Relevant to the Conservation and Sustainable Use of Biological Diversity* (Secretariat of the Convention on Biological Diversity, 2011).
32. United Nations. *United Nations Declaration on the Rights of Indigenous Peoples* (United Nations, 2007).
33. Forkel, R. et al. Cross-Linguistic Data Formats, advancing data sharing and re-use in comparative linguistics. *Sci. Data* **5**, 1–10 (2018).
34. Forkel, R. & Hammarström, H. Glottocodes: Identifiers linking families, languages and dialects to comprehensive reference information. *Semantic Web* **13**, 917–924 (2022).
35. Teixidor-Toneu, I. et al. Improving visibility for knowledge holders in ethnobiological and ethnopharmacological publications. *J. Ethnopharmacol.* **355**, part A, 120632 (2026).
36. Anderson, J. & Hudson, M. The biocultural labels initiative: supporting indigenous rights in data derived from genetic resources. *Biodivers. Inf. Sci. Stand.* **4**, e59230 (2020).
37. Prehn, J. et al. Indigenous Data Governance and Universities Communiqué. https://doi.org/10.6084/m9.figshare.24201585 (2023).
38. Garba, I. et al. Indigenous Peoples and research: self-determination in research governance. *Front. Res. Metr. Anal.* **8**, 1272318 (2023).
39. Tydecks, L., Jeschke, J., Wolf, M., Singer, G. & Tockner, K. Spatial and topical imbalances in biodiversity research. *PloS One* **13**, e0199327 (2018).
40. Meyer, C., Kreft, H., Guralnick, R. & Jetz, W. Global priorities for an effective information basis of biodiversity distributions. *Nat. Commun.* **6**, 1–18 (2015).
41. Lei, D. et al. Comparative analysis of four medicinal floras: phylogenetic methods to identify cross-cultural patterns. *Plants People Planet* **2**, 614–626 (2020).
42. Frainer, A. et al. Cultural and linguistic diversities are underappreciated pillars of biodiversity. *Proc. Natl. Acad. Sci.* **117**, 26539–26543 (2020).
43. Teixidor-Toneu, I. et al. Co-conserving Indigenous and local knowledge systems with seeds. *Trends Plant Sci.* **28**, 1370–1378 (2023).
44. Vandebroek, I., West, J., Otero-Walker, K. & Silvestrini, S. M. Fostering greater recognition of Caribbean traditional plant knowledge. *Trends Ecol. Evol.* **39**, 9–12 (2023).
45. Rist, S. & Dahdouh-Guebas, F. Ethnosciences—a step towards the integration of scientific and indigenous forms of knowledge in the management of natural resources for the future. *Environ., Dev. Sustain.* **8**, 467–493 (2006).
46. Hill, K. X. et al. Rematriation and climate justice: Intersections of indigenous health and place. *J. Clim. Change Health* **18**, 100314 (2024).
47. Martínez-Tagüeña, N. & Rentería-Valencia, R. F. Social cohesion and environmental governance among the Comcaac of Northern Mexico. In *Stewardship of Future Drylands and Climate Change in the Global South: Challenges and Opportunities for the Agenda 2030*, 321–335 (Springer International Publishing, 2019).
48. Bispo-dos-Santos, A. We belong to the land. In *Terra: Afro-Indigenous Anthology*, 7–18 (Piseagrama, 2023).
49. Ulrich, H. et al. Understanding the nature of metadata: systematic review. *J. Med. Internet Res.* **24**, e25440 (2020).
50. Douglas, C., Goulding, R., Farris, L. & Atkinson-Grosjean, J. Socio-cultural characteristics of usability of bioinformatics databases and tools. *Interdiscip. Sci. Rev.* **36**, 55–71 (2011).
51. Mulgan, R. 'Accountability': an ever-expanding concept?. *Public Adm.* **78**, 555–573 (2000).
52. Dechter, R. & Pearl, J. Structure identification in relational data. *Artif. Intell.* **58**, 237–270 (1992).
53. Hill, R. et al. Working with indigenous, local and scientific knowledge in assessments of nature and nature's linkages with people. *Curr. Opin. Environ. Sustain.* **43**, 8–20 (2020).
54. Berkes, F. *Sacred Ecology: Traditional Ecological Knowledge and Resource Management*, 4th ed. (Taylor and Francis, 2015).
55. Hanazaki, N. Local and traditional knowledge systems, resistance, and socioenvironmental justice. *J. Ethnobiol. Ethnomed.* **20**, 5 (2024).
56. Kimmerer, R. W. & Artelle, K. A. Time to support Indigenous science. *Science* **383**, 243–243 (2024).
57. Levis, C. et al. Indigenizing conservation science for a sustainable Amazon. *Science* **386**, 1229–1232 (2024).
58. Johnson, J. T. et al. Weaving Indigenous and sustainability sciences to diversify our methods. *Sustain. Sci.* **11**, 1–11 (2016).
59. Brink, M. & van Hintum, T. Practical consequences of digital sequence information (DSI) definitions and access and benefit-sharing scenarios from a plant genebank's perspective. *Plants People Planet* **4**, 23–32 (2022).
60. Scholz, A. H. et al. Multilateral benefit-sharing from digital sequence information will support both science and biodiversity conservation. *Nat. Commun.* **13**, 1086 (2022).

Acknowledgements

C.G.J., A.S.L., M.T.S., and S.Z. are supported by Serrapilheira Institute. G.O. has benefited from grants managed by the Agence Nationale de la Recherche (AIBSI: ANR-22-EXES-0005; CEBA: ANR-10-LABX-25-01). I.T.-T. has received financial support from the French government under the France 2030 investment plan, as part of the Initiative d'Excellence d'Aix (Marseille Université) AMIDEX (AMX-22-CPJ-05) and the IRD ANR-CPJ contract number 402299/00. F.D.-G. and J.H.'s academic and community collaborations have been supported by the EC-funded Erasmus Mundus Joint Master Degree in Tropical Biodiversity and Ecosystems (Contract number 2019–1451) and by the Belgian Federal Science Policy Office (BELSPO, ENVIRONED, BL/03/VT45). K.D. was financed by a development cooperation grant of Université libre de Bruxelles). F.D.-G. and C.N.M.'s academic and community collaborations have been supported by VLIRUOS Teams project SAVE-FISH (Contract number KE2022TEA517A105),

VLIRUOS ICP-Connect Mundus tropicalis (Contract number GO2024ICP222A105), and VLIRUOS Institutional University Cooperation (IUC) AFYA MOJA. S.G. is supported by the Spanish Ministry of Science under Ramón y Cajal research grant (RYC2023-042575-I). R.K. is supported by the Estonian Ministry of Culture (KUM-TA66). S.P. is supported by the Calleva Foundation. MTS received a CAPES-DS research scholarship. I.V.'s research and community collaborations have been supported by the National Geographic Society (Grant HJ-161R-17) and the Mohamed bin Zayed Species Conservation Fund. N.H. is supported by CNPq research productivity scholarship (306789/2022-1). Thanks to IRD (France) for funding the travel of 5 session participants in the 18th Congress of the International Society of Ethnobiology in Marrakech, May 2024.

Author contributions

C.J.P. conceived the original idea for the paper and led the writing effort, reviewed and approved the final manuscript. I.T.-T. conceived the original idea for the paper and led the writing effort, and coordinated the integration of all ideas and arguments, reviewed and approved the final manuscript. G.O. conceived the original idea for the paper and led the writing effort, and coordinated the integration of all ideas and arguments, reviewed and approved the final manuscript. F.A. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. S.O.B. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. Á.M.B.-B. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. F.J.B.-C. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. F.D.-G. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. E.D. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. Z.H.D.-C. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. K.D. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. H.T.D.-C. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. V.S.F.-K. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. S.G. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. S.C.G. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. A.J.H. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. J.H. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. F.M.J. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. R.K. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. M.N.I.K. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. I.K. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. C.L. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. A.S.L. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. G.M. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. W.M. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. C.N.M. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. N.E.N. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. M.N.N. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. S.A.O. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. S.S.P. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. N.P. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. S.P. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. J.Polanía contributed to the development of the concepts and writing, reviewed and approved the final manuscript. J.Prakofjewa contributed to the development of the concepts and writing, reviewed and approved the final manuscript. M.T.S. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. R.S. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. M.B.T. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. T.U. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. Y.U. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. I.V. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. A.C.X. contributed to the development of the concepts and writing, reviewed and approved the final manuscript. S.Z. provided major contributions to the development of the arguments, conceptual framework, and text editing; and reviewed and approved the final manuscript. N.H. conceived the original idea for the paper, led the writing effort, and coordinated the integration of all ideas and arguments, reviewed, approved and submitted the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44185-026-00121-0.

Correspondence and requests for materials should be addressed to Natalia Hanazaki.

[1]PPG Social Anthropology, Museu Nacional, Rio de Janeiro, Brazil. [2]Coordinator of the Sector Chamber of Guardians of Biodiversity (CGen/Ministry of the Environment), Brasília, Brazil. [3]IMBE, Aix Marseille Université, Avignon Université, CNRS, IRD, Marseille, France. [4]LEEISA, CNRS, Université de Guyane, IFREMER, Cayenne, French Guiana. [5]Moore Inclusive Conservation Lab, Department of Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, Vancouver, Canada. [6]Department of Biological Sciences, Eduardo Mondlane University, Maputo, Mozambique. [7]Lands & Livelihoods Lab, Department of Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, Vancouver, Canada. [8]Instituto de Estudios Amazónicos e Insulares & Facultad de Ciencias Agrícolas, Universidad Central del Ecuador, Quito, Ecuador. [9]Systems Ecology and Resource Management (SERM), Department of Organism Biology and Interfaculty Institute of Social-Ecological Transitions, Université Libre de Bruxelles, Bruxelles, Belgium. [10]Systems Ecology and Resource Management, Ecology, Evolution & Genetics (bDIV), Department of Biology, Vrije Universiteit Brussel, Brussels, Belgium. [11]Mangrove Specialist Group (MSG), Species Survival Commission (SSC), International Union for the Conservation of Nature (IUCN), London, UK. [12]Diretoria de Pesquisas, Jardim Botânico do Rio de Janeiro, Rio de Janeiro, Brazil. [13]Florida Museum of Natural History, University of Florida, Gainesville, Fl, USA. [14]Systems Ecology and Resource Management, Department of Organism Biology and Interfaculty Institute of Social-Ecological Transitions, Université Libre de Bruxelles, Bruxelles, Belgium. [15]Systems Ecology and Resource Management Research Unit, Department of Organism Biology and Interfaculty Institute of Social-Ecological Transitions, Université Libre de Bruxelles, Bruxelles, Belgium. [16]Diretoria de Pesquisas, Jardim Botânico do Rio de Janeiro (JBRJ), Rio de Janeiro, Brazil. [17]Institut Botànic de Barcelona (IBB), CSIC-CMCNB, Barcelona, Spain. [18]Laboratoire de Biomathématiques et d'Estimations Forestières, Faculty of Agronomic Sciences, University of Abomey-Calavi, Abomey Calavi, Benin. [19]Department of Geography, University of Namur, Namur, Belgium. [20]Department of Oceanography and Hydrography, Kenya Marine and Fisheries Research Institute, Mombasa, Kenya. [21]Department of Environmental Sciences, Open University of the Netherlands, Heerlen, Netherlands. [22]Marine Biology Research Group, Biology Department, Ghent University, Ghent, Belgium. [23]Department of Anthropology and Archaeology, University of Bristol, Bristol, UK. [24]Estonian Literary Museum, Tartu, Estonia. [25]Forestry and Wood Technology Discipline, Khulna University, Khulna, Bangladesh. [26]Independent ethnobotanist at Green Unfolding, Forest Row, UK. [27]Graduate Program in Ecology, Department of Ecology and Zoology, Universidade Federal de Santa Catarina, Florianópolis, Brazil. [28]Coordenação Nacional das Comunidades Tradicionais Caiçaras (CNCTC), Rede de Povos e Comunidades Tradicionais (PCTs), Peruíbe, Brazil. [29]Laboratori de Botànica–Unitat Associada CSIC, Facultat de Farmàcia i Ciències de l'Alimentació–Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Catalonia, Spain. [30]Royal Botanic Gardens, Kew, Richmond, UK. [31]Department of Chemistry & Biological Sciences, Marine & Fisheries Programme, Technical University of Mombasa, Mombasa, Kenya. [32]Centro de Estudios en Geografía Humana, El Colegio de Michoacán, Zamora, Mexico. [33]Horn Point Laboratory, University of Maryland Center for Environmental Science, Cambridge, MD, USA. [34]Department of Forestry, Mizoram University, Aizawl, India. [35]Universidade Federal de Santa Catarina, Florianópolis, Brazil. [36]School of Biological and Behavioural Sciences, Queen Mary University of London, London, UK. [37]IUCN Sustainable Use and Livelihoods Specialist Group (SULi), London, UK. [38]Departamento de Ciencias Forestales, Universidad Nacional de Colombia Sede Medellín, Medellín, Colombia. [39]Biocultural Diversity Lab, Department of Environmental Sciences, Informatics & Statistics, Università Ca' Foscari Venezia, Venezia, Italy. [40]Movimento dos Pequenos Agricultores (MPA); PPG História das Ciências e das Técnicas e Epistemologia, Universidade Federal do Rio de Janeiro, Janeiro, Brazil. [41]Laboratory of Human Ecology and Ethnobotany (ECOHE), Department of Ecology and Zoology,, Universidade Federal de Santa Catarina, Florianópolis, Brazil. [42]US Fulbright - Center of Excellence in Biodiversity and Natural Resource Management, National Herbarium of Rwanda, University of Rwanda, Venezia, Rwanda. [43]Royal Botanic Gardens, Kew, UK. [44]Central Department of Botany, Tribhuvan University, Kirtipur, Nepal. [45]Caribbean Centre for Research in Bioscience (CCRIB), Faculty of Science and Technology, The University of the West Indies, Mona, Jamaica. [46]Center for International Forestry Research and World Agroforestry Center (CIFOR-ICRAF), Bogor, Indonesia. [47]These authors contributed equally: Cristiane Julião Pankararu, Irene Teixidor-Toneu, Guillaume Odonne, Natalia Hanazaki. ✉e-mail: natalia.hanazaki@ufsc.br